Designing Stable Mechanisms for Economic Environments

Paul J. Healy[*] Laurent Mathevet[†]

Abstract

We study the design of mechanisms that Nash-implement Walrasian or Lindahl allocations *and* induce supermodular games for a wide class of economies. Such mechanisms are robust to the presence of myopic agents who use adaptive learning rules to choose their strategies. We proceed in three steps: First, we identify strong necessary conditions on the functional form of any mechanism that implement Walrasian or Lindahl equilibria. Second, we use these necessary conditions to identify impossibility results for mechanisms with small strategy spaces. Finally, we show how to use additional dimensions in the strategy space to turn any Walrasian or Lindahl mechanism into a supermodular mechanism.

DRAFT COPY: December 20, 2008

[*]Department of Economics, The Ohio State University, Columbus, OH 43210. healy.52@osu.edu.
[†]Department of Economics, University of Texas, Austin, TX 78712. lmath@mail.utexas.edu.

1 Introduction

In economic environments, the equilibrium concepts of Léon Walras and Erik Lindahl represent normatively appealing selections from the space of possible consumption bundles. A major criticism of any equilibrium-based concept, however, is that it lacks a foundation for how agents arrive at the equilibrium and, in the case of multiplicity, to which equilibrium they will converge. Much of the criticism lies in the lack of an explicit price formation process; Lindahl's concept additionally suffers from the fact that individuals must treat their individualized prices as given even though they exercise monopsony power in their own consumption of the public good.

Mechanism design (as formulated by Hurwicz, 1972) circumvents much of this criticism. Here, a central planner chooses a mechanism that consists of a set of allowable strategies and a mapping from strategies to outcomes. Agents, given their preferences, choose equilibrium strategies of this mechanism. Since mechanisms can be designed such that (Nash) equilibrium strategies map into Walrasian or Lindahl allocations, the ambiguities regarding a price formation process can be completely circumvented.

This shift from price-taking competitive environments to non-cooperative games, however, introduces a new ambiguity regarding equilibration: How do agents arrive at the (Nash) equilibrium of the chosen mechanism? If we lack of a theory of equilibration to Nash equilibrium, we still lack a complete description of how economies can arrive at Walrasian or Lindahl equilibria.

Fortunately, past theoretical and experimental work has provided some insight into the equilibration process for non-cooperative games; many models of 'adaptive' learning have been proposed and tested (see Fudenberg and Levine, 1998 for a review of the theoretical literature and Camerer, 2003 for a survey of experimental results). Most of the proposed models are 'adaptive learning rules' in which agents myopically move towards strategies that yield higher payoffs (see Milgrom and Roberts, 1990 for a formal definition). Examples of adaptive learning rules include Cournot dynamics, fictitious play and Bayesian learning.

Milgrom and Roberts (1990) show that if a game is *supermodular* (see Topkis, 1998) then myopic agents who use adaptive learning rules will eventually converge to choosing strategies from a set formed by the smallest and largest equilibria of the game (Milgrom and Roberts, 1990); if a game has a unique equilibrium then agents who use these rules will eventually converge to equilibrium play.

Using laboratory experiments to study convergence properties of mechanisms, Chen and Plott (1996), Chen and Tang (1998), Cabrales and Ponti (2000), Chen and Gazzale (2004), and Healy (2006) find that actual play converges to equilibrium in those mechanisms in which adaptive learning rules are predicted to converge to equilibrium. These results suggest that if a mechanism were to always induce a supermodular game then play would always converge toward Nash equilibrium and, consequently, Walrasian or Lindahl allocations would result.

Using these insights, Chen (2002) develops a family of mechanisms that always induce a supermodular game over a class of quadratic, quasilinear preferences. This proves that such mechanisms can be constructed.

Our goal with this paper is to step back from Chen's 'proof of concept' result and show *how* to develop supermodular mechanisms. We proceed in three steps: First, in Section 3 we show that there are fairly strong restrictions on what a continuously differentiable mechanism must look like if it is to implement Walrasian or Lindahl allocations. Second, in Section 5 we use these strong necessary conditions to provide certain impossibility results, such as the impossibility of finding a mechanism with one-dimensional strategy spaces that implements Walrasian allocations and the non-existence of supermodular one-dimensional mechanisms that implement Lindahl allocations. Finally, in Section 6 we exploit the fact that adding dimensions to the strategy space allows new freedom in the form of the mechanism; freedom enough to add a 'penalty' term that guarantees the mechanism induces supermodular games for a wide class of environments.

2 The Model

2.1 Economic Environments

Consider an economy that consists of n agents each endowed with a preference relation defined on a two-dimensional commodity space. For each trader i denote the individually feasible consumption set by $\mathcal{C}_i$ and the endowment vector by $\omega_i = (\omega_x^i, \omega_y^i)$. We require that $\omega_i \in \mathcal{C}_i$ for each i. The two-dimensional net trade vector of agent i is given by $z_i = (x_i, y_i)$ and the set of i's individually feasible net trades is $\mathcal{Z}_i := \mathcal{C}_i - \{\omega_i\}$. Thus, we can describe i's preferences over net trades by a preference relation defined on $\mathcal{Z}_i$. We assume that, for all $i \in \mathcal{I}$, if $z \in \mathcal{Z}_i$ and $z' \geq z$ then $z' \in \mathcal{Z}_i$. A net trade vector $z = (z_1, \ldots, z_n)$ is said to be *feasible* if $z_i \in \mathcal{Z}_i$ for each i and *balanced* if $\sum_i z_i = 0$.

For simplicity we assume that each agent's preferences over net trades are representable by a utility function of the form $u_i(x_i, y_i|\theta_i)$, where θ_i identifies i's *type* and is drawn from some set of admissible types Θ_i. Define $\Theta = \times_i \Theta_i$ to be the set of all admissible type profiles. We assume throughout that u_i is strictly increasing in x_i (the numéraire good) for all y_i and θ_i; when we describe results on stability we further assume that u_i is quasilinear in x_i. We let p represent the price of the second good, normalizing the numéraire price to one.

This model, as specified, describes an exchange economy with purely private goods; we can easily reinterpret the model to allow the second good to be a purely public good by making three changes to the model: (1) every feasible net trade vector must be such that $y_i = y_j$ for all agents i and j, (2) $\omega_y^i = \omega_y^j$ for all i and j, and (3) include a firm who can produce y units of the public good from $c(y)$ units of the numéraire and aims to maximize profit $(py - c(y))$. For simplicity we assume a constant marginal cost of production, so that $c(y) = \kappa y$ for all y, with $\kappa > 0$.

A Walrasian equilibrium of a private goods economy is a net trade vector z^* and a price p^*

such that z^* is balanced, feasible, and maximizes each agent i's utility among all feasible net trades satisfying i's budget constraint that $x_i + p^* y_i \leq 0$.

A Lindahl equilibrium of a public goods economy is a net trade vector z^* and a vector of individual prices $p^* = (p_1^*, \ldots, p_n^*)$ such that z^* is balanced, feasible (which guarantees that $y_i^* = y_j^* = y^*$ for all i and j), maximizes each agent i's utility among all feasible net trades satisfying i's budget constraint that $x_i + p_i^* y_i \leq 0$, and, among all feasible net trades, maximizes the firm's profit of $(\sum_i p_i^*)y - c(y)$.

Note that Lindahl equilibria are of the same dimensionality as Walrasian equilibria; the latter requires $2n$ quantities but only one price while the former requires only $n+1$ quantities but needs n prices.

2.2 Mechanisms & Implementation

A *social choice correspondence* $f : \Theta \twoheadrightarrow \mathcal{Z}$ maps type profiles into sets of allocations. For example, f might identify all Pareto optimal net trades for each θ (the *Pareto correspondence*), all net trades z for which there is some price p such that (z, p) constitutes a Walrasian equilibrium at θ (the *Walrasian correspondence*), or, in a public goods setting, all net trades z for which there is some price vector p such that (z, p) constitutes a Lindahl equilibrium at θ (the *Lindahl correspondence*).

A *mechanism* $\Gamma = (\mathcal{M}, h)$ is a *message space* $\mathcal{M} = \times_i \mathcal{M}_i$ and an *outcome function* $h : \mathcal{M} \twoheadrightarrow \mathcal{Z}$ mapping each message profile $m = (m_1, \ldots, m_n)$ into a net trade vector z. A Nash equilibrium message profile of the mechanism Γ at the type profile θ is an $m^* \in \mathcal{M}$ such that, for each i and $m_i' \in \mathcal{M}_i$,

$$u_i(h(m^*)|\theta_i) \geq u_i(h(m_i', m_{-i}^*)|\theta_i),$$

where (m_i', m_{-i}^*) represents the message vector where i chooses m_i' and each $j \neq i$ chooses m_j^*.

In the case of economic environments with two goods, the outcome function can equivalently be written as a vector of $2n$ functions of the form $x_i(m)$ and $y_i(m)$ for each $i \in \mathcal{I}$.

The *Nash correspondence* $\nu : \Theta \twoheadrightarrow \mathcal{M}$ identifies the set of Nash equilibrium message profiles for each environment θ. A mechanism $(\mathcal{M}, h)$ is said to *implement* a social choice correspondence f if, for all $\theta \in \Theta$,

$$h(\nu(\theta)) = f(\theta).$$

3 Necessary Conditions for Nash Implementation

Our ultimate goal is to describe a procedure for designing mechanisms that Nash implement Walrasian or Lindahl equilibria and have desirable stability properties. To do this we first identify what all mechanisms that Nash implement Walrasian or Lindahl equilibria must look like. Then, given this necessary condition on the form of the mechanism, we can show how to modify any such mechanism to guarantee the required stability properties. In this section we focus on the necessary condition; stability is covered in the following section.

In the most general form, a mechanism for implementing Walrasian or Lindahl equilibria can be specified by a mapping

$$\mathcal{M} \ni m \mapsto (x_i(m), y_i(m))_{i=1}^n \in \mathcal{Z}.$$

Given any function $y_i(m)$, it is without loss of generality that we can express i's net trade of the numéraire as

$$x_i(m) = -q_i(m_{-i})y_i(m) - g_i(m), \tag{1}$$

where the per-unit 'price' term q_i does not depend on m_i (and could be identically equal to zero) and the 'penalty' term $g_i(m)$ is arbitrary. Thus, a mechanism can equivalently be described by the mapping $m \mapsto (y_i(m), q_i(m_{-i}), g_i(m))_{i=1}^n$.

Our proofs rely heavily on differentiability arguments, so we restrict attention to those mechanisms whose messages are finite-dimensional vectors of real numbers and whose outcome functions are twice continuously differentiable in every dimension. To avoid explicitly the complications arising with boundary equilibria we further assume that all message spaces are open sets. In most applications the message spaces equal the entire space, but if a mechanism has a closed message space our characterization result simply applies to any open subset of the space.

Assumption 1 (Differentiability)**.** For each agent i the message space $\mathcal{M}_i$ is an open subset of $\mathbb{R}^{K_i}$, where $K_i \in \{1, 2, \ldots\}$ and, for each dimension $k \in \{1, \ldots, K_i\}$ and message vector $m \in \mathcal{M}$, the functions x_i and y_i are twice continuously differentiable in m_{ik} at m.

Our first proposition derives necessary conditions on the form of a mechanism that Nash implements Walrasian or Lindahl equilibria; however, these conditions apply at equilibrium points only. Since messages that cannot be equilibria for any type profile θ need not map to any particular allocation, mechanisms may be arbitrarily badly behaved at these off-equilibrium points. For the special case of one-dimensional message spaces ($K_i = 1$ for all i) we are able to prove that if the type space is rich enough then it must be that *every* message is a Nash equilibrium for some type profile. Thus, our necessary conditions apply to the entire strategy space and we can derive a fairly strong result about the required functional form of the mechanism. If the strategy spaces have more than one dimension then we cannot guarantee that all messages are equilibria and therefore we cannot derive necessary conditions that cover the entire strategy space.

As a simple example of how extra dimensions introduce freedom in the functional form of the mechanism, take any mechanism with one-dimensional message spaces in which agents send $m_i \in \mathbb{R}^1$ and the outcome is given by $\tilde{y}_i(m)$ and $\tilde{x}_i(m)$ for each i. Assume that this mechanism implements some social choice correspondence $f(\theta)$. Now consider a new mechanism in which agents send two-dimensional messages of the form $m_i = (r_i, s_i)$ and the outcomes are given by $\tilde{y}_i(r)$ and $\tilde{x}_i(r) - |s_i|$. Clearly, if m_i^* is a Nash equilibrium of the original mechanism at type profile θ then $(m_i^*, 0)$ is a

Nash equilibrium of the new mechanism at θ and the new mechanism also implements f. The extra dimension serves no purpose in this case, but the example shows how there can be message vectors (with $s_i \neq 0$ here) that cannot be Nash equilibria for any θ. Thus, the necessary conditions we derive at equilibrium points cannot apply everywhere. It is exactly this freedom that we have in the added dimensions, however, that allows us to modify existing mechanisms to guarantee additional off-equilibrium properties such as supermodularity.

