Theoretical Economics 5 (2010), 403–443 1555-7561/20100403

Supermodular mechanism design

Laurent Mathevet
Department of Economics, University of Texas

This paper introduces a mechanism design approach that allows dealing with the multiple equilibrium problem, using mechanisms that are robust to bounded rationality. This approach is a tool for constructing supermodular mechanisms, i.e., mechanisms that induce games with strategic complementarities. In quasilinear environments, I prove that if a social choice function can be implemented by a mechanism that generates bounded strategic substitutes—as opposed to strategic complementarities—then this mechanism can be converted into a supermodular mechanism that implements the social choice function. If the social choice function also satisfies some efficiency criterion, then it admits a supermodular mechanism that balances the budget. Building on these results, I address the multiple equilibrium problem. I provide sufficient conditions for a social choice function to be implementable with a supermodular mechanism whose equilibria are contained in the smallest interval among all supermodular mechanisms. This is followed by conditions for supermodular implementability in unique equilibrium. Finally, I provide a revelation principle for supermodular implementation in environments with general preferences.

Keywords. Implementation, mechanisms, multiple equilibrium problem, learning, strategic complementarities, supermodular games.

JEL classification. C72, D78, D83.

1. Introduction

For the economist designing contracts, taxes, or other institutions, there is a dilemma between the simplicity of a mechanism and the multiple equilibrium problem. On the one hand, simple mechanisms often restrict attention to the "right" equilibria. In public good contexts, for example, most tax schemes overlook inefficient equilibrium situations, provided that one equilibrium outcome is a social optimum. On the other hand,

Laurent Mathevet: lmath@mail.utexas.edu
I am profoundly grateful to Federico Echenique, Matt Jackson, Preston McAfee, and Leeat Yariv for their help and encouragement. Special thanks are due to Rabah Amir for his insightful comments as a discussant at the IESE Conference on Complementarities, and to Morgan Kousser, John Ledyard, Thomas Palfrey, Eran Shmaya, and David Young for helpful advice and conversations. I also wish to thank anonymous referees for thoughtful suggestions as well as Kim Border, Chris Chambers, Paul J. Healy, Bong Chan Koh, Serkan Kucuksenel, Paul Milgrom, and seminar/conference participants at Caltech, Stanford GSB, IESE, UBC, UC Irvine, UT Austin, Carnegie Mellon, Oxford, INSEAD, Autonoma de Barcelona, Pompeu Fabra, LBS, the 5th LGS Conference, and the Game Theory Conference at Stony Brook. The HSS Division at Caltech, Matt Jackson, and Andrea Mattozzi are gratefully acknowledged for financial support. Part of this paper was written while I was visiting Stanford Economics Department, and I am grateful for their hospitality.

DOI: 10.3982/TE604

elaborate mechanisms guarantee that all equilibria yield a desired outcome, but they often are unreasonably complex. Why would agents play the right equilibrium or how could they play an equilibrium of a game they may not understand? The question of equilibrium play in mechanism design combines the multiple equilibrium problem and bounded rationality. If agents are fully rational and knowledgeable, then equilibrium play is not an issue; however, there are many design problems where these assumptions are too strong. Unfortunately, while some mechanisms solve the multiplicity problem, many do not stand up to departures from full rationality.[1]

In this paper, I develop supermodular mechanism design. This approach is a tool for constructing mechanisms that lead the agents to play supermodular games. These mechanisms, which I call "supermodular mechanisms," are both equipped to handle the multiple equilibrium problem, and robust to bounded rationality.

Supermodular mechanisms provide a way to address the multiple equilibrium problem. In these mechanisms, agents' strategies are complements, meaning that an agent wants to take a higher strategy when others do the same. In view of Milgrom and Roberts (1990), supermodular mechanisms have extremal equilibria, and the interval in between gives the amplitude of the multiple equilibrium problem. Using this interval, it becomes possible to minimize the multiplicity problem, to measure it, and sometimes to solve it. This paper describes how to build supermodular mechanisms where this interval and so the multiplicity problem are minimized. It also uncovers a method for approximating and computing this interval, thereby measuring the multiplicity problem. This is particularly useful for measuring the welfare loss from agents playing an unintended equilibrium. Finally, it gives sufficient conditions for equilibrium uniqueness.

Supermodular mechanisms are robust to boundedly rational behaviors. The interval between the extremal equilibria contains all the iteratively undominated strategy profiles, and all the limit points of adaptive and sophisticated learning dynamics (Milgrom and Roberts 1990, 1991).[2] These theoretical properties are corroborated by strong experimental evidence, showing how convergence to the equilibrium is significantly better for supermodular games. Chen and Gazzale (2004) run experiments on a game for which they control supermodularity. They show how convergence in that game is strikingly better when it is supermodular. Healy (2006) tests five public goods mechanisms and observes that subjects learn to play the equilibrium only in those mechanisms that induce a supermodular game.[3] As such, supermodular mechanisms contribute to fill the gap in the literature emphasized by Jackson (2001): "Issues such as how well various mechanisms perform when players are not at equilibrium but learning or adjusting are

[1]Theoretical works by Muench and Walker (1983), Cabrales (1999), and Cabrales and Ponti (2000) show that learning and stability may be serious issues in the Groves–Ledyard (1977), Abreu–Matsushima (1994), and Sjöström (1994), mechanisms. Healy (2006) and Chen and Tang (1998) provide experimental evidence that convergence of learning dynamics may fail in various mechanisms, such as Proportional Tax or the paired-difference mechanism.

[2]Vives (1990) reports a related result for learning à la Cournot.

[3]Experiments on the Groves–Ledyard mechanism establish that convergence is far better when some parameter is high, values at which the mechanism is supermodular (Chen and Plott 1996, Chen and Tang 1998).

quite important [...] and yet have not even been touched by implementation theory. [This topic] has not been looked at from the perspective of designing mechanisms to have nice learning or dynamic properties." There are many examples where supermodular mechanisms could be used to approach an objective through iterations: a principal designing supermodular contracts to approach revenue maximization, a government applying a supermodular tax system to approach the efficient public goods level, the traffic authorities setting up toll systems (Sandholm 2002, 2005) to minimize congestion, and so forth.[4]

Supermodular games are also attractive in an implementation framework, because their mixed strategy equilibria are locally unstable under monotone adaptive dynamics, such as Cournot dynamics and fictitious play (Echenique and Edlin 2004). Ruling out mixed strategy equilibria is common in implementation theory and often arbitrary, but it is sensible in supermodular implementation. To the contrary, many pure-strategy equilibria are stable. In a parameterized supermodular game, all equilibria that are increasing in the parameter are stable, such as the extremal equilibria (Echenique 2000).

In quasilinear environments, I develop the theory of supermodular mechanism design along three types of results. First, I show that many social choice functions can be implemented with a supermodular mechanism. Essentially, all social choice functions for which strategic substitutability—as opposed to strategic complementarity—is bounded are supermodular implementable; this class includes all twice-continuously differentiable social choice functions and all social choice functions on finite type spaces. The result is established by turning an existing mechanism into one that induces a supermodular game. The transformation technique is constructive and simple, yet powerful. I explain it in the next section in a public goods example. A function is added to each agent's transfer. These functions turn the agents' announcements into complements and they vanish in expectation; thus the original equilibrium remains after adding the functions, and implementability prevails.

Second, I prove that (under bounded substitutes) if there are at least three agents and if the social choice function satisfies some efficiency criterion, then it is implementable by a supermodular mechanism that balances budget. Budget balancing requires that there be no transfers into or out of the system, which is important for full efficiency. Achieving budget balancing is difficult under dominant-strategy implementation (Green and Laffont 1979), but possible under Bayesian implementation (Arrow 1979, d'Aspremont and Gérard-Varet 1979). When there are three or more agents, balancing transfers under supermodular implementation is nearly as general as Bayesian implementation allows.

Interestingly, there are cases where dominant-strategy implementation cannot balance the budget, whereas it is possible to balance the budget and induce a supermodular game with a unique equilibrium.

Third, I deal with the multiple equilibrium problem. I show that if a social choice function satisfies some basic smoothness properties, then it is implementable with a supermodular mechanism whose equilibria are contained in the smallest interval among

[4]Other examples include a procurement department running an auction to allocate jobs, a group of scientists creating a control system for planetary exploration vehicles (see Parkes 2004, Tumer and Wolpert 2004 for issues related to cognitive intelligence).

all supermodular mechanisms. This result applies to all social choice functions that are twice-differentiable and depend on types through an aggregate. The proof relies on designing a supermodular mechanism that generates the weakest complementarities. The interval between the extremal equilibria will be shown to decrease with the amount of complementarities, hence such a mechanism produces the tightest interval. The main interest of a "small" interval is that it leads agents to play a profile whose outcome is close to the desired outcome. Furthermore, I give sufficient conditions for equilibrium uniqueness. The game is thus dominance-solvable and all learning dynamics converge to the equilibrium. As a by-product, it implies coalition-proof Nash implementation by Milgrom and Roberts (1996). Finally, this paper offers an explicit method for computing bounds on the equilibrium set, which is useful when the uniqueness conditions do not hold. This allows us to measure how serious is the multiple equilibrium.

The theory applies broadly to well known models with quasilinear preferences: Public goods (Section 3), principal multiagents (Section 8.1), auctions, and bargaining (Myerson and Satterthwaite 1983, Section 8.2). For example, I present an application to team production where a principal contracts with two agents to maximize net profits. The social choice function is uniquely supermodular implementable. This paper also considers (in Section 7) approximate implementation to extend the applicability of the framework. The objective becomes to supermodularly implement social choice functions that are arbitrarily close to the "target social choice function." It turns out that most bounded social choice functions admit nearby social choice functions that are supermodular implementable.

Finally, this paper provides a revelation principal for supermodular mechanism design. It says that if there is a mechanism that supermodularly implements a social choice function and if the range of the equilibrium strategies in the desired equilibrium has a particular structure—lattice—then there is a direct mechanism that supermodularly implements that social choice function truthfully. The existence of a revelation principle for supermodular implementation is a particularly relevant question to examine, because the paper limits attention to direct mechanisms and because the space of mechanisms to consider is very large. This allows us to identify situations where direct mechanisms cause no loss of generality and when they do, which restrictions they impose.

The remainder of the paper is organized as follows. The next section is a literature review. Section 3 presents the leading public goods example. Section 4 lays out the framework of supermodular mechanism design. Section 5 contains the main results. Section 8 provides several applications of the theory to traditional models and introduces approximate supermodular implementation. Section 9.1 presents the supermodular revelation principle. Finally, Section 9.2 discusses issues related to dominant strategies and the interpretation of learning in Bayesian games.

2. Literature review

There are a number of papers related to robustness issues in mechanism design such as equilibrium multiplicity or learning. Abreu and Matsushima (1992) show that for any social choice function, there exist arbitrarily close social choice functions that can be implemented by iterative deletion of strictly dominated strategies. Although their result is strong and general,[5] their mechanism remains too complex. As nearby social choice functions get closer to the desired one, the dimension of the message space goes to infinity, and more steps of iterated deletion are necessary to reach the equilibrium. Furthermore, the Abreu–Matsushima mechanism does not perform well in experiments (Sefton and Yavaş 1996). In contrast, this paper studies exact implementation with direct mechanisms for which there is experimental support. Supermodular mechanisms are, at least, a simple way to obtain dominance properties and a valuable alternative to more complex mechanisms in many situations.

Chen (2002) provides one of the first papers explicitly aimed at improving learning in mechanism design. In a complete information environment with quasilinear utilities, she constructs a mechanism that Nash implements Lindahl allocations and induces a supermodular game. My paper introduces supermodular implementation in Bayesian contexts and generalizes Chen's result in incomplete information. Cabrales (1999) and Serrano and Cabrales (2007) study implementation with boundedly rational agents. The learning dynamics they consider require players to strictly randomize over all improvements on past play.[6] This rules out many natural dynamics considered here. There are also general impossibility results related to stability in Nash implementation (Jordan 1986, Kim 1986).

The existing literature on robustness uses methodologies that differ from the methodology of this paper. Most approaches rely on equilibrium concepts, such as Eliaz (1999) and Saijo et al. (2005). Instead, this paper adopts the methodology advocated in Jackson (1992), which restricts the class of admissible mechanisms, given a solution concept. The solution concept adopted in the paper is weak—Bayesian equilibrium—and the focus is on supermodularity. Similarly, Sandholm (2002, 2005) uses implementation in potential games to obtain evolutionary properties of the mechanism.

3. Motivation and intuition

This section provides a simple economic example where the natural mechanism to use—the expected externality mechanism of Arrow (1979) and d'Aspremont and Gérard-Varet (1979)—has poor properties: It is not dominance-solvable and convergence to the equilibrium fails under many dynamics. In contrast, the resulting supermodular mechanism has a unique globally stable equilibrium and is dominance-solvable.

Consider a principal who needs to decide the level of a public good, such as public safety or the time or money spent on some event. Let $X = [0, 2]$ denote the possible values of the public good. There are two agents, 1 and 2, whose type spaces are $\Theta_1 = \Theta_2 \subset [0, 1]$. Types are independently uniformly distributed. The agents' preferences are quasilinear, $u_i(x, \theta_i) = V_i(x, \theta_i) + t_i$, where $x \in X$, $\theta_i \in \Theta_i$, and $t_i \in \mathbb{R}$ is the transfer from

[5] The solution concept is strong enough to predict convergence of many learning dynamics to the unique equilibrium outcome (see, e.g., Milgrom and Roberts 1991). Note, however, that some adaptive dynamics from Milgrom and Roberts (1990) do not converge to a uniquely rationalizable profile.

[6] This feature allows players to exit the integer or modulo game.

the principal to agent i. The valuation functions are $V_1(x, \theta_1) = \theta_1 x - x^2$ and $V_2(x, \theta_2) = \theta_2 x + x^2/2$.

The principal wishes to make an allocation-efficient decision, i.e., she aims to maximize the sum of the valuation functions by choosing $x^*(\theta) = \theta_1 + \theta_2$. To this end, she wants the agents to reveal their true type. She opts for the expected externality mechanism, which gives the transfers[7]

$$t_1(\hat{\theta}_1, \hat{\theta}_2) = \tfrac{1}{2} + \tfrac{1}{2}\hat{\theta}_2 + \hat{\theta}_2^2 + \hat{\theta}_1 + \tfrac{1}{2}\hat{\theta}_1^2, \qquad t_2(\hat{\theta}_1, \hat{\theta}_2) = -t_1(\hat{\theta}_1, \hat{\theta}_2).$$

Given these transfers, agent i's expected utility in the mechanism is

$$E[u_i] = -\tfrac{1}{2}\hat{\theta}_i^2 + (\theta_i + B_i)\hat{\theta}_i, \tag{1}$$

where B_i is the expectation of a function of j's strategy. From (1), i's best response is to announce $\theta_i + B_i$, where B_i can be interpreted as agent i's bias for over- or underreporting. The only strategic variable is the bias, because only biases depend on what the opponent does. It turns out that an agent's bias depends on her opponent's strategy through her bias only:

$$B_1 \approx -2B_2, \qquad B_2 \approx B_1. \tag{2}$$

Agent 2 wants the same bias as agent 1, whereas 1 wants an opposite bias. This game has a flavor of "matching pennies," because one agent essentially tries to match another agent who tries to mismatch.

Although truthtelling is an equilibrium,[8] it is unclear whether and how the agents arrive at that equilibrium. For the sake of argument, consider a game with integer biases and best responses described by (2). Such a game cannot be dominance-solvable. Thus the epistemic conditions that would guarantee that agents play truthfully are too strong for many design settings. Furthermore, a study of stability will reveal that truthtelling also has scant dynamic foundations, because many learning dynamics fail to converge.

Consider a learning model (Fudenberg and Levine 1998) applied to the game induced by the mechanism. Time proceeds in discrete periods. At each period, both agents look at the past history of play, use this to formulate beliefs on the other agent's future strategy, and then announce a best response (fully characterized by a bias) to their beliefs.[9] For simplicity, think of the equivalent situation where agents announce a bias at each period. For each $t \in \mathbb{N}$, a learning rule for agent i, such as fictitious play, takes as input the history of biases (of agent j) from 0 to $t-1$ and produces beliefs about which biases agent j will announce at t. Given these beliefs, agent i (myopically) chooses an optimal bias.

There are many learning rules for which, not only do the agents not converge to truth-revealing, but the play cycles forever. First, this is the case for weighted fictitious

[7]This mechanism allows truthful implementation of allocation-efficient decision rules (see Arrow 1979, d'Aspremont and Gérard-Varet 1979, or Section 23.D in Mas-Colell et al. 1995), i.e., truthtelling is a Bayesian equilibrium of the mechanism.

[8]Note that null biases, $B_1 = B_2 = 0$, are best responses to each other.

[9]Alternatively, we could think of a situation where types are drawn independently every period and agents announce a type rather than a strategy.

play (see, e.g., Ho 2008).[10] This is also the case for Cournot dynamics, where cycling prevails wherever the dynamic starts (except truthtelling); for this dynamic the slightest belief perturbations destabilize the truthful equilibrium. Cycling also plagues many families of dynamics with a longer memory. For example, consider dynamics where players remember the last T periods. They assign a probability π to the strategy played at $t-1$ and assign $(1-\pi)\delta^k/C$ to the strategy played at $t-k$, $k=2,\dots,T$, where C is normalized so that the probabilities add up to 1. Simulations reveal that learning can fail under many values of the parameters. For $T \in \{2,3\}$, $\delta = .9$, and $\pi \geq .5$, the process enters a cycle for many starting points. Increasing the memory size does not always improve learning. For $T=4$, $\delta = .8$, and $\pi \leq .65$, the profile converges to the truthful equilibrium, but it cycles for $\pi \geq .7$. A larger memory does not necessarily improve learning, because cycling reappears when $T = \{5,6\}$, $\delta = .8$, and $\pi \leq .65$. At last, cycling is also an issue for sophisticated learning dynamics (Milgrom and Roberts 1991).

To fix the problem, this paper advocates converting the mechanism into one that induces a supermodular game. In this example, the budget balancing assumption is dropped, but it will be treated in the main text.

