Theoretical Economics 7 (2012), 609–661 1555-7561/20120609

Designing stable mechanisms for economic environments

Paul J. Healy
Department of Economics, The Ohio State University

Laurent Mathevet
Department of Economics, University of Texas

We study the design of mechanisms that implement Lindahl or Walrasian allocations and whose Nash equilibria are dynamically stable for a wide class of adaptive dynamics. We argue that supermodularity is not a desirable stability criterion in this mechanism design context, focusing instead on *contractive* mechanisms. We provide necessary and sufficient conditions for a mechanism to Nash-implement Lindahl or Walrasian allocations, show that these conditions are inconsistent with the contraction property when message spaces are one-dimensional, and then show how to use additional dimensions to achieve dynamic stability while gaining budget balance out of equilibrium.

Keywords. Mechanism design, implementation, stability, learning.

JEL classification. C62, C72, C73, D02, D03, D51.

1. Introduction

It is well known that equilibrium outcomes are inefficient in economies with public goods when contributions are voluntary. The mechanism design literature provides various incentive schemes that solve these inefficiencies, assuming agents select equilibrium strategies when playing a mechanism. Historically, the general impossibility results when applying the weak requirement of dominant strategy equilibrium (Gibbard 1973, Satterthwaite 1975, Hurwicz and Walker 1990, Zhou 1991) led to the search for mechanisms that implement optimal public-goods allocations when players are assumed to select Nash equilibrium strategies. Groves and Ledyard (1977), Hurwicz (1979b), and Walker (1981), among others, all provide examples of such mechanisms. From a theoretical standpoint, these mechanisms completely solve the free-rider problem in public-goods economies.

Early laboratory tests, however, revealed that the empirical success of these mechanisms was limited because agents in fact do not play equilibrium strategies; rather, the play of these mechanisms can best be described using myopic learning dynamics, such

Paul J. Healy: healy.52@osu.edu
Laurent Mathevet: lmath@mail.utexas.edu

The authors thank Yaron Azrieli, Yan Chen, Vince Crawford, Federico Echenique, Guillaume Frechette, Matt Jackson, John Ledyard, Stephen Morris, Charlie Plott, William Sandholm, Steve Spear, Matt Van Essen, and Mark Walker for their helpful comments and conversations. Healy gratefully acknowledges the National Science Foundation for funding this research through Award SES-0847406.

Copyright © 2012 Paul J. Healy and Laurent Mathevet. Licensed under the Creative Commons Attribution-NonCommercial License 3.0. Available at http://econtheory.org.
DOI: 10.3982/TE898

as best-response play to a recent history of actions (Chen and Plott 1996, Chen and Tang 1998, Chen and Gazzale 2004, Healy 2006). Thus, mechanisms that induce dynamically stable games drive play to equilibrium, while mechanisms that induce unstable games do not. These "wind tunnel" tests suggest that theorists should add dynamic stability to the constraints of the mechanism design problem and focus on implementing optimal allocations in *stable* Nash equilibria.

In this paper, we take the next step, incorporating these experimental observations about stability back into the theory. We provide example mechanisms that are not only stable, but also budget-balanced both in and out of equilibrium. Budget balance out of equilibrium is vital when discussing dynamic stability, as it ensures that allocations are still feasible when play has not yet converged. We focus on Lindahl equilibrium allocations as our objective function for implementation in public-goods economies, and we extend our results to Walrasian equilibria when all goods are private. We also show how our mechanisms are constructed and we discuss the limitations of our procedure. Specifically, our paper proceeds in four steps.

1. After introducing the basic environment and notation, we provide our notion of dynamic stability: *contractive* mechanisms. A mechanism is contractive if, in every possible environment, it induces a game whose best-response functions are contraction mappings. We prove that a large family of learning dynamics is globally stable in contractive games (Theorem 1) and point out that the family of convergent dynamics is even larger if utilities are concave in one's own strategy. The result for contractive games is an analog of the Milgrom and Roberts (1990) stability result for supermodular games. We also argue that supermodularity—which has been suggested previously as a desirable stability notion—may not guarantee stability in mechanism design settings. This motivates our search for contractive mechanisms.

2. Assuming quasilinear preferences, we provide an example mechanism that fully implements Lindahl allocations, is contractive, gives concave utilities, and is budget balanced both in and out of equilibrium (Theorem 2). We also provide a contractive and budget-balanced mechanism with concave utilities that fully implements Walrasian allocations in private-goods economies (Theorem 3). Quasilinearity cannot be relaxed too far; the results of Jordan (1986) and Kim (1987) guarantee that no (well behaved) mechanism can be stable for general preferences.

3. Next, we show how such mechanisms are constructed. Ignoring all stability concerns and relaxing quasilinearity, we provide necessary and sufficient conditions for continuous mechanisms to fully implement Lindahl or Walrasian allocations in Nash equilibrium (Theorems 4, 5, and 7). These results provide an understanding of what types of mechanisms can be considered for implementation and how desiderata such as stability can then be added. As an example of the restrictiveness of the necessary and sufficient conditions, we prove that if a mechanism has a one-dimensional strategy space for each player, then it cannot fully implement Lindahl (or Walrasian) allocations and be contractive (Theorem 6). This explains the necessity of using two-dimensional message spaces in our example mechanisms.

4. Finally, we discuss the limitations and possible generalizations of our approach. For example, contractive mechanisms for more than two goods can be constructed, but rely on strong assumptions about complementarities. Additionally, no mechanism can be contractive if concavity of preferences becomes arbitrarily small.

Relative to the existing literature, our necessary and sufficient conditions for Nash-implementing Lindahl or Walrasian allocations (ignoring stability concerns) provide a new understanding about the types of mechanisms that can be used in general equilibrium (or "economic") environments.[1] With one-dimensional message spaces, the necessary condition is quite strong: Agents' announcements must represent individual purchases of the non-numeraire good at prices determined by others' messages. Thus, in the public-goods setting, the choice of announcement is equivalent to the choice of the public-goods level, taking prices as given. In this way, the mechanism must parallel the consumer's optimization problem given in the very definitions of Walrasian and Lindahl equilibrium. Sufficiency is obtained by assuming, in addition, that every possible Lindahl or Walrasian allocation can be reached by some announcement.[2]

The motivation for requiring dynamic stability in mechanism design is manifold. First and foremost, this paper continues the dialog between theory and data that hopefully will converge on acceptable mechanisms for real-world application. To aid in the continuation of this dialog, we provide a *recipe* for designing stable mechanisms. New ingredients can be added as new behavioral regularities are discovered. In contrast, a single example of a stable mechanism is less desirable because these new behavioral regularities may render that particular mechanism ineffective.

Dynamic stability of equilibrium also has appeal independent from the existing experimental results. If equilibrium is arrived at through iterated applications of best response in players' internal logic or through iterations of pre-play communication, then stable equilibria are the most likely to arise and are the most robust to perturbations in opponents' logic or pre-play communication. Thus, we also view stability as a device to make static Nash implementation more robust.

It is important to study stability under a wide class of admissible dynamics because experimental evidence suggests that the process of learning can vary dramatically from one environment to another. Existing work on economic environments by Vega-Redondo (1989), de Trenqualye (1989), and Kim (1993, 1996), for example, focuses on particular learning dynamics that may or may not be descriptive in various settings. In

[1]We borrow the phrase "economic environment" from Hurwicz's early work on mechanism design. Hurwicz and Reiter (2006, p. 14) describe economic environments as those concerned with production, consumption, and exchange. Specifically, an economic environment specifies the constraints on those three activities. These are typically a special case of the more general "social choice" environments studied by Arrow (1951) or Maskin (1999), for example.

[2]Our necessary condition was first suggested by Brock (1980) (see also Groves and Ledyard 1987), though not proved generally. Reichelstein and Reiter (1988) use differential geometry techniques to explore the minimal message space size needed for Walrasian implementation in Nash equilibrium. Their proof technique also suggests that "price-taking" is a necessary condition. We describe and prove this claim in a much more straightforward way, extend it to public-goods economies, and add sufficiency results that lead to a useful characterization of implementing mechanisms.

contrast, Chen (2002) develops a public-goods mechanism that is supermodular, following Milgrom and Roberts's (1990) result that supermodular games have dynamically stable equilibrium sets for a wide class of dynamics.[3]

We argue in Section 3 that supermodularity is not an appropriate stability concept for mechanism design in certain contexts. Milgrom and Roberts (1990) prove that if a game is supermodular, then any adaptive learning dynamic (which plays undominated strategies against not too distant histories) must converge to the smallest interval containing all Nash equilibria. But now imagine a supermodular game with Nash equilibria at the corners of the (compact) strategy space. The Milgrom–Roberts stability result is vacuous here since the smallest interval containing all equilibria is the entire strategy space. If, as in the current mechanism design literature, we make the strategy space unbounded, then the corner equilibria are eliminated, but nothing guarantees that the remaining interior equilibria are stable. Thus, supermodularity's stability properties can be highly ambiguous when the strategy space is unbounded or when little is known about the size of the equilibrium set.

Instead of requiring supermodularity, we design mechanisms with contractive best-response functions. We prove that this guarantees stability for many learning dynamics. Our example mechanisms are also fully budget-balanced. Requiring that mechanisms satisfy budget balance out of equilibrium is vitally important when admitting dynamic adjustment processes. If allocations are not balanced out of equilibrium, then the social planner will be required to fund the surplus (or absorb the shortage) in early periods when play has not yet converged. There is no guarantee that early surpluses are offset by later shortages, and the total subsidy required across time may vary greatly depending on the initial conditions and exact path of play. Practical applications, therefore, demand balanced budgets.

Requiring full implementation—where every equilibrium maps to a desirable outcome, and vice versa—is also important in a dynamic context, because it guarantees that agents do not settle on equilibria whose outcomes are not desirable. By dealing with full implementation, our paper presents an advantage over Mathevet (2010), who builds supermodular mechanisms in Bayesian environments but focuses on weak implementation and minimizing the size of the equilibrium set.

The structure of the paper is as follows: We review related literature in the following subsection. We introduce a two-good general-equilibrium model and the basic definitions of implementation in Section 2. In Section 3, we introduce contractiveness as our stability notion and discuss some of the dangers of focusing instead on supermodularity. We then provide two contractive, budget-balanced mechanisms: one for Lindahl allocations and one for Walrasian allocations. We provide necessary and sufficient conditions for a mechanism to implement Lindahl or Walrasian allocations in Section 4. We first study the case of mechanisms with one-dimensional strategy spaces for each agent, show that no one-dimensional mechanism that satisfies these conditions can be contractive, and then generalize the necessary and sufficient conditions to higher dimensional mechanisms. This provides an understanding of how our two mechanisms

[3]Technically, Chen's (2002) mechanism is *open*-supermodular since its strategy space is not compact; see Section 3.

are constructed and how other, similar mechanisms can be constructed. Finally, in Section 5, we discuss how our results can generalize to economies with multiple nonnumeraire goods, why it is difficult to relax our assumption of quasilinearity, and what future directions (and limitations) we foresee for this line of research.

Related literature

We focus attention on Nash implementation in economic environments, where the problem of stability is relatively longstanding. The first (and most well known) fully optimal public-goods mechanism is that of Groves and Ledyard (1977). Muench and Walker (1983) show that as an economy becomes large, the Groves–Ledyard mechanism either becomes highly unstable (in best-response dynamics) if the punishment parameter remains small or payoffs become arbitrarily "flat" if the punishment parameter grows large; in either case, attainability of equilibrium becomes a concern. If preferences are not quasilinear, then the Groves–Ledyard mechanism may have many undesirable equilibria (Bergstrom et al. 1983); however, these equilibria may not be a concern since they are unstable and disappear when the punishment parameter is sufficiently large (Page and Tassier 2004). Chen and Tang (1998) show that the mechanism also becomes supermodular in quasilinear environments with a large punishment parameter, though the critical requirement of a compact strategy space for supermodular games is omitted, leading to ambiguous predictions about stability.[4]

Regardless of its stability properties, a major drawback of the Groves–Ledyard mechanism is that it is not individually rational: Agents' final utility may be lower than that of their initial endowments. Hurwicz (1979a) proves that, under a mild continuity requirement, if one wants to implement Pareto optimal and individually rational outcomes in economic settings, then one must implement the Walrasian or Lindahl equilibrium allocations. From this view, the mechanisms of Hurwicz (1979b) and Walker (1981) that Nash-implement Lindahl allocations are preferable; we refer to such mechanisms as *Nash–Lindahl* mechanisms.

Unfortunately, the Nash–Lindahl mechanisms of Hurwicz and Walker are known to have poor stability properties, and experimental results (Chen and Tang 1998, Healy 2006) confirm that this severely hinders performance. Kim (1987) (following Jordan 1986) shows that for a certain class of preferences, *all* Nash–Lindahl mechanisms must be unstable for at least one preference profile in the class. As mentioned above, Vega-Redondo (1989), de Trenqualye (1989), and Kim (1993, 1996) all design Nash–Lindahl mechanisms that are stable for particular dynamics under various restrictions on preferences. The Kim and Jordan results also force us to restrict preferences in this paper; our stability results are proven assuming quasilinear preferences for the public good. Stability results for completely general preferences are impossible.

The first carefully controlled laboratory experiments of Nash–Lindahl mechanisms were performed by Chen and Plott (1996).[5] The subsequent experimental research

[4]We refer to such games as open-supermodular.

[5]Scherr and Babb (1975), Smith (1979), Harstad and Marrese (1982), and Tideman (1983) ran earlier experiments and many authors tested inefficient public-goods processes such as the voluntary contributions

(Chen and Tang 1998, Chen and Gazzale 2004, Healy 2006) has suggested that supermodularity is a sufficient (if not necessary) condition for subjects to converge to Nash equilibrium. Supermodularity (following Milgrom and Roberts 1990) in this context requires a compact strategy space and implies monotone best responses. Based on the learning result from Milgrom and Roberts (1990), Chen (2002) provides a family of supermodular Nash–Lindahl mechanisms, though their strategy spaces are not compact.