To make this issue of dimensions explicit, we let $\mathcal{M}_i = \mathcal{R}_i \times \mathcal{S}_i$, where, for each i, $\mathcal{R}_i \subseteq \mathbb{R}^{J_i}$ represents those dimensions $k \in \{1, \ldots, J_i\}$ for which there is some message vector $m = (r, s)$ such that $\partial y_i(r, s)/\partial r_{ik} \neq 0$. The set $\mathcal{S}_i \subseteq \mathbb{R}^{K_i - J_i}$ represents those dimensions $k \in \{1, \ldots, K_i - J_i\}$ for which $\partial y_i/\partial s_{ik} \equiv 0$.

With the partitioning of the strategy spaces into $\mathcal{R}_i$ and $\mathcal{S}_i$ we can modify equation (1) slightly and write any mechanism's numéraire outcome function as

$$x_i(r, s) = -q_i(r_{-i}, s)y_i(r) - g_i(r, s). \tag{2}$$

This formulation allows the 'price' term q_i to depend on the components of i's message that do not affect y_i, but not on the remaining dimensions.

Roughly speaking, our first proposition shows that, if a mechanism Nash implements Walrasian or Lindahl equilibria then the 'penalty' term $g_i(m)$ must equal zero at any Nash equilibrium message profile (r^*, s^*) such that small perturbations of r^* also lead to Nash equilibrium messages (possibly for different type profiles). Thus, each agent i is simply paying a per-unit price for the non-numéraire commodity without any extra 'penalty' terms, and this per-unit price cannot depend on i's announcement of r_i alone.

To capture formally this idea of small perturbations in r_i^* leading to other equilibrium messages, we employ the following definition:

Definition 1. Given a set $\mathcal{M}' \subseteq \mathcal{M}$, the $\mathcal{R}$-*interior* of $\mathcal{M}'$ is defined by

$$\mathcal{R}\text{int}(\mathcal{M}') = \left\{ m = (r, s) \in \mathcal{M} : (\exists \varepsilon > 0)(\forall i \in \mathcal{I}) \ ||r_i' - r_i|| < \varepsilon \Rightarrow (r_i', r_{-i}, s) \in \mathcal{M}' \right\}.$$

If $m^* \in \mathcal{R}\text{int}(\nu(\Theta))$, say that m^* is an $\mathcal{R}$-*interior Nash equilibrium.*

In words, the $\mathcal{R}$-interior of a set is simply those points in the set such that all small perturbations in r lead to points also within the set; the $\mathcal{R}$-interior is a superset of the interior since no restriction is made on perturbations in s. For example, figure 1 shows an example of an $\mathcal{R}$-interior Nash equilibrium m^* that is not in the interior of $\nu(\Theta)$.

Proposition 1. If a mechanism $\Gamma = (\mathcal{M}, (x_i(\cdot), y_i(\cdot))_{i=1}^n)$ is differentiable (assumption A1) and Nash implements either the Walrasian or Lindahl correspondence then at any $\mathcal{R}$-interior Nash

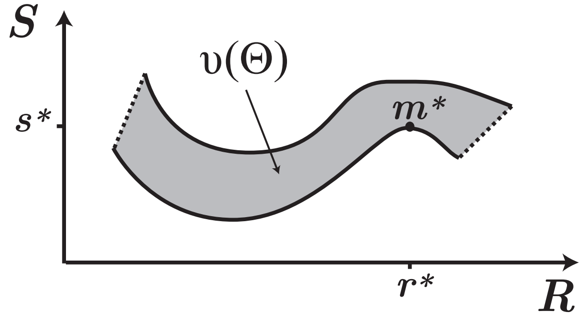

Figure 1: An $\mathcal{R}$-interior Nash equilibrium m^* that is not in the interior of $\nu(\Theta)$.

equilibrium $m^* = (r^*, s^*)$,

$$x_i(r^*, s^*) = -q_i(r^*_{-i}, s^*)y(r^*) \tag{3}$$

for each $i \in \mathcal{I}$, so that $g_i(m^*) = 0$.

Although proposition 1 gives a strong result, its scope of applicability is limited to those messages which can be an equilibrium. If the type space is rich enough, however, one might expect that any messages can be an equilibrium for some type profile and, therefore, that equation (12) would apply over the whole message space.

Unfortunately, we can identify two classes of messages that can *never* constitute a Nash equilibrium (under the maintained assumptions) for *any* type profile θ.

The first class of messages that cannot be Nash equilibria are those for which the s_i components are not all best responses. Since the s_i component of each agent's message enters only into the determination of the numéraire, the s_i components can be thought of as a strategy in a separate 'transfer-maximizing game' in which agents take r as fixed and simultaneously choose s_i to maximize $x_i(r, s_i, s_{-i})$. Define

$$\sigma_i(r, s_{-i}) := \arg\max_{s'_i \in S_i} x_i(r, s'_i, s_{-i}) \tag{4}$$

to be i's set of best responses in S_i to s_{-i} given r. The Nash equilibria of this transfer-maximizing game can then be identified by

$$\sigma(r) := \{s \in S : (\forall i \in \mathcal{I})\ s_i \in \sigma_i(r, s_{-i})\}\,. \tag{5}$$

Since preferences are strictly monotone in the numéraire, the payoffs of the transfer-maximizing game are independent of θ. Therefore it is possible for the mechanism designer to determine each σ_i and σ without knowledge of θ. Clearly, if s is not in $\sigma(r)$ then (r, s) cannot be an equilibrium of the mechanism for *any* type profile $\theta \in \Theta$.

The second class of messages that cannot be Nash equilibria are those for which perturbations in different dimensions of $\mathcal{R}_i$ lead to different rates of substitution between the private good and the public good. If this were the case the agent would prefer to deviate in (at least) one dimension.

To proceed, we assume that the mechanism's y_i function is always responsive in r_i for each i. This assumption is in fact relatively innocuous; all existing mechanisms of which we are aware satisfy this condition, and if a mechanism did not satisfy this condition at some point m then our necessary conditions would simply not apply at this point.

Assumption 2 (Responsive y_i)**.** $\partial y_i(r)/\partial r_{ik}$ is bounded away from zero for each $r \in \times_{i=1}^n \mathcal{R}_i$, $i \in \mathcal{I}$, and $k \in \{1, \ldots, J_i\}$.

To visualize this idea of different rates of substitution in different dimensions, it is useful to think of the set of (x_i, y_i) pairs that an agent i is capable of obtaining as he unilaterally varies r_{ik} away from some equilibrium point r_i^*. Given the above discussion, we need only to focus on those deviations in which the change in r_i is paired with an appropriate change in s_i; as any utility-maximizing player i deviates from r_i^* to r_i' he must also adjust s_i^* to some $s_i' \in \sigma_i(r_i', r_{-i}^*, s_{-i}^*)$. Thus, we graph out those (x_i, y_i) pairs that i can reach by deviating in r_i and matching this deviation with some s_i that continues to be a best response in the resulting transfer-maximizing game.

Since y_i is invertible in each r_{ik} (because $\partial y_i/\partial r_{ik}$ is bounded away from zero) we can define $y_{ik}^{-1}(\hat{y}|r_{-(ik)})$ as the inverse of $y_i(\cdot, r_{-(ik)})$, where $r_{-(ik)}$ is the vector r with component r_{ik} removed. Using this notation, the graph of (x, y) pairs that i can realize by varying r_{ik} from (r, s) is the graph of the function

$$\chi_{ik}(y|r, s) := x_i(\underbrace{y_{ik}^{-1}(y|r_{-(ik)})}_{r_{ik}}, r_{-(ik)}, \underbrace{\sigma_i(\underbrace{y_{ik}^{-1}(y|r_{-(ik)})}_{r_{ik}}, r_{-(ik)}, s_{-i})}_{s_i}, s_{-i}). \tag{6}$$

Note that χ_{ik} is the value function of a maximization problem (maximizing $x_i(r, \cdot, s_{-i})$ over $\mathcal{S}_i$) and is therefore single-valued even if the set of maximizers ($\sigma_i(r, s_{-i})$) is not. For the sake of exposition, assume that χ_{ik} is differentiable at some point $m^* = (r^*, s^*)$ (Milgrom and Segal, 2002 provide conditions for differentiability everywhere, though we do not require differentiability of χ_{ik} in our propositions). The downward slope of χ_{ik} at $m^* = (r^*, s^*)$ can then be described by

$$P_{ik}(r^*, s^*) := -\frac{\partial x_i(r^*, s^*)/\partial r_{ik}}{\partial y_i(r^*)/\partial r_{ik}}. \tag{7}$$

In economic terms, P_{ik} represents the *effective price* of y_i charged by the mechanism at m^* as r_{ik} moves in the direction of increasing y_i. This is shown in figure 2.

The effective price in a mechanism serves the same role *locally* as prices in a Walrasian or Lindahl equilibrium. To see this, write agent i's utility at θ_i in the game induced by some mechanism $(\mathcal{M}_i, x_i, y_i)_{i \in \mathcal{I}}$ as

$$\tilde{u}_i(r, s|\theta_i) := u_i(y_i(r), x_i(r, s)|\theta_i).$$

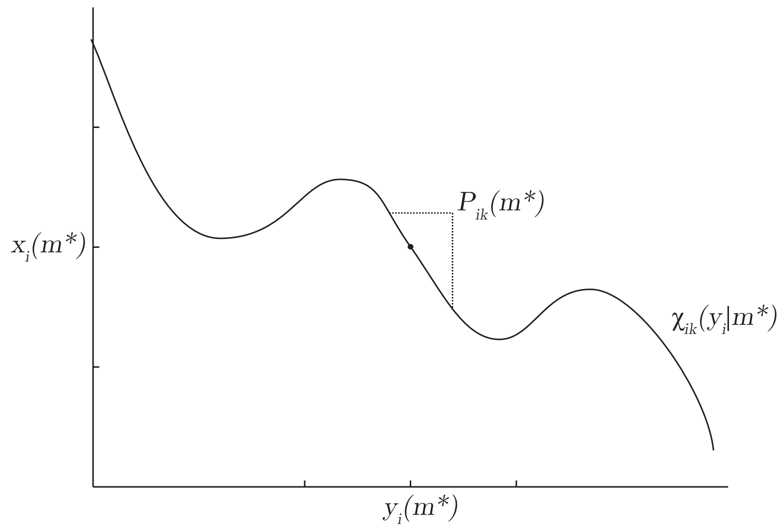

Figure 2: The effective price of changes in r_{ik}.

Since the message space is open, at any Nash equilibrium $m^* = (r^*, s^*)$ it must be that for each dimension $k \in \{1, \ldots, J_i\}$, m^* satisfies the first order condition

$$\frac{\partial \tilde{u}_i(m^*|\theta_i)/\partial y_i}{\partial \tilde{u}_i(m^*|\theta_i)/\partial x_i} = P_{ik}(m^*). \tag{8}$$

But if this mechanism implements a Walrasian or Lindahl equilibrium, the ratio of marginal utilities must also equal the Walrasian or Lindahl price. Thus, the effective prices at the equilibrium message profiles must match the Walrasian or Lindahl prices for each environment θ.

Given these two classes of messages that cannot be equilibria, define

$$\begin{aligned} \mathcal{M}^* \quad := \quad & \{m \in \mathcal{M} : (\forall i \in \mathcal{I})(\forall k, l \in \{1, \ldots, J_i\}) \; P_{ik}(m) = P_{il}(m)\} \\ & \cap \{m = (r, s) \in \mathcal{M} : s \in \sigma(r)\} \end{aligned} \tag{9}$$

as those messages that are not ruled out by these arguments. If the mechanism has a one-dimensional message space for each agent ($\mathcal{R}_i = \mathbb{R}^1$ and $\mathcal{S}_i = \emptyset$ for all i) then the two restrictions identified above are trivially satisfied and so $\mathcal{M}^* = \mathcal{M}$; in general, however, $\mathcal{M}^*$ is a strict subset of $\mathcal{M}$ and could even be empty.

We now explore how rich the type space must be so that all messages in $\mathcal{M}^*$ are equilibria for some admissible type profile.

To do so, we require one additional restriction on the form of the mechanism: it cannot be that deviations can cause arbitrarily large changes in x_i, relative to the corresponding change in y_i. If the ratio of these changes were unbounded as agents deviate from some point m^* then the function

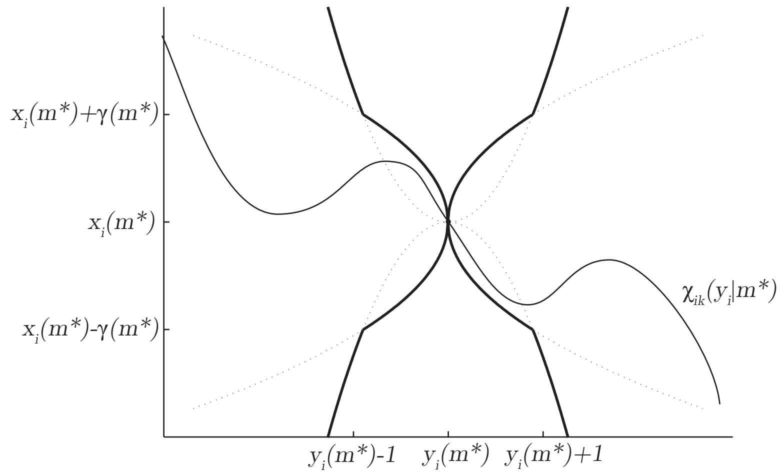

Figure 3: The bounds on $\chi_{ik}(y|m^*)$ imposed by assumption 3.