The main insight is to use transfers to align agents' biases by creating complementarities between their announcements. Say that from agent i's viewpoint, j reports a large type if j's reported type exceeds the average (truthful) type: $\hat{\theta}_j \geq E_{\theta_j}[\theta_j]$. By appending $\rho_i \hat{\theta}_i(\hat{\theta}_j - E_{\theta_j}[\theta_j])$ with $\rho_i > 0$ to the original transfers,

$$t_i^{\mathrm{SM}}(\hat{\theta}) = E_{\theta_j}[t_i(\hat{\theta}_i, \theta_j)] + \rho_i \hat{\theta}_i(\hat{\theta}_j - E_{\theta_j}[\theta_j]), \tag{3}$$

agent i receives an extra reward for announcing large types when j does so as well ($\hat{\theta}_j \geq E_{\theta_j}[\theta_j]$) and she is taxed if she still reports large types when j does not ($\hat{\theta}_j < E_{\theta_j}[\theta_j]$).

Parameter ρ_i captures the intensity of the punishment or reward. Intuitively, there should be ρ_i large enough such that both agents want to bias their report in the same direction, because the reward (punishment) for (not) conforming to their opponent becomes high, regardless of the original incentives. By doing so, the mechanism has actually become supermodular. Note $\partial^2 t_i^{\mathrm{SM}}(\hat{\theta})/\partial\hat{\theta}_i\,\partial\hat{\theta}_j = \rho_i$. If $\partial^2 V_i(x_i(\hat{\theta}), \theta_i)/\partial\hat{\theta}_i\,\partial\hat{\theta}_j$ is bounded below, a condition called *bounded substitutes*, then there is ρ_i large enough such that

$$\frac{\partial^2 V_i(x_i(\hat{\theta}_i, \hat{\theta}_j), \theta_i)}{\partial\hat{\theta}_i\,\partial\hat{\theta}_j} + \frac{\partial^2 t_i^{\mathrm{SM}}(\hat{\theta}_i, \hat{\theta}_j)}{\partial\hat{\theta}_i\,\partial\hat{\theta}_j} \geq 0 \quad \text{for all } \hat{\theta}, \theta_i, \tag{4}$$

which means that the game is supermodular.[11]

[10] To apply this model here, type spaces need to be finite, so that strategy spaces are also finite. The results are given for types in $\{0, .5, 1\}$. Strategies are initially assigned arbitrary weights, and beliefs are updated each period by multiplying all weights by $1-\pi$, $0 < \pi < 1$, and by adding 1 to the weight of the opponent's strategy that was played last.

[11] If the complete information payoffs define a supermodular game for each $\theta \in \Theta$, then the (ex ante) Bayesian game is supermodular. Loosely speaking, supermodular games are characterized by compact strategy spaces and utility functions whose cross-partial derivatives between own action and others' actions are positive.

Truthtelling was an equilibrium of the original mechanism, and so it must still be an equilibrium under the modified mechanism, because it has left the expected utility functions unchanged: $E_{\theta_j}[t_i^{\mathrm{SM}}(\cdot, \theta_j)] = E_{\theta_j}[t_i(\cdot, \theta_j)]$. Furthermore, there exist ρ_1 and ρ_2 for which this is the unique equilibrium (Theorem 5 of Section 6). Supermodular games with a unique equilibrium are dominance-solvable, most learning dynamics converge to the equilibrium, and the equilibrium is stable.

4. Supermodular mechanism design: The framework

4.1 *The environment*

Consider n agents, each endowed with quasilinear preferences over a set of alternatives. An alternative is a vector $(x, t) = (x_1, \ldots, x_n, t_1, \ldots, t_n)$, where x_i is an element of a compact set $X_i \subset \mathbb{R}^{m_i}$ and $t_i \in \mathbb{R}$ for all i. In this environment, x_i is interpreted as agent i's allocation and t_i is the money transfer i receives. Each agent i has a compact type space $\Theta_i \subset \mathbb{R}$ (finite or infinite) endowed with the usual order and information is incomplete. There is a common prior with density ϕ on Θ known to the mechanism designer. Types are assumed to be independently distributed and ϕ has full support. Each agent i's preferences over alternatives are represented by a bounded utility function $u_i(x_i, t_i, \theta_i) = V_i(x_i, \theta_i) + t_i$, where $V_i : X_i \times \Theta_i \to \mathbb{R}$ is a bounded function called i's valuation.

A mechanism designer wishes to implement an allocation for each realization of types. This objective is represented by a decision rule $x : \Theta \mapsto (x_i(\theta))_i$. To this end, the designer sets up a transfer scheme $t_i : \Theta \to \mathbb{R}$ for each i. A mechanism is denoted by $\Gamma = (\{\Theta_i\}, (x, t))$ and it describes the strategic situation into which agents are put. Agents are asked to announce a type; from the vector of announced types, an allocation and a transfer accrue to each agent.[12] The pair $f = (x, t)$ is called a social choice function. Letting $\hat{\theta}_i$ be i's announced type, i's (ex post) utility function in Γ is $u_i^{\Gamma}(\hat{\theta}, \theta_i) = V_i(x_i(\hat{\theta}), \theta_i) + t_i(\hat{\theta})$. A pure strategy for agent i under incomplete information is a function $\hat{\theta}_i : \Theta_i \to \Theta_i$ that maps true types into announced types. Strategy $\hat{\theta}_i(\cdot)$ is called a deception. Agent i's (ex ante) utility function in Γ is $U_i^{\Gamma}(\hat{\theta}_i(\cdot), \hat{\theta}_{-i}(\cdot)) = E_{\theta}[u^{\Gamma}(\hat{\theta}(\theta), \theta_i)]$.

This paper is concerned with supermodular mechanisms. To define these mechanisms, several definitions are in order. A function $g : \mathbb{R} \times \mathbb{R}^{n-1} \to \mathbb{R}$ such that $g : (y, z) \mapsto g(y, z)$ has increasing (decreasing) differences in (y, z) if, whenever $y \geq y'$ and $z \geq z'$, $g(y, z) - g(y', z) \geq (\leq)\, g(y, z') - g(y', z')$; g satisfies the single-crossing property in (y, z) if, whenever $y \geq y'$ and $z \geq z'$, $g(y, z') \geq g(y', z')$ implies $g(y, z) \geq g(y', z)$ and $g(y, z') > g(y', z')$ implies $g(y, z) > g(y', z)$. If g has decreasing differences in (y, z), then variables y and z are said to be substitutes. If g has increasing differences or satisfies the single-crossing property in (y, z), then y and z are said to be complements.

A game is a tuple $(N, \{S_i, u_i\})$, where N is a finite set of players; each $i \in N$ has a strategy space $S_i \subset \mathbb{R}$ and a payoff function $u_i : \prod_{i \in N} S_i \to \mathbb{R}$. Generic elements of S_i are denoted s_i and $s_{-i} = (s_1, \ldots, s_{i-1}, s_{i+1}, \ldots, s_n)$. Subsets of the real line are endowed with the Euclidean topology.

[12] Most of the paper is concerned with direct mechanisms.

Definition 1. A game $\mathcal{G} = (N, \{S_i, u_i\})$ is supermodular if, for all $i \in N$, the following statements hold.

(i) Space $S_i \subset \mathbb{R}$ is compact.

(ii) Payoff u_i is bounded, and has increasing differences in (s_i, s_{-i}).

(iii) Payoff u_i is upper semicontinuous in s_i for each s_{-i} and continuous in s_{-i} for each s_i.

There are three stages at which it is relevant to formulate the game induced by mechanism Γ: ex ante, interim, and ex post (complete information). Let $\mathcal{G}(\theta) = (N, \{\Theta_i, u_i^\Gamma(\cdot, \theta_i)\})$ be the game induced by mechanism Γ ex post. Let $\mathcal{G} = (N, \{\Theta_i^{\Theta_i}, U_i^\Gamma\})$ be the ex ante Bayesian game induced by Γ. Among these three formulations of the game, the paper considers supermodularity at the ex post level, because this is the strongest requirement. If the ex post game is always supermodular, then the game will be supermodular in its ex ante and interim formulations.

Definition 2. A social choice function $f = (x, t)$ is (truthfully) supermodular implementable if truthtelling, i.e., $\hat{\theta}_i(\theta_i) = \theta_i$ for all i, is a Bayesian equilibrium of $\mathcal{G}$ and if $\mathcal{G}(\theta)$ is supermodular for each θ.

4.2 *Payoff assumptions*

This section contains the main assumptions on payoffs that are used throughout the paper. Most definitions impose conditions on the composition of the valuation functions and the decision rule.

As suggested in Section 3, the intuition behind most results is that transfers should add complementarities to compensate for the substitute effects coming from the valuation functions. To be able to compensate, these substitute effects should be bounded: $\partial^2 V_i(x_i(\hat{\theta}), \theta_i)/\partial\hat{\theta}_i\,\partial\hat{\theta}_j \leq T_i$ for all $\hat{\theta}$ and θ_i. Given that the type sets are compact, this assumption is trivially satisfied if the valuation functions and the decision rule are twice-continuously differentiable.[13]

It will be useful to generalize the previous condition to nondifferentiable environments. For any $\theta_i'', \theta_i', \theta_i \in \Theta_i$ and $\theta_{-i} \in \Theta_{-i}$, let $\Delta V_i(\theta_i'', \theta_i', \theta_{-i}, \theta_i) = V_i(x_i(\theta_i'', \theta_{-i}), \theta_i) - V_i(x_i(\theta_i', \theta_{-i}), \theta_i)$, In general environments, say that the valuation functions and the decision rule have *bounded substitutes* if for each i and θ_i, there is a real number $T_i(\theta_i)$ such that $\Delta V_i(\theta_i'', \theta_i', \theta''_{-i}, \theta_i) - \Delta V_i(\theta_i'', \theta_i', \theta'_{-i}, \theta_i) \geq T_i(\theta_i)(\theta_i'' - \theta_i') \sum_{j\neq i}(\theta_j'' - \theta_j')$ for all $\theta_i'' \geq \theta_i'$, $\theta''_{-i} \geq \theta'_{-i}$, and θ_i. The valuations and the decision rule generate *uniformly* bounded substitutes if the previous inequality admits a uniform lower bound T_i. Note that the assumption of bounded substitutes is always satisfied when type sets are finite.

Unless otherwise stated, the valuation functions and the decision rule are assumed to form a *continuous family* throughout the paper: For all i and θ_i, $V_i(x_i(\hat{\theta}), \theta_i)$ is continuous in $\hat{\theta}_{-i}$ for fixed $\hat{\theta}_i$ and $V_i(x_i(\hat{\theta}), \theta_i)$ is upper semicontinuous in $\hat{\theta}_i$ for fixed $\hat{\theta}_{-i}$.

[13]The valuation functions and the decision rule are (twice-) continuously differentiable if for all i, there exist open sets $O_i \supset \Theta_i$ and $U_i \supset X_i$, such that $V_i : U_i \times O_i \to \mathbb{R}$ and $x_i : \prod_{i\in N} O_i \to U_i$ are (twice-) continuously differentiable.

5. Supermodular implementable social choice functions and budget balancing

This section delineates the set of supermodular implementable social choice functions with and without budget balancing. The paper builds on these results to address the multiple equilibrium problem in the next section.

5.1 *A general result*

The class of supermodular implementable social choice functions constitutes the starting point of the analysis. The first result identifies mild boundedness and continuity conditions under which the class of supermodular implementable social choice functions is the same as that of implementable social choice functions.

Theorem 1. *Suppose the valuation functions and decision rule x generate uniformly bounded substitutes. Assume $E_{\theta_{-i}}[t_i(\cdot, \theta_{-i})]$ is upper semicontinuous.*

(i) There exist transfers t such that $f = (x, t)$ is implementable if and only if there are transfers t^{SM} such that (x, t^{SM}) is supermodular implementable.

(ii) Agents receive the same interim expected utility in equilibrium under (x, t) and (x, t^{SM}), and $E_{\theta_{-i}}[t_i(\cdot, \theta_{-i})] = E_{\theta_{-i}}[t_i^{\mathrm{SM}}(\cdot, \theta_{-i})]$.

According to this theorem, if the decision rule and the utility functions are relatively well behaved in the sense of continuous families and bounded substitutes, then a social choice function is implementable if and only if it can be converted into a supermodular implementable social choice function. So the class of supermodular implementable social choice functions is large. Transfers are at the heart of the result: It is always possible to add complementarities into the transfers without affecting the incentives that appear in the expected value. It is also interesting to note that if the original social choice function f satisfies some ex ante or interim participation constraints, then so does (x, t^{SM}), because agents receive the same interim expected utility in equilibrium.

The assumptions of uniformly bounded substitutes and continuity are generally satisfied. This is the case when type sets are finite and in twice-continuously differentiable environments. This leads to the following important corollaries that cover many cases of interest.

Corollary 1. *Let type spaces be finite. For any valuation functions, if the social choice function $f = (x, t)$ is implementable, then there exist transfers t^{SM} such that (x, t^{SM}) is supermodular implementable.*

Corollary 2. *If $f = (x, t)$ is an implementable social choice function such that the decision rule x and the valuations are twice-continuously differentiable, then there exist transfers t^{SM} such that (x, t^{SM}) is supermodular implementable.*

Most results in the paper take implementable social choice functions as starting points, because there exist well known conditions under which a social choice function is implementable with continuous transfers.[14] The next proposition combines Theorem 1 and those implementability conditions.

Proposition 1. *Let type spaces be compact intervals. Suppose the decision rule x and the valuations generate uniformly bounded substitutes. If $E_{\theta_{-i}}[V_i(x_i(\hat{\theta}_i, \theta_{-i}), \theta_i)]$ is continuous in $(\hat{\theta}_i, \theta_i)$ and $\partial E_{\theta_{-i}}[V_i(x_i(\hat{\theta}_i, \theta_{-i}), \theta_i)]/\partial \theta_i$ is increasing in $\hat{\theta}_i$, then there are transfers t^{SM} such that (x, t^{SM}) is supermodular implementable.*

In smooth environments, supermodular implementable decision rules are those rules that lead each agent i's expected marginal valuation to be nondecreasing.

5.2 *Adding the budget constraint*

This section investigates supermodular implementation under budget balancing. In some design problems, the planner should not realize a net gain from the mechanism. Since the planner cannot sustain deficits, this implies that transfers are balanced: $\sum t_i = 0$. If a social choice function combines an efficient decision with balanced transfers, then it is *fully efficient*. That is, it maximizes the sum of the utility functions subject to feasibility $\sum t_i \leq 0$. The next example shows that complementarities between agents' announcements might be irreconcilable with budget balancing.

Example 1. Consider the public goods example of Section 3. In this example, if there exist transfers $\{t_i^{\mathrm{SM}}(\cdot)\}_{i=1,2}$ such that the resulting social choice function (x, t^{SM}) is supermodular implementable, then inequality (4) must hold for both agents. That is, the cross-partial derivatives of $t_1(\hat{\theta})$ must be greater than 2 and the cross-partial derivatives of $t_2(\hat{\theta})$ must be greater than -1; hence their sum will be strictly greater than 0. The budget balance condition requires $\sum_{i=1,2} t_i(\hat{\theta}) = 0$, so the sum of the cross-partial derivatives of the transfers must be null. As a result, budget balancing must be violated in this example, if there is supermodular implementation. ◊

This example points to the difficulty of balancing the budget in some situations with two players. The next theorem provides sufficient conditions for supermodular implementation with balanced transfers with at least three agents. Say that a decision rule x is *allocation-efficient* if $x(\theta) \in \arg\max_{x \in X} \sum_{i \in N} V_i(x_i, \theta_i)$ for all $\theta \in \Theta$.

Theorem 2. *Let $n \geq 3$. Suppose the valuation functions and the decision rule generate uniformly bounded substitutes. If the decision rule is allocation-efficient, then there are balanced transfers t^{BB} such that (x, t^{BB}) is supermodular implementable.*

Basically, if substitutes are bounded, any allocation-efficient decision rule can be paired with a transfer scheme to give a fully efficient supermodular implementable social choice function. The proof is constructive and appears in Appendix B. Transfers

[14] See, e.g., Proposition 23.D.2 in Mas-Colell et al. (1995) for linear utility functions.

t^{BB} correspond to a transformation of the transfers in the expected externality mechanism (Arrow 1979, d'Aspremont and Gérard-Varet 1979). This transformation is similar to that in Theorem 1, as it adds complementarity between agents' announcements in a pairwise fashion. It then becomes possible to subtract from each individual's transfer those complementarities that come from the other agents' transfers and that do not concern that individual, thus balancing the whole system. Note that this transformation does not use any agent as a "sink" that absorbs the others' deficit or surplus; the budget balancing effort is shared by all. Moreover, not all supermodular transformations would allow balancing the budget.[15]

It is important to keep in mind that these supermodular transformations, while leaving the (interim) expected utility unchanged, add risk into the transfers by creating interdependencies between agents. This is particularly noticeable in the above, because the expected externality mechanism has no interdependencies between agents.

6. Optimal and unique supermodular implementation

This section deals with the multiple equilibrium problem in supermodular implementation. In Section 3, the linear transformation in (3) is one possible way to obtain a supermodular mechanism. In general, among all the possible ways to turn a mechanism into a supermodular mechanism, that is, among all supermodular mechanisms that implement a decision rule, which one has the smallest equilibrium set? This question is important, because even if a mechanism has an equilibrium outcome with some desirable property, it may have other equilibrium outcomes that are undesirable. This is particularly relevant for supermodular implementation, because the truthful equilibrium is the only equilibrium known to have the desired outcome. This section provides a closed form for the supermodular mechanism whose equilibrium set size is minimal. If the size of the equilibrium set, called interval prediction, is small, then agents should play a profile close to truthtelling and to the desired outcome. This section also contains conditions for truth-revealing to be the unique equilibrium—a case in which supermodular implementation is particularly powerful.

6.1 *Optimal implementation*

I begin with an example that explains the foundations of optimal supermodular implementation. This example suggests that the equilibrium set enlarges as complementarities become stronger. This observation will be the centerpiece of the main result, because minimizing the size of the equilibrium set then comes down to minimizing the complementarities, which is more tractable. This key observation warrants a series of definitions and a proposition, which lead to the main result.