The closest work to ours is Van Essen (forthcoming). He also notes that supermodularity with unbounded strategy spaces does not imply stability and adopts the contraction approach as a "fix" for this instability. He then provides an example mechanism that is both supermodular and contractive. In this paper, we do not require supermodularity since it adds unnecessary constraints to the problem; this is shown more formally in Section 3. We also provide general characterization results for Nash implementation; prove that for a large class of learning dynamics, agents' choices converge to the Nash equilibrium in contractive games; demonstrate an impossibility result for stable one-dimensional mechanisms; focus on a general "recipe" for designing stable mechanisms; provide a Walrasian mechanism; and—most importantly—develop mechanisms that are budget-balanced out of equilibrium. Stable mechanisms that are not budget-balanced may require significant subsidies (or generate large surpluses) while strategies are adjusting toward equilibrium; our budget-balanced mechanisms never create a subsidy or a surplus at any time, guaranteeing feasible outcomes in every period. The stable mechanisms of Van Essen (forthcoming) and Chen (2002) are not budget-balanced out of equilibrium.

More recent experimental work on supermodular mechanisms suggests that the performance of mechanisms out of equilibrium may dramatically affect realized efficiency. Van Essen et al. (2012) and Van Essen (2010) show that the Chen (2002) mechanism is outperformed by the Kim (1993) and Van Essen (forthcoming) mechanisms, respectively, because the latter mechanisms give smaller out-of-equilibrium "punishments" and budget imbalances, resulting in higher overall efficiency. These results highlight the need for out-of-equilibrium budget balance, as well as the usefulness of providing a general recipe for designing mechanisms that can take into account such lessons in the design of future mechanisms.

There is comparatively little work on implementing Walrasian allocations, presumably because decentralized markets generally perform well. Yet an alternative mechanism may be desirable for several reasons. First, if the number of agents is small, then the price-taking assumption becomes tenuous; a mechanism with a game-theoretic foundation is more likely to succeed. Dynamic stability then guarantees that adaptively adjusting agents can still arrive at the Walrasian allocations. Second, adaptive learning models in the competitive mechanism focus on tâtonnement-like adjustment processes where stability is, in general, not guaranteed (Scarf 1960, Hirota 1985) and where feasibility (off-equilibrium) of the consumption plans is also a problem. We focus instead on designing game-theoretic mechanisms that are fully balanced—hence trades are feasible off-equilibrium—and stable under a family of learning dynamics that is known to be

mechanism (see Ledyard 1995), but Chen and Plott (1996) were the first to test directly a theoretically optimal mechanism without modifications in a controlled laboratory setting.

reasonably descriptive. To our knowledge, the only other paper to focus on the design of dynamically stable (approximate) Walrasian mechanisms is Walker (1984).

Various methods for generating stability directly through the solution concept have also been studied. For example, dominant strategy equilibria are certainly dynamically stable for nearly any reasonable learning process. Unfortunately, standard impossibility results severely limit its applicability (Green and Laffont 1977, Roberts 1979). Furthermore, if the dominant strategy is not strict, then myopically adapting agents may converge to undesirable Nash equilibria, as was observed in tests of the Vickrey–Clarke–Groves mechanism by Cason et al. (2006) and Healy (2006) (see also Saijo et al. 2007).

Abreu and Matsushima (1992) provide a mechanism that has a dominance-solvable equilibrium whose outcome is a lottery placing an arbitrarily large weight on the desired allocation. From a theoretical standpoint, the result is very strong; it implies convergence of a wide class of learning dynamics to the equilibrium point. Yet their mechanism is of limited practical use: As the mechanism becomes more precise (placing more weight on the desired allocation), the dimensionality of the message space diverges to infinity. Furthermore, laboratory tests of the mechanism (Sefton and Yavaş 1996, inspired by Glazer and Rosenthal 1992) find that subjects do not move toward the equilibrium over 14 periods of play. This suggests that the speed at which iterated dominance is respected by learning is slow, or nonexistent. These results are in line with the findings of McKelvey and Palfrey (1992), Stahl and Wilson (1995), Nagel (1995), and others, showing that subjects do not appear to learn to play iteratively undominated strategies. These results apparently limit the applicability of mechanisms that rely on iterated-deletion solution concepts.[6]

Sandholm (2002, 2005, 2007) studies stable Nash implementation of efficient resource utilization in congestion games. He uses transfers to convert externality problems with poor stability properties into potential games with excellent stability properties. In more general settings with continuous levels of public and private goods and rich type spaces, however, it is typically not possible to use transfers to create a potential game.

Cabrales (1999) shows that, in the canonical mechanism of Maskin (1999), adaptive Markovian dynamics (placing positive probability on better responses to opponents' last-period strategies) converge to and remain at the Nash equilibrium. Cabrales and Serrano (2011) extend this result, proving that a quasimonotonicity condition is necessary for implementation in the steady states of these dynamics; when a no-worst-alternative condition is also satisfied, implementation can be achieved using a variation on the canonical mechanism. Cabrales (1999) also shows that the Abreu–Matsushima mechanism is vulnerable to "drift" when agents use adaptive Markovian dynamics, since the equilibrium also admits nonequilibrium best responses for each agent.

[6]Bergemann and Morris (2011) consider rationalizable implementation, which is equivalent to iterated deletion of strictly dominated strategies when the strategy space is finite. They show that virtual implementation in iteratively undominated strategies requires a social choice function to select agents' favorite social choice outcome when preferences are identical.

2. The model

Economic environments

Consider a two-good general-equilibrium economy in which agents $i \in \{1, \dots, n\} = \mathcal{I}$ have endowments $\omega_i = (\omega_i^x, \omega_i^y) \in \mathbb{R}^2$, make net trades $z_i = (x_i, y_i) \in \mathbb{R}^2 - \{\omega_i\}$, and have preferences over net trades representable by a utility function $u_i(x_i, y_i|\theta_i)$, where $\theta_i \in \Theta_i$ identifies i's type drawn from i's type space Θ_i.[7] We assume that for each i and θ_i, u_i is increasing in x_i (the numeraire good) for all y_i and differentiable in both goods; in our discussion of stability, we restrict attention to the special case of quasilinear preferences where $u_i(x_i, y_i|\theta_i) = v_i(y_i|\theta_i) + x_i$. Let $\omega = (\omega_1, \dots, \omega_n)$, $z = (z_1, \dots, z_n)$, and $\theta = (\theta_1, \dots, \theta_n) \in \Theta = \bigtimes_i \Theta_i$, and let $p \in \mathbb{R}$ represent the price of the nonnumeraire good, normalizing the numeraire price to 1. A net trade vector $z \in \mathbb{R}^{2n}$ is *balanced* if $\sum_i z_i = 0$.

Unlike most general-equilibrium models, ours does not restrict the feasible consumption set to the positive orthant. Since no mechanism can Nash-implement Walrasian or Lindahl equilibria when boundary equilibria are permitted (see Hurwicz 1979a or Jackson 2001), it is necessary to rule out such equilibria, either by allowing unbounded consumption bundles or by restricting preferences so that boundary equilibria never obtain. The latter approach is more common (see Groves and Ledyard 1977, for example), but may be incompatible with our notion of dynamic stability.

As specified, the model describes an exchange economy with purely private goods. But we can easily reinterpret the model to allow the second good to be a purely public good by making four changes: (1) every feasible net trade must be such that $y_i = y_j$ for all agents i and j, (2) $\omega_i^y = \omega_j^y$ for all i and j, (3) there is a single firm, capable of producing y units of the public good from $c(y)$ units of the numeraire, that aims to maximize the profit function $(\sum_i p_i)y - c(y)$, where each p_i is the price paid by agent i, and (4) an allocation is now said to be balanced if $c(y) + \sum_i x_i = 0$. In this paper, we assume a constant marginal cost of production $\kappa > 0$ so that $c(y) = \kappa y$.

A Walrasian equilibrium of a private-goods economy at type vector θ is a net trade vector z^* and a price p^* such that z^* is balanced and maximizes each $u_i(\cdot, \cdot|\theta_i)$ subject to the budget constraint $y_i p^* + x_i \leq 0$. Here z^* is referred to as a Walrasian equilibrium allocation.

A Lindahl equilibrium of a public-goods economy is a net trade vector z^* (the Lindahl equilibrium allocation) and a vector of individual prices $(p_1^*, \dots, p_n^*)$ such that z^* is balanced, maximizes each $u_i(\cdot, \cdot|\theta_i)$ subject to the budget constraint $y_i p_i^* + x_i \leq 0$, and 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 consists of $2n$ quantities and only one price, while the former has $n + 1$ quantities but needs n prices.

[7] In Section 5, we generalize our results to multiple goods in additive environments.

Mechanisms and implementation

A *social choice correspondence* $f:\Theta \twoheadrightarrow \mathbb{R}^{2n}$ maps type profiles into sets of net trades. For example, f might identify all Pareto optimal net trades for each θ (the *Pareto correspondence*), all Walrasian equilibrium allocations (the *Walrasian correspondence*), or, in a public-goods setting, all Lindahl equilibrium allocations (the *Lindahl correspondence*).

A *mechanism* $\Gamma = (\mathcal{M}, h)$ consists of a *message space* $\mathcal{M} = \bigtimes_i \mathcal{M}_i$ and an *outcome function* $h:\mathcal{M} \to \mathbb{R}^{2n}$ mapping each message profile $m = (m_1, \ldots, m_n)$ into a net trade vector z. A mechanism Γ is also called a *game form*; when combined with a particular type profile θ, the mechanism induces a well specified game with strategy spaces $\mathcal{M}_i$ for each i and utilities over strategy profiles given by

$$U_i(m|\theta_i) := u_i(h(m)|\theta_i).$$

We let

$$\beta_i(m|\theta_i) = \{m_i \in \mathcal{M}_i : U_i(m_i, m_{-i}|\theta_i) \geq U_i(m_i', m_{-i}|\theta_i)\ \forall m_i' \in \mathcal{M}_i\}$$

represent i's best-response correspondence and define $\beta = (\beta_1, \ldots, \beta_n)$. The *Nash correspondence* $\nu:\Theta \twoheadrightarrow \mathcal{M}$ identifies the set of pure-strategy Nash equilibrium message profiles m^* of Γ at each θ; formally, $\nu(\theta) = \{m \in \mathcal{M} : m \in \beta(m|\theta)\}$. A mechanism $(\mathcal{M}, h)$ is said to *(Nash-) implement* a social choice correspondence f if, for all $\theta \in \Theta$,

$$h(\nu(\theta)) = f(\theta). \tag{1}$$

We sometimes refer to (1) as *full* implementation; if $h(\nu(\theta)) \subseteq f(\theta)$, we say that Γ *weakly* Nash-implements f, and if $h(\nu(\theta)) \cap f(\theta) \neq \varnothing$, then Γ *partially* Nash-implements f.

In the case of economic environments with two goods, the outcome function h can equivalently be written as a pair of functions of the form $x_i(m)$ and $y_i(m)$ for each $i \in \mathcal{I}$. In this paper, we consider mechanisms for which $\mathcal{M}_i \subseteq \mathbb{R}^{K_i}$ for some $K_i \in \{1, 2, \ldots\}$ for each i. When $\mathcal{M}_i$ has $J_i < K_i$ dimensions that enter into the y_i function and $K_i - J_i$ dimensions that do not, then we may, for notation's sake, partition agent i's strategy space into $\mathcal{M}_i = \mathcal{R}_i \times \mathcal{S}_i$ with $\mathcal{R}_i \subseteq \mathbb{R}^{J_i}$ and $\mathcal{S}_i \subseteq \mathbb{R}^{K_i - J_i}$. Letting $\mathcal{R} = \bigtimes_i \mathcal{R}_i$ and $\mathcal{S} = \bigtimes_i \mathcal{S}_i$, we have that $y_i:\mathcal{R} \times \mathcal{S}_{-i} \to \mathbb{R}$ and $x_i:\mathcal{R} \times \mathcal{S} \to \mathbb{R}$. In a public-goods setting, if the mechanism generates only feasible allocations, then it must be that $y_i \equiv y:\mathcal{R} \to \mathbb{R}$ for each i since $y_i = y_j$ for all $i \neq j$.

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

$$x_i(m) = -q_i(m_{-i})y_i(m) - g_i(m), \tag{2}$$

so that the per-unit "price" term q_i does not depend on m_i. Thus, any mechanism can be equivalently described by a list of functions of the form $q_i(m_{-i})$, $g_i(m)$, and $y_i(m)$ for each i. This formulation makes explicit the price and "penalty" components of $x_i(m)$.

3. Stable mechanisms

Contractiveness as a notion of stability

There are many possible notions of stability that one could apply when designing a mechanism. We focus here on *contractive* mechanisms whose best-response functions are, in every type profile θ, contraction mappings.[8] Formally, if $(\mathcal{M}, d)$ is a complete metric space with metric d, then a (single-valued) function $\beta : \mathcal{M} \to \mathcal{M}$ is a *d-contraction mapping* if there is some constant $\xi \in (0,1)$ such that for all $m, m' \in \mathcal{M}$,

$$d(\beta(m), \beta(m')) \leq \xi d(m, m').$$

When the metric d is understood, we simply refer to β as a *contraction mapping*. When β describes the (single-valued) best-response function of a particular game, we say that the game is *contractive*. The following useful lemma provides a simple sufficient condition for a continuously differentiable function β to be a contraction mapping.

Lemma 1. *If $\mathcal{M} \subseteq \mathbb{R}^K$ for some $K \in \{1, 2, \ldots\}$, then a continuously differentiable function $\beta : \mathcal{M} \to \mathcal{M}$ is a contraction mapping if $\sup_{m \in \mathcal{M}} \|D\beta(m)\| < 1$, where $D\beta(m)$ is the differential matrix of β and $\|\cdot\|$ is any matrix norm.*

The proof of this lemma follows easily from Conlisk (1973). Using the absolute row-sum norm, for example, one can show that β is a contraction mapping if $\sup_{m \in \mathbb{R}} \sum_l |\partial \beta_k(m)/\partial m_l| < 1$ at every m for each dimension k.

Since mechanisms induce different games for different type profiles $\theta \in \Theta$, we must extend our definition of a contractive game when describing mechanisms.

Definition 1. Let $(\mathcal{M}, d)$ be a complete metric space. A mechanism $\Gamma = (\mathcal{M}, h)$ with outcome function h is *d-contractive on Θ* if for every $\theta \in \Theta$, the induced game with preferences $U_i(m|\theta_i)$ has a single-valued best-response function $\beta(\cdot, \theta) : \mathcal{M} \to \mathcal{M}$ that is a d-contraction mapping.