χ_{ik} would become arbitrarily steep and, by monotonicity, all agents of every type would prefer to deviate to gain infinite quantities of x_i at a (relatively) infinitesimal cost in terms of y_i. Thus, there needs to be some joint condition on how dramatic changes in x_i can be (relative to changes in y_i) along with a condition on how 'steep' agents' indifference curves can become in (x_i, y_i)-space.

We choose to require a weak form of Hölder continuity of order 2 on χ_{ik} to rule out mechanisms in which χ_{ik} has a slope that becomes infinite as m approaches some m^* or diverges 'too quickly' as m grows without bound.

Assumption 3 (Weak Hölder Continuity)**.** For all $r \in \mathcal{R}$, all $s \in \sigma(r)$, and all $i \in \mathcal{I}$ there exists some finite $\gamma_i(r) > 0$ such that for all $r_i' \in \mathcal{R}_i$ and $s_i' \in \sigma_i(r_i', r_{-i}, s_{-i})$,

$$\left|x_i(r_i', r_{-i}, s_i', s_{-i}) - x_i(r, s)\right| \leq \gamma_i(r) \max\left\{\left|y_i(r_i', r_{-i}) - y_i(r)\right|^2, \left|y_i(r_i', r_{-i}) - y_i(r)\right|^{1/2}\right\}$$

To interpret Assumption 3, first suppose that r_i' leads to large changes in y_i (more than one unit in magnitude). Then the first term in the maximand applies and the associated change in x_i must be less than the squared change in y_i times some multiple that varies in r. If instead r_i' leads to small changes in y_i (less than one unit in magnitude) then the second term in the maximand applies and the associated change in x_i must be less than the square root of the change in y_i times some variable multiplier. In both cases, the requirement is strictly weaker than requiring that χ_{ik} be Hölder continuous of degree 2 or that χ_{ik} be Lipschitz continuous. The bounds on χ_{ik} imposed by this assumption are demonstrated in figure 3.

Given assumption 3, we then require that the type space at least contains all possible quasilinear preferences with a quadratic valuation of y_i. This allows for types with sufficient steepness in agents' indifference curves to prevent them from preferring deviations from any m^*. This is formalized in the following assumption, where $\mathbb{R}_{++}$ represents the set of strictly positive real numbers.

Assumption 4 (Rich Type Space)**.** For each vector $(\alpha_i, \beta_i)_{i=1}^n \in (\mathbb{R}_{++} \times \mathbb{R})^n$ there is some $\theta \in \Theta$ such that for each i,

$$u_i(x_i, y_i|\theta_i) = \left(-\alpha_i y_i^2 + \beta_i y_i\right) + x_i.$$

Under assumption 4 we can define Θ_2 as the subset of Θ containing all preferences of this form; formally, let

$$\Theta_2 := \left\{\theta \in \Theta : (\forall i \in \mathcal{I})\ u_i(x_i, y_i|\theta_i) = \left(-\alpha_i y_i^2 + \beta_i y_i\right) + x_i\right\}. \tag{10}$$

We now show that if the slope of χ_{ik} does not become arbitrarily steep (assumption 3) and the type space is rich enough (assumption 4) then $\mathcal{M}^* \subseteq \nu(\Theta)$, so the derived set $\mathcal{M}^*$ exactly corresponds to the set of all equilibrium message profiles.

Proposition 2. If a mechanism $\Gamma = (\mathcal{M}_i, x_i, y_i)_{i\in\mathcal{I}}$ satisfies assumptions 1, 2 and 3 and the type space Θ satisfies assumption 4 then every message profile $m \in \mathcal{M}^*$ (defined in equation 9) is a Nash equilibrium of Γ for some $\theta \in \Theta$. Thus, $\mathcal{M}^* = \nu(\Theta)$.

Corollary 1. Under assumptions 1 through 4, if $\mathcal{M}_i = \mathbb{R}^1$ for each i then every message is a Nash equilibrium for some environment ($\mathcal{M} = \nu(\Theta)$).

Combining proposition 1 and 2 gives the following theorem and corollary:

Theorem 1. If a mechanism $\Gamma = (\mathcal{M}_i, x_i, y_i)_{i\in\mathcal{I}}$ Nash implements the Lindahl or Walrasian correspondences and satisfies assumptions 1, 2 and 3 and the type space Θ satisfies assumption 4 then for every $i \in \mathcal{I}$ and every $m^* = (r^*, s^*)$ in the $\mathcal{R}$-interior of $\mathcal{M}^*$,

$$x_i(r^*, s^*) = -q_i(r^*_{-i}, s^*) y_i(r^*), \tag{11}$$

so that $g_i(m^*) = 0$.

For one-dimensional mechanisms and public goods environments this theorem provides a strong restriction on the functional form of any mechanism implementing the Lindahl correspondence.

Corollary 2. Under assumptions 1 through 4, if Γ Nash implements the Lindahl correspondence with $\mathcal{M}_i = \mathbb{R}^1$ for each i then for every $i \in \mathcal{I}$ and *every* $m \in \mathcal{M}$,

$$x_i(m) \equiv -q_i(m_{-i}) y_i(m), \tag{12}$$

so that $g_i(m) \equiv 0$.

Walker (1981) and Hurwicz (1979) provide one-dimensional mechanisms for implementing Lindahl allocations, both of which are generalized by Tian (1990). In each of these cases $q(m) = b + Am$

and $g_i(m) \equiv 0$, where b is a $1 \times n$ vector of coefficients and A is a $n \times n$ matrix of coefficients with $a_{ii} = 0$ for each i; thus, $q_i(m)$ does not depend on m_i and there is no penalty term, consistent with corollary 2. The Groves and Ledyard (1977) mechanism, on the other hand, uses a non-zero penalty term and consequently does not implement Lindahl allocations.

For the Walrasian correspondence the above result would hold, but since there can only be one price ($q_i \equiv q_j \equiv q$) and since each q_i cannot depend on m_i, the price function q must be constant. Thus, we have the following impossibility result:

Corollary 3. There does not exist a mechanism with $\mathcal{M}_i = \mathbb{R}^1$ for each i that Nash implements the Walrasian correspondence unless every $\theta \in \Theta$ has an identical and unique Walrasian equilibrium price.

4 Designing Supermodular Mechanisms

4.1 Lattice-Theoretic Definitions and Supermodular Games

The basic definitions of lattice theory in this section are discussed in Milgrom and Roberts (1990) and Topkis (1998).

In this paper, we endow each agent's message space $\mathcal{M}_i \subset \mathbb{R}^n$ with the usual order $\geq$. That is, $m_i \geq m_i'$ in $\mathcal{M}_i$ if and only if each dimension of m_i is larger than or equal to the corresponding dimension of m_i'. Since $\geq$ is transitive, reflexive, antisymmetric on $\mathcal{M}_i$, $(\mathcal{M}_i, \geq)$ is a lattice for all i. The set $\mathcal{M} = \prod_i \mathcal{M}_i$ is endowed with the product order. Each $\mathcal{M}_i$ is endowed with the Euclidean topology.[1] Let $(\mathcal{M}_{-i}, \geq)$ be a partially ordered set[2]. A function $g_i : \mathcal{M}_i \to \mathbb{R}$ is *supermodular* if, for all $m_i, m_i' \in \mathcal{M}_i$, $g_i(m_i) + g_i(m_i') \leq g_i(m_i \wedge m_i') + g_i(m_i \vee m_i')$; $g_i : \mathcal{M}_i \times \mathcal{M}_{-i} \to \mathbb{R}$ has *increasing differences* in (m_i, m_{-i}) if, whenever $m_i \geq m_i'$ and $m_{-i} \geq m_{-i}'$, $g_i(m_i, m_{-i}) - g_i(m_i', m_{-i}) \geq g_i(m_i, m_{-i}') - g_i(m_i', m_{-i}')$. If g_i has increasing differences or satisfies the single-crossing property in (m_i, m_{-i}), then m_i and m_{-i} are said to be complements.

A game is a tuple $(\mathcal{I}, \{(\mathcal{M}_i, \geq), u_i\})$, where $\mathcal{I}$ is a finite set of players; each $i \in \mathcal{I}$ has a strategy space $\mathcal{M}_i \subset \mathbb{R}^n$ endowed with the usual order; and each i has a payoff function $u_i : \mathcal{M}_i \times \mathcal{M}_{-i} \to \mathbb{R}$.

Definition 2. A game $\mathcal{G} = (\mathcal{I}, \{(\mathcal{M}_i, \geq), u_i\})$ is *supermodular* if for all $i \in \mathcal{I}$,

1. u_i is bounded, supermodular in m_i for each m_{-i}, and has increasing differences in (m_i, m_{-i});

2. u_i is upper-semicontinuous in m_i for each m_{-i} and continuous in m_{-i} for each m_i.

Consider a mechanism $\Gamma = (\mathcal{M}_i, x_i, y_i)_{i \in \mathcal{I}}$. Agent i's utility function in the mechanism given type profile $\theta \in \Theta$ is $u_i(y_i(m), x_i(m)|\theta_i)$.

[1] The *order-interval* topology on a lattice is the topology whose subbasis for the closed sets is the set of closed intervals. A *closed interval* $[x, y]$ in $\mathcal{M}$ is the set of $z \in \mathcal{M}$ such that $y \geq z \geq x$.

[2] A partially ordered set is a set with a binary relation that is transitive, reflexive, antisymmetric.

Definition 3. A mechanism $\Gamma = (\mathcal{M}_i, x_i, y_i)_{i \in \mathcal{I}}$ *supermodularly implements* a social correspondence $f : \Theta \twoheadrightarrow \mathcal{Z}$ if, for all $\theta \in \Theta$,

1. $(y_i(\nu(\theta)), x_i(\nu(\theta))) = f(\theta)$, where ν is the Nash correspondence, and

2. the game $(\mathcal{I}, \{(\mathcal{M}_i, \geq), u_i(y_i(.), x_i(.)|\theta)\})$ is supermodular.

5 Impossibility Results

To design supermodular mechanisms, we need to take into account the restrictions imposed on the transfers by Theorem 1. By corollary 3, there cannot be any one-dimensional mechanism that implements the Walrasian correspondence, much less a supermodular one. Although there do exist one-dimensional mechanisms that implement the Lindahl correspondence (see Walker, 1981, for example), we now show that no one-dimensional mechanism can *supermodularly* implement the Lindahl correspondence when the function $y(m)$ takes its usual form of $y(m) = \sum_i m_i$.

Theorem 2. Under assumptions 1, 2, 3' and 4', there cannot be any one-dimensional mechanism with allocation function $y(m) = \sum_i m_i$ that supermodularly implements the Lindahl correspondence.

Most existing Lindahl mechanisms use $y(m) = \sum_i m_i$, so the result applies to all of them. But most of these mechanisms also employ a symmetric price function, in which case, the result extends to all symmetric allocation functions y.

Definition 4. For clarity, the ith component of any message profile gives agent i's message. Let π be any bijection from $\{1, \ldots, n\}$ onto itself. A mechanism is symmetric if $y(m) = y(m_{\pi(1)}, \ldots, m_{\pi(I)})$ and $q_i(m) = q_{\pi(i)}(m_{\pi(1)}, \ldots, m_{\pi(I)})$ for all $i \in \mathcal{I}$.

Theorem 2'. Under assumptions 1, 2, 3' and 4', there does exist any symmetric one-dimensional mechanism that supermodularly implements the Lindahl correspondence.

6 Supermodular Mechanisms

6.1 Lindahl Mechanisms

In this section, we are concerned with the stability of mechanisms that implement the Lindahl correspondence. Although there already exist stable Lindahl mechanisms, the existing literature does not provide a methodology for building such mechanisms. Our results, on the other hand, describe a way for stabilizing any mechanism with certain properties, and so they should be seen as a tool for constructing new mechanisms with desirable stability properties. We adopt the methodology introduced in Mathevet (2007) for weak Bayesian implementation, which exploits strategic complementarities to obtain nice stability properties. More precisely, we will provide a method for converting mechanisms into supermodular mechanisms. This method consists of adding complementarities until the mechanism becomes supermodular. So, if the original mechanism produces strategic substitutes—as opposed to strategic complementarities—then it must do so only to a bounded degree, for otherwise we cannot compensate for 'infinite substitutes'. This motivates the following assumption.