Example 2. Consider the public goods example of Section 3. If transfers t_i^{SM}, $i = 1, 2$, are set with $\rho_1 = 2\frac{1}{2}$ and $\rho_2 = -\frac{1}{2}$, the game induced by the mechanism is supermodular

[15] Theorem 2 can be modified to apply to situations where, for every realization of types, enough taxes need to be raised to pay the cost of x. This constraint takes the form $\sum_{i\in N} t_i(\theta) \geq C(x(\theta))$ for all θ (see Ledyard and Palfrey 2007).

and truthtelling is the unique equilibrium. For $\rho_1 = 3\frac{1}{5}$ and $\rho_2 = \frac{1}{2}$, the supermodular induced game now has a smallest and a largest equilibrium. In the smallest equilibrium, agent 1 announces 0 for any type below $c_1 \approx .47$ and announces $\theta_1 - c_1$ for types above; agent 2 announces 0 for any type below $c_2 \approx .55$ and announces $\theta_2 - c_2$ for types above. In the largest equilibrium, agent 1 announces $\theta_1 + c_1$ for any type below $1 - c_1$ and announces 1 for types above; agent 2 announces $\theta_2 + c_2$ for any type below $1 - c_2$ and announces 1 for types above. By increasing ρ_1 to 4 and ρ_2 to 1, the following situation occurs: The smallest equilibrium is now the smallest profile of the entire space (each agent always announces her smallest type), and the largest equilibrium is the largest profile (each agent always announces her largest type). Increasing ρ_1 and ρ_2 has had three negative consequences: (i) By increasing these parameters above $5/2$ and $-1/2$, two new equilibria have been generated. By increasing them more, (ii) the interval prediction has enlarged to become the whole space, and (iii) the truthful equilibrium has become locally unstable. To see this, think of a (symmetric) scenario with an increasing best response function from $[-1, 1]$ into $[-1, 1]$—a mapping from biases into biases—that intersects the 45-degree line three times. The middle intersection occurs at 0 and represents the truthtelling equilibrium. Then the extremal equilibria are stable, but the truthtelling equilibrium is unstable. Most adaptive dynamics starting off the truthful equilibrium move away from it toward an extremal equilibrium. ◊

One way to measure the degree of complementarity between the variables of a function is by looking at its cross-partial derivatives. Large cross-partials mean that the degree of complementarity is high and vice versa. In Example 2, the degree of complementarities of the transfers is ρ_i. Optimal supermodular implementation involves designing a mechanism whose induced supermodular game has the weakest complementarities among all supermodular mechanisms. The rationale behind it is clear from Example 2, as there is evidence that extremal equilibria move apart with the degree of complementarities. Thus, optimal implementation should provide the best compromise between complementarities and equilibrium set size.

As mentioned above, the cross-partial derivatives offer a way to measure complementarities in twice-differentiable environments. It is natural to say that a transfer function $\tilde{t}$ generates larger complementarities than t, denoted $\tilde{t} \succeq_{\mathrm{ID}} t$, if $\partial^2 \tilde{t}_i(\hat{\theta})/\partial\hat{\theta}_i \partial\hat{\theta}_j \geq \partial^2 t_i(\hat{\theta})/\partial\hat{\theta}_i \partial\hat{\theta}_j$ for all $\hat{\theta}$, j, and i. The next definition formalizes this idea and extends it to nondifferentiable transfers. Let $>_{-i}$ be the product order on $\mathbb{R}^{n-1}$.

Definition 3. Define the ordering relation $\succeq_{\mathrm{ID}}$ on the space of transfer functions such that $\tilde{t} \succeq_{\mathrm{ID}} t$ if, for all $i \in N$ and for all $\theta_i'' > \theta_i'$ and $\theta_{-i}'' >_{-i} \theta_{-i}'$, $\tilde{t}_i(\theta_i'', \theta_{-i}'') - \tilde{t}_i(\theta_i'', \theta_{-i}') - \tilde{t}_i(\theta_i', \theta_{-i}'') + \tilde{t}_i(\theta_i', \theta_{-i}') \geq t_i(\theta_i'', \theta_{-i}'') - t_i(\theta_i'', \theta_{-i}') - t_i(\theta_i', \theta_{-i}'') + t_i(\theta_i', \theta_{-i}')$.

While $\succeq_{\mathrm{ID}}$ is transitive and reflexive on the space of transfer functions, it is not antisymmetric. Consider the set of $\succeq_{\mathrm{ID}}$ equivalence classes of transfers, denoted $\mathcal{T}$.[16]

The next proposition provides the foundations for the definition of optimal implementation. It is also interesting in itself for the theory of supermodular games.[17] It

[16]Any quasi-order is transformed into a partial order using equivalence classes.
[17]See Milgrom and Roberts (1995, pp. 189–190) for a related result.

shows that if a transfer function t'' generates more complementarities than a transfer function t', then t'' induces a game whose interval prediction is larger than the interval prediction of the game induced by t'. As such, the objective of minimizing the complementarities coincides with the objective of minimizing the interval prediction.

For any $t \in \mathcal{T}$ and supermodular implementable $f = (x, t)$, let $\overline{\theta}^t(\cdot)$ and $\underline{\theta}^t(\cdot)$ denote the extremal (Bayesian) equilibria of the game induced by the mechanism.

PROPOSITION 2. *Let the decision rule and the valuation functions be such that* $E_{\theta_{-i}}[V_i(x_i(\hat{\theta}_i, \theta_{-i}), \theta_i)]$ *is continuous in* $(\hat{\theta}_i, \theta_i)$. *For any supermodular implementable social choice functions* (x, t'') *and* (x, t'), *if* $t'' \succeq_{\text{ID}} t'$, *then* $[\underline{\theta}^{t'}(\cdot), \overline{\theta}^{t'}(\cdot)] \subset [\underline{\theta}^{t''}(\cdot), \overline{\theta}^{t''}(\cdot)]$.

A social choice function will be optimally supermodular implementable if it is supermodular implementable with transfers that generate the weakest complementarities. By the previous proposition, this gives the tightest interval prediction around the truthful equilibrium among all supermodular mechanisms.

DEFINITION 4. A social choice function $f = (x, t^*)$ is optimally supermodular implementable if it is supermodular implementable and $t \succeq_{\text{ID}} t^*$ for all transfers $t \in \mathcal{T}$ such that (x, t) is supermodular implementable.

The main theorem determines which decision rules are optimally supermodular implementable. The conclusion is rather powerful: In twice-continuously differentiable environments, all implementable social choice functions whose decision rule satisfies some dimensionality condition are optimally supermodular implementable. This dimensionality condition is defined as follows. A decision rule $x : \Theta \mapsto (x_i(\theta))$ is *dimensionally reducible* if, for each $i \in N$, there are twice-continuously differentiable functions $h_i : \mathbb{R}^2 \to X_i$ and $r_i : \Theta_{-i} \to \mathbb{R}$ such that $r_i(\cdot)$ is increasing and $x_i(\theta) = h_i(\theta_i, r_i(\theta_{-i}))$ for all $\theta \in \Theta$. The condition is trivially true when there are two individuals. If there are more than two, a player's decision rule can depend on her own type directly, but it must depend on her opponents' types indirectly through a real-valued aggregate.[18] Examples 3 and 7 provide examples of dimensionally reducible (and efficient) decision rules.

THEOREM 3. *Let the valuation functions be twice-continuously differentiable and let* $f = (x, t)$ *be a social choice function where* x *is dimensionally reducible. If* f *is implementable, then there exist* t^* *such that* (x, t^*) *is optimally supermodular implementable.*

There are many possible ways to transform a mechanism into a supermodular mechanism. This section elicits a particular transformation that generates the smallest equilibrium set among all possible conversions, hence all supermodular mechanisms.

It is worth pointing out that weakening the complementarities does not always slow down learning; actually the opposite is often true. A "flatter" best response function increases the rate of convergence toward the equilibrium set.

[18]Taking types in $[0, 1]$, it excludes, for example, x for which $x_1(\theta) = \theta_1^{\theta_2 \theta_3} + \theta_1 + \theta_2 + \theta_3$.

6.2 *Unique implementation*

In this section, I provide sufficient conditions for a social choice function to be uniquely supermodular implementable. After giving conditions for minimizing the interval prediction, it is natural to study situations where truthtelling is the unique equilibrium. The induced game is then dominance-solvable and all learning dynamics converge to the equilibrium. As a by-product, coalition-proof Nash implementation is implied via Milgrom and Roberts (1996).

Definition 5. A social choice function f is uniquely supermodular implementable if it is supermodular implementable and the truthful equilibrium is the unique Bayesian equilibrium.

Before providing the sufficient conditions, some definitions are in order. The valuations and the decision rule have *strong differences* if, for each i and θ_i, there is a real number $\gamma_i(\theta_{-i})$ such that $\partial^2 V_i(x_i(\hat{\theta}_i, \theta_{-i}), \theta_i)/\partial\hat{\theta}_i\,\partial\theta_i \geq \gamma_i(\theta_{-i})$ for all $\hat{\theta}_i$, θ. In the nondifferentiable case, the condition becomes that for all $\hat{\theta}''_i \geq \hat{\theta}'_i$ and $\theta''_i \geq \theta'_i$, $\Delta V_i(\hat{\theta}''_i, \hat{\theta}'_i, \theta_{-i}, \theta''_i) - \Delta V_i(\hat{\theta}''_i, \hat{\theta}'_i, \theta_{-i}, \theta'_i) \geq \gamma_i(\theta_{-i})(\hat{\theta}''_i - \hat{\theta}'_i)(\theta''_i - \theta'_i)$.

The previous condition is derived exogenously from the primitives of the model. The next condition is determined endogenously from the transfers in the mechanism. A mechanism $\Gamma = (\{\Theta_i\}, (x, t))$ generates *bounded complements* if, for each i and θ_i, there is a real number $K_i(\theta_i)$ such that $\Delta u_i(\theta''_i, \theta'_i, \theta''_{-i}, \theta_i) - \Delta u_i(\theta''_i, \theta'_i, \theta'_{-i}, \theta_i) \leq K_i(\theta_i) \times (\theta''_i - \theta'_i)\sum_{j\neq i}(\theta''_j - \theta'_j)$ for all $\theta''_i \geq \theta'_i$ and $\theta''_{-i} \geq \theta'_{-i}$. The mechanism generates *uniformly* bounded complements if there is a uniform upper bound K_i. The valuation functions and the decision rule are said to generate bounded complements if the mechanism $\Gamma = (\{\Theta_i\}, (x, t^0))$ generates bounded complements, where t^0 is the transfer function that is identically zero.

6.2.1. *Sufficient conditions*

The sufficient conditions for unique supermodular implementation are imposed on a matrix, which I call the *matrix of complementarities*. To define this matrix, assume strong differences and suppose that the mechanism $(\{\Theta_i\}, f)$ generates bounded complements (see Section 4). The matrix of complementarities is the $n \times n$ matrix whose ith row contains $-E_{\theta_{-i}}[\gamma_i(\theta_{-i})]$ as the ith entry and $E_{\theta_i}[K_i(\theta_i)]$ as the other entries:

$$\mathcal{C} = \begin{pmatrix} -E_{\theta_{-1}}[\gamma_1(\theta_{-1})] & E_{\theta_1}[K_1(\theta_1)] & \cdots & E_{\theta_1}[K_1(\theta_1)] \\ E_{\theta_2}[K_2(\theta_2)] & -E_{\theta_{-2}}[\gamma_2(\theta_{-2})] & \cdots & E_{\theta_2}[K_2(\theta_2)] \\ \vdots & \vdots & \ddots & \vdots \\ E_{\theta_n}[K_n(\theta_n)] & E_{\theta_n}[K_n(\theta_n)] & \cdots & -E_{\theta_{-n}}[\gamma_n(\theta_{-n})] \end{pmatrix}.$$

The matrix of complementarities is determined endogenously by the mechanism through the bounds on complements (K_i's). The matrix indicates how sensitive agents are as a whole to their own type versus their opponents' announcements. On the one hand, when the complementarities between own announcement and type are strong (large $E[\gamma(\cdot)]$), agents tend to announce high types regardless of their opponents' strategies. This favors uniqueness. On the other hand, when the complementarities between agents' announcements are strong (large $E[K(\cdot)]$), it is source of multiplicity. The dominating effect is captured by the definiteness of the matrix of complementarities. If it is negative-definite, the first effect is stronger. For example, if the sum of the entries on each row is negative, then the matrix is negative-definite.

Proposition 3. *Let f be a supermodular implementable social choice function. Let the valuation functions and the decision rule be continuously differentiable. If the matrix of complementarities is negative-definite, then f is uniquely supermodular implementable.*

The previous proposition is mostly useful a posteriori, that is, after building it, the designer can check whether the mechanism has a unique equilibrium. But one would also like to know about the size of the equilibrium set beforehand, based on the primitives of the design problem. The following corollary accomplishes this goal by using the optimal transfers. Since the optimal transfers are built from the primitives of the model, they can be used within Proposition 3. Since they give the smallest equilibrium set, they are the natural candidates to lead to unique implementation.

Corollary 3. *Let the valuation functions be twice-continuously differentiable. Let $f = (x, t)$ be an implementable social choice function with a dimensionally reducible decision rule. Denote*

$$K_i(\theta_i) = \max_{j \in N, \hat{\theta} \in \Theta} \left(\frac{\partial^2 V_i(x_i(\hat{\theta}), \theta_i)}{\partial \hat{\theta}_i \, \partial \hat{\theta}_j} - \min_{\theta_i \in \Theta_i} \frac{\partial^2 V_i(x_i(\hat{\theta}), \theta_i)}{\partial \hat{\theta}_i \, \partial \hat{\theta}_j} \right). \tag{5}$$

If the matrix of complementarities obtained from (5) is negative-definite, then there are transfers t^* such that (x, t^*) is uniquely supermodular implementable.

This corollary points to the heterogeneity of complementarities across types (within the valuations) as being responsible for multiplicity of equilibria. Since the designer has to make the induced game supermodular for each realization,[19] she may have to add complementarities for some realizations of types that are unnecessarily high for others, thus generating multiple equilibria.

Negative-definiteness of the matrix of complementarities is a convenient condition for unique implementation, but it is not necessary. The gap between necessary and sufficient conditions is essentially a computational matter. To help close the gap, I provide a way to compute bounds on the equilibrium set. This gives an alternative method to check for uniqueness, but it goes beyond, since it can be used to get an idea of the size of the equilibrium set when the uniqueness condition fails. Finding these bounds involves

[19] It is sufficient but not necessary that the ex post game be supermodular for each realization so that the ex ante Bayesian game is supermodular. For example, if the prior is mostly concentrated on some subset Θ' of Θ, it may not be necessary to make the ex post payoffs supermodular for types in $\Theta \backslash \Theta'$. Of course, the possibility of neglecting $\Theta \backslash \Theta'$ depends on how unlikely that set is compared to how submodular the utility function may be for types in that set.

solving systems of equations formed exclusively from the primitives of the model. For example, the ability to compute bounds on the equilibrium set is attractive for welfare analysis, as it becomes possible to measure the maximal loss in efficiency caused by the existence of many equilibria.

Proposition 4. *Let $f = (x, t)$ be a supermodular implementable social choice function. Let the valuations and the decision rule be continuously differentiable. Assume bounded substitutes and suppose mechanism $(\{\Theta_i\}, f)$ generates bounded complements. Then there exist two systems of $2n$ equations: one whose lowest solution bounds the equilibrium set from below; the other whose largest solution bounds the equilibrium set from above.*

These two systems of equations, which figure in (25) and (27) in Appendix B, are constructed by approximating the game induced by the mechanism. The approximated game is a supermodular game whose interval prediction includes the equilibrium set of the original game. In the approximated game, agents have utility functions that are quadratic approximations of their real utility functions. Hence these conditions are necessary for uniqueness if utilities are indeed quadratic. If types are distributed uniformly, these systems are simple systems of $2n$ quadratic equations. Example 3 is an illustration how useful this method can be.

Example 3. Consider the public goods setting of Section 3 with a third player whose type is uniformly and independently distributed in $[0, 1]$. Her valuation function is $V_3(x, \theta_3) = \theta_3 x - x^3/10$. The decision rule $x(\theta) = 5/3(\sqrt{1 + 6/5(\sum_i \theta_i)} - 1)$ is allocation-efficient and dimensionally reducible. The designer can choose the budget balancing transfers t^{BB} or the optimal transfers t^*. If she prefers full efficiency, then she chooses $\rho_1 \geq 8$, $\rho_2 \geq 5$, and $\rho_3 \geq 6$, and transfers t^{BB}. However, the interval prediction of the resulting game is always the entire space, so she may prefer to minimize the interval prediction by using t^*. These transfers give $K_i(\theta_i) = 3/5(1 - \theta_i)$ and $\gamma_i(\theta_{-i}) = 1/\sqrt{1 + 6/5(1 + \sum \theta_j)}$. It is easy to check that the corresponding matrix of complementarities is not negative-definite. Proposition 4 becomes a crucial tool to find bounds on the equilibrium set. This is done by solving two simple systems of $2n$ quadratic equations. The following system corresponds to (27) (see Appendix B) and should be solved for $(A_i, B_i)_{i=1}^n$:

$$A_i = 1 - 1.1\left(\sum_{j \neq i} \frac{(A_j - B_j)^2}{2A_j} - 1\right), \quad i = 1, \ldots, n$$

$$B_i = 1.1\left(\sum_{j \neq i} \frac{(A_j - B_j)^2}{2A_j} - 1\right), \quad i = 1, \ldots, n.$$

The solution defines piecewise affine functions according to (26) that are bounds on the equilibrium set. The solution of the above is roughly $(1.1, .1)_i$ and that of (25) is $(1, 0)_i$, which corresponds to truthtelling. Thus, all the equilibria are contained in a small strip below truthtelling. Welfare analysis reveals that profiles in that interval prediction can at most result in a .03 loss in total utility under the optimal transfers (compared to truthtelling). ◊

Next I apply the uniqueness results to the leading public goods example, where any mechanism can be converted into a supermodular mechanism with a unique equilibrium, even if the original mechanism has many equilibria. This shows how weak implementation can be turned into strong implementation.