Drop the reference to d when the metric is understood. Contractiveness is a strong property to require of a mechanism; by the Banach fixed point theorem, it guarantees the existence of a unique Nash equilibrium of Γ at θ. This equilibrium is globally stable under the Cournot best-response dynamic. This means that if Γ also Nash-implements some social choice function f, then the outcome $f(\theta)$ will in fact be realized in the limit, when agents' play is described by Cournot best response.

Clearly, the processes that best describe dynamic human behavior are more complex and subtle than the simple Cournot best-response dynamic, so guaranteeing stability for a larger family of dynamics is desirable. In this vein, we provide a contraction-mapping analog of the Milgrom–Roberts stability result for supermodular games: There is a family of *adaptive best-response dynamics* (ABR dynamics) such that every dynamic in this family is globally stable in any game with a contractive best-response function.

[8] Our reasons for using contractive mechanisms instead of supermodular mechanisms are discussed at the end of this section.

Formally, a learning dynamic is a function $\mu:\{1,2,\ldots\} \to \bigtimes_i \Delta(\mathcal{M}_i)$ specifying a mixed strategy profile $\mu(t)$ for each point in time $t \in \{1,2,\ldots\}$.[9] Let $S(\mu(t)) \subseteq \mathcal{M}$ be the support of each $\mu(t)$ and let $m(t) \in S(\mu(t))$ be the realized action at time t. For example, a Cournot best-response dynamic would be a function satisfying $S(\mu(t)) = m(t) \in \beta(m(t-1)|\theta)$ for all $t > 1$. To describe ABR dynamics formally, let $H(t',t) = \{m(s) : t' \leq s < t\}$ denote the realized history of play from time t' up to (but not including) t and let m^* denote the unique Nash equilibrium of the game under consideration. Fix a metric d. For any $r \geq 0$, let $B(r|m^*) = \{m \in \mathcal{M} : d(m,m^*) \leq r\}$ be the closed ball with center m^* and radius r. Given any bounded set $\mathcal{M}' \subset \mathcal{M}$, define

$$B(\mathcal{M}') = \bigcap \{B(r|m^*) : \mathcal{M}' \subseteq B(r|m^*)\}$$

to be the smallest closed ball centered at m^* that includes $\mathcal{M}'$.

Definition 2. A learning dynamic $\{m(t)\}$ is an *adaptive best-response dynamic* (*ABR dynamic*) if $\forall t'$, $\exists \hat{t} > t'$, $\forall t \geq \hat{t}$, $S(\mu(t)) \subseteq B(\beta(B(H(t',t))))$.

To understand this definition, consider first two points in time t' and t. Take the point $m' \in H(t',t)$ that is farthest from the equilibrium m^* and consider *all* points in $\mathcal{M}$ that are closer to the equilibrium than m'. Calculate the best response to each of those points, and among those calculated best responses, let m'' be the farthest from the equilibrium. The requirement that $S(\mu(t)) \subseteq B(\beta(B(H(t',t))))$ simply states that the date-t mixed strategy cannot put positive weight on strategies that are farther from m^* than m''. Thus, players observe history $H(t',t)$, form a "belief" that the next profile will be in $B(H(t',t))$, and choose any profile that is either a best response to this belief or, at least, mixes over actions that are no farther from equilibrium than any best response to this belief.

The quantifiers then say that for any date t', there is some later date $\hat{t}$, after which the dynamic ignores the history of play prior to t'. Thus, the effect of early strategies must eventually vanish.

Theorem 1. *If a game is contractive, then all adaptive best-response dynamics converge to the unique Nash equilibrium.*

Formal proofs appear in the Appendix.

Whether a given dynamic is an ABR dynamic may depend on the contractive game under consideration. Simple Cournot dynamics are always in the ABR class. So too is the family of k-period best-response dynamics suggested by Healy (2006) to be a reasonable description of subjects' play in experiments on repeated public-goods mechanisms.[10] In fact, any dynamic where players best respond to some pure-strategy belief formed from a convex combination of the not too distant history of play is always an ABR dynamic and thus is convergent in contractive games.

[9] This definition could be generalized to allow for continuous or finite time intervals.

[10] In these dynamics, each $m(t)$ is a best response to the strategy $(1/k)\sum_{s=t-k}^{t} m(s)$. Empirically, $k = 5$ fits the Healy (2006) data best.

Dynamics based on (mixtures of) best responses to beliefs (i.e., probability distributions) over $B(H(t',t))$—such as fictitious play—are not always in the ABR class, but they are ABR dynamics in the contractive games induced by our mechanisms. As we illustrate next, the key property for these dynamics to be ABR is concavity of the utility functions. Consider a two-player game with $\mathcal{M}_i = [-1,1]$ and $U_i(m) = \max\{1 - |m_i - m_j/2|, 1 - \varepsilon - |m_i - 1|\}$ for each i, where $\varepsilon > 0$ is small. Player i's best response to m_j is $m_j/2 \leq 1/2$, it yields utility 1 and the game is contractive. Playing $m_i = 1$ yields utility $1 - \varepsilon$ regardless of m_j. If i best responds to a belief distribution that puts nontrivial weight on multiple actions m_j, then $m_i = 1$ becomes the unique best response. In other words, i's best response to her beliefs is not in the convex hull of the best responses to the support of her beliefs. In this contractive game, fictitious play is not an ABR dynamic and it may not converge to the unique equilibrium $m^* = (0,0)$. However, if U_i is also concave in m_i, then best responses to beliefs must be in the convex hull of best responses to the support of the beliefs. In this case, fictitious play becomes an ABR dynamic and converges to the unique equilibrium.

Given these results, we focus primarily on designing contractive mechanisms. But we also verify that our particular mechanisms induce concave utility functions U_i, so that the set of convergent ABR dynamics is made even larger.

Supermodularity as a notion of stability

The existing literature takes the approach of requiring a mechanism to induce *supermodular* games, which are also known to have certain desirable stability properties. Consider the game induced by some mechanism $\Gamma = (\mathcal{M}, h)$ at type profile θ. Recall that each $\mathcal{M}_i$ has $K_i \geq 1$ dimensions and m_{ik} represents the kth dimension of m_i. If each U_i is twice differentiable everywhere, then, following Milgrom and Roberts (1990), this game is said to be supermodular if the following properties hold:

1. For all $i \in \mathcal{I}$ and $k \neq l \in \{1,\dots,K_i\}$, $\partial^2 U_i/\partial m_{ik}\,\partial m_{il} \geq 0$.
2. For all $i \neq j \in \mathcal{I}$, $k \in \{1,\dots,K_i\}$, and $l \in \{1,\dots,K_j\}$, $\partial^2 U_i/\partial m_{ik}\,\partial m_{jl} \geq 0$.
3. For all i, $\mathcal{M}_i$ is a compact interval in $\mathbb{R}^{K_i}$.

Properties 1 and 2 guarantee that β_i is increasing in others' strategies. If conditions 1 and 2 are satisfied but 3 is not, then we say the game is *open-supermodular*. Milgrom and Roberts (1990) prove that for every supermodular game, there is a smallest and largest Nash equilibrium, denoted here by $\underline{m}^*$ and $\overline{m}^*$, and if a given learning dynamic is "adaptive"—roughly, if it selects undominated strategies against a not too distant history of past play—then that dynamic converges to the interval $[\underline{m}^*, \overline{m}^*]$. If the game has a unique equilibrium ($\underline{m}^* = \overline{m}^*$), then the equilibrium point is globally stable under all adaptive learning dynamics.

Unfortunately, the usefulness of this stability result is sometimes quite limited. Since the strategy space is required to be compact, then $\underline{m}^*$ and $\overline{m}^*$ may well be corner equilibria. In this case, the stability result may be vacuous. To illustrate, consider a simple

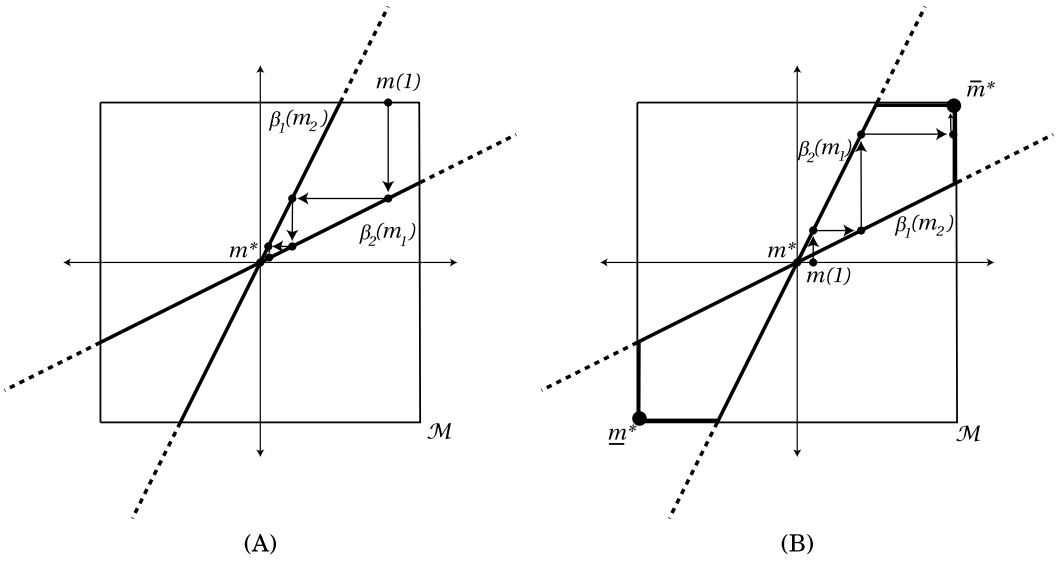

FIGURE 1. Examples of Cournot best-response dynamics in (A) a supermodular and contractive game ($\alpha = 0.5$), and in (B) a supermodular game that is not contractive ($\alpha = 2$). Dashed lines represent the best-response functions if the strategy spaces were unbounded.

two-player game where $\mathcal{M}_i = [-100, 100]$ and $\beta_i(m) = \alpha m_j$ for each $i \in \{1, 2\}$. Two examples of such games are shown in Figure 1. If $\alpha \geq 0$, then this game is supermodular. If $\alpha \in (-1, 1)$, then the game is contractive.

In panel A of Figure 1 the game is both supermodular and contractive since $\alpha = 0.5$. There is a unique Nash equilibrium at $m^* = (0, 0)$ and, by Milgrom and Roberts (1990), all adaptive dynamics converge to m^*. The figure illustrates a typical path for Cournot best-response dynamics, starting at $m(1)$ and converging monotonically to m^*.

Panel B shows the best-response functions when $\alpha = 2$. The game is still supermodular ($\alpha > 0$), but now has three Nash equilibria: $\underline{m}^* = (-100, -100)$, $m^* = (0, 0)$, and $\overline{m}^* = (100, 100)$. In this case, the stability theorem of Milgrom and Roberts (1990) is vacuous; the bounds on the limits of adaptive dynamics are the entire strategy space. Furthermore, the interior equilibrium is unstable under most adaptive dynamics; a simple Cournot best-response process initiated away from m^* converges monotonically to either $\underline{m}^*$ or $\overline{m}^*$.

Now consider open-supermodular versions of these games, where the strategy space is unbounded. The best-response functions are now represented by the dashed lines in the figure. When $\alpha = 0.5$ (panel A), stability of the unique equilibrium is maintained. When $\alpha = 2$ (panel B), however, the now-unique equilibrium m^* continues to be unstable and the best-response dynamic diverges quickly.

Clearly, the stability of the interior equilibrium is driven by the magnitude of the best-response slopes, but not their sign. In other words, contractiveness is the more appropriate notion of stability in games with unbounded strategy spaces or when convergence to corner equilibria is considered undesirable.

Most existing work on supermodular mechanism design fails to appreciate this issue.[11] The following example demonstrates, however, that the type of instability for supermodular games shown in panel B of Figure 1 can also occur in a mechanism that implements a desirable social choice function.

EXAMPLE 1. Let $\mathcal{M}_i = \mathbb{R}^1$ for each i and suppose n is even. Take any quasilinear public-goods environment of the form $v_i(y|\theta_i) + x_i$, where there exists some $\eta > 1$ such that $v_i'' \in (-\eta, -1/\eta)$ for every i and θ_i. Consider the mechanism with free parameter $\gamma > 0$ given by

$$q_i(m) = \begin{cases} \dfrac{\kappa}{n} - \gamma \displaystyle\sum_{j \neq \{i, i+n/2\}} m_j & \text{if } i \leq n/2 \\ \dfrac{\kappa}{n} + \gamma \displaystyle\sum_{j \neq \{i, i-n/2\}} m_j & \text{if } i > n/2 \end{cases}$$

$$g_i(m) \equiv 0$$

and

$$y(m) = \sum_{i=1}^{n/2} m_i - \sum_{i=n/2+1}^{n} m_i$$

for each i.

One can show that this mechanism implements the Lindahl correspondence using a proof very similar to that of Walker (1981). Calculating the slopes of the individual best-response functions, however, gives

$$\frac{\partial \beta_i}{\partial m_j} = \frac{\gamma}{-v_i''} - 1 \geq \frac{\gamma}{\eta} - 1$$

for each $j \notin \{i, i + n/2\}$ and $\partial \beta_i / \partial m_{i+n/2} = 0$. If $\gamma > \eta$, then β_i is nondecreasing in m_j for all θ_i and so the game is open-supermodular. If $\gamma > 2\eta$, however, then the slopes of the best-response functions are greater than 1, and the Cournot dynamics are unstable. ◊

Figure 2 shows the path of best-response dynamics for a particular example with $n = 4$.[12] The dynamic is initiated very close to equilibrium, but diverges exponentially and is unstable.

Chen (2002) designs a family of mechanisms that she shows to be (open-) supermodular under certain parameter restrictions. Given the above discussion, it is unclear whether this implies dynamic stability. Van Essen (2009) shows, however, that the mechanism is also contractive under the same parameter restrictions, and so stability is guaranteed.

[11] See Chen and Tang (1998), Chen (2002), and Healy (2006), for example, or Chen (2008) for a survey. Mathevet (2010) is an exception; he considers Bayesian implementation and studies supermodular mechanisms with "small" equilibrium sets.

[12] Specifically, $v_i(y) = -(1/2)(y-2)^2$ for each i, $\kappa = 4$, and $\gamma = 2$. The Nash equilibrium profile is $m^* = (1/4, 1/4, 1/4, 1/4)$ and the dynamic is initiated at $m(0) = (0, 0, 0, 0)$. At all points in time there are two pairs of players choosing the same strategies, resulting in just two paths in the figure.