Assumption 5. (Boundedly-Convex Quasilinear Preferences). For each $\theta \in \Theta$ each agent i has preferences given by

$$u_i(x_i, y_i|\theta_i) = v_i(y_i|\theta_i) + x_i$$

where $\partial^2 v_i/\partial y_i^2 \in [M, 0)$ for some $M < 0$.

The next theorem shows that if there is a mechanism that implements the Lindahl correspondence, then there is also a supermodular mechanism that implements the Lindahl correspondence. This theorem works by appending a term to the transfer function x_i that vanishes in equilibrium and provides sufficient complementarities to make the mechanism induce a supermodular game. This provides supermodularity without affecting the desirable equilibrium properties of the original mechanism. This transformation technique has a similar flavor to Mathevet (2007).

Theorem 3. Let $\Gamma = (\mathcal{M}_i, (y, x_i))_{i\in\mathcal{I}}$ be a one-dimensional differentiable (1) mechanism that implements the Lindahl correspondence such that $y(m) = \sum_i m_i$ and $q_i(\cdot)$ has bounded first-derivatives. If Θ satisfies assumption 5, then there exists a supermodular mechanism $(\mathcal{A}_i, (y, x_i^{\text{SM}}))_{i\in\mathcal{I}}$ that implements the Lindahl correspondence.

This result identifies conditions on two functions so that, if transfers are appended to these functions, they generate complementarities while maintaining the original incentives. We can verify that the following functions satisfy all these conditions (see the proof of the theorem):

$$q_i^{\text{SM}}(m_{-i}, a_{-i}) = \tau(\sum_{j\neq i} m_j) - \frac{1}{n}\sum_{j\neq i}\psi_j(a_j)$$

and

$$H_i(m, a, \rho_i) = -\frac{1}{2}(\psi_i(a_i) - \tau(\sum_{k=1}^{n} m_k))^2 - \rho_i\frac{1}{2}\sum_{j\neq i}(a_j - \tau(\sum_{k=1}^{n} m_k))^2$$

where τ is any increasing affine function, and ψ_i, $i = 1, \ldots, n$ are increasing functions with slope bounded above zero. These functions can be used to turn the Walker mechanism into a supermodular mechanism, which can be useful, given that this mechanism is known to have stability issues. Note that when τ and ψ are the identify function we recover a piece of Chen's mechanism.

The next proposition shows that Theorem 3 implies that any supermodular Lindahl mechanism has a unique Nash equilibrium under assumption 5. This is a positive result, because such mechanisms, for which we give examples below, have strong robustness properties.

Proposition 3. Suppose Θ contains only boundedly-convex quasilinear preferences (A5). Any differentiable mechanism with a responsive y function (A1 and A2) that supermodularly implements the Lindahl correspondence has a unique Nash equilibrium.

Proof. We proceed by showing that, under assumption 5, there can be only one Lindahl allocation for each environment θ. Then, we prove that it implies that a supermodular mechanism has a unique Nash equilibrium. If preferences are strictly convex, then there is a unique efficient public good level $y^*(\theta)$, because

$$\text{argmax}_{y\in\mathbb{R}} \sum_{i\in\mathcal{I}} v_i(y|\,\theta) = \{y^*(\theta)\}.$$

Similarly, there is a unique allocation of the numeraire,

$$x_i^*(\theta) = p_i^*(\theta)y^*(\theta),$$

where $p_i^*(\theta) = \partial v_i(y^*(\theta)|\theta)/\partial y$. By way of contradiction, suppose that a supermodular mechanism implements the Lindahl correspondence and has several equilibria. Then, its largest equilibrium must be strictly greater than its smallest one. Since the public good function is strictly monotonic, these two equilibria cannot correspond to the same public good level, hence one is inefficient (because only one public good level is efficient). This is a contradiction, because this mechanism implements the Lindahl correspondence. □

6.2 The Walrasian Correspondence

6.2.1 Supermodular Implementation of the Walrasian Correspondence

In this section, we are concerned with the stability of mechanisms that implement the (constrained) Walrasian correspondence. We present a procedure to stabilize smooth Walrasian mechanisms by turning them into supermodular mechanisms. The analysis applies for any finite number of private goods as long as preferences are strictly convex.

The mechanisms that we study satisfy some feasibility properties. The first property is *balancedness*, which is some inter-agent feasibility condition, essentially requiring that net trades add up to zero. The second property is *individual feasibility* (in equilibrium), which requires the mechanism to allocate consumption bundles within the consumption sets of the agents.

Definition 5. A mechanism $(\mathcal{M}_i, (y_i, x_i))_{i\in\mathcal{I}}$ is *balanced* if $\sum_{i\in\mathcal{I}} y_i(m) = \sum_{i\in\mathcal{I}} x_i(m) = 0$, for all $m \in \mathcal{M}$. Given an environment $\{(u_i(y_i, x_i|\theta))_{i\in\mathcal{I}} : \theta \in \Theta\}$, the mechanism is *balanced in equilibrium* if for all θ, $(y_i(m^*), x_i(m^*)) \in \mathcal{Z}_i$ for all $m^* \in \nu(\theta)$.

Definition 6. Given an environment $\{(u_i(y_i, x_i|\theta))_{i\in\mathcal{I}} : \theta \in \Theta\}$, a mechanism $(\mathcal{M}_i, (y_i, x_i))_{i\in\mathcal{I}}$ is individually feasible (in equilibrium) if for all θ, $(y_i(m^*), x_i(m^*)) \in \mathcal{Z}_i$ for all $m^* \in \nu(\theta)$.

Now we provide our main result.

Theorem 4. Under assumption 5, suppose there is a $2n$-dimensional, twice-differentiable, balanced, and individually feasible mechanism $(\mathcal{M}_i, (y_i, x_i))_{i \in \mathcal{I}}$ that implements the Walrasian correspondence such that:

1. $y_i(s) = s_i - \frac{\sum_{j \neq i} s_j}{n-1}$, and
2. $x_i(m) = q(r_{-i}) \cdot y_i(s) + g_i(r) + S_i(m_{-i})$ where $q(\cdot)$ has first-derivatives which are bounded below zero, and $\partial^2 q(\cdot)/\partial r_{i,k} \partial r_{i,z} = 0$ for any distinct dimensions k, z.

Then there exists a $2n$-dimensional, twice-differentiable, individually feasible, supermodular mechanism $(\mathcal{M}_i, (y_i, x_i^{\text{SM}}))_{i \in \mathcal{I}}$ that implements the Walrasian correspondence and is balanced in equilibrium.

In this theorem, we have given the simplest way of transforming a Walrasian mechanism into a supermodular mechanism. There are, of course, many possible transformations; in particular, we can generate complementarities by adding dimensions but we did not do it here.

Hurwicz proposed a smooth mechanism which implements the Walrasian correspondence and satisfies all the conditions of Theorem 4; so it follows as a corollary that the Walrasian correspondence is supermodular implementable. Besides, the price function in Hurwicz's mechanism has further properties that allow us to turn it into a supermodular mechanism, while guaranteeing balancedness off-equilibrium.

Proposition 4. The Walrasian correspondence is supermodular implementable with a balanced mechanism.

In our mechanisms, agents essentially choose a price-quantity pair. Interestingly, while the quantity they choose is their own consumption bundle, the price they choose is the price vector for the other agents. Supermodular mechanisms derive their nice properties from the fact that agents want to increase their price-quantity announcement, as others do the same. As a result, dynamics inherit a certain monotonicity which leads play towards equilibrium.

6.2.2 Mechanism Design as a Game-Theoretic Foundation for Walrasian Equilibria

General equilibrium theory has long been criticized for providing poor justifications of the Walrasian equilibrium. There is no plausible theory of how economies attain competitive equilibrium (see Kirman, 1989), and some researchers also question the rationality postulate on the agents (Simon, 1978, e.g.).

Our paper provides mechanisms that help guide boundedly rational agents to play equilibrium profiles whose outcomes are Walrasian allocations. So, we offer a game-theoretic explanation of how competitive equilibrium can emerge using mechanisms that differ from the standard (though vaguely-defined) competitive mechanism. Our explanation uses as its foundations the literature on bounded rationality and learning (Fudenberg and Levine, 1998). While viewing the economy as a non-cooperative game is not new (Shapley and Shubik, 1977; Chatterji and Ghosal, 2004; Gul and Stacchetti, 1999; and Milgrom and Strulovici, 2009), there have been few studies concerned with bounded rationality and convergence in competitive economies (see Crockett et al., 2008, for one exception).

Our contribution stems from the limits of Walrasian tâtonnement, which is the first well-defined dynamics to formalize a market equilibration process. Despite its intuitive appeal—operating through excess demand—it lacks certain foundational components. For example, the forces at work behind the price adjustment and the behavioral assumptions on the agents are unclear. Moreover, many economies may not be able to attain a competitive equilibrium via such a process; although we focus on quasilinear economies where stability and uniqueness are guaranteed when preferences for the non-numéraire good are strictly convex and monotonic (see Brown and Calsamiglia, 2007 or Hildenbrand, 1983), we do not require monotonicity and so the known stability results do not apply. Without monotonicity, stability—even existence—is not guaranteed; e.g., goods may not be gross substitutes even though there are no income effects. Furthermore, equilibrium uniqueness only means that there is a unique price ratio for which excess demands are null; this price ratio could correspond to several Walrasian allocations. A dynamic adjustment process can help in this situation to determine to which allocations consumers will converge.

A Proofs

A.1 Proposition 1

Proof of Proposition 1. For any $\theta \in \Theta$ let $p_i(\theta)$ be agent i's price for good y_i at the Walrasian or Lindahl equilibrium for environment θ. For any $m \in \nu(\Theta)$ let $\phi(m) \in \Theta$ identify an environment θ for which m is an equilibrium. Thus, $p_i(\phi(m))$ is the Walrasian or Lindahl price that must be charged to agent i in the environment $\phi(m)$. Pick any $m^* = (r^*, s^*)$ in the relative interior of $\nu(\Theta|s^*)$ and, for notational simplicity, let $y_i^* = y_i(r^*)$ and $x_i^* = x_i(m^*)$. The proof then follows from three important observations that must be true at m^* for each $i \in \mathcal{I}$:

1. Because m^* is a Nash equilibrium for some $\theta \in \Theta$ the following first-order condition is satisfied for each $k \in \{1, \ldots, J_i\}$:

$$\frac{\partial u_i(x_i^*, y_i^*|\theta_i)}{\partial y_i} \frac{\partial y_i(r_i^*)}{\partial r_{ik}} = \frac{\partial u_i(x_i^*, y_i^*|\theta_i)}{\partial x_i} \left[q_i(r_{-i}^*, s^*) \frac{\partial y_i(r_i^*)}{\partial r_{ik}} + \frac{\partial g_i(r^*, s^*)}{\partial r_{ik}} \right]. \tag{13}$$

2. If m^* maps to a Walrasian or Lindahl equilibrium for some $\theta \in \Theta$ then it must be that the transfers collected by the mechanism equals the transfers of the numéraire required by the

Walrasian or Lindahl equilibrium:

$$q_i(r^*_{-i}, s^*)y_i(r^*) + g_i(r^*, s^*) = p_i(\phi(r^*, s^*))y_i(r^*). \tag{14}$$

3. If m^* maps to a Walrasian or Lindahl equilibrium for some $\theta \in \Theta$ then the Walrasian or Lindahl price must equal the marginal rate of substitution of y_i in terms of x_i:

$$\frac{\partial u_i(x^*_i, y^*_i|\theta_i)/\partial y_i}{\partial u_i(x^*_i, y^*_i|\theta_i)/\partial x_i} = p_i(\phi(r^*, s^*)). \tag{15}$$

Dividing both sides of (13) by $\partial u_i/\partial x_i$, inserting equation (15), and rearranging gives

$$\frac{\partial g_i(r^*, s^*)}{\partial r_{ik}} = \left[p_i(\phi(r^*, s^*)) - q_i(r^*_{-i}, s^*)\right] \frac{\partial y_i(r^*)}{\partial r_{ik}}. \tag{16}$$

for each i and k.

By assumption 1 and the fact that m^* is in the relative interior of $\nu(\Theta|s^*)$ these equations must also hold for every m in some open neighborhood in $\nu(\Theta|s^*)$ around m^*. Thus, we can differentiate equation (14) with respect to each r_{ik} to get

$$\frac{\partial g_i(r^*, s^*)}{\partial r_{ik}} = \left[p_i(\phi(r^*, s^*)) - q_i(r^*_{-i}, s^*)\right] \frac{\partial y_i(r^*)}{\partial r_{ik}} + \frac{dp_i(\phi(r^*, s^*))}{dr_{ik}} y_i(r^*). \tag{17}$$

Compare this to equation (16); it must be true that either $y_i(r^*) = 0$ or $dp_i(\phi(r^*, s^*)/dr_{ik} = 0$ for all k.