EXAMPLE 4. Consider the model of Section 3. If the designer had used transfers $\hat{t}_i(\hat{\theta}) = t_i(\hat{\theta}) - 3\hat{\theta}_i\hat{\theta}_j + \frac{3}{2}\hat{\theta}_i$, $i = 1, 2$, instead of starting with the expected externality transfers, then she would have induced a game with many equilibria. There are two equilibria where one agent always reports 0 while the other always reports 1. Furthermore, truthtelling is an unstable equilibrium; any perturbation results in a departure from it. Yet (x, t^*) is uniquely supermodular implementable. To see why, note $\partial x_i(\theta)/\partial\theta_i = 1$ and $\partial^2 V_i(x, \theta_i)/\partial x\,\partial\theta_i = 1$, which implies $\gamma_i(\theta_{-i}) = 1$, $i = 1, 2$. Since $\partial^2 V_i(x(\hat{\theta}), \theta_i)/\partial\hat{\theta}_1\,\partial\hat{\theta}_2$ is constant in θ_i for $i = 1, 2$, $K_i(\theta_i) = 0$. Proposition 3 completes the proof. ◊

As a final remark on uniqueness, it is worth mentioning that if types are assumed to be independent, which is standard in mechanism design, then the information structure can hardly be used as a tool to obtain uniqueness.[20] So one has to impose conditions on payoffs, which necessarily imply a trade-off involving complementarities. The way to measure complementarities creates a gap between necessity and sufficiency. This gap depends on how tight the bounds on complementarities (K_i's and γ_i's) are.

6.2.2. *Unique implementation and full efficiency*

This section deals with the multiple equilibrium problem under the budget balance condition. As previous results show, making sure that the equilibrium set is minimal requires some flexibility in designing the transfers. But this flexibility is not always compatible with balancing the budget, as suggested by the functional form of the balanced transfers. As a result, in some environments (see Example 3 below), there will be a conflict between full efficiency and robustness. Example 1 already delivered the message: Sometimes the designer must sacrifice one or the other. Either the designer uses the supermodular mechanism with a unique equilibrium at the price of a balanced budget (full efficiency) or she loses robustness by balancing the budget via the expected externality mechanism. Nonetheless, there are situations where the design problem admits a supermodular mechanism with unique equilibrium, and this robustness can be autofinanced by the agents. This is formalized in the next proposition.

PROPOSITION 5. *Let $n \geq 3$. Consider valuation functions and an allocation-efficient decision rule that are continuously differentiable, and generate strong differences. Suppose that these valuations and the decision rule produce substitutes and complements that*

[20] For example, we cannot use the global games argument of Frankel et al. (2003), which heavily relies on correlation across types.

are uniformly bounded, respectively, by T_i and τ_i. If $\tau_i - T_i < E_{\theta_i}[\gamma_i(\theta_i)]$, then there are balanced transfers t^{BB} such that (x, t^{BB}) is uniquely supermodular implementable.

The proposition gives sufficient conditions for the balanced transfers identified in Theorem 2 to yield truthtelling as a unique equilibrium. The following example is an application.

Example 5. Consider the same setting as the public goods example of Section 3 with an additional player, player 3, whose type is independently distributed from the other players' types in $\Theta_3 = [0, 1]$. Player 3's valuation function is $V_3(x, \theta_3) = \theta_3 x$. Let $X = [0, 3]$ and $x(\theta) = \sum_i \theta_i$; x is allocation-efficient and $\gamma_i(\theta_i) = 1$ for all i. The valuations and the decision rule produce complements and substitutes that admit the same bounds. Proposition 5 says that there exist $\{\rho_i\}$, for example $\rho_i = 1/2$ for all i, such that (x, t^{BB}) is uniquely supermodular implementable with a balanced budget. ◊

This paper does not answer, in general, the question of reaching uniqueness and supermodularity with the budget constraint. This is a difficult issue that requires the use of more general mechanisms.

7. Approximate supermodular implementation

I generalize some results within the context of approximate (or virtual) implementation.[21] In well behaved environments, the results so far have been quite general. They apply to a variety of contexts, such as principal multiagent (Section 8) and public goods models. However, there are interesting situations with discontinuities that fall outside the scope of the current results. A way around this problem is approximate implementation; the objective becomes to supermodularly implement (well behaved) social choice functions that are arbitrarily close to a "target social choice function." After describing the failure of bounded substitutes in the basic auction setting, I present results that accommodate these situations.

Consider the following auction model. There is one unit of an indivisible good to be allocated among two buyers $\{1, 2\}$ whose types lie in $[\underline{\theta}, \overline{\theta}]$. An outcome is represented by the vector (x_1, x_2), where $x_i = 1$ if i gets the good and 0 otherwise. Buyer i's utility function is $u_i(x_i, \theta_i) = \theta_i x_i + t_i$. The allocation-efficient decision rule x^* attributes the good to the buyer with the highest type:

$$x_1^*(\theta) = \begin{cases} 1 & \text{if } \theta_1 \geq \theta_2 \\ 0 & \text{otherwise} \end{cases} \quad \text{and} \quad x_2^*(\theta) = 1 - x_1^*(\theta). \tag{6}$$

Note that for any types such that $\theta_2'' > \theta_1'' > \theta_2' > \theta_1'$, $x_1(\theta_1'', \theta_2'') - x_1(\theta_1', \theta_2'') - x_1(\theta_1'', \theta_2') + x_1(\theta_1', \theta_2') = -1$. Hence, the assumption of bounded substitutes requires the existence of T such that $-\theta_1 \geq T(\theta_1'' - \theta_1')(\theta_2'' - \theta_2')$ for all $\theta_1 \in \Theta_1$. This is clearly impossible, as the above order of types can be maintained while $\theta_1' \uparrow \theta_2'$ and $\theta_1'' \downarrow \theta_2'$. Substitutes are unbounded and none of the results applies.

[21] See Abreu and Matsushima (1992) and Duggan (1997).

Clearly, the problem is caused by the lack of smoothness of the decision rule. So the idea is to approximate the social choice function by smooth implementable functions that are known to satisfy the desired conditions. Let $\|\cdot\|_p$ be the L_p norm.

Definition 6. A decision rule x is approximately (optimally) supermodular implementable if there exists a sequence of (optimally) supermodular implementable social choice functions $\{(x_n, t_n)\}$ such that, for $1 \leq p < \infty$, $\lim_{n\to\infty} \|x_{n,i} - x_i\|_p = 0$ for all i.

The next results provide conditions so that a social choice function can be approached by a sequence of supermodular implementable social choice functions. The intuition is that the space of smooth functions is dense in L_p spaces, hence a smooth approximation exists for any L_p function. Smooth social choice functions satisfy the bounded substitutes assumption, so there remains only to establish that incentive compatibility can be preserved along the sequence. Furthermore, if the decision rule also satisfies the dimensionality condition from Section 6.1, then it is approachable by social choice functions whose supermodular game form gives the tightest interval prediction.

Proposition 6. *Let the valuation functions be twice-continuously differentiable such that $\partial V_i(x_i, \theta_i)/\partial\theta_i$ is increasing in x_i for all i. If the decision rule is such that $x_i \in L_p$ is increasing in $\hat{\theta}_i$ for all i, then it is approximately supermodular implementable.*

Proposition 7. *Let the valuation functions be twice-continuously differentiable such that $\partial V_i(x_i, \theta_i)/\partial\theta_i$ is increasing in x_i for all i. If the decision rule is such that, for all i, there exist $h_i : \mathbb{R}^2 \to X_i$ and $r_i : \Theta_{-i} \to \mathbb{R}$ such that*

(i) h_i is bounded and increasing in its first variable

(ii) r_i is continuous and strictly increasing[22]

(iii) $x_i(\theta) = h_i(\theta_i, r_i(\theta_{-i}))$,

then it is approximately optimally supermodular implementable.

These results apply to many discontinuous models of interest such as public goods, auctions, and bilateral trading (Section 8). In the above auction, for example, let $r_i(\theta_{-i}) = \max\{\theta_j : j \neq i\}$, and let $h_i(\theta_i, r_i) = 1$ if $\theta_i > r_i$, and 0 otherwise. Since all the conditions are satisfied, Proposition 7 applies. In addition, these results suggest a dilemma between close implementability and robustness of equilibrium play. I develop this idea in Section 10; Section 8.2 provides an illustration.

8. Applications

8.1 *Principal–multiagent problem*

This section applies the theory to the traditional principal–multiagent problem with hidden information. A principal contracts with n agents. Agent i's type lies in $[\underline{\theta}_i, \overline{\theta}_i]$.

[22] Function r_i is strictly increasing if $r_i(\theta''_{-i}) > r_i(\theta'_i)$ whenever $\theta''_{-i} \gg \theta'_{-i}$.

Types are independently distributed according to a common prior. Each agent i exerts some observable effort $x_i \in X_i$ and bears a cost $c_i(x_i, \theta_i)$ when her type is θ_i. From all the efforts $x = (x_1, \dots, x_n)$ and types, the principal receives utility $w(x, \theta)$. The principal faces the problem of designing a profit-maximizing contract, subject to incentive and reservation-utility constraints. A contract is a function that maps types into effort and transfer levels for each agent. The principal's problem can be stated as

$$(\hat{x}, \hat{t}) \in \arg\max_{f=(x,t)} E_\theta \left[w(x(\theta), \theta) - \sum_{i=1}^{n} t_i(\theta) \right] \tag{7}$$

subject to

$$E_{\theta_{-i}}\big[t_i(\theta_i, \theta_{-i}) - c_i(x_i(\theta_i, \theta_{-i}), \theta_i)\big] \geq E_{\theta_{-i}}\big[t_i(\theta_i', \theta_{-i}) - c_i(x_i(\theta_i', \theta_{-i}), \theta_i)\big] \tag{8}$$

for all θ_i and θ_i', and

$$E_{\theta_{-i}}\big[t_i(\theta_i, \theta_{-i}) - c_i(x_i(\theta_i, \theta_{-i}), \theta_i)\big] \geq \overline{u}_i \tag{9}$$

for all θ_i. Condition (8) requires truthtelling to be an equilibrium. Condition (9) is an interim participation constraint, as agents may opt out of the mechanism if it does not meet their reservation utility.

If the underlying functions w and c_i, and the prior are smooth and guarantee the existence of a (dimensionally reducible) solution, then the contract is (optimally) supermodular implementable. In words, if the principal is in a position to engage in a smooth revenue-maximizing, and incentive-compatible contract, which allows voluntary participation, then she can turn that contract into a supermodular contract that retains properties (7), (8), and (9), and minimizes the size of the equilibrium set.

8.2 *Auctions, bilateral trading, and allocation schemes*

Auctions are notorious for lacking complementarities. Consider the auction setting of Section 7. Two buyers, 1 and 2, have utility function $v(\theta_i)x_i + t_i$, $i = 1, 2$, where v is strictly increasing. The allocation-efficient decision rule x^* is unchanged and equal to (6). According to Proposition 7,[23] it is approachable by a sequence of optimally supermodular implementable decision rules. An example of such a sequence is

$$x_{1,\epsilon}(\hat{\theta}) = c \int_0^{\sum_i (a_i \hat{\theta}_i + b_i)/\epsilon} \frac{1}{1+t^2}\, dt + k \quad \text{and} \quad x_{2,\epsilon} = 1 - x_{1,\epsilon},$$

where constants c, k, a_i, and b_i, $i = 1, 2$, are chosen appropriately. The allocation of the good under the approximate mechanism is probabilistic. The smaller is ϵ, the larger is the probability that the highest announcement receives the good. This is illustrated in Figure 1. The step function represents the efficient allocation rule x_1 (assuming agent 2's type is θ_2) and the curve is the approximation.

The transfer $t^*_{i,\epsilon}$ to agent i is the expected transfer that i would receive under a second-price auction—if i made the same announcement and others played truthfully—

[23] In the case with $n \geq 3$, let $h_i(\theta_i, r_i) = 1$ if $\theta_i > r_i$ and 0 otherwise for all i. Note that h_i is bounded and increasing in θ_i. Now choose $r_i(\theta_{-i}) = \max\{\theta_j : j \neq i\}$, which is continuous and strictly increasing.

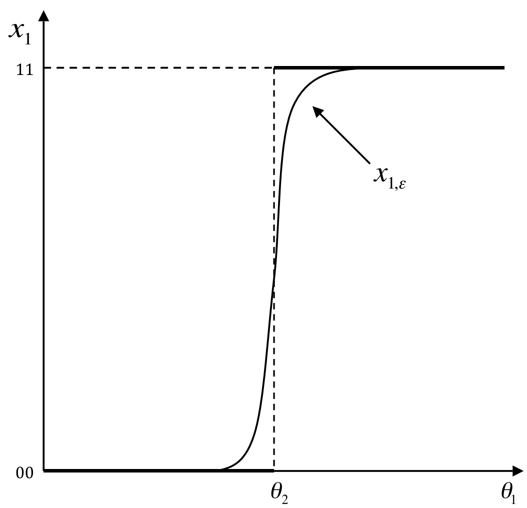

FIGURE 1. Approximate allocation.

plus the optimal supermodular transformation (see (14) and (15) in Appendix B). The matrix of complementarities $\mathcal{C}_\epsilon$ can be formed. Since substitutes are unbounded in the exact case, it is not surprising that they have a lower bound that decreases to infinity as $\epsilon \to 0$ in the approximate case. To compensate for this, the transfers add more complementarities as ϵ vanishes, which leads complementarities to explode on those parts of the space where reports were already complements. As a consequence, the interval prediction of the game induced by $\{x_\epsilon, t^*_\epsilon\}$ must be the entire space in the limit. What is interesting, however, is to find the smallest amount of inefficiency $\underline{\epsilon}$ for which $\mathcal{C}_{\underline{\epsilon}}$ is still negative-definite and uniqueness is preserved. This illustrates nicely the forces at work in Proposition 3. Letting $v(\theta) = \theta^\omega$ with $\omega > 0$, it always takes a larger ω to obtain a smaller $\underline{\epsilon}$. That is, the more sensitive an agent is to her own type, the higher is the degree of complementarities that can be compatible with uniqueness; hence the lower ϵ can be. In the standard case with $\omega = 1$, $\underline{\epsilon} \approx .45$. When $\omega = 6$, $\underline{\epsilon} \approx .15$. Though uniqueness fails for some ϵ, the approximate approach is useful when the equilibrium set is small (Theorem 3).

The results for approximate implementation also apply to public goods situations where a society of agents decide whether to undertake a public project. The bargaining mechanism of Myerson and Satterthwaite (1983, p. 274) also satisfies the assumptions of Proposition 7 and, as such, the decision rule is approachable by a sequence of optimally supermodular implementable decision rules. The expected gains from trade along the sequence converge to the maximal expected gains.

9. Discussion

9.1 *A revelation principle for general preferences*

I present a revelation principle for supermodular mechanism design. This paper has limited attention to direct supermodular mechanisms, and they widely apply in quasilinear environments. For general preferences, however, how restrictive are direct mechanisms?

To study this question, one has to extend the framework of Section 4 to environments with general preferences, type spaces, and mechanisms (see Mathevet 2007). This generality carries over to the definition of a supermodular game, because the message spaces are no longer subsets of the real line, hence an ordering relation—beyond the usual order on the reals—is required.

For general supermodular design problems, the challenge is to specify an ordered message space and an outcome function so that agents adopt monotone best responses. The existence of a revelation principle for supermodular design is a relevant question, because the set of all possible message spaces and orders on those spaces is so large that the task seems difficult. This principle gives conditions so that, if a social choice function is supermodular implementable, then there exists a direct-revelation mechanism that supermodularly implements this social choice function truthfully.

The question pertains to the existence of orders on type spaces that make the (induced) direct game supermodular. Although the resulting orders on type sets may be very complex, the agents do not need to know them to play the mechanism.[24] The designer herself only needs to know that some orderings exist, not their definition. A proof of this result can be found in Mathevet (2007).

Theorem 4 (The supermodular revelation principle for finite types). *Let Θ_i be finite for all i. If there exists a mechanism $(\{M_i\}, g)$ that supermodularly implements the social choice function f such that there is an equilibrium $m^*(\cdot)$ for which $g \circ m^* = f$ and $m_i^*(\Theta_i)$ is a lattice, then f is truthfully supermodular implementable.*[25]

According to the supermodular revelation principle, limiting attention to direct mechanisms amounts to restricting one's scope to mechanisms where the equilibrium strategies are lattice-ranged. The proof works by constructing an order on each player's type space that is order-isomorphic to the range of her equilibrium strategy.

How tight is the condition imposed on the range of the equilibrium strategies? Example 8 in Appendix A suggests that it is somewhat minimally sufficient. The example presents a supermodular implementable social choice function where this range is not a lattice and that cannot be supermodularly implemented by any direct mechanism. At the same time, it shows that the revelation principle fails to hold, in general, for supermodular mechanism design.

The theorem also states conditions that are verifiable a posteriori; it may indeed be useful to know when a complex mechanism can be replaced with a simpler direct mechanism. The next corollary identifies situations where direct mechanisms cause no loss of generality.

[24] Orders are useful for the analyst, but not for the players, much in the same way as differentiability of the utility functions.

[25] General—possibly non-Euclidean—message spaces are endowed with an order and $m_i^*(\Theta_i)$ is a lattice under that order. Mathevet (2007) generalizes the theorem to infinite type spaces.

COROLLARY 4. *Let Θ_i be finite for all $i \in N$. If there is a mechanism $(\{M_i\}, g)$ with totally ordered message sets that supermodularly implements a social choice function f, then f is truthfully supermodular implementable.*

If the designer is interested only in mechanisms where the message spaces are totally ordered, then she can look at direct mechanisms without loss of generality.

Theorem 4 gives only sufficient conditions for revelation, but in those cases where a supermodular direct mechanism exists, while the lattice condition is violated, the existence of an order has little or nothing to do with a revelation principle, because supermodularity cannot be shown to be transmitted from the indirect mechanism.

9.2 *Dominant-strategy versus supermodular implementation*

This section highlights the advantage of supermodular implementation over dominant-strategy implementation. One of the main arguments in favor of supermodular mechanisms is the existence of a subset of profiles, the interval prediction, which provides a robust forecast of what agents will play. But equilibrium play is also quite robust under dominant-strategy implementation, although it, too, is not immune to the multiple equilibrium problem.[26]

First, strategy proofness, requiring truthtelling to be a dominant-strategy, is not always possible, even in smooth environments. Mookherjee and Reichelstein (1992, Proposition 2 and Definition 5) have shown that a sufficient and "nearly" necessary condition for strategy-proofness is that the valuation functions and the decision rule satisfy some single-crossing property. They make an assumption, the one-dimensional condensation property, which makes it easier to satisfy that necessary condition.