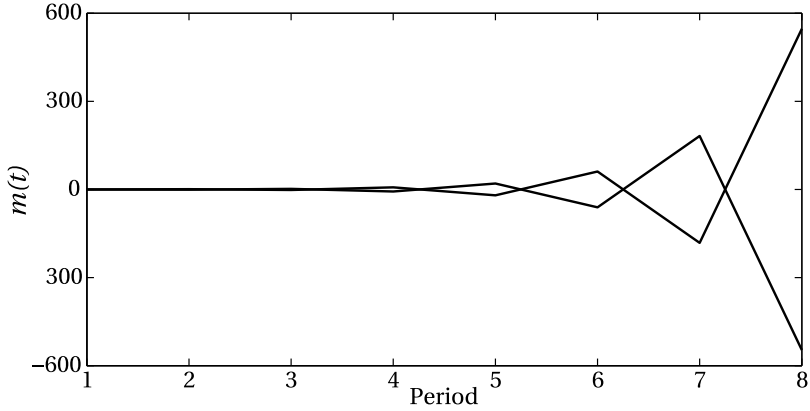

FIGURE 2. Simulated best-response dynamics for an unstable supermodular mechanism.

In fact, Van Essen's (2009) supermodularity-implies-contractiveness result appears to be a more general phenomenon with a fairly intuitive explanation. Most existing public-goods mechanisms feature a public-goods function $y(m)$ and a payment function $x_i(m)$ of a form similar to

$$x_i(m) = -\frac{\kappa}{n} y(m) - \gamma g_i(m),$$

where $\gamma > 0$ is a free parameter and $g_i(m)$ is increasing in the level of "disagreement" between players' messages. Examples include the mechanisms of Groves and Ledyard (1977), Chen (2002), Van Essen (forthcoming), and our example Lindahl mechanism below. For γ near zero, the mechanism is essentially a voluntary contribution mechanism with equal cost sharing. This game has a best-response function whose slope is $-(n-1)$ and is, therefore, highly unstable. For very high values of γ, the mechanism induces a dynamically stable coordination game with a best-response slope of $+1$. As γ is increased, the best-response slope typically increases monotonically from $-(n-1)$ to $+1$, as in Figure 3. At some threshold γ', the game becomes contractive, with a slope of at least -1. Beyond a second threshold of $\gamma'' > \gamma'$, the slope becomes positive, inducing a supermodular game. Thus, choosing a high enough γ to guarantee supermodularity (as in Chen and Tang 1998, Chen 2002, and Van Essen forthcoming) is more than sufficient to also guarantee that the mechanism is contractive.

Figure 3 shows the slope of the linear best-response function in the Groves–Ledyard mechanism for various values of γ using the utility parameters from the experiments of Chen and Tang (1998) and Arifovic and Ledyard (2011), where $n = 5$.[13] Contractiveness obtains for $\gamma \geq 30$ and supermodularity obtains for $\gamma \geq 80$. Experimental results suggest no convergence to equilibrium at $\gamma = 1$, very slow convergence for $\gamma = 10$, and rapid convergence for $\gamma \geq 30$.[14] Convergence is generally faster and postconvergence behavior is more stable for higher values of γ. These results suggest that contractiveness is a

[13]Specifically, $v_i(y|\theta_i) = A_i - B_i y^2$ with $B_i \in [1, 8]$. The values of γ tested are 1, 30, 50, 100, and 260.

[14]The slow convergence for $\gamma = 10$ is intriguing; it suggests that subjects may follow a particular dynamic that can be stable even in games that are not contractive. Arifovic and Ledyard (2011) provide a dynamic that fits well the broad patterns of the experimental data.

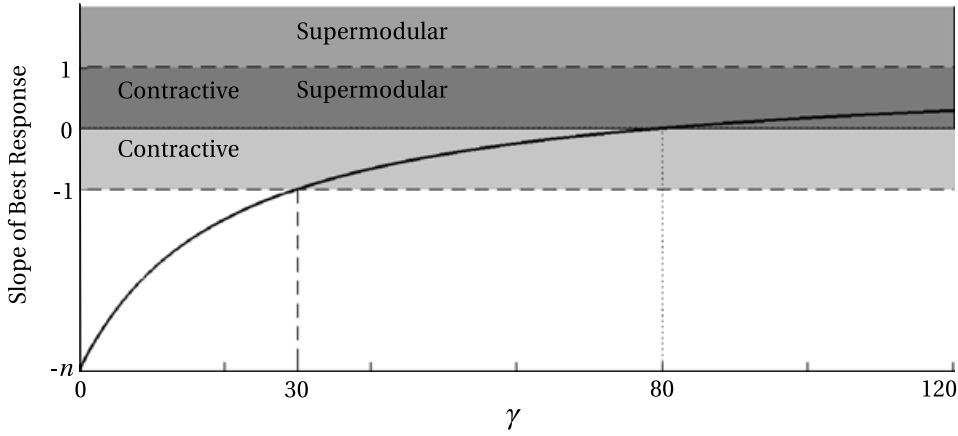

FIGURE 3. The slope of the best-response functions in the Groves–Ledyard mechanism as the parameter γ varies.

useful predictor of stability in the laboratory, and supermodularity may be an excessive requirement.[15]

There is some justification, however, for requiring *both* supermodularity and contractiveness. First, convergence to equilibrium is fastest (in theory) when best-response slopes are close to zero. The Arifovic and Ledyard (2011) experimental results suggest that a positive slope leads to faster convergence than a negative slope.[16] Requiring supermodularity guarantees that best-response slopes fall in this range.

One major difference between supermodular and contractive games is that contractive games have unique equilibria, while supermodular games may not. Each has its drawbacks and benefits in the context of mechanism design: Contractive mechanisms can only implement single-valued social choice functions (or single-valued selections from social choice correspondences), since multiple equilibria would be required to implement multiple outcomes. This may be beneficial, however, because it avoids the ambiguity and selection problems of games with multiple equilibria. Supermodular games, on the other hand, can implement multivalued objectives, but suffer from indeterminacies regarding equilibrium selection. In our view, mechanism design in practice most likely focuses on single-valued selections of social choice correspondences, in which case the benefits of contractiveness prevail.

Given the theoretical considerations and experimental evidence, we focus here on contractiveness as our notion of stability.

[15] Chen and Gazzale (2004) study the compensation mechanism of Varian (1994) in the laboratory. When the punishment parameter β is increased, the best-response slope increases from $-1/2$ to $+1$. As in the Arifovic and Ledyard (2011) experiments, convergence is stronger for higher values of β.

[16] A slope very near $+1$ is undesirable, however, because convergence slows again. This begins to appear in Arifovic and Ledyard's (2011) data when $\gamma = 260$.

Two stable mechanisms

In this section, we present mechanisms that have nearly all the features one might ask; they implement Pareto optimal and individually rational allocations for a wide range of economic environments, they are dynamically stable for a large family of adaptive learning dynamics, they balance the budget both in and out of equilibrium, and the individual message spaces are of minimal dimension necessary for dynamic stability.

To obtain stability results, we restrict attention to concave, quasilinear preferences of the form $u_i(x_i, y_i|\theta_i) = v_i(y_i|\theta_i) + x_i$ that satisfy the next assumption.

ASSUMPTION 1. *For all types $\theta \in \Theta$, all agents i have quasilinear preferences of the form $v_i(y_i|\theta_i) + x_i$, where $v_i' > 0$ and there is some $\eta > 0$ such that $v_i'' \in (-\eta, -1/\eta)$.*[17]

We justify our assumptions on preferences in a later section by showing that it is nearly necessary for stability under well behaved mechanisms. The argument is based on the instability theorems of Jordan (1986) and Kim (1987). These results imply that stability cannot be achieved for completely general preferences, so we assume quasilinearity of preferences. Moreover, Jordan remarks that instability crucially relies on the range of second-order preference behavior, hence the bound on concavity. Our Proposition 2 in Section 5 demonstrates the necessity of these bounds. Also, the exact values of the bounds need to be known by the designer so as to set mechanism parameters that guarantee contractiveness. Thus, a designer must have substantial knowledge about the space of possible preferences.

A contractive mechanism for Lindahl allocations

The description of our stable mechanism for Lindahl allocations follows. Let $\mathcal{M}_i = \mathcal{R}_i \times \mathcal{S}_i$ for each i with $\mathcal{R}_i = \mathcal{S}_i = \mathbb{R}^1$, choose $\delta > 0$ and $\gamma > 0$, and set

$$y(r) = \frac{1}{n}\sum_i r_i \tag{3}$$

$$q_i(m_{-i}, s_{-i}) = \frac{\kappa}{n} + \delta(n-1)\left(s_{i-1} - \frac{\gamma}{n-1}\sum_{j\neq i} r_j\right) \tag{4}$$

and

$$g_i(r, s) = \tfrac{1}{2}(s_i - \gamma r_{i+1})^2 + \tfrac{1}{2}\delta(s_{i-1} - \gamma r_i)^2, \tag{5}$$

where $i+1$ and $i-1$ are taken to be modulo n.

THEOREM 2. *The mechanism defined by (3)–(5) fully Nash-implements the Lindahl correspondence. Under Assumption 1, if*

$$\gamma \in \left(\frac{\sqrt{(n+\eta^2)^2 + 4n(n-1)\eta^4} - (n+\eta^2)}{2n\eta^2}, \frac{n-1}{n}\right) \tag{6}$$

[17] These bounds are inconsistent with Assumption 6(A) in the Appendix, but, depending on the mechanism, may or may not be consistent with Assumptions 4 or 4′.

and

$$\delta \in \left(\frac{\eta}{n\gamma}, \frac{1}{\left(\frac{n-1}{n} - \gamma\right)(1+\gamma)n\eta} \right), \tag{7}$$

then the mechanism is contractive but not supermodular. If

$$\gamma \in \left(\frac{n-1}{n}, 1 \right) \tag{8}$$

and

$$\delta \in \left(\frac{\eta}{n\gamma}, \frac{1}{\left(\gamma - \frac{n-1}{n}\right)(1-\gamma)n\eta} \right), \tag{9}$$

then the mechanism is both contractive and supermodular.[18] *In either case, the mechanism is contractive and each U_i is concave in (r_i, s_i).*

The messages in this mechanism can be interpreted in the following way: Agents are ordered in a circle. Each r_i represents agent i's requested level of the public good. The actual level chosen is the average of the requests. The message s_i represents i's guess of his neighbor's request, adjusted by γ. Agents are penalized for the error in their guess $(s_i - \gamma r_{i+1})$, as well as for the error in their neighbor's guess about them $(s_{i-1} - \gamma r_i)$. In equilibrium, guesses are accurate, meaning $s_i = \gamma r_{i+1}$ and $s_{i-1} = \gamma r_i$, and so no penalties are realized. Prices q_i are simply an equal share of the marginal cost (κ/n) plus a linear term that increases in the difference between $i-1$'s guess of i's request and the average of the others' actual requests. In equilibrium, guesses are correct, so agents who request higher public-good levels are forced to pay higher prices. The sum of prices in equilibrium always equals the marginal cost, as is required at Lindahl allocations.

To gain intuition for the stability result, suppose that some agent $j \neq i+1$ increases r_j by one unit. For agent i, this increase has a quantity effect of increasing $y(r)$ by $1/n$ and a price effect of reducing $q_i(r_{-i}, s_{-i})$ by $\delta\gamma$. Agent i's response to the quantity effect is to reduce r_i by an equal amount, returning $y(r)$ to his most preferred level. But the lower personal price for the public good induces agent i to increase $y(r)$ by $\delta\gamma/v_i''$. Finally, the resulting net change in r_i is further tempered by a coordination effect, since changing r_i increases the penalty term $\delta(s_{i-1} - \gamma r_i)^2$. The parameter values are then chosen so that these effects nearly cancel out and the resulting total change in r_i is sufficiently close to zero, regardless of v_i''.

A unit increase in r_{i+1} not only has these same price, quantity, and coordination effects, but also a coordination effect on s_i. Agent i's optimal response is to increase s_i

[18] If $n \leq \eta^2/(2\eta - 1)$, then such a value of δ exists only if γ is sufficiently close to $(n-1)/n$ or 1. Specifically, a value exists if and only if

$$\gamma \in \left(\frac{n-1}{n}, 1 - \frac{(n+\eta^2) + \sqrt{(n+\eta^2)^2 - 4n^2\eta^2}}{2n\eta^2} \right) \cup \left(1 - \frac{(n+\eta^2) - \sqrt{(n+\eta^2)^2 - 4n^2\eta^2}}{2n\eta^2}, 1 \right),$$

both of which are nonempty intervals. If $n > \eta^2/(2\eta - 1)$, then a value of δ exists for every $\gamma \in ((n-1)/n, 1)$. Contractiveness obtains on the boundary of supermodularity (when $\gamma = (n-1)/n$) only in the very special case where $\eta \in (1, n/(n-1))$. Thus, we rule out $\gamma = (n-1)/n$ generally.

by γ, keeping his guess accurate. As long as $\gamma < 1$, this added coordination effect is also contractive.

Finally, a unit increase in s_{i-1} has similar offsetting effects: Agent i increases r_i by $1/\gamma$ due to the coordination effect, but reduces r_i by $\delta(n-1)/v_i''$ due to the increased price effect. Again, these responses are tempered by the penalty from $i-1$'s guess becoming inaccurate, and the parameter restrictions ensure that the net effects almost perfectly offset.

Consider now large economies or economies with very rich type spaces. If n or η grows large, then γ must approach 1. For large η (fixing n), δ must become large. For large n, however, δ could in fact become quite small, as the lower bound $\eta/(n\gamma)$ from (9) is decreasing in n. Thus, increasing the richness of the type space necessitates extreme penalties and price reactions, but simply increasing the population size does not.

One downside of the above mechanism is that it fails to balance the budget at certain out-of-equilibrium message profiles. This occurs both because the penalty terms (g_i) may all be strictly positive and because the price terms (q_i) may sum to something other than the marginal cost. Both of these sources of imbalance can be corrected, however, by adding appropriate terms to the penalty functions. For the case of $n \geq 4$, this can be done by taking the g_i function from (5) and modifying it to

$$\hat{g}_i(r,s) = g_i(r,s) + d_i^L(r_{-i}, s_{-i}), \tag{10}$$

where $d_i^L(r_{-i}, s_{-i})$ is a polynomial given by (30) in the Appendix.