If $y_i(r^*) = 0$ then by equation (14) we have $g_i(r^*, s^*) = 0$, giving the result.

If $y_i(r^*) \neq 0$ but $dp_i(\phi(r^*, s^*)/dr_{ik} = 0$ for all k then $g_i(r^*, s^*)$ can be expressed as $h_i(r^*_{-i}, s^*)y_i(r^*)$ for some function h_i that does not depend on r_i. But then $x_i(r^*, s^*)$ can be re-written as:

$$x_i(r^*, s^*) = -\left[q_i(r^*_{-i}, s^*) + h_i(r^*_{-i}, s^*)\right] y_i(r^*).$$

Label the bracketed term as $\tilde{q}_i(r^*_{-i}, s^*)$ and we have that

$$x_i(r^*, s^*) = -\tilde{q}_i(r^*_{-i}, s^*)y_i(r^*),$$

and again no 'penalty' term appears. □

A.2 Proposition 2

Proposition 2 can be weakened to allow for mechanisms that fail the Hölder continuity requirement (assumption 3), which would happen if the x_i functions use polynomials of order more than twice as large as the y_i functions. This is done by modifying assumptions 3 and 4 to:

Assumption 3' (Weak Hölder Continuity)**.** Associated with the mechanism Γ is some $\rho \in \{1, 2, \ldots\}$ such that for all $r \in \mathcal{R}$, all $s \in \sigma(r)$, and all $i \in \mathcal{I}$ there exists some finite $\gamma_i(r) > 0$ such that for all $r_i' \in \mathcal{R}_i$ and $s_i' \in \sigma_i(r_i', r_{-i}, s_{-i})$,

$$\left|x_i(r_i', r_{-i}, s_i', s_{-i}) - x_i(r, s)\right| \leq \gamma_i(r) \max\left\{\left|y_i(r_i', r_{-i}) - y_i(r)\right|^{\rho}, \left|y_i(r_i', r_{-i}) - y_i(r)\right|^{1/\rho}\right\},$$

and

Assumption 4' (Rich Type Space)**.** Let $\hat{\rho} \in \{2, 4, 6, \ldots\}$ be the smallest even value of ρ satisfying assumption 3'. For each vector $(\alpha_i, \beta_i)_{i=1}^n \in (\mathbb{R}_{++} \times \mathbb{R})^n$ there is some $\theta \in \Theta$ such that for each i,

$$u_i(x_i, y_i|\theta_i) = \left(-\alpha_i y_i^{\hat{\rho}} + \beta_i y_i\right) + x_i.$$

When $\rho = 2$, assumptions 3' and 4' are identical to assumptions 3 and 4, respectively. As ρ increases assumption 3' becomes strictly weaker. Technically, assumption 4' becomes neither stronger or weaker as ρ changes, but in practical terms a higher ρ requires preferences that are more 'exotic' (using higher-order polynomials) and may therefore be viewed as less desirable.

Given these modified assumptions, we can now prove a generalization of proposition 2 for any $\rho \geq 1$; again, the case of $\rho = 2$ reduces to the original statement of proposition 2.

Proposition 2'. Take any mechanism $\Gamma = (\mathcal{M}_i, x_i, y_i)_{i \in \mathcal{I}}$ satisfying assumptions 1, 2 and 3' for some $\rho \in \{1, 2, 3, \ldots\}$ and any type space Θ satisfying assumption 4'. If $\rho \leq 2$ then every $m \in \mathcal{M}^*$ is a Nash equilibrium of Γ for some $\theta \in \Theta$. If $\rho > 2$ then every $m \in \mathcal{M}^*$ such that $y_i(r) \neq 0$ for all i is a Nash equilibrium of Γ for some $\theta \in \Theta$.

Proof of Proposition 2'. Let $\hat{\rho}$ be the smallest even number weakly greater than ρ. Define $\mathcal{M}^{**}$ by

$$\mathcal{M}^{**} = \left\{(r, s) \in \mathcal{M}^* : (\forall i \in \mathcal{I})\ y_i(r)^{\hat{\rho}-2} \neq 0\right\}.$$

Note that if $\rho \in \{1, 2\}$ then $\hat{\rho} = 2$ and $\mathcal{M}^* = \mathcal{M}^{**}$ (using the convention that $0^0 = 1$). Proposition 2' can then be proven by showing that $\mathcal{M}^{**} \subseteq \nu(\Theta)$. This is done by constructing a mapping $\phi : \mathcal{M}^{**} \to \Theta_{\hat{\rho}}$ (where $\Theta_{\hat{\rho}} \subseteq \Theta$ by assumption 4) such that $m \in \nu(\phi(m))$ for all $m \in \mathcal{M}^{**}$. Thus,

$$\mathcal{M}^{**} \subseteq \nu(\phi(\mathcal{M}^{**})) = \nu(\Theta_{\hat{\rho}}) \subseteq \nu(\Theta),$$

giving the result.

Specifically, consider the mapping $\phi : \mathcal{M}^{**} \to \Theta_{\hat{\rho}}$ such that $\phi_i(m^*) = (\alpha_i(m^*), \beta_i(m^*)) \in \mathbb{R}_+ \times \mathbb{R}$ for each $m^* \in \mathcal{M}^{**}$ and

$$u_i(x_i, y_i|\phi_i(m^*)) = v_i(y_i|\phi_i(m^*)) + x_i,$$

where

$$v_i(y_i|\phi_i(m^*)) = -\frac{\alpha_i(m^*)}{\hat{\rho}} y_i^{\hat{\rho}} + \beta_i(m^*) y_i$$

and, for a given value of $\alpha_i(m^*)$ (to be determined later in the proof), $\beta_i(m^*)$ is given by

$$\beta_i(r^*, s^*) := \alpha_i(r^*, s^*) y_i^{\hat{\rho}-1}(r) + P_{ik}(r^*, s^*) \tag{18}$$

(recall that P_{ik} is the effective price function defined in equation 7 and does not depend on k since $m^* \in \mathcal{M}^{**}$).

We now fix an arbitrary $m^* = (r^*, s^*) \in \mathcal{M}^{**}$ and show that m_i^* is a best response to m_{-i}^* for each i in environment $\phi(m^*) = (\alpha_i(m^*), \beta_i(m^*))_{i \in I}$. This is done in two steps; first we verify that m_i^* is a local optimum in response to m_{-i}^* for each i and then we show m_i^* can be made a global optimum by increasing $\alpha_i(m^*)$ sufficiently (allowing $\beta_i(m^*)$ to adjust appropriately as $\alpha_i(m^*)$ changes).

Given $\phi_i(m^*)$, i's objective is to choose (r_i, s_i) to maximize

$$-\frac{\alpha_i(m^*)}{\hat{\rho}} y_i(r_i, r_{-i}^*)^{\hat{\rho}} + \beta_i(m^*) y_i(r_i, r_{-i}^*) + x_i(r_i, s_i, r_{-i}^*, s_{-i}^*). \tag{19}$$

For local optimality, the first-order conditions for each s_{ik} are already satisfied at m^* by the construction of $\mathcal{M}^{**}$ (see equation 9). As for r_{ik}, agent i's first-order condition for utility maximization at (r^*, s^*) with respect to each r_{ik} is

$$\left[-\alpha_i(r^*, s^*) y_i^{\hat{\rho}-1}(r) + \beta_i(r^*, s^*)\right] \frac{\partial y_i(r)}{\partial r_{ik}} + \frac{\partial x_i(r, s)}{\partial r_{ik}} = 0.$$

But the construction of β_i (equation 18) guarantees that this is satisfied at $(r, s) = (r^*, s^*)$ for any $\alpha_i(r^*, s^*)$, so the first-order conditions are satisfied for all $m^* \in \mathcal{M}^{**}$.

To describe the second-order conditions for local optimality, we show that the matrix of second-partial derivatives of i's objective function will be negative definite for sufficiently large $\alpha_i(m^*)$. Shortening notation, let $\mathbf{X_r}$ and $\mathbf{X_s}$ be the column vectors of partial derivatives of x_i with respect to r_i and s_i, respectively, and let $\mathbf{X_{rr}}$, $\mathbf{X_{rs}}$, and $\mathbf{X_{ss}}$ represent the matrices of cross-partial derivatives of x_i. Similarly define $\mathbf{Y_r}$ and $\mathbf{Y_{rr}}$ as the partial and cross-partial derivatives of y_i, respectively. Using this notation, the matrix of second partial derivatives of the objective function (19) (after inserting the definition of $\beta_i(m^*)$ from equation 18) is given by the $K_i \times K_i$ matrix

$$\mathbf{H_i} = \left[\begin{array}{c:c} -\alpha_i(m^*)(\hat{\rho}-1) y_i(r^*)^{\hat{\rho}-2} \left(\mathbf{Y_r} \cdot \mathbf{Y_r}^{\mathrm{T}}\right) + P_{ik}(m^*) \mathbf{Y_{rr}} + \mathbf{X_{rr}} & \mathbf{X_{rs}} \\ \hdashline \mathbf{X_{rs}}^{\mathrm{T}} & \mathbf{X_{ss}} \end{array}\right],$$

where again $P_{ik}(m^*)$ does not depend on k since $m^* \in \mathcal{M}^{**}$. Now take any direction $(\mathbf{d_r}, \mathbf{d_s}) \neq \mathbf{0}$ of deviation from m_i^*. Since $m^* \in \mathcal{M}^{**}$ implies $s^* \in \sigma(r^*)$, we know that any deviation with $\mathbf{d_r} = \mathbf{0}$ will not yield strictly higher utility, hence $(\mathbf{0}, \mathbf{d_s})^{\mathrm{T}} \cdot \mathbf{H_i} \cdot (\mathbf{0}, \mathbf{d_s}) \leq 0$. For any direction $(\mathbf{d_r}, \mathbf{d_s})$ with

$\mathbf{d_r} \neq \mathbf{0}$ we have

$$\begin{aligned}(\mathbf{d_r}, \mathbf{d_s})^{\mathrm{T}} \cdot \mathbf{H_i} \cdot (\mathbf{d_r}, \mathbf{d_s}) &= -\alpha_i(m^*)(\hat{\rho}-1)y_i(r^*)^{\hat{\rho}-2}\, \mathbf{d_r}^{\mathrm{T}} \left(\mathbf{Y_r} \cdot \mathbf{Y_r}^{\mathrm{T}}\right) \mathbf{d_r} + K_i(m^*) \\ &= -\alpha_i(m^*)(\hat{\rho}-1)y_i(r^*)^{\hat{\rho}-2} \left(\mathbf{d_r}^{\mathrm{T}}\mathbf{Y_r}\right)^2 + K_i(m^*)\end{aligned}$$

where

$$K_i(m^*) = \mathbf{d_r}^{\mathrm{T}} \left[P_{ik}(m^*)\mathbf{Y_{rr}} + \mathbf{X_{rr}}\right] \mathbf{d_r} + 2\mathbf{d_r}^{\mathrm{T}}\mathbf{X_{rs}}\mathbf{d_s} + \mathbf{d_s}^{\mathrm{T}}\mathbf{X_{ss}}\mathbf{d_s}.$$

Since x_i and y_i are continuously differentiable and $\partial y_i/\partial r_i$ is bounded away from zero, $K_i(m^*)$ is finite for all m^*. Because $y_i(r^*)^{\hat{\rho}-2} \neq 0$, α_i can be chosen to be any function satisfying

$$\alpha_i(m^*) > K_i(m^*) \left((\hat{\rho}-1)y_i(r^*)^{\hat{\rho}-2}\right)^{-1} \left(\mathbf{d_r}^{\mathrm{T}}\mathbf{Y_r}\right)^{-2}$$

for all $m^* \in \mathcal{M}^{**}$, so that $(\mathbf{d_r}, \mathbf{d_s})^{\mathrm{T}} \cdot \mathbf{H_i} \cdot (\mathbf{d_r}, \mathbf{d_s}) < 0$. Thus, m_i^* is a local best response to m_{-i}^* for large enough $\alpha_i(m^*)$.

We now construct $\phi_i(m^*)$ by increasing $\alpha_i(m^*)$ until m_i^* is a *global* best response to m_{-i}^*. Since m_i^* is a local best response, there is some neighborhood $\mathcal{N}_i(m^*)$ of m_i^* on which m_i^* maximizes i's utility given $\alpha_i(m^*)$. Although increasing α_i may change the neighborhood around m^* on which m_i^* is a local best response, the neighborhood can only increase in size as α_i is increased. Thus, we ignore this dependence of $\mathcal{N}_i(m^*)$ on α_i and show that any $m_i' \notin \mathcal{N}_i(m^*)$ yields a lower payoff than m_i^* when α_i is sufficiently large.

To proceed, pick any m_i' and m_i'' such that $m_i^* \in (m_i', m_i'') \subset \mathcal{N}_i(m^*)$ and, to shorten notation, let $y_i^* = y_i(r^*)$, $x_i^* = x_i(m^*)$, $y_i' = y_i(r_i', r_{-i}^*)$, $x_i' = x_i(m_i', m_{-i}^*)$, $y_i'' = y_i(r_i'', r_{-i}^*)$, and $x_i'' = x_i(m_i'', m_{-i}^*)$.