The next example, inspired by McAfee and McMillan (1991), violates the one-dimensional condensation property, and the necessary condition for strategy-proofness. So the decision rule is not dominant-strategy implementable, yet it is uniquely supermodular implementable.

EXAMPLE 6. Two agents, 1 and 2, whose types are independently and uniformly distributed in [0, 3], exert some effort to produce an observable contribution x_i. The amount of effort e_i necessary for x is $e_1(x, \theta_1) = (3 - \theta_1)(x_1 - x_2) + x_1 + \frac{9}{2}$ and $e_2(x, \theta_2) = (3 - \theta_2)(x_2 + x_1)$. Agent 2 has positive externalities on her counterpart, whereas agent 1 has negative externalities. Before transfers, the principal has utility $w(x, \theta) = u(x, \theta) - c_p(x, \theta)$, where c_p represents the production costs. The principal's objective is to solve (7) subject to (8) and the ex ante minimum wage $E_\theta[t_i(\theta)] \geq 0$ on the economy. Let function u and the production costs be such that the optimal decision rule is $x^*(\theta) = (\theta_2\theta_1 - 3/2\theta_1, \theta_2 - \theta_1)$.[27] Agent i's valuation is $V_i(x, \theta_i) = -e_i(x, \theta_i)$. Proposition 8 in Appendix B applies, so there exist transfers that implement x^*. Constructing optimal transfers from (14) and (15) gives $t_1^*(\hat{\theta}) = -\frac{1}{2}\hat{\theta}_1^2 - 3\hat{\theta}_1 + 2\hat{\theta}_2\hat{\theta}_1 + \frac{27}{18}$ and

[26] Dominant strategies do not rule out the existence of weakly dominated strategies. So agents can very well play another dominant-strategy equilibrium or a Nash equilibrium whose outcome differs from the social choice function. Also, learning dynamics may converge to those "unwanted" equilibria. Saijo et al. (2005) report situations where dominant-strategy implementation has serious drawbacks.

[27] For example, let $u(x, \theta) = \theta_2(\theta_1 x_1 + x_2)$ and $c_p(x, \theta) = (x_1^2 + x_2^2)/2 + \theta_1(3/2x_1 + x_2)$.

$t_2^*(\hat{\theta}) = -\frac{5}{4}\hat{\theta}_2^2 + 3\hat{\theta}_2 + 3\hat{\theta}_2\hat{\theta}_1 - \frac{675}{72}$. Since $K_i(\theta_i) = \theta_i$, $\gamma_1(\theta_2) = \theta_2 - \frac{1}{2}$, and $\gamma_2(\theta_1) = \theta_1 + 1$, we can check that the matrix of complementarities

$$\mathcal{C} = \begin{pmatrix} -1 & 3/2 \\ 3/2 & -5/2 \end{pmatrix}$$

is negative-definite and so truthtelling is the unique equilibrium. ◊

Second, it is well known that dominant-strategy implementation is often incompatible with balancing budget (Green and Laffont 1979, Laffont and Maskin 1980). The next example depicts a situation where unique supermodular implementation allows balancing the budget in a case where dominant strategies cannot. In other words, by weakening the solution concept and putting some structure on the game form, it is possible to balance the budget and maintain a high likelihood of equilibrium play.

Example 7. In the public goods example of Section 3, let $\Theta_1 = \Theta_2 = [2, 3]$. Add a third player, player 3, whose type is independently distributed from the other players' types in $\Theta_3 = [2, 3]$. Player 3's valuation function is $V_3(x, \theta_3) = \theta_3 x - \ln x$. Letting $X = [5, 10]$, the allocation-efficient decision rule is $x(\theta) = \frac{1}{2}(\sum_i \theta_i + \sqrt{(\sum_i \theta_i)^2 - 4})$. By Theorem 3.1 in Laffont and Maskin (1980), the decision rule is dominant-strategy implementable only if transfers are of the Groves form. However, these transfers cannot balance the budget, because they violate the necessary condition from Theorem 4.1 in Laffont and Maskin (1980). Nevertheless, since $\tau_i - T_i < .03$ and $\gamma_i > 1$ for all i, Proposition 5 implies that x is uniquely supermodular implementable with a balanced budget. ◊

Ideally, one would like all strategy-proof social choice functions to be uniquely supermodular implementable, but such a result is not known.

10. Conclusion

This paper introduces a mechanism design approach that allows dealing with the multiple equilibrium problem by using simple mechanisms that are robust to bounded rationality. The main motivation behind this approach is the question of the likelihood of the desired equilibrium outcome. If the design problem satisfies the rationality and the epistemic conditions that ensure equilibrium play, then this question is irrelevant, as there already exist elaborate mechanisms that solve the multiple equilibrium problem. However, the more complex the mechanism, the more unlikely it is to satisfy these conditions. In most design problems, these conditions are unrealistic, hence robustness to less than full rationality is crucial. Yet, this has been neglected in the mechanism design and implementation literature. This paper presents a possible answer. The methodology consists in inducing supermodular games. While supermodularity induces properties of dominance-solvability, it has stronger theoretical[28] and experimental implications (Camerer 2003, Sefton and Yavaş 1996).

[28] See footnote 5.

Beyond the results, this paper brings out basic questions about robustness and the design problem. Robustness here means the existence of a small subset of strategy profiles that provide a good prediction of what agents will play. In view of Section 6.2, one may wonder whether there is a price to pay for robustness in terms of efficiency. The trade-off appears quite clearly in this framework; sometimes the designer must sacrifice robustness for full efficiency or vice versa. In the public goods example, the designer can modify the expected externality mechanism and secure dominance-solvability at the price of a balanced budget or she can use the expected externality mechanism to balance the budget but she loses dominance-solvability. This may be related to the specifics of supermodular implementation, but it is an interesting issue. One may also wonder whether there is a price to pay for robustness in terms of closeness of the decision rule implemented. This has obvious implications in terms of efficiency. If one accepts some imprecision around the social choice function, then supermodular implementability extends, even in its optimal version. This evokes the trade-off between close implementability and stability raised by Cabrales (1999), regarding the Abreu–Matsushima (1992, 1994) mechanism.

This paper raises issues that have not been discussed. The multiple equilibrium problem in supermodular mechanisms suggests an alternative solution, namely strong implementation. Strong implementation requires all equilibria of the mechanisms to yield desired outcomes. Healy and Mathevet (2008) show that the Lindahl and the Walrasian social choice functions are uniquely implementable with a contractive mechanism.

Like many Bayesian mechanisms, the present mechanisms are parametric in the sense that they rely on agents' prior beliefs. Thus the designer uses information other than that received from the agents (Hurwicz 1972). It may be interesting to design nonparametric supermodular mechanisms. This is yet another justification for indirect mechanisms, as nonparametric direct Bayesian mechanisms impose dominant-strategy incentive compatibility (Ledyard 1978).

Finally, it is important to pursue testing supermodular games. Since this paper provides a general framework, it is a good candidate for experimental tests. From a practical viewpoint, discretizing type spaces may simplify the players' task of announcing deceptions at each round. There are also simple environments with continuous types where announcing a strategy is equivalent to choosing a real number, such as the leading public goods.[29]

Appendix A

This example shows that the revelation principle fails to hold in general for supermodular Bayesian implementation.

Example 8. Consider two agents, 1 and 2, with type spaces $\Theta_1 = \{\theta_1^1, \theta_1^2\}$ and $\Theta_2 = \{\theta_2^1, \theta_2^2, \theta_2^3\}$. The prior assigns probability 1/6 to each $\theta \in \Theta$. Let $X = \{x_1, \ldots, x_{12}\}$ be

[29] In the public goods example of Section 3, announcing best reply comes down to choosing an intercept in a compact set (see (1)).

the outcome space. Agent 1's preferences are given by utility function $u_1(x_n, \theta_1)$:

	x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8	x_9	x_{10}	x_{11}	x_{12}
$u_1(\cdot, \theta_1^1)$	−10	0	16	−13	−2	33	−21	−2	18	−19	0	36
$u_1(\cdot, \theta_1^2)$	−10	0	16	−21	−2	18	−13	−2	33	−19	0	36

Agent 2 has constant utility function u_2. Let the social choice function f be defined as

$f(\cdot,\cdot)$	θ_2^1	θ_2^2	θ_2^3
θ_1^1	x_4	x_5	x_6
θ_1^2	x_7	x_8	x_9

Consider the mechanism $\Gamma = (\{M_1, M_2\}, g)$. Agent 1's message space is $M_1 = \{\underline{m}_1, m_1^1, m_1^2, \overline{m}_1\}$. Endow M_1 with $\succeq_1$ such that m_1^1 and m_1^2 are unordered, $\overline{m}_1$ is the greatest element, and $\underline{m}_1$ is the smallest element. Agent 2's message space is $M_2 = \{\underline{m}_2, m_2^1, \overline{m}_2\}$. Endow M_2 with $\succeq_2$ such that $\overline{m}_2 \succeq_2 m_2^1 \succeq_2 \underline{m}_2$. The outcome function g is given by

$g(\cdot,\cdot)$	$\underline{m}_2$	m_2^1	$\overline{m}_2$
$\underline{m}_1$	x_1	x_2	x_3
m_1^1	$f(\theta_1^1, \theta_2^1)$	$f(\theta_1^1, \theta_2^2)$	$f(\theta_1^1, \theta_2^3)$
m_1^2	$f(\theta_1^2, \theta_2^1)$	$f(\theta_1^2, \theta_2^2)$	$f(\theta_1^2, \theta_2^3)$
$\overline{m}_1$	x_{10}	x_{11}	x_{12}

I show that mechanism Γ supermodularly implements f. Consider strategy $m_2^*(\cdot)$ such that $m_2^*(\theta_2^1) = \underline{m}_2$, $m_2^*(\theta_2^2) = m_2^1$, and $m_2^*(\theta_2^3) = \overline{m}_2$. Given u_2 is constant, this strategy is a best response to any strategy of 1. For all m_1, we have

$$\sum_{m_2} u_1(g(m_1^1, m_2), \theta_1^1) > \sum_{m_2} u_1(g(m^1, m_2), \theta_1^1)$$

$$\sum_{m_2} u_1(g(m_1^2, m_2), \theta_1^2) > \sum_{m_2} u_1(g(m_1, m_2), \theta_1^2).$$

Thus 1's best response $m_1^*(\cdot)$ to $m_2^*(\cdot)$ is such that $m_1^*(\theta_1^1) = m_1^1$ and $m_1^*(\theta_1^2) = m_1^2$. So $(m_1^*(\cdot), m_2^*(\cdot))$ is a Bayesian equilibrium and $g \circ m^* = f$. Moreover, for each θ_1, $u_1(g(m_1, m_2), \theta_1)$ is supermodular in m_1 and has increasing differences in (m_1, m_2). So mechanism Γ supermodularly implements f.

Does this imply that there exists a direct mechanism $(\{\Theta_i\}, f)$ that truthfully implements f in supermodular game form? By means of contradiction, suppose there is such a mechanism. Then, for $(\Theta_1, \succeq_1)$ to be a lattice, it must be totally ordered. Assume $\theta_1^2 >_1 \theta_1^1$. Let $\theta_i^k(\cdot) = \theta_i^k$ regardless of i's true type. Let $\theta_1^T(\cdot)$ be the truthful strategy for 1 and let $\theta_1^L(\cdot)$ be constant lying. Note $\theta_1^1(\cdot) <_1 \theta_1^T(\cdot), \theta_1^L(\cdot)$. Since Θ_2 must be a lattice, θ_2^1 and θ_2^2 must be ordered. Hence $\theta_2^1(\cdot)$ and $\theta_2^2(\cdot)$ are ordered.

Let $u_1^f(\hat{\theta}_1(\cdot), \hat{\theta}_2(\cdot)) = E_\theta[u_1(f((\hat{\theta}_1(\cdot), \hat{\theta}_2(\cdot)), \theta_1)]$. Since the direct mechanism must induce a supermodular game, u_1^f must satisfy the single-crossing property in

$(\hat{\theta}_1(\cdot), \hat{\theta}_2(\cdot))$. Given

$$-2 = u_1^f(\theta_1^T(\cdot), \theta_2^2(\cdot)) \geq u_1^f(\theta_1^1(\cdot), \theta_2^2(\cdot)) = -2$$
$$-13 = u_1^f(\theta_1^T(\cdot), \theta_2^1(\cdot)) > u_1^f(\theta_1^1(\cdot), \theta_2^1(\cdot)) = -17,$$

u_1^f satisfies the single-crossing property in $(\hat{\theta}_1(\cdot), \hat{\theta}_2(\cdot))$ only if $\theta_2^1 >_2 \theta_2^2$. But

$$-2 = u_1^f(\theta_1^L(\cdot), \theta_2^2(\cdot)) \geq u_1^f(\theta_1^1(\cdot), \theta_2^2(\cdot)) = -2$$

does not imply $-21 = u_1^f(\theta_1^L(\cdot), \theta_2^1(\cdot)) \geq u_1^f(\theta_1^1(\cdot), \theta_2^1(\cdot)) = -17$. The single-crossing property is violated. A similar contradiction is reached if, instead, we assume $\theta_1^1 >_1 \theta_1^2$. Consequently, the single-crossing property is violated. The social choice function f is not truthfully supermodular implementable, although it is supermodular implementable. This example suggests that the conditions of Theorem 4 are somewhat minimally sufficient. ♢

Appendix B

Proof of Theorem 1. Sufficiency is immediate. So suppose that $f = (x, t)$ is (Bayesian) implementable. Since truthtelling is an equilibrium, it implies

$$\begin{aligned} &E_{\theta_{-i}}[V_i(x_i(\theta_i, \theta_{-i}), \theta_i)] + E_{\theta_{-i}}[t_i(\theta_i, \theta_{-i})] \\ &\qquad\qquad \geq E_{\theta_{-i}}[V_i(x_i(\hat{\theta}_i, \theta_{-i}), \theta_i)] + E_{\theta_{-i}}[t_i(\hat{\theta}_i, \theta_{-i})] \end{aligned} \tag{10}$$

for all $\hat{\theta}_i$. For $\rho_i \in \mathbb{R}$, let

$$\delta_i(\hat{\theta}_i, \hat{\theta}_{-i}) = \sum_{j \neq i} \rho_i \hat{\theta}_i \hat{\theta}_j$$

and define

$$t_i^{\mathrm{SM}}(\hat{\theta}_i, \hat{\theta}_{-i}) = \delta_i(\hat{\theta}_i, \hat{\theta}_{-i}) + E_{\theta_{-i}}[t_i(\hat{\theta}_i, \theta_{-i})] - E_{\theta_{-i}}[\delta_i(\hat{\theta}_i, \theta_{-i})].$$

Note that transfers t_i and t_i^{SM} have the same expected value: $E_{\theta_{-i}}[t_i^{\mathrm{SM}}(\cdot, \theta_{-i})] = E_{\theta_{-i}}[t_i(\cdot, \theta_{-i})]$. Thus, (x, t^{SM}) is implementable by (10). The payoffs $V_i + t_i^{\mathrm{SM}}$ satisfy the continuity requirements, because $\delta_i : \Theta \to \mathbb{R}$ is continuous and bounded, and $E_{\theta_{-i}}[t_i(\cdot, \theta_{-i})]$ is upper semicontinuous. Next, I show that it is possible to choose ρ_i so that $V_i(\cdot, \theta_i) + t_i^{\mathrm{SM}}(\cdot)$ has increasing differences in $(\hat{\theta}_i, \hat{\theta}_{-i})$ for all θ_i. Since substitutes are uniformly bounded, there exists T_i such that, for all $\theta_i'' \geq \theta_i'$ and $\theta_{-i}'' \geq \theta_{-i}'$,

$$\Delta V_i(\theta_i'', \theta_i', \theta_{-i}'', \theta_i) - \Delta V_i(\theta_i'', \theta_i', \theta_{-i}', \theta_i) \geq T_i(\theta_i'' - \theta_i') \sum_{j \neq i} (\theta_j'' - \theta_j') \tag{11}$$

for all $\theta_i \in \Theta_i$. Set $\rho_i > -T_i$. It follows from (11) that

$$\Delta V_i(\theta_i'', \theta_i', \theta_{-i}'', \theta_i) - \Delta V_i(\theta_i'', \theta_i', \theta_{-i}', \theta_i) + \rho_i(\theta_i'' - \theta_i') \sum_{j \neq i} (\theta_j'' - \theta_j') \geq 0$$

or, equivalently,

$$\Delta V_i(\theta_i'', \theta_i', \theta_{-i}'', \theta_i) - \Delta V_i(\theta_i'', \theta_i', \theta_{-i}', \theta_i) + \sum_{j\neq i} \rho_i(\theta_i''\theta_j'' + \theta_i'\theta_j' - \theta_i''\theta_j' - \theta_i'\theta_j'') \geq 0.$$

Thus, $V_i(\cdot, \theta_i) + t_i^{\mathrm{SM}}(\cdot)$ has increasing differences in $(\hat{\theta}_i, \hat{\theta}_{-i})$ for all θ_i. Finally, since Θ_i is a chain, u_i^f is supermodular in $\hat{\theta}_i(\cdot)$. □

The next proposition is a standard result, so the proof is omitted (see Proposition 23.D.2 in Mas-Colell et al. 1995).

Proposition 8. *Consider valuation functions and a decision rule x such that $E_{\theta_{-i}}[V_i(x_i(\hat{\theta}_i, \theta_{-i}), \theta_i)]$ is continuous in $(\hat{\theta}_i, \theta_i)$.*

(i) If the social choice function $f = (x, t)$ is implementable, then for all $\hat{\theta}_i$,

$$E_{\theta_{-i}}[t_i(\hat{\theta}_i, \theta_{-i})] = -E_{\theta_{-i}}[V_i(x_i(\hat{\theta}_i, \theta_{-i}), \hat{\theta}_i)] + \int_{\underline{\theta}_i}^{\hat{\theta}_i} \frac{\partial E_{\theta_{-i}}[V_i(x_i(s, \theta_{-i}), s)]}{\partial \theta_i}\, ds + \epsilon(\underline{\theta}_i). \tag{12}$$

(ii) Let the decision rule be such that $\partial E_{\theta_{-i}}[V_i(x_i(\hat{\theta}_i, \theta_{-i}), \theta_i)]/\partial\theta_i$ is increasing in $\hat{\theta}_i$ for each θ_i and i. If transfers t satisfy (12), then $f = (x, t)$ is implementable.