Corollary 1. *Suppose $n \geq 4$ and Assumption 1 holds. If γ and δ satisfy (6) and (7) or (8) and (9), then the mechanism defined by (3)–(5) and (10) fully Nash-implements the Lindahl correspondence, is contractive on Θ, and is budget-balanced for all $m \in \mathcal{M}$. Furthermore, each U_i is concave in (r_i, s_i).*

Beyond this specific mechanism, our approach shows how other contractive mechanisms can be constructed. For example, we can show that Chen's (2002) open-supermodular mechanism is also contractive, which Van Essen (2009) has verified. Van Essen (forthcoming) also provides a mechanism that is both contractive and open-supermodular. Both Chen's and Van Essen's mechanisms are not balanced out of equilibrium. To our knowledge, the mechanism in (3)–(5) and (10) is the first contractive mechanism that is also budget-balanced out of equilibrium. Since out-of-equilibrium learning is an important motivation for this research, we view budget balance as an important requirement. Although our example mechanism is somewhat complex—especially with the budget-balancing adjustment—our impossibility result for one-dimensional mechanisms (Theorem 6) suggests that complexity cannot be improved substantially.

A contractive mechanism for Walrasian allocations

The process of designing a contractive Walrasian mechanism is nearly identical to the process of designing a contractive Lindahl mechanism, though the exact functional forms obviously must differ.

The example mechanism we provide is two dimensional with $\mathcal{M}_i = \mathbb{R}_i \times \mathcal{S}_i = \mathbb{R}^2$ for each i. Fix any $\gamma \in (1/(n-1), 1)$. The outcome functions are then given by

$$y_i(r, s_{-i}) = r_i - \frac{1}{n-1} \sum_{j \neq i} r_j \tag{11}$$

$$q_i(r_{-i}, s_{-i}) = \frac{1}{\delta} \left(s_{i-1} + \gamma \sum_{j \neq i} r_j \right) \tag{12}$$

and

$$g_i(r, s) = (s_i - \gamma r_{i+1})^2 \tag{13}$$

for each i, with $i-1$ and $i+1$ taken to be modulo n.

THEOREM 3. *The mechanism defined by (11)–(13) fully Nash-implements the Walrasian correspondence. For any $\gamma \in (1/(n-1), 1)$ and $\delta > \gamma\eta(n-1)$ (where η is the bound on each v_i''), it is contractive on Θ and each U_i is concave in (r_i, s_i).*

This mechanism is not supermodular for any parameters, since increases in s_{i-1} generate a price effect that leads i to reduce y_i by reducing r_i.

Here, agents are submitting a suggested net trade quantity (r_i) and a guess of their neighbor's suggested net trade (s_i). Their actual net trade is the amount by which their suggested trade is above or below the average suggested trade. They receive a penalty for incorrect guesses of their neighbor, which, along with the price term q_i, disciplines the mechanism to generate dynamic stability in much the same way that the Lindahl mechanism was made stable through penalties.

The mechanism is stable because it creates inertia in each dimension of the message. Each agent is a price-taker and chooses her level of private good y_i by using the first dimension of her message. If agents $j \neq i$ change their messages, then agent i will try to match the average variation of the r_j's to restore her preferred net trade. But she does not want to match it completely because of the price effect. Notice, indeed, that an increase in $\sum_{j \neq i} r_j$ increases the price q_i, hence it moderates i's response. In the end, the change in agent i's first dimension is less than the variation in the average of others. Out of equilibrium, the mechanism is not anonymous in the sense that each agent is possibly offered a different price for the same goods. So the second dimension of the message guarantees that the price is the same for everyone in equilibrium. But this has to be done while preserving stability, which is why i only chooses s_i to match a fraction ($\gamma < 1$) of her neighbor's suggested trade.

As with the Lindahl mechanism, the lower bound on δ need not grow as the economy becomes large. Since γ can be kept arbitrarily close to $1/(n-1)$, the lower bound on δ can be kept arbitrarily close to (and just above) η. But, as in the Lindahl mechanism, δ must grow large as the type space becomes rich, since its lower bound grows linearly in η. With a large δ, the price function converges to zero, so agents choose their consumption bundle as though the nonnumeraire good were effectively free. This highlights a key difference between Lindahl and Walrasian mechanisms: If agents ignore

prices and penalties in the Lindahl mechanism, then the game is naturally unstable, and so prices and penalties are needed to restore stability. If agents ignore prices and penalties in the Walrasian mechanism, the game is (approximately) stable since each r_j enters into y_i negatively. In this case, the prices and penalties are needed only to ensure that the mechanism implements Walrasian allocations; the magnitude of the price function is made small so that the price effect does not interfere with stability. The penalty term is needed only to pin down the optimal choice of s_i, which can then be used to ensure that the agents' prices are always equal in equilibrium without letting the price function depend on r_i or s_i.

The above mechanism fails to balance the budget at certain out-of-equilibrium message profiles, because the nonnumeraire good is not balanced out of equilibrium. As with the Lindahl correspondence, the sources of imbalance can be corrected by adding appropriate terms to the penalty functions. For the case of $n \geq 4$, this can be done by taking the g_i function from (13) and modifying it to

$$\hat{g}_i(r, s) = g_i(r, s) + d_i^W(r_{-i}, s_{-i}), \tag{14}$$

where $d_i^W(r_{-i}, s_{-i})$ is given by (35) in the Appendix.

COROLLARY 2. *If $n \geq 4$ and $\delta > \gamma\eta(n-1)$, then the mechanism defined by (11)–(14) fully Nash-implements the Lindahl correspondence, is contractive on Θ, and is budget-balanced for all $m \in \mathcal{M}$. Furthermore, each U_i is concave in (r_i, s_i).*

4. CHARACTERIZATIONS OF IMPLEMENTING MECHANISMS

We now describe the process by which these mechanisms were constructed. A general method for constructing stable mechanisms is useful because example mechanisms may not pass the test of time as behavioral research progresses. Other restrictions on mechanisms may be discovered, and this method for constructing new mechanisms can easily be adapted as new restrictions are added.

Our method has two steps: First, we characterize the "shape" of implementing mechanisms. Then we add our extra requirement of stability as a further restriction on the shape of the mechanisms. This leads us to rule out one-dimensional mechanisms as potential candidates. We, therefore, conclude that a stable mechanism is a multidimensional contractive price–quantity mechanism, as in the example mechanisms given above. These characterizations also have implications about the complexity that is necessary for a mechanism to Nash-implement Walrasian or Lindahl allocations.

We no longer apply Assumption 1 (quasilinearity); in this section, we require only that preferences be differentiable and strictly increasing in the numeraire.

Characterization of one-dimensional mechanisms

For clarity, we first restrict attention to *one-dimensional* mechanisms where $\mathcal{M}_i = \mathbb{R}^1$ for each $i \in \mathcal{I}$. The characterization theorems are more transparent and intuitive in the

one-dimensional case; the case of higher-dimensional mechanisms is briefly covered in a later subsection.

We begin by assuming twice-differentiable mechanisms. This assumption is mainly for technical convenience and does not substantially hinder our ability to design stable mechanisms (Theorems 2 and 3). Furthermore, we believe that highly discontinuous mechanisms (such as Maskin's canonical mechanism or Abreu–Matsushima's dominance-solvable mechanism) may be too complex for real-world application.

ASSUMPTION 2 (Differentiability). *For each agent i, the message space $\mathcal{M}_i$ equals $\mathbb{R}^1$, and the functions x_i and y_i are twice continuously differentiable in m_i on $\mathcal{M}$.*

Our next assumption explicitly rules out cases where agent i's outcome function becomes arbitrarily flat. This does not rule out any existing mechanisms in the literature; most use linear functions such as $y_i(m) = \sum_j m_j$.

ASSUMPTION 3 (Responsive y_i). *For each i, there exists some $\varepsilon_i > 0$ such that for all $m \in \mathcal{M}$, $|\partial y_i(m)/\partial m_i| \geq \varepsilon_i$.*

Under Assumption 3, y_i becomes bijective in m_i. This guarantees a form of citizen sovereignty wherein each agent is able to select any $\hat{y}_i \in \mathbb{R}$ through his or her choice of m_i.[19] It also means y_i is invertible for each m_{-i}, enabling us to view agent i's attainable consumption bundles as the graph of a single-valued function from y_i into x_i. We denote this by

$$\chi_i(\hat{y}_i|m_{-i}) := x_i(y_i^{-1}(\hat{y}_i|m_{-i}), m_{-i}),$$

where $y_i^{-1}(\hat{y}_i|m_{-i})$ identifies the unique m_i such that $y_i(m_i, m_{-i}) = \hat{y}_i$. Thus $\chi_i(\hat{y}_i|m_{-i})$ is the amount of good x that i must choose if he wants $\hat{y}_i$ units of good y, given m_{-i}.

We show an example of $\chi_i(y_i|m_{-i})$ in Figure 4. At the point m^*, the outcome $(x_i(m^*), y_i(m^*))$ is realized by agent i. As i differentially changes his message m_i, he differentially changes his allocation (x_i, y_i) along the graph of χ_i. The downward slope of this graph at m^*—which we label $P_i(m^*)$—represents the differential change in x_i per unit of y_i. We call this the *effective price* of y_i charged by the mechanism at m^*. Formally,

$$P_i(m) := -\frac{\partial x_i(m)/\partial m_i}{\partial y_i(m)/\partial m_i}. \tag{15}$$

If m^* is a Nash equilibrium, then the standard first-order conditions imply that

$$\frac{\partial u_i(x_i, y_i|\theta_i)/\partial y_i}{\partial u_i(x_i, y_i|\theta_i)/\partial x_i} = P_i(m^*),$$

so that the marginal rate of substitution between y_i and x_i equals the effective price of the mechanism at m^*.[20]

[19] This is reminiscent of Novshek (1985), for example, who views firms in an oligopoly market as choosing aggregate output rather than individual production.

[20] Recall that $\mathcal{M}$ is open, so there are no boundary Nash equilibria.

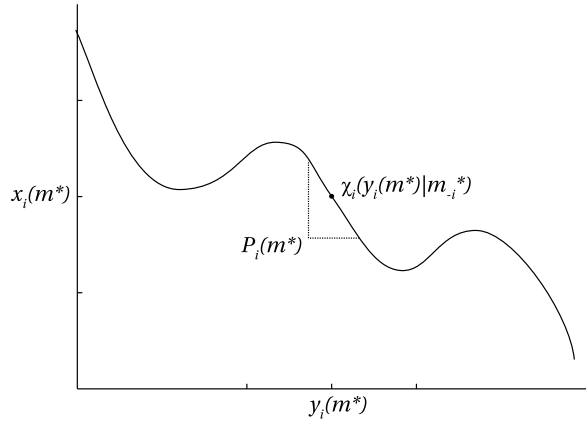

FIGURE 4. The mapping $\chi_i(y_i|m_{-i})$ and the effective price $P_i(m)$ at m^*.

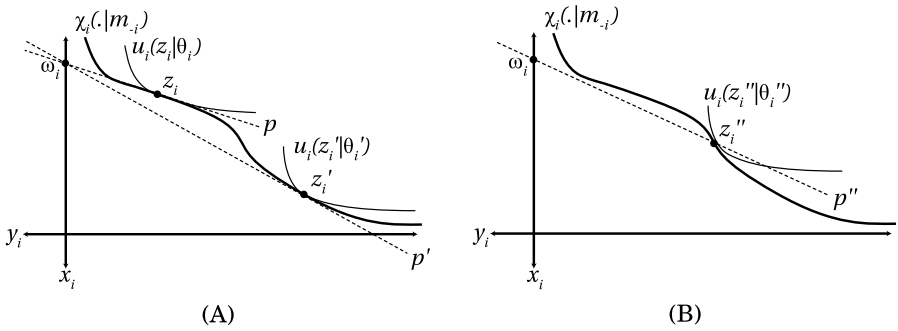

FIGURE 5. (A) The triple tangency property and (B) a bad Nash equilibrium.

If this mechanism Nash-implements a Walrasian or Lindahl equilibrium, then the marginal rate of substitution must also equal the Walrasian or Lindahl price. Thus, the effective prices at the equilibrium message profile m^* must also match the Walrasian or Lindahl price for each environment θ. This leads to the following observation:

OBSERVATION (The triple tangency property). If a mechanism Nash-implements Walrasian or Lindahl allocations, then at any Nash equilibrium m^*, each agent's indifference curve in (x_i, y_i) space must be tangent to both the mechanism's outcome manifold $\chi_i(\cdot|m^*_{-i})$ and the corresponding Walrasian or Lindahl equilibrium price hyperplane.

The triple tangency property is illustrated in panel A of Figure 5; for type θ_i, the point z_i is both a Nash equilibrium outcome and a Walrasian allocation at price p. Similarly, z_i' is a Nash equilibrium and a Walrasian allocation (at price p') for type θ_i'.

Now consider panel B of Figure 5. If the type space is sufficiently rich—meaning that every outcome z is a Nash equilibrium outcome for some environment—then there exists some $\theta'' \in \Theta$ such that the point z_i'' is also a Nash equilibrium outcome. This must be a "bad" Nash equilibrium, however, because the outcome z_i'' cannot possibly be a Walrasian equilibrium allocation; the indifference curve is not tangent to the

price hyperplane p'' that connects ω_i and z_i''. Because of this bad Nash equilibrium, the mechanism represented by χ_i does not (weakly) implement the Walrasian or Lindahl correspondence.

If the type space is rich, then *every* point along χ_i can be made into a Nash equilibrium outcome by selecting an appropriate type profile. If we require weak implementation—and, therefore, no bad equilibria—then χ_i must be linear and pass through ω_i. Any nonlinearity creates a bad equilibrium. But a linear χ_i function means that i is forced to act as if he is choosing an optimal level of y_i, taking as given a fixed per-unit price. This price—the slope of χ_i—may still vary in m_{-i}, but not in m_i. This simple observation generates our key necessary condition.

Assumption 4 (Rich type space). *We have* $\nu(\Theta) = \mathcal{M}$.

This assumption places restrictions on the equilibrium set rather than on the primitives of the model; in the Appendix, we provide two linked assumptions on the primitives that together imply Assumption 4.