To show that $u_i(x_i^*, y_i^*) - u_i(x_i', y_i') \geq 0$ for some α_i', we expand this expression to get

$$\alpha_i' \left[\left(\frac{\hat{\rho}-1}{\hat{\rho}} y_i^{*\hat{\rho}} + \frac{1}{\hat{\rho}} y_i'^{\hat{\rho}}\right) - \left(y_i^{*\hat{\rho}}\right)^{\frac{\hat{\rho}-1}{\hat{\rho}}} \left(y_i'^{\hat{\rho}}\right)^{\frac{1}{\hat{\rho}}}\right] + P_{ik}(m^*) \left(y_i^* - y_i'\right) \geq \left(x_i' - x_i^*\right),$$

which, by assumption 3', is true if

$$\begin{aligned}\alpha_i' \left[\left(\frac{\hat{\rho}-1}{\hat{\rho}} y_i^{*\hat{\rho}} + \frac{1}{\hat{\rho}} y_i'^{\hat{\rho}}\right) - \left(y_i^{*\hat{\rho}}\right)^{\frac{\hat{\rho}-1}{\hat{\rho}}} \left(y_i'^{\hat{\rho}}\right)^{\frac{1}{\hat{\rho}}}\right] + P_{ik}(m^*) \left(y_i^* - y_i'\right) \geq \\ \gamma_i(m^*)\hat{\rho} \max\left\{|y_i^* - y_i'|^{\hat{\rho}}, |y_i^* - y_i'|^{\frac{1}{\hat{\rho}}}\right\}\end{aligned} \tag{20}$$

(the extra $\hat{\rho}$ before the maximizing operator is needed for a later step). But the term in square brackets is the difference between the weighted arithmetic mean and the weighted geometric mean of the two points $y_i^{*\hat{\rho}}$ and $y_i'^{\hat{\rho}}$; by the AM-GM inequality this difference is positive. Thus, there is some finite α_i' at which inequality (20) is true. Similarly, there is some finite α_i'' at which the expression $u_i(x_i^*, y_i^*) - u_i(x_i'', y_i'') \geq 0$ is true. Let $\alpha_i(m^*) = \max\{\alpha_i', \alpha_i''\}$ and now fix $\phi_i(m^*) = (\alpha_i(m^*), \beta_i(m^*))$.

Suppose that $y_i' < y_i''$ (the proof for the case where $y_i'' < y_i'$ is symmetric) and pick any $y_i \geq y_i''$. Suppose that

$$\alpha_i(m^*)\left[\left(\frac{\hat{\rho}-1}{\hat{\rho}}y_i^{*\hat{\rho}}+\frac{1}{\hat{\rho}}y_i^{\hat{\rho}}\right)-\left(y_i^{*\hat{\rho}}\right)^{\frac{\hat{\rho}-1}{\hat{\rho}}}\left(y_i^{\hat{\rho}}\right)^{\frac{1}{\hat{\rho}}}\right]+P_{ik}(m^*)\left(y_i^*-y_i\right) \qquad (21)$$
$$-\gamma_i(m^*)\hat{\rho}\max\left\{|y_i^*-y_i|^{\hat{\rho}},|y_i^*-y_i|^{\frac{1}{\hat{\rho}}}\right\}\geq 0,$$

which is true for $y_i = y_i''$ (see inequality (20)). Then the derivative of the left-hand side of this inequality is positive, implying that the inequality is true *for all* $y_i \geq y_i''$; to see this, take the derivative of the left-hand side and multiply by $(y_i - y_i^*) > 0$ to get either

$$\alpha_i(m^*)\left[y_i^{*\hat{\rho}}-y_i^{*\hat{\rho}-1}y_i+y_i^{\hat{\rho}}-y_i^*y_i^{\hat{\rho}-1}\right]+P_{ik}(m^*)(y_i^*-y_i)-\gamma_i(m^*)\hat{\rho}(y_i-y_i^*)^{\hat{\rho}} \qquad (22)$$

or

$$\alpha_i(m^*)\left[y_i^{*\hat{\rho}}-y_i^{*\hat{\rho}-1}y_i+y_i^{\hat{\rho}}-y_i^*y_i^{\hat{\rho}-1}\right]+P_{ik}(m^*)(y_i^*-y_i)-\gamma_i(m^*)\frac{1}{\hat{\rho}}(y_i-y_i^*)^{1/\hat{\rho}}. \qquad (23)$$

In either case, the expression is greater than the left-hand side of (21) because

$$\left[y_i^{*\hat{\rho}}-y_i^{*\hat{\rho}-1}y_i+y_i^{\hat{\rho}}-y_i^*y_i^{\hat{\rho}-1}\right]\geq\left[\left(\frac{\hat{\rho}-1}{\hat{\rho}}y_i^{*\hat{\rho}}+\frac{1}{\hat{\rho}}y_i^{\hat{\rho}}\right)-\left(y_i^{*\hat{\rho}}\right)^{\frac{\hat{\rho}-1}{\hat{\rho}}}\left(y_i^{\hat{\rho}}\right)^{\frac{1}{\hat{\rho}}}\right]$$

reduces to

$$\left(\frac{\hat{\rho}-1}{\hat{\rho}}y_i^{*\hat{\rho}}+\frac{1}{\hat{\rho}}y_i^{\hat{\rho}}\right)\geq\left(y_i^{*\hat{\rho}}\right)^{\frac{\hat{\rho}-1}{\hat{\rho}}}\left(y_i^{\hat{\rho}}\right)^{\frac{1}{\hat{\rho}}},$$

which is just the AM-GM inequality again. Thus, both (22) and (23) are positive. By continuity, (21) is positive for all $y_i \geq y_i''$ and so deviations resulting in $y_i \geq y_i''$ are not profitable. A symmetric argument shows that deviations to $y_i \leq y_i'$ are also not profitable. Since we already know that deviations resulting in $y_i \in (y_i', y_i'')$ are unprofitable, the proof is complete. □

Proof of Theorem **??**. By means of contradiction, suppose that there exists a one-dimensional mechanism that supermodularly implements the Lindahl correspondence. By Theorem **??**, the transfer function in this mechanism must take the following form:

$$x_i(m) = q_i(m_{-i})y(m)$$

where q_i is a price function that does not depend on i's message. Consider a particular environment θ', where $v_i(y|\theta') = v_i(y)$ for all y and i. Since the mechanism supermodularly implements the Lindahl correspondence, we have for any i and $j \neq i$,

$$\frac{d^2v_i(y(m))}{dy^2} \geq \frac{\partial q_i(m_{-i})}{\partial m_j}. \qquad (24)$$

By strict convexity of preferences, (24) implies that

$$0 > \frac{\partial q_i(m_{-i})}{\partial m_j} \tag{25}$$

for any i and $j \neq i$. It follows from the Lindahl requirement that $\sum_i q_i(m^*_{-i}) = \kappa$ for every Nash equilibrium m^*. We know from Corollary **??** that every message profile is an equilibrium when the type space is rich enough. So, $\sum_i q_i(m_{-i}) = \kappa$ for all m_{-i}, and thus

$$\frac{\partial}{\partial m_k} \sum_i q_i(m_{-i}) = \sum_i \frac{\partial q_i(m_{-i})}{\partial m_k} = 0.$$

This is a contradiction, because (25) implies

$$0 > \sum_i \frac{\partial q_i(m_{-i})}{\partial m_k}.$$

□

Proof of Proposition **??**. By way of contradiction, suppose that there exists a symmetric one-dimensional mechanism that supermodularly implements the Lindahl correspondence. By Theorem **??**, the transfer function in this mechanism must take the following form:

$$x_i(m) = q_i(m_{-i})y(m)$$

where q_i is the symmetric price function that does not depend on i's message. Consider a particular environment θ', where $v_i(y|\theta') = v_i(y)$ for all y and i. Since the mechanism supermodularly implements the Lindahl correspondence, we have for any i and $j \neq i$,

$$\frac{d^2 v(y(m))}{dy^2} \frac{\partial y(m)}{\partial m_i} \frac{\partial y(m)}{\partial m_j} + \frac{dv(y(m))}{dy} \frac{\partial^2 y(m)}{\partial m_i \partial m_j} \geq \frac{\partial q_i(m)}{\partial m_j} \frac{\partial y(m)}{\partial m_i} + q_i(m_{-i}) \frac{\partial^2 y(m)}{\partial m_i \partial m_j} \tag{26}$$

for all m. Since the mechanism is symmetric, it induces a symmetric supermodular game, which guarantees that a symmetric equilibrium exists.[3] Denote this equilibrium by $m^* = (m^*, \ldots, m^*)$. Symmetry also implies $\partial y(m^*)/\partial m_i = \partial y(m^*)/\partial m_j$ for all distinct i and j, hence

$$\frac{\partial y(m)}{\partial m_i} \frac{\partial y(m)}{\partial m_j} > 0.$$

[3]In particular, the extremal equilibria in symmetric supermodular games are symmetric.

By strict convexity of preferences, (26) implies

$$\frac{\partial^2 y(m^*)}{\partial m_i \partial m_j}\left(\frac{dv(y(m^*))}{dy} - q_i(m^*_{-i})\right) > \frac{\partial q_i(m^*_{-i})}{\partial m_j}\frac{\partial y(m^*)}{\partial m_i}. \tag{27}$$

In any equilibrium,[4]

$$\frac{dv(y(m^*))}{dy} = q_i(m^*_{-i})$$

so that (27) becomes

$$0 > \frac{\partial q_i(m^*_{-i})}{\partial m_j}\frac{\partial y(m^*)}{\partial m_i}. \tag{28}$$

It follows from the Lindahl requirement that $\sum_i q_i(m^*_{-i}) = \kappa$ for every Nash equilibrium m^*. We know from Corollary **??** that every message profile is an equilibrium when the type space is rich enough. So, $\sum_i q_i(m_{-i}) = \kappa$ for all m_{-i}, and thus

$$0 > \sum_i \frac{\partial q_i(m^*_{-i})}{\partial m_k}\frac{\partial y(m^*)}{\partial m_i} = \frac{\partial y(m^*)}{\partial m_i}\sum_i \frac{\partial q_i(m_{-i})}{\partial m_k} = \frac{\partial y(m^*)}{\partial m_i}\frac{\partial}{\partial m_k}\sum_i q_i(m_{-i}) = 0.$$

So there does exist any symmetric one-dimensional mechanism that supermodularly implements the Lindahl correspondence. □

Proof of Theorem 3. Define transfers x_i^{SM} by modifying the original transfers x_i as follows:

$$x_i^{\text{SM}}(m,a) = x_i(m) + \gamma_i q_i^{\text{SM}}(m_{-i}, a_{-i})y(m) + H_i(m,a,\rho_i)$$

where

$$H_i(m,a,\rho_i) = h_i(m,a_i) + \rho_i \sum_{j\neq i} h_j(m,a_j).$$

Functions h_i and q_i^{SM}, $i = 1,\ldots,n$, are chosen to satisfy several conditions. First, for each $i \in \mathcal{I}$, $h_i(m,a_i)$ is a twice-differentiable function with $m \in \mathcal{M}$ and $a_i \in \mathbb{R}$ such that

$$\frac{\partial^2 h_i(m,a)}{\partial m_i \partial a_j} > 0,\ \frac{\partial^2 h_i(m,a)}{\partial a_i \partial m_k} \geq 0,\ \frac{\partial^2 h_i(m,a)}{\partial m_i \partial m_j} \geq \underline{B} > -\infty$$

for all m, a, $j \neq i$ and $k, i \in \mathcal{I}$. Furthermore, for all m and i,

$$\max_{a_i \in \mathbb{R}} h_i(m,a_i) = h_i(m,a^*(y(m))) = 0.$$

Let $\mathbf{a}^*(y(m))$ denote the n-dimensional vector of identical entries $a^*(y(m))$, and let $\mathbf{a}^*_{-i}(m)$ be its

[4]By the first-order conditions, $\frac{dv(y(m^*))}{dy}\frac{\partial y(m^*)}{\partial m_i} = q_i(m^*_{-i})\frac{\partial y(m^*)}{\partial m_i}$. Since $\frac{\partial y(m^*)}{\partial m_i}$ is non-zero by assumption, the equality follows.