Proof of Proposition 2. Let (x, t'') and (x, t') be any supermodular implementable social choice functions such that $t'', t' \in \mathcal{T}$ and $t'' \succeq_{\mathrm{ID}} t'$. For any supermodular implementable social choice function, the induced game has a smallest and a greatest equilibrium along with a truthful equilibrium in between. Denote the truthtelling strategy by $\theta_i^T(\cdot)$, i.e., $\theta_i^T(\theta_i) = \theta_i$, and abusing notation, let $\underline{\theta}_i$ and $\overline{\theta}_i$ be constant strategies where agent i always announces his lowest or highest type. Let $\mathcal{G}_\ell$ be the same game as $\mathcal{G}$ except that the strategy spaces are restricted from $\Theta_i^{\Theta_i}$ to $[\underline{\theta}_i, \theta_i^T(\cdot)]$. Likewise, let $\mathcal{G}_u$ be the game $\mathcal{G}$, where the strategy spaces are restricted from $\Theta_i^{\Theta_i}$ to $[\theta_i^T(\cdot), \overline{\theta}_i]$. Since $\mathcal{G}$ is supermodular, by definition, those modified games $\mathcal{G}_\ell$ and $\mathcal{G}_u$ are also supermodular. As such, each of $\mathcal{G}_\ell$ and $\mathcal{G}_u$ has a largest and a smallest equilibrium. Furthermore, $\mathcal{G}_\ell$ has the same least equilibrium as game $\mathcal{G}$, and the largest equilibrium of $\mathcal{G}_\ell$ is truthtelling. Likewise, $\mathcal{G}_u$ has the same largest equilibrium as $\mathcal{G}$, and the smallest equilibrium of $\mathcal{G}_u$ is truthtelling. Define $U_i(\hat{\theta}(\cdot), t) = E_\theta[V_i(x_i(\hat{\theta}(\theta)), \theta_i) + t_i(\hat{\theta}(\theta))]$. I show that (i) $U_i(\hat{\theta}(\cdot), t)$ has decreasing differences in $(\hat{\theta}_i(\cdot), t)$ in game $\mathcal{G}_\ell$; (ii) $U_i(\hat{\theta}(\cdot), t)$ has increasing differences in $(\hat{\theta}_i(\cdot), t)$ in game $\mathcal{G}_u$. By Theorem 6 in Milgrom and Roberts (1990), this shows how the extremal equilibria in each modified game vary in response to change in transfers with respect to $\succeq_{\mathrm{ID}}$. In the end, since one extremal equilibrium is always truthtelling in each modified game, this shows how the untruthful equilibrium—which is actually an extremal equilibrium of $\mathcal{G}$—varies with respect to $\succeq_{\mathrm{ID}}$. Before proving (i) and (ii), note that Proposition 8 implies that all transfers t_i such that (x, t) is implementable have the same expected value $E_{\theta_{-i}}[t_i(\hat{\theta}_i, \theta_{-i})]$ up to a constant. Therefore, all implementing transfers can be written $\delta_i(\hat{\theta}_i, \hat{\theta}_{-i}) - E_{\theta_{-i}}[\delta_i(\hat{\theta}_i, \theta_{-i})] + E_{\theta_{-i}}[t_i(\hat{\theta}_i, \theta_{-i})]$ for some function $\delta_i : \Theta \to \mathbb{R}$, where t is any arbitrary transfers such that (x,t) is implementable. First, consider $\mathcal{G}_\ell$. Let $\delta'_i(\cdot)$ be the function such that $t'_i(\hat\theta) = \delta'_i(\hat\theta) - E_{\theta_{-i}}[\delta'_i(\hat\theta_i, \theta_{-i})] + E_{\theta_{-i}}[t_i(\hat\theta_i, \theta_{-i})]$ and define $\delta''_i(\cdot)$ similarly for t''. Choose any deceptions $\theta''_i(\cdot) > \theta'_i(\cdot)$ and any $\hat\theta_{-i}(\cdot)$ such that $\hat\theta_j(\theta_j) \le \theta_j$ for all θ_j and $j \ne i$, so we are in $\mathcal{G}_\ell$. Note $t'' \succeq_{\mathrm{ID}} t'$ implies $\delta'' \succeq_{\mathrm{ID}} \delta'$. That is, for all i and all θ,

$$\begin{aligned}&\delta''_i(\theta''_i(\theta_i), \theta_{-i}) - \delta''_i(\theta''_i(\theta_i), \hat\theta_{-i}(\theta_{-i})) - \delta''_i(\theta'_i(\theta_i), \theta_{-i}) + \delta''_i(\theta'_i(\theta_i), \hat\theta_{-i}(\theta_{-i}))\\ &\quad - \delta'_i(\theta''_i(\theta_i), \theta_{-i}) + \delta'_i(\theta''_i(\theta_i), \hat\theta_{-i}(\theta_{-i})) + \delta'_i(\theta'_i(\theta_i), \theta_{-i}) - \delta'_i(\theta'_i(\theta_i), \hat\theta_{-i}(\theta_{-i})) \ge 0,\end{aligned}$$

which implies

$$\begin{aligned}&E_\theta\big[\delta''_i(\theta''_i(\theta_i), \theta_{-i}) - \delta''_i(\theta''_i(\theta_i), \hat\theta_{-i}(\theta_{-i}))\big] - E_\theta\big[\delta''_i(\theta'_i(\theta_i), \theta_{-i}) - \delta''_i(\theta'_i(\theta_i), \hat\theta_{-i}(\theta_{-i}))\big]\\ &\quad - E_\theta\big[\delta'_i(\theta''_i(\theta_i), \theta_{-i}) - \delta'_i(\theta''_i(\theta_i), \hat\theta_{-i}(\theta_{-i}))\big] + E_\theta\big[\delta'_i(\theta'_i(\theta_i), \theta_{-i}) - \delta'_i(\theta'_i, \hat\theta_{-i}(\theta_{-i}))\big]\\ &\ge 0.\end{aligned} \tag{13}$$

But (13) is equivalent to

$$U_i(\theta''_i(\cdot), \hat\theta_{-i}(\cdot), t'') + U_i(\theta'_i(\cdot), \hat\theta_{-i}(\cdot), t') - U_i(\theta''_i(\cdot), \hat\theta_{-i}(\cdot), t') - U_i(\theta'_i(\cdot), \hat\theta_{-i}(\cdot), t'') \le 0$$

for all $\hat\theta_{-i}(\cdot)$. That is, $U_i(\hat\theta(\cdot), t)$ has decreasing differences in $(\hat\theta_i(\cdot), t)$ for each $\hat\theta_{-i}(\cdot)$. It follows from Theorem 6 in Milgrom and Roberts (1990) that the smallest equilibrium in $\mathcal{G}_\ell$ is decreasing in t. The same argument applies to $\mathcal{G}_u$. There, we look at all strategies such that $\hat\theta_j(\theta_j) \ge \theta_j$ for all θ_j and $j \ne i$. As a result, the sign in (13) is reversed, which implies $U_i(\hat\theta(\cdot), t)$ has increasing differences in $(\hat\theta_i(\cdot), t)$ for each $\hat\theta_{-i}(\cdot)$. The greatest equilibrium in $\mathcal{G}_u$ is thus increasing in t. □

Proof of Theorem 3. Suppose $f = (x,t)$ is implementable and x is dimensionally reducible. Letting

$$\delta_i(\hat\theta_i, \hat\theta_{-i}) = -\int_{\underline{\theta}_i}^{\hat\theta_i} \int_{r_i(\underline{\theta}_{-i})}^{r_i(\hat\theta_{-i})} \min_{\theta_i \in \Theta_i} \frac{\partial^2 V_i(h_i(s_i, r_i), \theta_i)}{\partial r_i\, \partial s_i}\, dr_i\, ds_i \tag{14}$$

for all $\hat\theta \in \Theta$, I show that

$$t^*_i(\hat\theta_i, \hat\theta_{-i}) = \delta_i(\hat\theta_i, \hat\theta_{-i}) - E_{\theta_{-i}}[\delta_i(\hat\theta_i, \theta_{-i})] + E_{\theta_{-i}}[t_i(\hat\theta_i, \theta_{-i})] \tag{15}$$

is well defined and that (x, t^*) is optimally supermodular implementable. By Proposition 1, $E_{\theta_{-i}}[t_i(\hat\theta_i, \theta_{-i})]$ is well defined and given by (12). Since V_i and h_i are C^2 on an open set containing compact set Θ_i, $\min_{\theta_i\in\Theta_i} \partial^2 V_i(h_i(s_i, r_i), \theta_i)/\partial r_i\, \partial s_i$ exists, is continuous in (r_i, s_i) by the Maximum Theorem, and is bounded; hence it is integrable and $\delta_i : \Theta \to \mathbb{R}$ is continuous. Since δ_i is also bounded, $E_{\theta_{-i}}[\delta_i(\cdot, \theta_{-i})]$ is well defined and so is $t^*_i : \Theta \to \mathbb{R}$. The next step is to verify the continuity requirements. As a continuous function on a compact set, δ_i is uniformly continuous in $\hat\theta$, and so $E_{\theta_{-i}}[\delta(\hat\theta_i, \theta_{-i})]$

is continuous in $\hat{\theta}_i$. Since V_i is C^2, $E_{\theta_{-i}}[t_i(\hat{\theta}_i, \theta_{-i})]$, defined as (12), is upper semicontinuous in $\hat{\theta}_i$. Put together, the utility functions satisfy all the continuity requirements. Finally, I prove that (x, t^*) is optimally supermodular implementable. Note that $E_{\theta_{-i}}[t_i^*(\hat{\theta}_i, \theta_{-i})] = E_{\theta_{-i}}[t_i(\hat{\theta}_i, \theta_{-i})]$ and thus (x, t^*) is implementable. By construction, t_i^* is twice-differentiable and

$$\begin{aligned}\frac{\partial^2 t_i^*(\hat{\theta}_i, \hat{\theta}_{-i})}{\partial\hat{\theta}_i\,\partial\hat{\theta}_j} &= \frac{\partial^2 \delta_i(\hat{\theta}_i, \hat{\theta}_{-i})}{\partial\hat{\theta}_i\,\partial\hat{\theta}_j}\\ &= \frac{\partial}{\partial\hat{\theta}_j}\int_{r_i(\underline{\theta}_{-i})}^{r_i(\hat{\theta}_{-i})} -\min_{\theta_i\in\Theta_i}\frac{\partial^2 V_i(h_i(\hat{\theta}_i, r_i), \theta_i)}{\partial r_i\,\partial s_i}\,dr_i\\ &= -\left(\min_{\theta_i\in\Theta_i}\frac{\partial^2 V_i(h_i(\hat{\theta}_i, r_i(\hat{\theta}_{-i})), \theta_i)}{\partial r_i\,\partial s_i}\right)\frac{\partial r_i(\hat{\theta}_{-i})}{\partial\hat{\theta}_j}.\end{aligned}$$

Since $r_i(\cdot)$ is an increasing function,

$$\begin{aligned}-\min_{\theta_i\in\Theta_i}\frac{\partial^2 V_i(x_i(\hat{\theta}_i, \hat{\theta}_{-i}), \theta_i)}{\partial\hat{\theta}_i\,\partial\hat{\theta}_j} &\equiv -\min_{\theta_i\in\Theta_i}\left(\frac{\partial^2 V_i(h_i(\hat{\theta}_i, r_i(\hat{\theta}_{-i})), \theta_i)}{\partial r_i\,\partial s_i}\frac{\partial r_i(\hat{\theta}_{-i})}{\partial\hat{\theta}_j}\right)\\ &= -\left(\min_{\theta_i\in\Theta_i}\frac{\partial^2 V_i(h_i(\hat{\theta}_i, r_i(\hat{\theta}_{-i})), \theta_i)}{\partial r_i\,\partial s_i}\right)\frac{\partial r_i(\hat{\theta}_{-i})}{\partial\hat{\theta}_j}\\ &= \frac{\partial^2 t_i^*(\hat{\theta}_i, \hat{\theta}_{-i})}{\partial\hat{\theta}_i\,\partial\hat{\theta}_j}.\end{aligned}$$

Therefore,

$$\begin{aligned}\frac{\partial^2 (V_i(x_i(\hat{\theta}), \theta_i) + t_i^*(\hat{\theta}))}{\partial\hat{\theta}_i\,\partial\hat{\theta}_j} &\equiv \frac{\partial^2 V_i(x_i(\hat{\theta}_i, \hat{\theta}_{-i}), \theta_i)}{\partial\hat{\theta}_i\,\partial\hat{\theta}_j} + \frac{\partial t_i^*(\hat{\theta})}{\partial\hat{\theta}_i\,\partial\hat{\theta}_j}\\ &= \frac{\partial^2 V_i(x_i(\hat{\theta}_i, \hat{\theta}_{-i}), \theta_i)}{\partial\hat{\theta}_i\,\partial\hat{\theta}_j} - \min_{\theta_i\in\Theta_i}\frac{\partial^2 V_i(x_i(\hat{\theta}_i, \hat{\theta}_{-i}), \theta_i)}{\partial\hat{\theta}_i\,\partial\hat{\theta}_j} \geq 0\end{aligned}$$

for all $\hat{\theta}$, θ_i and j, i, and so (x, t^*) is supermodular implementable. Moreover, for all transfers $t \in \mathcal{T}$ such that (x, t) is supermodular implementable, it must be that

$$\frac{\partial^2 t_i(\hat{\theta})}{\partial\hat{\theta}_i\,\partial\hat{\theta}_j} \geq -\min_{\theta_i\in\Theta_i}\frac{\partial^2 V_i(x_i(\hat{\theta}_i, \hat{\theta}_{-i}), \theta_i)}{\partial\hat{\theta}_i\,\partial\hat{\theta}_j} = \frac{\partial^2 t_i^*(\hat{\theta})}{\partial\hat{\theta}_i\,\partial\hat{\theta}_j}$$

for all $\hat{\theta}$ and j, i. This implies that (x, t^*) is optimally supermodular implementable. □

Proof of Theorem 3. (i) Since the social choice function is supermodular implementable, there exist a greatest and a smallest equilibrium in the game induced by the mechanism. By way of contradiction, suppose that the truthful equilibrium is not the unique Bayesian equilibrium. Then one of the extremal equilibria must be strictly greater/smaller than the truthful one. Suppose that the greatest equilibrium, denoted

$(\overline{\theta}_i(\cdot))_{i\in N}$, is strictly greater than the truthful equilibrium. That is, for all i, $\overline{\theta}_i(\theta_i) \geq \theta_i$ for almost everywhere θ_i and there exists $N^* \neq \varnothing$ such that, for all $i \in N^*$, $\overline{\theta}_i(\theta_i) > \theta_i$ for all θ_i in some subset of types with positive measure.

I evaluate the first-order condition of agent i's maximization program at the greatest equilibrium. Then, I bound it from above by an expression that cannot be positive for all players (hence the contradiction). Consider player i's interim utility at type θ_i against $\overline{\theta}_{-i}(\cdot)$:

$$E_{\theta_{-i}}\big[V_i(x_i(\hat{\theta}_i, \overline{\theta}_{-i}(\theta_{-i})), \theta_i)\big] + E_{\theta_{-i}}[t_i(\hat{\theta}_i, \overline{\theta}_{-i}(\theta_{-i}))]. \tag{16}$$

Since V_i, x_i, and t_i (by Proposition 8) are continuously differentiable, we can show that for any deception $\hat{\theta}_{-i}(\cdot)$, the first derivative of (16) with respect to $\hat{\theta}_i$ is

$$E_{\theta_{-i}}\bigg[\frac{\partial V_i(x_i(\hat{\theta}_i, \overline{\theta}_{-i}(\theta_{-i})), \theta_i)}{\partial \hat{\theta}_i}\bigg] + E_{\theta_{-i}}\bigg[\frac{\partial t_i(\hat{\theta}_i, \overline{\theta}_{-i}(\theta_{-i}))}{\partial \hat{\theta}_i}\bigg]. \tag{17}$$

By assumption, the utility functions and the decision rule produce bounded complements, so we have

$$\begin{aligned} E_{\theta_{-i}}\bigg[&\frac{\partial V_i(x_i(\hat{\theta}_i, \overline{\theta}_{-i}(\theta_{-i})), \theta_i)}{\partial \hat{\theta}_i} + \frac{\partial t_i(\hat{\theta}_i, \overline{\theta}_{-i}(\theta_{-i}))}{\partial \hat{\theta}_i} \\ &\qquad - \frac{\partial V_i(x_i(\hat{\theta}_i, \theta_{-i}), \theta_i)}{\partial \hat{\theta}_i} - \frac{\partial t_i(\hat{\theta}_i, \theta_{-i})}{\partial \hat{\theta}_i}\bigg] \\ &\leq \int_{\Theta_{-i}} K_i(\theta_i) \sum_{j\neq i} (\overline{\theta}_j(\theta_j) - \theta_j)\phi_{-i}(\theta_{-i})\, d\theta_{-i} = K_i(\theta_i) \sum_{j\neq i} E_{\theta_j}[\overline{\theta}_j(\theta_j) - \theta_j]. \end{aligned} \tag{18}$$

By (18),

$$(17) \leq K_i(\theta_i)\sum_{j\neq i} E_{\theta_j}[\overline{\theta}_j(\theta_j) - \theta_j] + E_{\theta_{-i}}\bigg[\frac{\partial V_i(x_i(\hat{\theta}_i, \theta_{-i}), \theta_i)}{\partial \hat{\theta}_i}\bigg] + E_{\theta_{-i}}\bigg[\frac{\partial t_i(\hat{\theta}_i, \theta_{-i})}{\partial \hat{\theta}_i}\bigg]. \tag{19}$$

By part (i) of Proposition 8,

$$E_{\theta_{-i}}\bigg[\frac{\partial t_i(\hat{\theta}_i, \theta_{-i})}{\partial \hat{\theta}_i}\bigg] = -E_{\theta_{-i}}\bigg[\frac{\partial V_i(x_i(\theta_i', \theta_{-i}), \hat{\theta}_i)}{\partial \theta_i'}\bigg|_{\theta_i' = \hat{\theta}_i}\bigg].$$

Therefore, (19) implies

$$\begin{aligned} (17) &\leq K_i(\theta_i)\sum_{j\neq i} E_{\theta_j}[\overline{\theta}_j(\theta_j) - \theta_j] + E_{\theta_{-i}}\bigg[\frac{\partial V_i(x_i(\hat{\theta}_i, \theta_{-i}), \theta_i)}{\partial \hat{\theta}_i} - \frac{\partial V_i(x_i(\hat{\theta}_i, \theta_{-i}), \hat{\theta}_i)}{\partial \theta_i'}\bigg] \\ &\leq K_i(\theta_i)\sum_{j\neq i} E_{\theta_j}[\overline{\theta}_j(\theta_j) - \theta_j] + E_{\theta_{-i}}[\gamma_i(\theta_{-i})](\theta_i - \hat{\theta}_i), \end{aligned} \tag{20}$$

where the last inequality follows from strong differences.