Theorem 4 (Necessity). *Under Assumptions 2, 3, and 4, if a mechanism* $\Gamma = (\mathcal{M}_i, q_i, g_i, y_i)_i$ *weakly Nash-implements the Walrasian or Lindahl correspondence, then for every* $i \in \mathcal{I}$ *and every* $m \in \mathcal{M}$,

$$x_i(m) \equiv -q_i(m_{-i})y_i(m), \tag{16}$$

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

Thus, $P_i(m) = q_i(m_{-i})$. Since q_i now represents the per-unit price paid by agent i, an immediate but useful corollary of Theorem 4 follows.

Corollary 3. *Under Assumptions 2, 3, and 4, if a mechanism* $\Gamma = (\mathcal{M}_i, q_i, g_i, y_i)_i$ *weakly Nash-implements the Walrasian correspondence, then for every* $m \in \mathcal{M}$,

$$q_i(m_{-i}) = q_j(m_{-j}) \quad \forall i, j \tag{17}$$

and, therefore,

$$\sum_i y_i(m) = 0. \tag{18}$$

If Γ *weakly Nash-implements the Lindahl correspondence, then for every* $m \in \mathcal{M}$,

$$y_i(m) = y_j(m) \quad \forall i, j \tag{19}$$

and, therefore,

$$\sum_i q_i(m_{-i}) = \kappa. \tag{20}$$

Theorem 4 is stated for weak implementation, but obviously applies to full implementation as well. This theorem gives a strong but intuitive result: If a mechanism is to Nash-implement the Walrasian or Lindahl correspondence, then each agent's message-choosing problem in the mechanism (taking others' messages as fixed) must be identical to the quantity-choosing problem in an exchange economy when prices are taken as given. In the case of a mechanism, the quantity y_i is chosen indirectly through the choice of m_i, and the price is determined endogenously as a function of m_{-i}. In exchange economies, agents choose y_i directly and face exogenously given prices.

For the case of public-goods economies, compare Theorem 4 with the mechanisms of Groves and Ledyard (1977), Walker (1981), and Tian (1990). All three are one-dimensional mechanisms in which q_i depends only on m_{-i}, but the Groves–Ledyard mechanism has a nontrivial penalty function g_i, while the latter two do not. Consequently, Walker's and Tian's mechanisms Nash-implement the Lindahl correspondence, while the Groves–Ledyard mechanism does not.

For private-goods economies, Theorem 4 is very strong. If no agent is allowed to affect his or her own per-unit price, if all agents must have the same price at every equilibrium message, and if every message is an equilibrium message for some type profile, then the only admissible price function q_i is a constant function that depends on *no* agents' reports. But clearly such a mechanism cannot fully implement the Walrasian correspondence on a rich type space, so we arrive at a contradiction. This proves the following corollary.

Corollary 4. *Under Assumptions 2–4, a one-dimensional mechanism that Nash implements the Walrasian correspondence does not exist.*

Corollary 4 was first proven by Reichelstein and Reiter (1988) using substantially different mathematical techniques.

Finally, we show that the necessary conditions of Theorem 4 are also sufficient for weak implementation. Full implementation is achieved if, for each Lindahl or Walrasian equilibrium point, there is some message $m' \in \mathcal{M}$ that maps to it.

Assumption 5. *For every $(x_i^*, y_i^*, p_i^*)_i \in \mathbb{R}^{2n+1}$ that is a Walrasian or Lindahl allocation for some $\theta \in \Theta$, there is a message $m' \in \mathcal{M}$ such that*

$$(x_i(m'), y_i(m'), q_i(m'))_i = (x_i^*, y_i^*, p_i^*)_i.$$

Note that Assumption 5 is slightly stronger than requiring $f(\Theta) \subseteq h(\mathcal{M})$, because it also requires that every possible Lindahl price be achievable by the q_i functions.

Theorem 5 (Sufficiency). *If a mechanism Γ satisfies Assumptions 2 and 3, and (16), (17), and (18), then Γ weakly Nash-implements the Walrasian correspondence. If (17) and (18) are replaced with (19) and (20), then Γ weakly Nash-implements the Lindahl correspondence. If, in addition, Γ satisfies Assumption 5, then Γ fully Nash-implements the Walrasian or Lindahl correspondence.*

Under Assumption 5, the above necessary conditions become sufficient. The proof of sufficiency for weak implementation is intuitive. Take the case of a private-goods economy. Since $y_i(m_i, m_{-i})$ is bijective in m_i, choosing m_i is equivalent to choosing y_i with $q_i(m_{-i})$ fixed. Hence, choosing a message is similar to maximizing utility subject to the budget constraint. (17) and (18) ensure market clearing.

Assumption 5 guarantees that any Walrasian equilibrium allocation can be reached by some message $m' \in \mathcal{M}$ with $q_i(m'_{-i})$ equalling the Walrasian price for each i. Because the Walrasian equilibrium allocation is budget-constrained optimal, y_i is bijective, and the linear mechanism mimics this budget constraint, the m'_i mapping to $y_i(m'_i, m'_{-i})$ must be a best response for agent i. Thus, every Walrasian allocation is a Nash equilibrium and full implementation is achieved.

To demonstrate the gap between weak implementation and full implementation, consider the equal-tax voluntary contribution mechanism, where $\mathcal{M}_i = \mathbb{R}^1$ for each i, $y(m) = \sum_i m_i$, and $x_i(m) = -\kappa y(m)/n$ (see Groves and Ledyard 1980 or Healy 2006). The hypotheses of the first part of Theorem 5 are satisfied, so this mechanism weakly Nash-implements the Lindahl correspondence. But Assumption 5 fails at any θ that has a Lindahl equilibrium $(x_i^*, y_i^*, p_i^*)_i$ with $p_i^* \neq p_j^*$ for some i, j (which is true generically), because $q_i(m_{-i}) = q_j(m_{-j})$ for every $m \in \mathcal{M}$. In these environments, the mechanism has no Nash equilibrium and full implementation fails.

The impossibility of contractive one-dimensional mechanisms

We now show that there cannot exist a mechanism with one-dimensional strategy spaces ($\mathcal{M}_i = \mathbb{R}^1$ for each i) that Nash-implements the Lindahl or Walrasian correspondence under our maintained assumptions.

THEOREM 6. *Under Assumptions 2–4 and 1, a mechanism with $\mathcal{M}_i = \mathbb{R}^1$ for each i that is both contractive and Nash-implements the Lindahl or Walrasian correspondence does not exist.*

Therefore, the contractiveness property requires slightly more complex mechanisms.

Note that the proof for the Walrasian correspondence is trivial, since no one-dimensional mechanism that implements the Walrasian allocations exists.

Inspection of the proof reveals that Theorem 6 holds true even if v_i'' can take any value in $(-\infty, 0)$; the bounds on v_i'' from Assumption 1 are needed in the sequel to generate higher-dimensional mechanisms that are contractive.

Characterization of higher-dimensional mechanisms

Multidimensional mechanisms are indispensable for stability, since one-dimensional mechanisms are necessarily unstable. We derive several useful conditions on implementing mechanisms by extending ideas from the previous section. We relegate part of the argument to the Appendix.

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 that affect $y_i(r, s_{-i})$ and let $\mathcal{S}_i \subseteq \mathbb{R}^{K_i - J_i}$ be those dimensions that do not. With the partitioning of the strategy spaces into $\mathcal{R}_i$ and $\mathcal{S}_i$, we can modify (2) slightly and write any mechanism's numeraire outcome function as

$$x_i(r, s) = -q_i(r, s) y_i(r, s_{-i}) - g_i(r, s).$$

(In a public-goods environment, y depends only on r.) Unlike (2), this formulation allows the price term q_i to depend on agent i's message. We now reformulate our previous assumptions for the case of multiple dimensions.

Assumption 1′ (Differentiability). *For each agent i and each message vector $m \in \mathcal{M}$, the functions x_i and y_i are twice continuously differentiable in every dimension of m_i.*

Assumption 2′ (Responsive y_i). *For each i, there exists some $\varepsilon_i > 0$ such that for all $(r, s) \in \mathcal{M}$ and all dimensions $k \in \{1, \ldots, J_i\}$, $|\partial y_i(r, s_{-i})/\partial r_{ik}| \geq \varepsilon$.*

Define

$$\sigma_i(r, s_{-i}) := \arg\max_{s_i' \in \mathcal{S}_i} x_i(r, s_i', s_{-i})$$

for each i and

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

Thus, $\sigma(r)$ represents the pure-strategy equilibria of a transfer-maximizing game in which agents pick s_i to maximize x_i given r. If $s \notin \sigma(r)$, then the pair (r, s) cannot be a Nash equilibrium of the mechanism for any θ.

For each agent i and dimension k, define the effective price along dimension k at message (r, s) by

$$P_{ik}(r, s) := -\frac{\partial x_i(r, s)/\partial r_{ik}}{\partial y_i(r, s_{-i})/\partial r_{ik}}$$

and note that by the same argument as in the one-dimensional case, it must be that $P_{ik}(r, s)$ equals i's marginal rate of substitution between y_i and x_i at any equilibrium (r, s). So a point (r', s') such that $P_{ik}(r', s') \neq P_{il}(r', s')$ cannot be a Nash equilibrium. Given these two restrictions, we now define

$$\mathcal{M}^* := \left\{m = (s, r) \in \mathcal{M} : (\forall i \in \mathcal{I})\ (\forall k, l \in \{1, \ldots, J_i\})\ P_{ik}(m) = P_{il}(m) \text{ and } s \in \sigma(r)\right\}$$

to be the set of "candidate equilibrium" points in $\mathcal{M}$. Note that if each r_i is one-dimensional—as is true in our example mechanisms from Section 3—then $\mathcal{M}^*$ is simply those points satisfying $s \in \sigma(r)$. We obtain the following necessary conditions.[21] If a mechanism Γ weakly Nash-implements the Walrasian correspondence, then for every $(r^*, s^*) \in \mathcal{M}^*$,

$$\sum_i y_i(r^*, s_{-i}^*) = 0 \tag{21}$$

[21] These conditions emerge as a direct corollary of Theorem 8 (see the Appendix).

and

$$q_i(r^*, s^*) = q_j(r^*, s^*) \quad \forall i, j. \tag{22}$$

If Γ weakly Nash-implements the Lindahl correspondence, then for every $m \in \mathcal{M}^*$,

$$\sum_i q_i(r^*, s^*) = \kappa \tag{23}$$

and

$$y_i(r^*, s^*_{-i}) = y_j(r^*, s^*_{-j}) \quad \forall i, j. \tag{24}$$

We use these conditions in the next result to build our mechanisms.

Assumption 3′. *For every $(x_i^*, y_i^*, p_i^*)_i \in \mathbb{R}^{2n+1}$ that is a Walrasian or Lindahl allocation for some $\theta \in \Theta$, there is a message $m' \in \mathcal{M}^*$ such that*

$$(x_i(m'), y_i(m'), q_i(m'))_i = (x_i^*, y_i^*, p_i^*)_i.$$

Theorem 7 (Sufficiency). *If Γ is a mechanism satisfying Assumptions 1′ and 2′, equalities (21) and (22), and*

(i) $q_i(r, s) \equiv q_i(r_{-i}, s_{-i})$

(ii) $g_i(r, s) \geq 0$ *for every* $(r, s) \in \mathcal{M}$

(iii) $g_i(r, s) = 0$ *if* $(r, s) \in \mathcal{M}^*$,

then Γ weakly implements the Walrasian correspondence. If (21) and (22) are replaced with (23) and (24), then Γ weakly implements the Lindahl correspondence. If, in addition, Γ satisfies Assumption 3′, then Γ fully Nash-implements the Walrasian or Lindahl correspondence.

This theorem implies that the mechanism from Theorem 2 implements the Lindahl correspondence. We can verify that $g_i \geq 0$ with $g_i = 0$ when $s \in \sigma(r)$, q_i depends only on m_{-i}, $\sum_i q_i = \kappa$ if $s \in \sigma(r)$, y is bijective in r_i, and for every Lindahl equilibrium, there is some message $m' \in \mathcal{M}$ that maps to the Lindahl equilibrium allocation and prices.

In summary, higher-dimensional mechanisms may allow agents to affect their own prices and face nontrivial penalty functions, but the penalty function must equal zero on the equilibrium set, and each agent's price must not change as the agent unilaterally changes r_i and adjusts s_i appropriately. At out-of-equilibrium or nonregular equilibrium points, however, we derive no restrictions on the shape of the mechanism. It is this freedom that allows us to introduce global stability properties into a mechanism. Intuitively, one should be able to take a mechanism satisfying the restrictions of Theorem 8 (page 653) and alter the mechanism on $\mathcal{M} \setminus \mathcal{M}^*$ so that any adaptive dynamic process that wanders off of $\mathcal{M}^*$ eventually returns to the appropriate point in $\mathcal{M}^*$, restoring the equilibrium.

How to construct contractive mechanisms

Given a mechanism, the slopes of the best-response functions can be calculated using one of two different methods.

If the value of $\sigma_i(r, s_{-i})$ can be computed from the mechanism, then the direct method can be used. First, it must be verified directly that σ_i is contractive; using the row-sum norm and Lemma 1, this means we must verify that

$$\sup_{(r,s)} \sum_{j \neq i} \left(\left| \frac{\partial \sigma_i}{\partial r_j} \right| + \left| \frac{\partial \sigma_i}{\partial s_j} \right| \right) < 1. \tag{25}$$

Next, the first-order condition of utility maximization with respect to r_i is calculated. Replacing s_i in that condition with $\sigma_i(r, s_{-i})$ describes the first-order condition in r_i as s_i adjusts optimally in response. Taking derivatives of this first-order condition with respect to each r_j and s_j identifies how r_i responds to changes in each other variable. Given these slopes, the row-sum condition

$$\sup_{(r,s)} \sum_{j \neq i} \left(\left| \frac{\partial \rho_i}{\partial r_j} \right| + \left| \frac{\partial \rho_i}{\partial s_j} \right| \right) < 1 \tag{26}$$

is sufficient.