$(n-1)$-dimensional analog. The function $q_i^{\text{SM}}(m_{-i}, a_{-i})$ is a twice-differentiable function such that

$$\frac{\partial q_i^{\text{SM}}(m_{-i}, a_{-i})}{\partial m_j} \geq \underline{b} > 0, \sum_i q_i^{\text{SM}}(m_{-i}, \mathbf{a}^*_{-i}(y(m))) = 0,$$

for all m_{-i}, a_{-i}, $j \neq i$, $i \in \mathcal{I}$. Finally, assume there is $\Delta > 0$ for which

$$\frac{\displaystyle\sum_{k\neq i} \frac{\partial^2 h_k(m, a_k)}{\partial m_i \partial m_j}}{\dfrac{\partial q_i^{\text{SM}}(m_{-i}, a_{-i})}{\partial m_j}} > \frac{\dfrac{\partial^2 h_j(m, a_j)}{\partial m_i \partial a_j}}{\dfrac{\partial q_i^{\text{SM}}(m_{-i}, a_{-i})}{\partial a_j}} + \Delta, \tag{29}$$

for all (m, a). Under the above transfers, agent i's utility function is

$$u_i^{\text{SM}}(m, a_i, a_{-i}) = v_i(y_i(m)|\theta) + q_i(m_{-i})y(m) + \gamma_i q_i^{\text{SM}}(m_{-i}, a_{-i})y(m) + H_i(m, a, \rho_i).$$

We first show that there exist values of γ_i and ρ_i for which mechanism $(\mathcal{A}_i, y, x_i^{\text{SM}})_{i\in\mathcal{I}}$ induces a supermodular game. Computing the cross-partial derivatives with respect to (m_i, m_j) and (m_i, a_j) gives

$$\frac{\partial^2 u_i^{\text{SM}}(m, a_i, a_{-i})}{\partial m_i \partial m_j} \geq T_i + \gamma_i \frac{\partial q_i^{\text{SM}}(m_{-i}, a_{-i})}{\partial m_j} + \frac{\partial^2 h_i(m, a_i)}{\partial m_i \partial m_j} + \rho_i \sum_{j\neq i} \frac{\partial^2 h_j(m, a_j)}{\partial m_i \partial m_j}$$

where[5]

$$\frac{d^2 v_i(y(m)|\theta)}{dy^2} + \frac{\partial q_i(m_{-i})}{\partial m_j} \geq T_i$$

and

$$\frac{\partial^2 u_i^{\text{SM}}(m, a_i, a_{-i})}{\partial m_i \partial a_j} = \gamma_i \frac{\partial q_i^{\text{SM}}(m_{-i}, a_{-i})}{\partial a_j} + \rho_i \sum_{j\neq i} \frac{\partial^2 h_j(m, a_j)}{\partial m_i \partial a_j}.$$

So, the utility function u_i^{SM} has increasing differences in (m_i, m_j) and (m_i, a_j) if

$$\gamma_i > \rho_i \left(-\frac{\displaystyle\sum_{k\neq i} \frac{\partial^2 h_k(m, a_k)}{\partial m_i \partial m_j}}{\dfrac{\partial q_i^{\text{SM}}(m_{-i}, a_{-i})}{\partial m_j}} \right) - \frac{T_i + \dfrac{\partial h_i(m_{-i}, a_{-i})}{\partial m_j \partial m_i}}{\dfrac{\partial q_i^{\text{SM}}(m_{-i}, a_{-i})}{\partial m_j}}$$

$$\gamma_i < \rho_i \left(-\frac{\dfrac{\partial^2 h_j(m, a_j)}{\partial m_i \partial a_j}}{\dfrac{\partial q_i^{\text{SM}}(m_{-i}, a_{-i})}{\partial a_j}} \right)$$

[5] The existence of a lower bound follows by assumption **??** and because $\frac{\partial q_i(m_{-i})}{\partial m_j}$ is bounded below.

which holds by (29), because the second member on the RHS of the first inequality is bounded. There are actually infinitely many solutions to this system of inequalities and, γ_i and ρ_i can be chosen arbitrarily large. So, $u_i^{\text{SM}}(m,a)$ has increasing differences in (m_i, m_j). Since H_i has increasing differences in (a_i, a_j) and in (a_i, m_k) for all k (including $k = i$) and ρ_i, so does $u_i^{\text{SM}}(m,a)$. We conclude that the game induced by the mechanism is supermodular.

Second, we prove that $(\mathcal{A}_i, (y, x_i^{\text{SM}}))_{i\in\mathcal{I}}$ strongly implements the Lindahl correspondence. To this end, we first show that, for any type θ and Lindahl allocation $(x_i^*(\theta), y^*(\theta))$, there exists a Nash equilibrium (m^*, a^*) such that $(x_i^{\text{SM}}(m^*, a^*), y(m^*, a^*)) = (x_i^*(\theta), y^*(\theta))$. Consider the following system of equations:[6]

$$\begin{cases} y(m) & = & y^*(\theta) \\ q_i(m_{-i}) + \gamma q_i^{\text{SM}}(m_{-i}, \mathbf{a}^*(y(\theta))) & = & -p_i(\theta), \text{ for } i = 1, \ldots, n-1. \end{cases}$$

where $(p_i(\theta))_{i\in\mathcal{I}}$ is the price vector associated with the above Lindahl allocation. Since q_i has bounded derivatives and $\partial q_i^{\text{SM}}/\partial m_j$ is bounded below by $\underline{b} > 0$, there is γ^* large enough so that $\partial(q_i(\cdot) + \gamma^* q_i^{\text{SM}}(\cdot, \mathbf{a}^*(y(\theta)))/\partial m_j$ is bounded above zero for each $j \neq i$. So, $q_i(\cdot) + \gamma^* q_i^{\text{SM}}(\cdot, \mathbf{a}^*(y(\theta)))$ is a surjection from $\mathbb{R}$ onto $\mathbb{R}$ in m_j for all $j \neq i$. As a result, the above system of equations has a solution.[7] For notational purposes, denote that solution m^*. Letting $a_i^* = a^*(y(m^*))$ for all i, we show next that (m^*, a^*) is a Nash equilibrium. From the system of equations, we obtain

$$x_i^{\text{SM}}(m^*, a^*) = -p_i(\theta) y^*(\theta). \tag{30}$$

Agent i has no incentive to unilaterally deviate from a_i^*, because, by definition of h_i,

$$u_i^{\text{SM}}(m^*, a^*) - u_i^{\text{SM}}(m^*, z_i, a_{-i}^*) = h_i(m^*, a^*) - h_i(m^*, a_i, a_{-i}^*) \geq 0.$$

Since $(x_i(\theta), y^*(\theta))_{i\in\mathcal{I}}$ is a Lindahl allocation with price $(p_i(\theta))_{i\in\mathcal{I}}$,

$$v_i(y^*(\theta)|\theta_i) - p_i(\theta) y^*(\theta) \geq v_i(y|\theta_i) - p_i(\theta) y$$

for all y. Therefore

$$v_i(y(m^*)|\theta_i) - p_i(\theta) y(m^*) \geq v_i(y(m_i, m_{-i}^*)|\theta_i) - p_i(\theta) y(m_i, m_{-i}^*) \tag{31}$$

for all m_i. It follows from our assumptions on H_i that $H_i(m_i, m_{-i}^*, a_i^*, a_{-i}^*) \leq 0$ for all m_i and a_{-i}.

[6]Note that if this system admits a solution, then it implies $-q_n(m_{-n}) - \gamma q_n^{\text{SM}}(m_{-n}, \mathbf{a}^*(y(\theta)) = x_n^*(\theta)$.

[7]This holds with more generality, because the statement that "if $f : X \to Y$ is surjective then there is $g : X \to Y$ such that $f(g(x)) = x$" is equivalent to the Axiom of Choice.

From the system of equations, we obtain

$$-p_i(\theta) = q_i(m^*_{-i}) + \gamma q_i^{\text{SM}}(m^*_{-i}, a^*_{-i}),$$

and since $H_i(m_i, m^*_{-i}, a^*_i, a^*_{-i}, \rho_i) \leq 0$ for all m_i and ρ_i,

$$\begin{aligned} -p_i(\theta)y(m_i, m^*_{-i}) &\geq (q_i(m^*_{-i}) + \gamma q_i^{\text{SM}}(m^*_{-i}, z^*))y(m_i, m^*_{-i}) + H_i(m_i, m^*_{-i}, a^*, \rho_i) \\ &= x_i^{\text{SM}}(m_i, m^*_{-i}, a^*) \end{aligned}$$

for all m_i. By (30), (31) implies

$$v_i(y(m^*)|\theta_i) + x_i^{\text{SM}}(m^*, z^*) \geq v_i(y(m_i, m^*_{-i})|\theta_i) + x_i^{\text{SM}}(m_i, m^*_{-i}, z^*)$$

for all m_i, and so,

$$u_i^{\text{SM}}(m^*, a^*) \geq u_i^{\text{SM}}(m_i, m^*_{-i}, a^*)$$

for all m_i. That is, agent i has no incentive to unilaterally deviate from m^*_i. To complete the proof, we need to establish that all Nash equilibria of the mechanism produce Lindahl allocations. In any Nash equilibrium (m^*, a^*), the FOC implies:

$$\frac{\partial v_i(y(m^*)|\theta_i)}{\partial m_i} + (q_i(m^*_{-i}) + \gamma q_i^{\text{SM}}(m^*_{-i}, a^*)) + \rho_i \frac{\partial h_i(m^*, a^*)}{\partial m_i} = 0. \tag{32}$$

We know that $H_i(m'_i, m_{-i}, \mathbf{a}^*(y(m))) \leq h_i(m, \mathbf{a}^*(y(m)))$ for all m'_i, and so

$$\frac{\partial h_i(m, \mathbf{a}^*(y(m)))}{\partial m_i} = 0.$$

Since a^* must be equal to $\mathbf{a}^*(y(m^*))$, (32) implies

$$\frac{dv_i(y(m^*))}{dy} = q_i(m^*_{-i}) + \gamma q_i^{\text{SM}}(m^*_{-i}, a^*). \tag{33}$$

Summing up the RHS of (33) gives

$$\gamma \sum_i q_i^{\text{SM}}(m^*_{-i}, a^*_{-i}) + \sum_i q_i(m^*_{-i}) = \kappa, \tag{34}$$

because $\sum_i q_i^{\text{SM}}(m_{-i}, \mathbf{a}^*_{-i}(y(m))) = 0$ by definition. Given

$$x_i^{\text{SM}}(m^*, a^*) = (q_i(m^*_{-i}) + \gamma q_i^{\text{SM}}(m^*_{-i}, a^*))y(m^*),$$

it follows from (33) and (34) that $(y(m^*), x^*_i(m^*_{-i}, a^*_{-i}))$ is a Lindahl allocation with price vector

$(q_i^{\mathrm{SM}}(m_{-i}^*, a_{-i}^*) + \sum_i q_i(m_{-i}^*))_{i\in\mathcal{I}}$. ☐

Proof of Theorem 4. In what follows, products should be understood as scalar product, and squared expressions as the scalar product of a vector with itself. Denote i's message by $m_i = (s_i, r_i) \in \mathbb{R}^n \times \mathbb{R}^n$. In what follows, products should be understood as scalar product. Define:

1. The allocation function
$$y_i(m) = s_i - \frac{\sum_{j\neq i} s_j}{n-1}.$$

2. The transfers
$$x_i^{\mathrm{SM}}(m) = x_i(m) - \beta q(r_{-i})y_i(s) - \frac{\gamma_i}{2}\left(r_i - \frac{\sum_{j\neq i} r_j}{n-1}\right)^2 - g_i(r) - S_i(m_{-i}). \tag{35}$$

Let
$$u_i^{\mathrm{SM}}(m) = V_i(y_i(m)|\theta) + x_i^{\mathrm{SM}}(m) = (1-\beta)q(r_{-i})y(s) - \frac{\gamma_i}{2}\left(r_i - \frac{\sum_{j\neq i} r_j}{n-1}\right)^2.$$

Step 1. We verify supermodularity. Since agents have $2n$-dimensional message spaces, there are $(n-1)(2n)^2$ *inter*-player pairs for which we must check increasing differences, as well as $\begin{pmatrix} 2 \\ 2n \end{pmatrix}$ *intra*-player pairs. By virtue of the scalar product and by assumption, there is no interaction between the k-th dimension of s_i or r_i and the z-th dimension of s_i or r_i when $z \neq k$. Therefore, we only need to check one of the $\begin{pmatrix} 2 \\ 2n \end{pmatrix}$ *intra*-player pairs. The problem comes down to checking pair (s_i, r_i):[8]
$$\frac{\partial^2 u_i^{\mathrm{SM}}(m)}{\partial s_i \partial r_i} = 0.$$
So, there are increasing differences in (s_i, r_i). Now, we check all inter-player pairs. Since there is no interaction across the message dimensions, the dimension subscript is implicit. Taking $\beta > 1$ and $\gamma_i > 0$ for all i, we check for increasing differences between the strategic variables of agents i and j:

1. Pair (s_i, s_j):
$$\frac{\partial^2 u_i^{\mathrm{SM}}(m)}{\partial s_i \partial s_j} = \frac{\partial^2 V_i(x_i(m)|\theta)}{\partial s_i \partial s_j} = \frac{\partial^2 V_i\left(s_i - \frac{\sum_{j\neq i} s_j}{n-1}|\theta\right)}{\partial s_i \partial s_j} = -\frac{d^2 V_i\left(s_i - \frac{\sum_{j\neq i} s_j}{n-1}|\theta\right)}{dy_i^2(n-1)} > 0,$$

[8] It is implicit here that we are looking at the same dimension of s_i and r_i, say (s_i^1, r_i^1).

because preferences are strictly convex (assumption **??**).