Since it is optimal for each player i to play $\overline{\theta}_i(\theta_i)$ for almost everywhere type θ_i, then the right-hand side of (20), evaluated at $\hat{\theta}_i = \overline{\theta}_i(\theta_i)$, must be positive for almost everywhere θ_i and all i. To see why, let $\Theta_i^* \subset \Theta_i$ be the set of θ_i for which the right-hand side of (20) is strictly negative when playing $\overline{\theta}_i(\theta_i)$. Note that Θ_i^* is measurable by definition, because the right-hand side of (20) is a measurable function in θ_i when plugging in $\overline{\theta}_i(\theta_i)$. If there were a player i for whom Θ_i^* had strictly positive measure, then playing $\overline{\theta}_i(\theta_i)$ would lead (17) to be strictly negative for all $\theta_i \in \Theta_i^*$. But for types in Θ_i^*, the player can announce types in $[\underline{\theta}_i, \overline{\theta}_i(\theta_i)]$ and so she would strictly prefer playing $\theta_i^*(\theta_i) = \overline{\theta}_i(\theta_i) - \varepsilon \mathbf{1}_{\Theta_i^*}$ for some small ε.[30]

Since the right-hand side of (20) is positive for almost everywhere θ_i when playing $\overline{\theta}_i(\theta_i)$, then it must be true in expectation for all i that

$$0 \le E_{\theta_i}[K_i(\theta_i)] \sum_{j \neq i} E_{\theta_j}[\overline{\theta}_j(\theta_j) - \theta_j] + E_{\theta_{-i}}[\gamma_i(\theta_{-i})] E_{\theta_i}[\theta_i - \overline{\theta}_i(\theta_i)]. \tag{21}$$

Letting

$$\mathcal{C} = \begin{pmatrix} -E_{\theta_{-1}}[\gamma_1(\theta_{-1})] & E_{\theta_1}[K_1(\theta_1)] & \cdots & E_{\theta_1}[K_1(\theta_1)] \\ E_{\theta_2}[K_2(\theta_2)] & -E_{\theta_{-2}}[\gamma_2(\theta_{-2})] & \cdots & E_{\theta_2}[K_2(\theta_2)] \\ \vdots & \vdots & \ddots & \vdots \\ E_{\theta_n}[K_n(\theta_n)] & E_{\theta_n}[K_n(\theta_n)] & \cdots & -E_{\theta_{-n}}[\gamma_n(\theta_{-n})] \end{pmatrix},$$

(21) implies the existence of a positive solution w^* to the system $\mathcal{C} \cdot w \ge 0$. But then it must be that $w^{*T} \mathcal{C} \cdot w^* \ge 0$, a contradiction because $\mathcal{C}$ is the matrix of complementarities and it is negative-definite. The same argument applies to show that there is no equilibrium that is smaller than the truthful equilibrium.

(ii) Let a_i and b_i be real numbers such that $K_i(\theta_i) \le a_i\theta_i + b_i$ for all θ_i. Define

$$Q_i(\hat{\theta}_i, \hat{\theta}_{-i}(\cdot), \theta_i) = (a_i\theta_i + b_i)\hat{\theta}_i \sum_{j \neq i} E_{\theta_j}[\hat{\theta}_j(\theta_j) - \theta_j] + E_{\theta_{-i}}[\gamma_i(\theta_{-i})]\hat{\theta}_i\Big(\theta_i - \tfrac{1}{2}\hat{\theta}_i\Big). \tag{22}$$

From now on, only consider strategies where all agents announce types above their true types (see the proof of Proposition 2 for a similar restriction). Following the same reasoning as the one leading to (20), we have

$$\frac{\partial Q_i(\hat{\theta}_i, \hat{\theta}_{-i}(\cdot), \theta_i)}{\partial \hat{\theta}_i} \ge \frac{\partial u_i(\hat{\theta}_i, \hat{\theta}_{-i}(\cdot), \theta_i)}{\partial \hat{\theta}_i}.$$

This implies that the game $(N, \{\Theta_i^{\Theta_i}, Q_i\})$ has a greatest equilibrium[31] that is larger than the greatest equilibrium of $\mathcal{G}$ (the argument is the same as in the proof of Proposition 2). However, given the quadratic nature of the payoffs in $(N, \{\Theta_i^{\Theta_i}, Q_i\})$, the greatest equilibrium in this game is easier to compute than in $\mathcal{G}$. Now I show how to compute it. Let

$$A_i = 1 + \frac{a_i(\sum_{j \neq i} E_{\theta_j}[\hat{\theta}_j(\theta_j)] - E_{\theta_j}[\theta_j])}{E_{\theta_{-i}}[\gamma_i(\theta_{-i})]}, \qquad B_i = \frac{b_i(\sum_{j \neq i} E_{\theta_j}[\hat{\theta}_j(\theta_j)] - E_{\theta_j}[\theta_j])}{E_{\theta_{-i}}[\gamma_i(\theta_{-i})]}. \tag{23}$$

[30] Note that $\theta_i^*(\cdot) \in \Sigma_i(\Theta_i)$ because $\overline{\theta}_i(\cdot) \in \Sigma_i(\Theta_i)$, so $\theta_i^*(\cdot)$ is a possible choice of deception.

[31] From (22), this game is obviously supermodular.

From (22) (and since $K_i(\theta_i) \geq 0$), i's best response to $\hat{\theta}_{-i}(\cdot)$ in $(N, \{\Theta_i^{\Theta_i}, Q_i\})$ is

$$\text{br}_i[\hat{\theta}_{-i}(\cdot)] = \begin{cases} A_i\theta_i + B_i & \text{if } \underline{\theta}_i \leq \theta_i \leq \dfrac{\overline{\theta}_i - B_i}{A_i} \\ \overline{\theta}_i & \text{otherwise.} \end{cases} \tag{24}$$

Let

$$\overline{e}_j[A_j, B_j] = A_j \int_{\underline{\theta}_j}^{(\overline{\theta}_j - B_j)/A_j} \theta_j \phi_j(\theta_j)\, d\theta_j + \Phi_j\left(\frac{\overline{\theta}_j - B_j}{A_j}\right)(B_j - \overline{\theta}_j) + \overline{\theta}_j.$$

From (24), compute the expected value of j's best response, and note that

$$E_{\theta_j}[\text{br}_j[\hat{\theta}_{-j}(\cdot)]] = \overline{e}_j[A_j, B_j].$$

All equilibria in $(N, \{\Theta_i^{\Theta_i}, Q_i\})$ are piecewise affine functions of the form (24) and each is characterized by a vector $(A_i^*, B_i^*)_i$ that satisfies (23):

$$A_i^* = 1 + \frac{a_i(\sum_{j\neq i} \overline{e}_j[A_j^*, B_j^*] - E_{\theta_j}[\theta_j])}{E_{\theta_{-i}}[\gamma_i(\theta_{-i})]}, \qquad B_i^* = \frac{b_i(\sum_{j\neq i} \overline{e}_j[A_j^*, B_j^*] - E_{\theta_j}[\theta_j])}{E_{\theta_{-i}}[\gamma_i(\theta_{-i})]}.$$

To find the upper bound on the equilibrium set of $\mathcal{G}$, we need to solve the following system of $2n$ equations for $(A_i, B_i)_i$:

$$\begin{aligned} A_i &= 1 + \frac{a_i(\sum_{j\neq i} \overline{e}_j[A_j, B_j] - E_{\theta_j}[\theta_j])}{E_{\theta_{-i}}[\gamma_i(\theta_{-i})]} \\ B_i &= \frac{b_i(\sum_{j\neq i} \overline{e}_j[A_j, B_j] - E_{\theta_j}[\theta_j])}{E_{\theta_{-i}}[\gamma_i(\theta_{-i})]}. \end{aligned} \tag{25}$$

This system has at least one solution where $A_i = 1$ and $B_i = 0$ for $i = 1, \ldots, n$ and it corresponds to the truthtelling equilibrium. There also exists a solution that corresponds to the greatest equilibrium of $(N, \{\Theta_i^{\Theta_i}, Q_i\})$. This solution defines strategies that bound the equilibrium set of $\mathcal{G}$ from above. To finish the proof, we apply the same argument, but we restrict agents to play strategies below truthtelling. That is, we consider only those strategies where all agents announce types below their true types. We do so, because the smallest equilibrium of $(N, \{\Theta_i^{\Theta_i}, Q_i\})$ is smaller than the smallest equilibrium of $\mathcal{G}$. Consider best responses of the form

$$\text{br}_i[\hat{\theta}_{-i}(\cdot)] = \begin{cases} \underline{\theta}_i & \text{if } \underline{\theta}_i < \theta_i < \underline{\theta}_i - \dfrac{B_i}{A_i} \\ A_i\theta_i + B_i & \text{otherwise,} \end{cases} \tag{26}$$

where $B_i \leq 0$. Let

$$\underline{e}_j[A_j, B_j] = A_j \int_{\underline{\theta}_j - B_j/A_j}^{\overline{\theta}_j} \theta_j \phi_j(\theta_j)\, d\theta_j + B_j + \Phi_j\left(\underline{\theta}_j - \frac{B_j}{A_j}\right)(\underline{\theta}_j - B_j).$$

A similar argument to the above shows that the smallest solution to the following system provides a lower bound for the equilibrium set of $\mathcal{G}$: For $i = 1, \dots, n$,

$$\begin{aligned} A_i &= 1 + \frac{a_i(\sum_{j\neq i} \underline{e}_j[A_j, B_j] - E_{\theta_j}[\theta_j])}{E_{\theta_{-i}}[\gamma_i(\theta_{-i})]} \\ B_i &= \frac{b_j(\sum_{j\neq i} \underline{e}_j[A_j, B_j] - E_{\theta_j}[\theta_j])}{E_{\theta_{-i}}[\gamma_i(\theta_{-i})]}. \end{aligned} \tag{27}$$

□

Proof of Corollary 3. By assumption, $E_{\theta_{-i}}[V_i(x_i(\hat{\theta}_i, \theta_{-i}), \theta_i)]$ is continuous in $(\hat{\theta}_i, \theta_i)$. So Proposition 8 and Theorem 3 imply that (x, t^*) is supermodular implementable, where t_i^*, $i = 1, \dots, n$, is defined by (15) using (12). By construction, the mechanism $(\{\Theta_i\}, (x, t^*))$ produces bounded complements, where each bound κ_i on complements is computed as

$$\kappa_i = \max_{j\neq i} \max_{(\hat{\theta},\theta_i)\in\Theta\times\Theta_i} \left(\frac{\partial^2 V_i(x_i(\hat{\theta}), \theta_i)}{\partial\hat{\theta}_i\,\partial\hat{\theta}_j} - \min_{\theta_i\in\Theta_i} \frac{\partial^2 V_i(x_i(\hat{\theta}), \theta_i)}{\partial\hat{\theta}_i\,\partial\hat{\theta}_j} \right).$$

Since the valuation functions and the decision rule are twice-continuously differentiable, the assumption of strong differences holds. Theorem 3 completes the proof. □

Proof of Theorem 2. Let

$$H_i(\hat{\theta}_{-i}) = -\left(\frac{1}{n-1}\right) \sum_{j\neq i} E_{\tilde{\theta}_{-j}}\left[\sum_{k\neq j} V_k(x_k(\hat{\theta}_j, \tilde{\theta}_{-j}), \tilde{\theta}_k)\right]$$

and, for $\rho_i \in \mathbb{R}$, let

$$\delta_i(\hat{\theta}_i, \hat{\theta}_{-i}) = \sum_{j\neq i} \rho_i \hat{\theta}_i \hat{\theta}_j.$$

Define

$$\begin{aligned} t_i^{\mathrm{BB}}(\hat{\theta}_i, \hat{\theta}_{-i}) = {} & \delta_i(\hat{\theta}_i, \hat{\theta}_{-i}) - E_{\theta_{-i}}[\delta_i(\hat{\theta}_i, \theta_{-i})] + E_{\tilde{\theta}_{-i}}\left[\sum_{j\neq i} V_j(x_j(\hat{\theta}_i, \tilde{\theta}_{-i}), \tilde{\theta}_j)\right] \\ & + H_i(\hat{\theta}_{-i}) - \frac{1}{n-2} \sum_{j\neq i} \sum_{k\neq i,j} \rho_j \hat{\theta}_j \hat{\theta}_k + \frac{1}{n-2} \sum_{j\neq i} \sum_{k\neq i,j} \rho_j \hat{\theta}_j E(\theta_k). \end{aligned} \tag{28}$$

First, (x, t^{BB}) is implementable, because x is allocation-efficient and

$$E_{\theta_{-i}}[t_i^{\mathrm{BB}}(\hat{\theta}_i, \theta_{-i})] = E_{\tilde{\theta}_{-i}}\left[\sum_{j\neq i} V_j(x_j(\hat{\theta}_i, \tilde{\theta}_{-i}), \tilde{\theta}_j)\right] + E_{\theta_{-i}}[H_i(\theta_{-i})],$$

which is the expectation of the transfers in the expected externality mechanism. Second, since for all θ, we have

$$\sum_{i\in N}\left(\delta_i(\theta_i, \theta_{-i}) - \frac{1}{n-2} \sum_{j\neq i} \sum_{k\neq i,j} \rho_j \theta_j \theta_k\right) = \sum_{i\in N} \delta_i(\theta_i, \theta_{-i}) - \frac{1}{n-2} \sum_{i\in N} \sum_{j\neq i} (n-2) \rho_i \theta_i \theta_j = 0$$

and

$$\sum_{i\in N}\left(\frac{1}{n-2}\sum_{j\neq i}\sum_{k\neq i,j}\rho_j\theta_j E(\theta_k)-E_{\theta_{-i}}[\delta_i(\theta_i,\theta_{-i})]\right)$$
$$=\frac{1}{n-2}\sum_{i\in N}\sum_{j\neq i}(n-2)\rho_i\theta_i E(\theta_j)-\sum_{i\in N}E_{\theta_{-i}}[\delta_i(\theta_i,\theta_{-i})]=0,$$

then

$$\sum_{i\in N}t_i^{\mathrm{BB}}(\theta)=\sum_{i\in N}E_{\tilde\theta_{-i}}\left[\sum_{j\neq i}V_j(x_j(\theta_i,\tilde\theta_{-i}),\tilde\theta_j)\right]+\sum_{i\in N}H_i(\theta_{-i})=0,$$

where the last equality follows from balancedness of transfers in the expected externality mechanism. The continuity requirements are clearly satisfied, as t_i^{BB} is continuous in $\hat\theta_{-i}$ for each $\hat\theta_i$ and is upper semicontinuous in $\hat\theta_i$ for each $\hat\theta_{-i}$. Next, I show how to pick ρ_i to have increasing differences in $(\hat\theta_i,\hat\theta_{-i})$. Since substitutes are uniformly bounded, there exists T_i such that, for all $\theta_i''\geq\theta_i'$ and $\theta_{-i}''\geq\theta_{-i}'$, $\Delta V_i(\theta_i'',\theta_i',\theta_{-i}'',\theta_i)-\Delta V_i(\theta_i'',\theta_i',\theta_{-i}',\theta_i)\geq T_i(\theta_i''-\theta_i')\sum(\theta_j''-\theta_j')$ for all $\theta_i\in\Theta_i$. Set $\rho_i>-T_i$. Choose any $\theta_{-i}''\geq_{-i}\theta_{-i}'$ and $\theta_i''>\theta_i'$. From (28), note that

$$t_i^{\mathrm{BB}}(\theta_i'',\theta_{-i}'')-t_i^{\mathrm{BB}}(\theta_i'',\theta_{-i}')-t_i^{\mathrm{BB}}(\theta_i',\theta_{-i}'')+t_i^{\mathrm{BB}}(\theta_i',\theta_{-i}')$$
$$=\delta_i(\theta_i'',\theta_{-i}'')-\delta_i(\theta_i'',\theta_{-i}')-\delta_i(\theta_i',\theta_{-i}'')+\delta_i(\theta_i',\theta_{-i}').$$

Therefore, $V_i(\cdot,\theta_i)+t_i^{\mathrm{BB}}(\cdot)$ has increasing differences in $(\hat\theta_i,\hat\theta_{-i})$ if for all θ_i,

$$\Delta V_i(\theta_i'',\theta_i',\theta_{-i}'',\theta_i)-\Delta V_i(\theta_i'',\theta_i',\theta_{-i}',\theta_i)+\sum_{j\neq i}\rho_i(\theta_i''\theta_j''+\theta_i'\theta_j'-\theta_i''\theta_j'-\theta_i'\theta_j'')\geq 0.$$

That this inequality holds follows similarly to the proof of Theorem 1. □

Proof of Proposition 5. Take any $\epsilon>0$ such that $\epsilon<E_{\theta_{-i}}[\gamma_i(\theta_{-i})]+T_i-\tau_i$. Define balanced transfers t_i^{BB} as in (28) with $\rho_i=-T_i+\epsilon$. Since substitutes are bounded by T_i, $\partial^2V_i(\cdot,\theta_i)+t_i^{\mathrm{BB}}(\cdot)/\partial\hat\theta_i\,\partial\hat\theta_j\geq\epsilon$; hence (x,t^{BB}) is supermodular implementable. Since the valuation functions and the decision rule yield complements that are bounded by τ_i, the mechanism $(\{\Theta_i\},(x,t^*))$ generates complements that are bounded by $\tau_i-T_i+\epsilon$. The matrix of complementarities produced by this mechanism must be negative-definite, because $\tau_i-T_i+\epsilon<E_{\theta_{-i}}[\gamma_i(\theta_{-i})]$ by assumption. Proposition 3 completes the proof. □