If σ_i cannot be solved directly then the implicit function theorem can still be used to derive closed-form expressions for the slopes of σ_i and ρ_i. These are given by the solution to the system of equations.

$$\begin{bmatrix} \frac{\partial^2 U_i}{\partial r_i^2} & \frac{\partial^2 U_i}{\partial r_i \partial s_i} \\ \frac{\partial^2 U_i}{\partial s_i \partial r_i} & \frac{\partial^2 U_i}{\partial s_i^2} \end{bmatrix} \begin{bmatrix} \frac{\partial \rho_i}{\partial r_j} & \frac{\partial \rho_i}{\partial s_j} \\ \frac{\partial \sigma_i}{\partial r_j} & \frac{\partial \sigma_i}{\partial s_j} \end{bmatrix} = - \begin{bmatrix} \frac{\partial^2 U_i}{\partial r_i r_j} & \frac{\partial^2 U_i}{\partial r_i s_j} \\ \frac{\partial^2 U_i}{\partial s_i r_j} & \frac{\partial^2 U_i}{\partial s_i s_j} \end{bmatrix}.$$

There is a unique solution to these equations if the leftmost matrix is invertible. In that case, closed-form expressions for the slopes can be derived. Given these slopes, the contraction conditions (25) and (26) can be checked directly. Parameters guaranteeing supermodularity can be derived similarly, by guaranteeing that each slope is positive.

These methods are only useful for checking the stability of an existing mechanism. The following procedure shows how our example Lindahl mechanism was constructed following the sufficiency conditions of Theorem 7. This provides a general blueprint for how other, similar mechanisms could be constructed.

Step 1. Only two dimensions are needed for stability, so let r_i and s_i each be one dimensional. This reduces $\mathcal{M}^*$ to those messages where $s = \sigma(r)$.

Step 2. For simplicity, let $y(r, s)$ depend only on r. Since $y(r)$ must be responsive (Assumption 2′), an obvious choice is $y(r) = \alpha \sum_i r_i$, with $\alpha > 0$.

Step 3. Since prices must sum to κ, let $q_i(r_{-i}, s_{-i}) = \kappa/n + \hat{q}_i(r_{-i}, s_{-i})$, where $\sum_i \hat{q}_i(r_{-i}, s_{-i}) = 0$ whenever $s \in \sigma(r)$. Its exact form will be determined in later steps.

Step 4. The penalty term must satisfy $g_i \geq 0$ with $g_i = 0$ when $s \in \sigma(r)$. Quadratic terms such as $(s_i - \gamma r_{i+1})^2$ are a simple choice and, with a linear $y(r)$ function, guarantee that U_i is concave in (r_i, s_i). This means $\sigma_i(r) = \gamma r_{i+1}$, and so $\gamma < 1$ guarantees that stability condition (25) is satisfied.

Step 5. To determine $\hat{q}_i$, note that an increase in any r_j generates a quantity effect on agent i through increasing $y(r)$. This prompts i to reduce r_i. This can be offset by having r_j enter negatively into $\hat{q}_i$, generating a positive price effect on r_i. For example, consider $\hat{q}_i(r_{-i}, s_{-i}) = \delta(-\sum_{j \neq i} r_j)$.

Step 6. Since we must have $\sum_i \hat{q}_i = 0$ when $s = \sigma(r)$, use the s_{-i} terms to balance $\hat{q}_i$. Specifically, let $\hat{q}_i(r_{-i}, s_{-i}) = \delta((n-1)s_{i-1} - \gamma \sum_{j \neq i} r_j)$, which now sums to zero when $s_{i-1} = \gamma r_i$.

Step 7. The price effect from including s_{i-1} in $q_i(r_{-i}, s_{-i})$ must be offset by a coordination effect in g_i. This is done by letting $g_i(r, s) = (s_i - r_{i+1})^2 + \delta(s_{i-1} - r_i)^2$.

Step 8. Now take the resulting mechanism, use the direct method for deriving the contraction conditions (25) and (26) above, and find parameter values α, γ, and δ that satisfy those two conditions. The resulting mechanism is contractive and has concave utilities over strategies.

The mechanisms of Chen (2002) and Van Essen (forthcoming) also can be thought of as being constructed through this procedure. Both choose $(s_i - y(r))^2$ instead of $(s_i - \gamma r_{i+1})^2$ at Step 4. Van Essen adds s_{i-1} to $\hat{q}_i$ in Step 6, but Chen instead adds $\sum_{j \neq i} s_j$. This necessitates adding $\sum_{j \neq i} \delta(s_j - y(r))^2$ to the penalty function at Step 7. Obviously, many other mechanisms could be constructed by varying these choices.

5. Discussion

Generalizing to multiple goods

Thus far our focus has been limited to economies with only one nonnumeraire good, and our stability result has been limited only to economies with preferences that are quasilinear in the numeraire. Here we discuss various ways in which our results can—and cannot—be generalized.

Consider now a $(K+1)$-good economy with one private numeraire good and $K \geq 1$ nonnumeraire goods. Agent i's net consumption of the kth nonnumeraire good is denoted y_i^k (or y^k if the good is public) and her consumption of the numeraire is x_i. We say preferences are *quasilinear-additive* if there exist functions $\{v_i^k\}_{k=1}^K$ such that $u_i(x_i, y_i|\theta_i) = x_i + \sum_k v_i^k(y_i^k|\theta_i)$.

Quasilinear-additive preferences allow us to easily describe K distinct two-good "sub-economies" in which $K-1$ of the nonnumeraire goods are held fixed and only the numeraire and kth nonnumeraire goods vary. The lack of complementarities guarantees that the fixed level of the other $K-1$ goods does not affect preferences in the kth subeconomy. Given any mechanism Γ defined for two-good economies, we can define the *K-fold extension* of Γ to be the mechanism in a $(K+1)$-good economy where

$\Gamma = (\mathcal{M}_i, x_i, y_i)$ is applied simultaneously to all K two-good subeconomies. Thus, agents submit messages $m_i = (m_i^1, \ldots, m_i^K) \in \mathcal{M}_i^K$, the kth nonnumeraire quantity is determined by $y_i(m^k)$ (where $m^k = (m_1^k, \ldots, m_n^k)$), and the numeraire quantity for agent i is determined by $\sum_k x_i(m^k)$. The induced utility function over the message space of the K-fold extension of Γ is simply the sum of induced utilities over each subeconomy:

$$U_i(m) = \sum_k \big[v_i^k(y_i(m^k)|\theta_i) + x_i(m^k)\big].$$

Proposition 1. *Take any mechanism Γ defined for two-good economies. In a $(K+1)$-good economy with quasilinear-additive preferences, the K-fold extension of Γ is contractive if Γ is contractive on every two-good subeconomy.*

Proposition 1 guarantees that the contractive mechanisms developed in Theorems 2 and 3 can be applied good-by-good in larger economies when preferences are quasilinear additive. The proof is simple: Changes to m_j^k for some j and k affect only y_i^k and x_i for agent i. But for $l \neq k$, this does not change i's marginal rate of substitution between y_i^l and x_i. Thus, player i's best response is affected only in the m_i^k component, and so stability in the $(K+1)$-good economy is equivalent to stability in each subeconomy.

Relaxing quasilinearity and bounded convexity of preferences

Although it is restrictive to limit attention to quasilinear-additive environments with bounded concavity, earlier results by Jordan (1986) and Kim (1987) suggest that it is difficult to go far beyond this with well behaved mechanisms. Jordan's result shows that in private-good economies, any well behaved mechanism that strongly Nash-implements the Walrasian correspondence admits an environment, with a unique Walrasian allocation, such that the corresponding Nash equilibrium is not stable under a wide class of continuous-time dynamics. Kim (1987) extends this result to public-good economies. Jordan remarks that his instability theorem crucially relies on the range of second-order preference behavior present in his environments. Our environments, instead, aim to turn off or bound these second-order effects. This explains why the stability problem becomes very difficult, even in quasilinear environments, when bounded concavity or additive separability is relaxed. Consider the case of bounded concavity first. Suppose that v_i'' is arbitrarily close to zero. Nash-implementing a Walrasian or Lindahl allocation requires the first-order condition $v_i'(y_i(m)|\theta_i) = q_i(m)$ to be satisfied. Consider any change in m_j for some $j \neq i$ that alters q_i. Agent i's best response to this change must alter v_i' by an equal amount to restore the first-order condition. When v_i'' is very small, however, this requires a very large shift in $y(m)$, which can only be accomplished by a large change in m_i.[22] Since the response in m_i is larger than the original shift in m_j, the mechanism cannot be contractive.

[22] Clearly there is an offsetting effect if $\partial y_i/\partial m_i$ is large, but this derivative is fixed for any given $y_i(m^*)$, while v_i'' is moving arbitrarily close to zero; eventually the v_i'' effect must dominate.

PROPOSITION 2. *Suppose $u_i(x_i, y_i|\theta_i) = v_i(y_i|\theta_i) + x_i$. If for every $\varepsilon > 0$, there is some i, θ_i, and y such that $v_i''(y|\theta_i) \in (-\varepsilon, 0)$, then no mechanism that implements Walrasian or Lindahl allocations is contractive on Θ.*

In the presence of multiple private (public) goods, the strength of complementarities between the various goods can have destabilizing effects that are hard to accommodate. Requiring contractiveness in each good separately is no longer sufficient for stability, and requiring contractiveness for all goods simultaneously becomes a strong condition as the number of goods is increased. Bounding concavity serves a similar purpose. Without it, a small change in some agents' messages could lead other agents to overreact greatly, as we argued in Proposition 2. This fundamental overreaction would not be overcome by appealing to a different notion of stability; bounded concavity appears to be necessary for any stability concept.

Summary and future directions

From a theoretical perspective, these newly constructed mechanisms have nearly all the features one might ask: they implement Pareto optimal and individually rational allocations for a wide range of economic environments, they are dynamically stable for a large family of adaptive learning dynamics, they balance the budget both in and out of equilibrium, and the individual message spaces are of minimal dimension necessary for dynamic stability.

The theorems in this paper make heavy use of the rich type space assumption. Sufficiently weakening this assumption opens the door for mechanisms to have q_i depend on r_i, or for g_i to be nonzero, which in turn makes dynamic stability an easier requirement to satisfy. For example, with only two possible type profiles (each with a unique Lindahl equilibrium) the triple tangency property needs to be satisfied only at two points; away from those two points, the mechanism can be "bent" arbitrarily to satisfy the desired stability properties. As the type space grows, this flexibility clearly diminishes.

The obvious next step for future research is to return to the lab with these newly constructed mechanisms to understand what additional requirements they should be asked to satisfy. Perhaps bounds on mechanism complexity or the limits on the magnitude of out-of-equilibrium punishments will be identified as the next important factor for the theory to incorporate. Eventually these mechanisms can be field-tested on a small scale and the theory will be refined further as a result.

APPENDIX

PROOF OF THEOREM 1. The proof follows by induction. Pick a starting time t_0. By definition of an ABR dynamic, for each point in time t_n there exists some later point in time $t_{n+1} > t_n$ such that for all $t \geq t_{n+1}$, $S(\mu(t)) \subseteq B(\beta(B(H(t_n, t))))$. For each $n \in \{1, 2, \ldots\}$, let $\mathcal{M}_n = H(t_n, t_{n+1})$ be the history of play from t_n to t_{n+1}.

For any metric d on $\mathcal{M}$, any set $\mathcal{M}' \subseteq \mathcal{M}$, and any point $m' \in \mathcal{M}$, the d-Hausdorff distance between $\mathcal{M}'$ and the singleton set $\{m'\}$ is given by

$$h_d(\mathcal{M}', m') = \sup_{m \in \mathcal{M}'} d(m, m').$$

Therefore, for any set $\mathcal{M}' \subseteq \mathcal{M}$, $h_d(\mathcal{M}', m^*) = h_d(B(\mathcal{M}'), m^*)$. Thus,

$$\begin{aligned}
\xi h_d(\mathcal{M}_1, m^*) &= \xi h_d(B(\mathcal{M}_1), m^*) \\
&\geq h_d(\beta(B(\mathcal{M}_1)), m^*) \\
&= h_d(B(\beta(B(\mathcal{M}_1))), m^*) \\
&\geq h_d(\mathcal{M}_2, m^*),
\end{aligned}$$

where the first inequality comes from the contraction property of β and the last inequality follows from the fact that $\mathcal{M}_2 \subset B(\beta(B(\mathcal{M}_2)))$. Taking any n and $n+1$, we can use a similar argument to show that $\xi h_d(\mathcal{M}_n, m^*) \geq h_d(\mathcal{M}_{n+1}, m^*)$. Therefore, for all $n > 1$,

$$\xi^n h_d(\mathcal{M}_1, m^*) \geq h_d(\mathcal{M}_n, m^*),$$

which implies that the sequence $\mathcal{M}_n$ converges to $\{m^*\}$, and so any ABR dynamics converges to m^*. □

Proof of Theorem 2. Recall that the mechanism is given by

$$\begin{aligned}
y(r) &= \frac{1}{n}\sum_i r_i \\
q_i(r_{-i}, s_{-i}) &= \frac{\kappa}{n} + \delta\left((n-1)s_{i-1} - \gamma\sum_{j\neq i} r_j\right) \\
g_i(r, s) &= \tfrac{1}{2}(s_i - \gamma r_{i+1})^2 + \tfrac{1}{2}\delta(s_{i-1} - \gamma r_i)^2.
\end{aligned}$$

Step 1. We first prove that the mechanism is contractive on the given parameter ranges.

We use the "direct method" for calculating the slopes of the best-response functions and verifying contractiveness for the given parameter restrictions.

The induced utility function over strategies for agent i is, therefore,

$$U_i(r, s) = v_i(y(r)) - q_i(r_{-i}, s_{-i})y(r) - g_i(r, s).$$

Let $\rho_i(r_{-i}, s_{-i})$ and $\sigma_i(r_{-i}, s_{-i})$ be i's best-response values of r_i and s_i, respectively.

Since s_i enters only into $g_i(r, s)$, it is clear that $\sigma_i = \gamma r_{i+1}$.