2. Pair (s_i, r_j):
$$\frac{\partial^2 u_i^{\text{SM}}(m)}{\partial s_i \partial r_j} = (1-\beta)\frac{\partial q(r_{-i})}{\partial r_j} > 0.$$

3. Pair (r_i, r_j):
$$\frac{\partial^2 u_i^{\text{SM}}(m)}{\partial r_i \partial r_j} = \gamma_i > 0$$

4. Pair (r_i, s_j):
$$\frac{\partial^2 u_i^{\text{SM}}(m)}{\partial r_i \partial s_j} = 0$$

Step 2. We want to show that if m^* is a Nash equilibrium of $(\mathcal{M}_i, (y_i, x_i^{\text{SM}}))_{i\in\mathcal{I}}$, then $(y_i(m^*), x_i^{\text{SM}}(m^*))_{i\in\mathcal{I}}$ is a Walrasian allocation. For any Nash equilibrium m^* of the mechanism, the first-order conditions must be satisfied:
$$\frac{\partial u_i^{\text{SM}}(m)}{\partial r_i} = -\gamma_i \left(r_i + \frac{\sum_{j\neq i} r_j}{n-1} \right) = 0,\ (i = 1, \dots, n).$$

Hence,

$$r_i^* \quad = \quad \frac{\sum_{j\neq i} r_j^*}{n-1},\ (i = 1, \dots, n), \tag{36}$$

and so $r_i^* = P$ for all $i = 1, \dots, n$. Therefore, the third member of the transfers (RHS of (35)) vanish in equilibrium. Since m^* is an equilibrium, we have

$$V_i(y_i(m^*)|\theta) + x_i^{\text{SM}}(m^*) \geq V_i(y_i(m_i, m_{-i}^*)|\theta) + x_i^{\text{SM}}(m_i, m_{-i}^*) \tag{37}$$

for all m_i. Letting $P^* = -(1-\beta)q(\mathbf{P})$ where $\mathbf{P}$ is a vector of identical entries P, (37) implies,

$$V_i(y_i(m^*)|\theta) - P^* y_i(m_i^*, m_{-i}^*) \geq V_i(y_i(m_i, m_{-i}^*)|\theta) - P^* y_i(m_i, m_{-i}^*)$$

for all m_i. Since $y_i(\cdot, m_{-i}^*) = y_i(\cdot, s_{-i}^*)$ is a surjection from $\mathbb{R}$ onto $\mathbb{R}$,

$$V_i(y_i(m^*)|\theta) - P^* y_i(m_i^*, m_{-i}^*) \geq V_i(y_i|\theta) - P^* \cdot y_i \tag{38}$$

for all y_i. Since $x_i^{\text{SM}}(m^*) = -P^* y_i(m^*)$, the budget equation $P^* y_i(m^*) + x_i^{\text{SM}}(m^*) = 0$ is satisfied at allocation $(y_i(m^*), x_i^{\text{SM}}(m^*))$. By (38), this allocation maximizes the utility subject to the budget equation. It only remains to show that the Nash equilibrium allocation is individually feasible. This allocation is preferred by agent i to her endowment, which is itself strictly preferred to all infeasible allocations. So, $(y_i(m^*), x_i^{\text{SM}}(m^*))$ must be individually feasible.

Step 3. We want to show that if $[(Y_i^*, X_i^*)_{i\in\mathcal{I}}, P^*]$ is a Walrasian equilibrium, then it is the outcome of a Nash equilibrium m^*. Define $Q : r \mapsto q(\mathbf{r})$ where $\mathbf{r}$ is the $(n-1)$-dimensional vector of identical entries r. We construct the following m^*: For each $i = 1, \ldots, n$, take

$$\begin{aligned} s_i^* &= (1 - 1/n)X_i^* \\ r_i^* &= r^* = Q^{-1}\left(\tfrac{-P^*}{1-\beta}\right) \end{aligned}$$

where Q^{-1} exists because q is strictly decreasing in each variable. Because $[(Y_i^*, X_i^*)_{i\in\mathcal{I}}, P^*]$ is a Walrasian equilibrium,

$$V_i(Y_i^*) - P^*Y_i^* \geq V_i(y_i) - P^*y_i$$

for all y_i. As a result, we have

$$V_i(y(m_i^*, m_{-i}^*)|\theta) - P^* \cdot y_i(m_i^*, m_i^*) \geq V_i(y(m_i, m_{-i}^*)|\theta) - P^* \cdot y_i(m_i, m_i^*)$$

for all m_i, and so

$$\begin{aligned} V_i(y(s_i^*, s_{-i}^*)|\theta) + (1-\beta)q(\mathbf{r}^*) \cdot y_i(s_i^*, s_i^*) - \frac{\gamma_i}{2}\left(r_i^* - \frac{\sum_{j\neq i} r_j^*}{n-1}\right) \\ \geq V_i(y(s_i, s_{-i}^*)|\theta) + (1-\beta)q(\mathbf{r}^*) \cdot y_i(s_i^*, s_i^*) - \frac{\gamma_i}{2}\left(r_i - \frac{\sum_{j\neq i} r_j^*}{n-1}\right) \end{aligned} \tag{39}$$

for all (s_i, r_i), because

$$0 = \frac{\gamma_i}{2}\left(r_i^* - \frac{\sum_{j\neq i} r_j^*}{n-1}\right) > \frac{\gamma_i}{2}\left(r_i - \frac{\sum_{j\neq i} r_j^*}{n-1}\right) \geq 0.$$

Hence m^* is a Nash equilibrium whose outcome is $(Y_i^*, X_i^*)_{i\in\mathcal{I}}$.

Step 4. We establish balancedness in equilibrium. From Step 1, we know that any equilibrium must be such that $r_i^* = r^*$. Hence, the third member of the transfers vanishes in equilibrium, and we are left with

$$x_i^{\text{SM}}(m^*) = Py_i(m^*),$$

for some P. So,

$$\sum_{i=1}^{n} x_i^{\text{SM}}(m^*) = P\sum_{i=1}^{n} y_i(m^*) = 0,$$

because $\sum_{i=1}^{n} y_i(m) = 0$ for all m. The mechanism is balanced in equilibrium. $\square$

Proof of Proposition 4. Hurwicz's mechanism is defined as:

1. The allocation function
$$y_i(m) = s_i - \frac{\sum_{j\neq i} s_j}{n-1}.$$

2. The transfers
$$x_i(m) = -\left(\frac{\sum_{j\neq i} r_j}{n-1}\right) y_i(s) - \frac{\gamma_i}{2}\left(r_i - \frac{\sum_{j\neq i} r_j}{n-1}\right)^2 + S_i(m_{-i}). \tag{40}$$

   where S_i is such that $\sum_{i\in\mathcal{I}} x_i(m) = 0$ for all m.

Consider the following transformation of Hurwicz's mechanism where $\beta > 1$:

1. The allocation function
$$y_i(m) = s_i - \frac{\sum_{j\neq i} s_j}{n-1}.$$

2. The transfers
$$x_i^{\text{SM}}(m) = (1-\beta)\left(\frac{\sum_{j\neq i} r_j}{n-1}\right) y_i(s) - \frac{\gamma_i}{2}\left(r_i - \frac{\sum_{j\neq i} r_j}{n-1}\right)^2 + S_i(m_{-i}) + B_i(m_{-i}) \tag{41}$$

   where
$$B_i(m_{-i}) = -\frac{(\sum_{j\neq i} \beta r_j)(\sum_{j\neq i} s_j)}{(n-1)^2} + \frac{\sum_{j\neq i} s_j \sum_{k\neq i,j} \beta r_k}{(n-1)(n-2)}.$$

We omit to prove that the game induced by this mechanism is supermodular, because it is similar to that of Theorem 4. Indeed, the main difference is that there is an extra term, B_i, but it does not depend on i's strategic variables. So, we check balancedness only:

$$\sum_{i\in\mathcal{I}} x_i^{\text{SM}}(m) = -\sum_{i\in\mathcal{I}} \beta\left(\frac{\sum_{j\neq i} r_j}{n-1}\right) y_i(s) + \sum_{i\in\mathcal{I}} B_i(m_{-i}), \tag{42}$$

by virtue of balancedness of Hurwicz's mechanism. Since

$$\begin{aligned}
\sum_{i\in\mathcal{I}} B_i(m_{-i}) &= -\sum_{i\in\mathcal{I}} \beta\left(\frac{\sum_{j\neq i} r_j}{n-1}\right)\left(\frac{\sum_{j\neq i} s_j}{n-1}\right) + \frac{\displaystyle\sum_{i\in\mathcal{I}}\sum_{j\neq i} s_j \sum_{k\neq i,j} \beta r_k}{(n-1)(n-2)} \\
&= -\sum_{i\in\mathcal{I}} \beta\left(\frac{\sum_{j\neq i} r_j}{n-1}\right)\left(\frac{\sum_{j\neq i} s_j}{n-1}\right) + \frac{\displaystyle\sum_{i\in\mathcal{I}}\sum_{j\neq i} s_i \beta r_j}{(n-1)},
\end{aligned}$$

the RHS of (42) is null, establishing that the mechanism is balanced. □

References

Brown, D. J., Calsamiglia, C., 2007. Marshall's theory of value and the strong law of demand, cowles Foundation Discussion Paper 1615.

Cabrales, A., Ponti, G., 2000. Implementation, elimination of weakly dominated strategies and evolutionary dynamics. Journal of Economic Dynamics and Control 3, 247–282.

Camerer, C. F., 2003. Behavioral Game Theory. Princeton University Press, Princeton, NJ.

Chatterji, S., Ghosal, S., 2004. Local coordination and market equilibria. Journal of Economic Theory 114 (2), 255–279.

Chen, Y., 2002. A family of supermodular Nash mechanisms implementing Lindahl allocations. Economic Theory 19, 773–790.

Chen, Y., Gazzale, R., 2004. When does learning in games generate convergence to Nash equilibria? The role of supermodularity in an experimental setting. American Economic Review 95 (5), 1505–1535.

Chen, Y., Plott, C. R., 1996. The Groves-Ledyard mechanism: An experimental study of institutional design 59, 335–364.

Chen, Y., Tang, F.-F., 1998. Learning and incentive-compatible mechanisms for public goods provision: An experimantal study. Journal of Political Economy 106, 633–662.

Crockett, S., Spear, S. E., Sunder, S., 2008. Learning competitive equilibrium. Journal of Mathematical Economics 44, 651–671.

Fudenberg, D., Levine, D. K., 1998. The Theory of Learning in Games. MIT Press.

Groves, T., Ledyard, J. O., 1977. Optimal allocation of public goods: A solution to the 'free-rider' problem. Econometrica 45, 783–809.

Gul, F., Stacchetti, E., 1999. Walrasian equilibrium with gross substitutes. Journal of Economic Theory 87, 95–124.

Healy, P. J., 2006. Learning dynamics for mechanism design: An experimental comparison of public goods mechanisms. Journal of Economic Theory 129 (1), 114–149.

Hildenbrand, W., 1983. On the law of demand. Econometrica 51, 997–1019.

Hurwicz, L., 1972. On informationally decentralized systems. In: McGuire, C. B., Radner, R. (Eds.), Decision and Organization: A Volume in Honor of Jacob Marschak. North-Holland, Amsterdam.

Hurwicz, L., 1979. Outcome functions yielding Walrasian and Lindahl allocations at Nash equilibrium points. Review of Economic Studies 46, 217–225.

Kirman, A. P., 1989. The intrinsic limits of modern economic theory: The emperor has no clothes. Economic Journal 99 (395), 126–139.

Mathevet, L. A., 2007. Supermodular bayesian implementation: Learning and incentive design, mimeo.

Milgrom, P., Roberts, J., 1990. Rationalizability, learning, and equilibrium in games with strategic complementarities. Econometrica 58, 1255–1277.

Milgrom, P., Strulovici, B., 2009. Substitute goods, auctions, and equilibrium. Journal of Economic Theory 144, 212–247.

Milgrom, P. R., Segal, I., 2002. Envelope theorems for arbitrary choice sets. Econometrica 72, 583–601.

Shapley, L., Shubik, M., 1977. Trading using one commodity as a means of payment. Journal of Political Economy 84, 937–968.

Simon, H. A., 1978. Rationality as process and as product of thought. American Economic Review 68 (2), 1–16.

Tian, G., 1990. Completely feasible and continuous implementation of the lindahl correspondence with a message spce of minimal dimension. Journal of Economic Theory 51, 443–452.

Topkis, D. M., 1998. Supermodularity and Complementarity. Princeton University Press, Princeton, NJ.

Walker, M., 1981. A simple incentive compatible scheme for attaining Lindahl allocations. Econometrica 49, 65–71.