Proof of Proposition 6. Let $O\supset\Theta$ be some open set. Define the extension of $x(\cdot)$ from Θ to O. For any $\theta\in O$, let $N^*(\theta)=\{j\in N:\theta_j\in[\underline\theta_j,\overline\theta_j]\}$, $\underline N(\theta)=\{j\in N:\theta_j<\underline\theta_j\}$, and $\overline N(\theta)=\{j\in N:\theta_j>\overline\theta_j\}$. Let $p_i:O\to\Theta_i$ be defined as

$$p_i(\theta)=\begin{cases}\theta_i & \text{if } i\in N^*(\theta)\\ \overline\theta_i & \text{if } i\in\overline N(\theta)\\ \underline\theta_i & \text{if } i\in\underline N(\theta).\end{cases}$$

The extension of $x(\cdot)$, denoted x^e, is the function $x^e_{(i,k)}(\theta) = x_{(i,k)}(p_1(\theta), \dots, p_n(\theta))$ for all $\theta \in O$, dimension k, and $i \in N$. Note that $x^e_{(i,k)} \in L_p(O)$ and it is increasing in $\hat{\theta}_i$ because $x_{(i,k)}$ is increasing in $\hat{\theta}_i$. By Theorem 12.10 in Aliprantis and Border (1999), the space of C^2 functions on O is norm dense in $L_p(O)$; hence there exists a sequence $\{x_n\}$ of C^2 functions from O into $\mathbb{R}$ such that $\lim_{n\to\infty}(\int_O |x_{n,(i,k)} - x^e_{(i,k)}|^p)^{1/p} = 0$ for all k and i. This implies $\lim_{n\to\infty}(\int_\Theta |x_{n,(i,k)} - x_{(i,k)}|^p)^{1/p} = 0$ for all k and all i. Moreover, we can take $\{x_n\}$ such that $x_{n,(i,k)}$ is increasing in θ_i on O_i for all k and i.[32] Since each Θ_i is compact, and V_i and x_n are C^2, then they form a continuous family, $\partial E_{\theta_{-i}}[V_i(x_{n,(i,k)}(\hat{\theta}), \theta_i)]/\partial\theta_i = E_{\theta_{-i}}[\partial V_i(x_{n,(i,k)}(\hat{\theta}), \theta_i)/\partial\theta_i]$ is increasing in $\hat{\theta}_i$ on Θ_i, and substitutes are bounded. Proposition 1 and Theorem 1 imply that, for all n, there exist t_n^{SM} such that $f = (x_n, t_n^{\text{SM}})$ is supermodular implementable. □

Proof of Proposition 7. I start by approximating the functions $h_{(i,k)}: \mathbb{R}^2 \to \mathbb{R}$ and $r_i: \mathbb{R}^n \to \mathbb{R}$ by C^2 functions. Then I study the convergence of the resulting composite function. Let μ^n denote the Lebesgue measure on $\mathbb{R}^n$. Because type sets are compact and h_i is bounded, Theorems 12.10 and 12.6 in Aliprantis and Border (1999) guarantee that there exists a sequence $\{h^m_{(i,k)}\}$ of C^2 functions that converges pointwise to $h_{(i,k)}$ for μ^2-almost all (θ_i, r_i). Consider function $r_i(\cdot)$ for any $i \in N$. By the Stone–Weierstrass Theorem, there exists a sequence of C^2-increasing functions $\{r_i^q\}$ that uniformly converges to r_i.[33] The triangle inequality gives

$$\begin{aligned}\int_\Theta \left|h^m_{(i,k)}(\theta_i, r_i^q(\theta_{-i})) - h_{(i,k)}(\theta_i, r_i(\theta_{-i}))\right| d\mu^n &\\ \leq \int_\Theta \left|h^m_{(i,k)}(\theta_i, r_i^q(\theta_{-i})) - h^m_{(i,k)}(\theta_i, r_i(\theta_{-i}))\right| d\mu^n &\qquad (29)\\ + \int_\Theta \left|h^m_{(i,k)}(\theta_i, r_i(\theta_{-i})) - h_{(i,k)}(\theta_i, r_i(\theta_{-i}))\right| d\mu^n. &\end{aligned}$$

The next step is to demonstrate that the second integral on the right-hand side of (29) converges to zero, as a result of the μ^2-almost everywhere convergence of $h^m_{(i,k)}$.[34] Note that

$$\begin{aligned}\int_\Theta \left|h^m_{(i,k)}(\theta_i, r_i(\theta_{-i})) - h_{(i,k)}(\theta_i, r_i(\theta_{-i}))\right| d\mu^n &\\ = \int_{\Theta_i \times r_i(\Theta_{-i})} |h^m_{(i,k)}(\theta_i, t) - h_{(i,k)}(\theta_i, t)|\, d\mu \times \mu_{r_i}, &\qquad (30)\end{aligned}$$

[32] Since $x_{i,k}$ is increasing in θ_i, it is always possible to take the members of the approximating sequence to be increasing (see Mas-Colell 1974).

[33] Since r_i is increasing, recall that we can take the members of the approximating sequence to be increasing.

[34] This is indeed not immediate. Suppose $\lim_{m\to\infty} h^m_{(i,k)}(\theta_i, r_i) = h_{(i,k)}(\theta_i, r_i)$ for μ^2-almost everywhere points in $\mathbb{R}^2$, except on the zero-measure set $\{(\theta_i, r_i^*): \theta_i \in I\}$, where I is some interval. If $r_i(\theta_i) = r_i^*$ for all $\theta_i \in I$, then $\int_\Theta |h^m_{(i,k)}(\theta_i, r_i(\theta_{-i})) - h_{(i,k)}(\theta_i, r_i(\theta_{-i}))|\, d\mu^n$ does not converge to 0, while we had convergence μ^2-almost everywhere.

where $\mu_{r_i} = \mu^{n-1} \circ r_i^{-1}$. One way to proceed is to apply the Radon–Nikodym Theorem. To this end, I show that μ_{r_i} is absolutely continuous with respect to μ. By way of contradiction, suppose that there is set A such that $\mu(A) = 0$ and $\mu_{r_i}(A) > 0$. This means that, for each j, there exist countable unions of intervals, $\bigcup_k I_j^k \subset \mathbb{R}$, such that $r_i(\theta_{-i}) \in A$ for all $\theta_{-i} \in \prod_j(\bigcup_k I_j^k)$. Since $r_i(\cdot)$ is continuous and strictly increasing, the set $r_i(\prod_j(\bigcup_k I_j^k))$ must contain some interval I. Since $I \subset A$ and $\mu(I) > 0$, it must be that $\mu(A) > 0$. This is a contradiction, so μ_{r_i} is absolutely continuous with respect to μ. Clearly, both μ_{r_i} and μ are (totally) finite on $r_i(\Theta_{-i})$. By the Radon–Nikodym Theorem, there exists f on $r_i(\Theta_{-i})$ such that $\mu_{r_i}(A) = \int_A f\, d\mu$ for every measurable set $A \subset r_i(\Theta_{-i})$. From (30), it gives

$$\begin{aligned}\int_\Theta \big|h^m_{(i,k)}(\theta_i, r_i(\theta_{-i})) - h_{(i,k)}(\theta_i, r_i(\theta_{-i}))\big|\, d\mu^n \\ = \int_{\Theta_i \times r_i(\Theta_{-i})} |h^m_{(i,k)}(\theta_i, t) - h_{(i,k)}(\theta_i, t)| f(t)\, d\mu^2.\end{aligned} \tag{31}$$

Since $|h^m_{(i,k)}(\theta_i, t) - h_{(i,k)}(\theta_i, t)| f(t)$ is integrable and dominated almost everywhere by $Hf(t)$ for $H > 0$ sufficiently large, we can apply the Bounded Convergence Theorem to show that (31) is zero as $m \to \infty$. This result allows the construction of the following subsequence from $\{h_i^m(\theta_i, r_i^q(\theta_{-i}))\}$.

- For each m, take $\alpha(m)$ such that $\int_\Theta |h^{\alpha(m)}_{(i,k)}(\theta_i, r_i(\theta_{-i})) - h_{(i,k)}(\theta_i, r_i(\theta_{-i}))|\, d\mu^n < 1/2m$.
- Since each $h^{\alpha(m)}$ is C^2, we know that $h^{\alpha(m)}_{(i,k)}(\theta_i, r_i^q(\theta_{-i}))$ converges uniformly to $h_i^{\alpha(m)}(\theta_i, r_i(\theta_{-i}))$ as $q \to \infty$. Thus, choose $\beta(m)$ such that $\int_\Theta |h_i^{\alpha(m)}(\theta_i, r_i^{\beta(m)}(\theta_{-i})) - h_i(\theta_i, r_i(\theta_{-i}))|\, d\mu^n < 1/2m$.

Along the subsequence $\{h_i^{\alpha(m)}(\cdot, r_i^{\beta(m)}(\cdot))\}$ so constructed, the left-hand side of (29) is less than $1/m$ for all m and thus it converges to $h_i(\cdot, r_i(\cdot))$ in L_1 norm. In other words, there is a sequence of dimensionally reducible decision rules $\{x_i^m\}$ that converges to x_i in L_1 space. Each x^m is implementable, because $\partial V_i(x_i, \theta_i)/\partial\theta_i$ is increasing in x_i and $x_i^m(\cdot)$ is increasing in $\hat{\theta}_i$ for each m. Theorem 3 completes the proof. □

References

Abreu, Dilip and Hitoshi Matsushima (1992), "Virtual implementation in iteratively undominated strategies: Complete information." *Econometrica*, 60, 993–1008. [406, 421, 428]

Abreu, Dilip and Hitoshi Matsushima (1994), "Exact implementation." *Journal of Economic Theory*, 64, 1–19. [404, 428]

Aliprantis, Charalambos D. and Kim C. Border (1999), *Infinite Dimensional Analysis*, second edition. Springer, Berlin. [439]

Arrow, Kenneth J. (1979), "The property rights doctrine and demand revelation under incomplete information." In *Economics and Human Welfare* (Michael J. Boskin, ed.), Academic Press, New York. [405, 407, 408, 414]

Cabrales, Antonio (1999), "Adaptive dynamics and the implementation problem with complete information." *Journal of Economic Theory*, 86, 159–184. [404, 407, 428]

Cabrales, Antonio and Giovanni Ponti (2000), "Implementation, elimination of weakly dominated strategies and evolutionary dynamics." *Review of Economic Dynamics*, 3, 247–282. [404]

Camerer, Colin F. (2003), *Behavioral Game Theory*. Princeton University Press, Princeton, New Jersey. [427]

Chen, Yan (2002), "A family of supermodular Nash mechanisms implementing Lindahl allocations." *Economic Theory*, 19, 773–790. [407]

Chen, Yan and Robert Gazzale (2004), "When does learning in games generate convergence to Nash equilibria? The role of supermodularity in an experimental setting." *American Economic Review*, 94, 1505–1535. [404]

Chen, Yan and Charles R. Plott (1996), "The Groves–Ledyard mechanism: An experimental study of institutional design." *Journal of Public Economics*, 59, 335–364. [404]

Chen, Yan and Fang-Fang Tang (1998), "Learning and incentive-compatible mechanisms for public goods provision: An experimental study." *Journal of Political Economy*, 106, 633–662. [404]

d'Aspremont, Claude and Louis-Andre Gérard-Varet (1979), "Incentives and incomplete information." *Journal of Public Economics*, 11, 25–45. [405, 407, 408, 414]

Duggan, John (1997), "Virtual Bayesian implementation." *Econometrica*, 65, 1175–1199. [421]

Echenique, Federico (2000), "Comparative statics by adaptive dynamics and the correspondence principle." *Econometrica*, 70, 833–844. [405]

Echenique, Federico and Aaron Edlin (2004), "Mixed equilibria are unstable in games of strategic complements." *Journal of Economic Theory*, 118, 61–79. [405]

Eliaz, Kfir (1999), "Fault tolerant implementation." *Review of Economic Studies*, 69, 589–610. [407]

Frankel, David M., Stephen Morris, and Ady Pauzner (2003), "Equilibrium selection in global games with strategic complementarities." *Journal of Economic Theory*, 108, 1–44. [420]

Fudenberg, Drew and David K. Levine (1998), *The Theory of Learning in Games*. MIT Press, Cambridge, Massachusetts. [408]

Green, Jerry R. and Jean-Jacques Laffont (1979), *Incentive in Public Decision-Making*. North-Holland, Amsterdam. [405, 427]

Groves, Theodore and John O. Ledyard (1977), "Optimal allocation of public goods: A solution to the 'free rider' problem." *Econometrica*, 45, 783–809. [404]

Healy, Paul J. (2006), "Learning dynamics for mechanism design: An experimental comparison of public goods mechanisms." *Journal of Economic Theory*, 129, 114–149. [404]

Healy, Paul J. and Laurent Mathevet (2008), "Designing stable mechanisms for economic environments". Unpublished paper, Department of Economics, Ohio State University. [428]

Ho, Teck-Hua (2008), "Individual learning in games." In *The New Palgrave Dictionary of Economics*, Volume 4, second edition (Lawrence E. Blume and Steven N. Durlauf, eds.), 230–235, Palgrave Macmillan, Basingstoke. [409]

Hurwicz, Leonid (1972), "On informationally decentralized systems." In *Decision and Organization* (Roy Radner and C. B. McGuire, eds.), 297–336, North-Holland, Amsterdam. [428]

Jackson, Matthew O. (1992), "Implementation in undominated strategies: A look at bounded mechanisms." *Review of Economic Studies*, 59, 757–775. [407]

Jackson, Matthew O. (2001), "A crash course in implementation theory." *Social Choice and Welfare*, 18, 655–708. [404]

Jordan, James S. (1986), "Instability in the implementation of Walrasian allocations." *Journal of Economic Theory*, 39, 301–328. [407]

Kim, Taesung (1986), "On the nonexistence of a stable Nash mechanism implementing Lindahl allocations." Unpublished paper, University of Minnesota. [407]

Laffont, Jean-Jacques and Eric Maskin (1980), "A differential approach to dominant strategy mechanisms." *Econometrica*, 48, 1507–1520. [427]

Ledyard, John O. (1978), "Incentive compatibility and incomplete information." *Journal of Economic Theory*, 18, 171–189. [428]

Ledyard, John O. and Thomas R. Palfrey (2007), "A general characterization of interim efficient mechanisms for independent linear environments." *Journal of Economic Theory*, 133, 441–466. [414]

Mas-Colell, Andreu (1974), "Continuous and smooth consumers: Approximation theorems." *Journal of Economic Theory*, 8, 305–336. [439]

Mas-Colell, Andreu, Michael D. Whinston, and Jerry R. Green (1995), *Microeconomic Theory*. Oxford University Press, Oxford. [408, 413, 431]

Mathevet, Laurent (2007), "Supermodular Bayesian implementation: Learning and incentive design." Working Paper 1265, California Institute of Technology, Division of the Humanities and Social Sciences. [425]

McAfee, R. Preston and John McMillan (1991), "Optimal contracts for teams." *International Economic Review*, 32, 561–577. [426]

Milgrom, Paul and John Roberts (1990), "Rationalizability, learning, and equilibrium in games with strategic complementarities." *Econometrica*, 58, 1255–1277. [404, 407, 431, 432]

Milgrom, Paul R. and John Roberts (1991), "Adaptive and sophisticated learning in normal form games." *Games and Economic Behavior*, 3, 82–100. [404, 407, 409]

Milgrom, Paul and John Roberts (1995), "Complementarities and fit: Strategy, structure, and organizational change in manufacturing." *Journal of Accounting and Economics*, 19, 179–208. [415]

Milgrom, Paul and John Roberts (1996), "Coalition-proofness and correlation with arbitrary communication possibilities." *Games and Economic Behavior*, 17, 113–128. [406, 417]

Mookherjee, Dilip and Stefan Reichelstein (1992), "Dominant strategy implementation of Bayesian incentive compatible allocation rules." *Journal of Economic Theory*, 56, 378–399. [426]

Muench, Thomas and Mark Walker (1983), "Are Groves–Ledyard equilibria attainable?" *Review of Economic Studies*, 50, 393–396. [404]

Myerson, Roger B. and Mark A. Satterthwaite (1983), "Efficient mechanisms for bilateral trading." *Journal of Economic Theory*, 29, 265–281. [406, 424]

Parkes, David C. (2004), "On learnable mechanism design." In *Collectives and the Design of Complex Systems* (K. Tumer and D. Wolpert, eds.), 107–131, Springer-Verlag, New York. [405]

Saijo, Tatsuyoshi, Tomas Sjöström, and Takehiko Yamato (2005), "Secure implementation." Working Paper 0056, Institute for Advanced Study, School of Social Science. [407, 426]

Sandholm, William H. (2002), "Evolutionary implementation and congestion pricing." *Review of Economic Studies*, 69, 667–689. [405, 407]

Sandholm, William H. (2005), "Negative externalities and evolutionary implementation." *Review of Economic Studies*, 72, 885–915. [405, 407]

Sefton, Martin and Abdullah Yavaş (1996), "Abreu–Matsushima mechanisms: Experimental evidence." *Games and Economic Behavior*, 16, 280–302. [407, 427]

Serrano, Roberto and Antonio Cabrales (2007), "Implementation in adaptive better-response dynamics." Working Paper 2007-16, Instituto Madrile no de Estudios Avanzados (IMDEA), Ciencias Sociales. [407]

Sjöström, Tomas (1994), "Implementation in undominated Nash equilibria without integer games." *Games and Economic Behavior*, 6, 502–511. [404]

Tumer, Kagan and David H. Wolpert (2004), "A survey of collectives." In *Collectives and the Design of Complex Systems* (K. Tumer and D. Wolpert, eds.), 107–131, Springer-Verlag, New York. [405]

Vives, Xavier (1990), "Nash equilibrium with strategic complementarities." *Journal of Mathematical Economics*, 19, 305–321. [404]

Submitted 2009-8-18. Final version accepted 2010-4-6. Available online 2010-4-6.