The first-order condition on r_i is given by

$$\frac{1}{n}v_i'(y(\rho_i, r_{-i})) - \frac{\kappa}{n^2} - \frac{1}{n}\delta\left((n-1)s_{i-1} - \gamma\sum_{j\neq i} r_j\right) + \delta\gamma(s_{i-1} - \gamma\rho_i) = 0.$$

Differentiating with respect to r_j $(j \neq i)$ gives

$$\frac{1}{n^2}v_i'' + \frac{1}{n^2}v_i''\frac{\partial \rho_i}{\partial r_j} + \delta\gamma\frac{1}{n} - \delta\gamma^2\frac{\partial \rho_i}{\partial r_j} = 0.$$

Therefore,

$$\frac{\partial \rho_i}{\partial r_j} = \frac{v_i'' + \delta\gamma n}{\delta\gamma^2 n^2 - v_i''}.$$

Differentiating with respect to s_{i-1} gives

$$\frac{1}{n^2}v_i''\frac{\partial \rho_i}{\partial s_{i-1}} - \delta\frac{n-1}{n} + \delta\gamma - \delta\gamma^2\frac{\partial \rho_i}{\partial s_{i-1}} = 0,$$

and so

$$\frac{\partial \rho_i}{\partial s_{i-1}} = \frac{\delta n(\gamma n - (n-1))}{\delta\gamma^2 n^2 - v_i''}.$$

Differentiating with respect to any other s_j $(j \notin \{i-1, i\})$ gives

$$\frac{1}{n^2}v_i''\frac{\partial \rho_i}{\partial s_j} - \delta\gamma^2\frac{\partial \rho_i}{\partial s_j} = 0,$$

which means

$$\frac{\partial \rho_i}{\partial s_j} = 0.$$

Finally, we know that $\sigma_i = \gamma r_{i+1}$, so

$$\frac{\partial \sigma_i}{\partial r_{i+1}} = \gamma$$

and all other slopes of σ_i are zero.

The contractiveness condition on σ_i (using the row-sum norm) is

$$\sum_{j\neq i}\left|\frac{\partial \sigma_i}{\partial r_j}\right| + \sum_{j\neq i}\left|\frac{\partial \sigma_i}{\partial s_j}\right| < 1$$

for all (r, s). This reduces here to

$$|\gamma| < 1. \tag{27}$$

The stability condition on ρ_i is

$$\sum_{j\neq i}\left|\frac{\partial \rho_i}{\partial r_j}\right| + \sum_{j\neq i}\left|\frac{\partial \rho_i}{\partial s_j}\right| < 1$$

for all (r, s). For the current mechanism, this becomes

$$(n-1)\left|\frac{v_i'' + \delta\gamma n}{\delta\gamma^2 n^2 - v_i''}\right| + \left|\frac{\delta n(\gamma n - (n-1))}{\delta\gamma^2 n^2 - v_i''}\right| < 1$$

or

$$(n-1)|v_i'' + \delta\gamma n| + \left|\delta n(\gamma n - (n-1))\right| < \delta\gamma^2 n^2 - v_i''.$$

Let

$$\delta > \frac{\eta}{\gamma n}, \tag{28}$$

so that the first absolute value in the stability condition is positive. We now consider three cases, depending on whether the argument of the second absolute value term is positive, negative, or zero. Note that in the first case, the mechanism is supermodular since all slopes are positive, while in the second case, it is not. The third case represents the boundary of supermodularity.

Case 1. $\gamma > (n-1)/n$. Given the stability condition σ_i from (27), we must have

$$\gamma \in \left(\frac{n-1}{n}, 1\right).$$

The stability condition then becomes

$$(n-1)(\delta\gamma n + v_i'') + \delta n(\gamma n - (n-1)) < \delta\gamma^2 n^2 - v_i''.$$

Solving this for δ gives

$$\delta < \frac{-v_i''}{n\left(\gamma - \frac{n-1}{n}\right)(1-\gamma)}.$$

Since $-v_i'' > 1/\eta$, this condition on δ is satisfied if

$$\delta < \frac{1}{\left(\gamma - \frac{n-1}{n}\right)(1-\gamma)n\eta}.$$

Recalling (28), contractiveness is satisfied if

$$\gamma \in \left(\frac{n-1}{n}, 1\right)$$

and

$$\delta \in \left(\frac{\eta}{n\gamma}, \frac{1}{\left(\gamma - \frac{n-1}{n}\right)(1-\gamma)n\eta}\right).$$

Note that this interval for δ is nonempty only if γ is close enough to either $(n-1)/n$ or 1. This is automatically satisfied whenever $n > \eta^2/(2\eta - 1)$, since $(n-1)/n$ and 1 become sufficiently close together in that case. If $n < \eta^2/(2\eta-1)$, then we must have

$$\gamma \in \left(\frac{n-1}{n}, 1 - \frac{(n+\eta^2) + \sqrt{(n+\eta^2)^2 - 4n^2\eta^2}}{2n\eta^2}\right) \cup \left(1 - \frac{(n+\eta^2) - \sqrt{(n+\eta^2)^2 - 4n^2\eta^2}}{2n\eta^2}, 1\right)$$

to guarantee that a required value of δ exists. The right interval is always nonempty. A bit of algebra reveals that the left interval is nonempty when $n < \eta^2$, which is true here since $n < \eta^2/(2\eta - 1)$. Thus, there are always values of γ close enough to $(n-1)/n$ or 1 to guarantee a choice of δ that ensures the mechanism is contractive.

Case 2. $\gamma < (n-1)/n$. In this case, (27) is automatically satisfied, and so stability of σ_i is guaranteed. The stability condition on ρ_i becomes

$$(n-1)(\delta\gamma n + v_i'') + \delta n((n-1) - \gamma n) < \delta\gamma^2 n^2 - v_i''.$$

Solving for δ gives

$$\delta < \frac{-v_i''}{n(1+\gamma)\left(\frac{n-1}{n} - \gamma\right)}.$$

Since $-v_i'' < 1/\eta$, this is satisfied whenever

$$\delta < \frac{1}{\left(\frac{n-1}{n} - \gamma\right)(1+\gamma)n\eta}.$$

Thus, contractiveness is satisfied if

$$\gamma < \frac{n-1}{n}$$

and

$$\delta \in \left(\frac{\eta}{n\gamma}, \frac{1}{\left(\frac{n-1}{n} - \gamma\right)(1+\gamma)n\eta}\right).$$

Note that such a δ exists only if γ is sufficiently close to $(n-1)/n$. Specifically, it must be that

$$\gamma \in \left(\frac{\sqrt{(n+\eta^2)^2 + 4n(n-1)\eta^4} - (n+\eta^2)}{2n\eta^2}, \frac{n-1}{n}\right),$$

which is always nonempty for $n \geq 3$.

Case 3. $\gamma = (n-1)/n$. In this case the second absolute value becomes zero (r_i does not respond to s_{i-1}) and so the stability condition reduces to

$$(n-1)(\delta\gamma n + v_i'') < \delta\gamma^2 n^2 - v_i''.$$

Plugging in $\gamma = (n-1)/n$ and solving for δ gives

$$\delta < -v_i'' \frac{n^2}{(n-1)^3},$$

which is satisfied for all v_i'' if

$$\delta < \frac{1}{\eta} \frac{n^2}{(n-1)^3}.$$

Thus, contractiveness obtains if

$$\delta \in \left(\frac{\eta}{n-1}, \frac{1}{\eta} \frac{n^2}{(n-1)^3} \right),$$

but this is nonempty only if

$$\eta < \frac{n}{n-1}.$$

Since $\eta > 1$ by definition, this case applies only when

$$\eta \in \left(1, \frac{n}{n-1} \right),$$

which is a very narrow range.

Step 2. We now prove that the mechanism Nash-implements the Lindahl correspondence.

To see that a unique Lindahl equilibrium exists for all θ, note the following necessary and sufficient conditions for any Lindahl equilibrium.

(i) Given p_i^* and y^*, it must be that $x_i^* = -p_i^* y^*$ for all i.

(ii) This implies that $\partial v_i(y^*)/\partial y = p_i^*$ for each i.

(iii) Linearity of the firm's profit function then implies that $\sum_i \partial v_i(y^*)/\partial y = \sum_i p_i^* = \kappa$.

Using these conditions, we can derive the unique Lindahl equilibrium in three steps:

(i) Since $v_i'' \in (-\eta, -1/\eta)$ for each i, there is one unique y^* satisfying the third necessary condition.

(ii) Given the unique y^*, there is one unique p_i^* for each i satisfying the second condition.

(iii) Given y^* and p_i^*, there is one unique x_i^* for each i satisfying the first condition.

Since the mechanism is contractive, it also has a unique Nash equilibrium (r^*, s^*) for every θ. Now take the equilibrium message (r^*, s^*) and let $p_i^* = q_i(r_{-i}^*, s_{-i}^*)$. Then $x_i(r^*, s^*) = -p_i^* y(r^*)$ for each i, satisfying the first condition. Furthermore, the first-order condition for maximization in r_i at an equilibrium point implies that

$$v_i'(y(r^*)) \frac{\partial y_i(r^*)}{\partial r_i} = p_i^* \frac{\partial y_i(r^*)}{\partial r_i} + \frac{\partial g_i(r^*, \sigma(r^*))}{\partial r_i}.$$

Since $\partial y_i(r)/\partial r_i \neq 0$ and $\partial g_i/\partial r_i = 0$ at the equilibrium point, we have that $v_i'(y(r^*)) = p_i^*$, satisfying the second condition. Finally, it is easy to check that $\sum_i q_i(r_{-i}^*, s_{-i}^*) = \kappa$ at the equilibrium point since $s_i^* = \gamma r_{i+1}^*$ for each i and so the third condition is satisfied. Thus, the unique equilibrium point is equal to the unique Lindahl allocation. □

PROOF OF COROLLARY 1. Consider the mechanism without $d_i^L(r_{-i}, s_{-i})$ added. In general, the excess numeraire collected at any message profile (r, s) equals

$$\sum_i q_i(r_{-i}, s_{-i})y(r) + \sum_i g_i(r, s) - \kappa y(r).$$

For the current mechanism, this equals

$$\sum_i \left[\frac{\kappa}{n} + \delta\left((n-1)s_{i-1} - \gamma\sum_{j\neq i} r_j\right)\right]y(r) + \sum_i \left[\frac{1}{2}(s_i - \gamma r_{i+1})^2 + \frac{\delta}{2}(s_{i-1} - \gamma r_i)^2\right] - \kappa y(r),$$

which reduces to

$$\left[\delta\sum_i \left((n-1)s_{i-1} - \gamma\sum_{j\neq i} r_j\right)y(r)\right] + \left[\sum_i \frac{1}{2}(s_i - \gamma r_{i+1})^2 + \sum_i \frac{\delta}{2}(s_{i-1} - \gamma r_i)^2\right]. \quad (29)$$

Now consider the function

$$d_i^L(r_{-i}, s_{-i}) = \delta\frac{n-1}{n}\left(\left(\sum_{j\neq i} r_j\right)[s_{i+1} - \gamma r_{i+2}] + r_{i+1}[s_{i+2} - \gamma r_{i+3}]\right) + \frac{1}{2}(s_{i+1} - \gamma r_{i+2})^2 + \frac{\delta}{2}(s_{i-2} - \gamma r_{i-1})^2. \quad (30)$$

Let the new mechanism's penalty function be $g_i(r, s) - d_i^L(r_{-i}, s_{-i})$.

The function $d_i^L(r_{-i}, s_{-i})$ has three properties. First, in any equilibrium $d_i^L = 0$, since $s_i = \gamma r_{i+1}$ for every i. Thus, total penalties remain zero in equilibrium. Second, d_i^L does not depend on i's announcement; therefore, neither the equilibrium nor the contractiveness of the mechanism is affected by the addition of this term. Finally, the sum of the d_i^L terms always equals the excess numeraire collected, so the new mechanism is always exactly balanced.

To see that the sum of d_i^L terms always equals the excess numeraire, note that

$$\sum_i d_i^L(r_{-i}, s_{-i}) = \delta\frac{n-1}{n}\sum_i\left(\left(\sum_{j\neq i} r_j\right)[s_{i+1} - \gamma r_{i+2}] + r_{i+1}[s_{i+2} - \gamma r_{i+3}]\right) + \sum_i\left[\frac{1}{2}(s_{i+1} - \gamma r_{i+1})^2 + \frac{\delta}{2}(s_{i-2} - \gamma r_{i-1})^2\right].$$

The second term here equals the second term in the expression of excess numeraire, (29). The first term can be rewritten as

$$\delta\frac{n-1}{n}\left(\sum_i\left(\sum_{j\neq i} r_j\right)[s_{i+1} - \gamma r_{i+2}] + \sum_i r_{i+1}[s_{i+2} - \gamma r_{i+3}]\right).$$

By shifting the indices of the final sum, this is equal to

$$\delta\frac{n-1}{n}\left(\sum_i\left(\sum_{j\neq i} r_j\right)[s_{i+1} - \gamma r_{i+2}] + \sum_i r_i[s_{i+1} - \gamma r_{i+2}]\right).$$

But now that reduces to

$$\delta\frac{n-1}{n}\Big(\sum_i r_i\Big)\Big[\Big(\sum_i s_i\Big)-\gamma\Big(\sum_i r_i\Big)\Big],$$

which is exactly the first term in (29). Thus, the resulting mechanism is always balanced. □

PROOF OF THEOREM 3.

Step 1. We first show that the mechanism is contractive. Agent i's utility over strategies in this mechanism is given by

$$U_i(r,s|\theta)=v_i(y_i(r,s_{-i})|\theta)-\delta\Big(s_{i-1}+\gamma\sum_{j\neq i}r_j\Big)y_i(r,s_{-i})-(s_i-\gamma r_{i+1})^2.$$

To find i's best-response function, we compute the first-order conditions

$$\frac{\partial U_i(r,s|\theta)}{\partial r_i}\equiv v_i'(y_i(\cdot))\frac{\partial y_i}{\partial r_i}-\frac{1}{\delta}\Big(s_{i-1}+\gamma\sum_{j\neq i}r_j\Big)\frac{\partial y_i}{\partial r_i}=0$$

$$\frac{\partial U_i(r,s|\theta)}{\partial s_i}\equiv s_i-\gamma r_{i+1}=0.$$

So

$$\rho_i(r_{-i},s_{-i})=\frac{1}{n-1}\sum_{j\neq i}r_j+v_i'^{-1}\Big(\frac{1}{\delta}\Big(s_{i-1}+\gamma\sum_{j\neq i}r_j\Big)\Big|\theta\Big)$$

$$\sigma_i(r_{-i},s_{-i})=\gamma r_{i+1}.$$

The sufficient conditions for the mechanism to be contractive are then

$$\sum_{j\neq i}\left|\frac{\partial\rho_i(r_{-i},s_{-i})}{\partial r_j}\right|+\sum_{j\neq i}\left|\frac{\partial\rho_i(r_{-i},s_{-i})}{\partial s_j}\right|=\Big(\sum_{j\neq i}\left|\frac{1}{n-1}+\frac{\gamma}{\delta v_i''(\cdot|\theta)}\right|\Big)+\left|\frac{1}{\delta v_i''(\cdot|\theta)}\right|<1 \tag{31}$$

and

$$\sum_{j\neq i}\left|\frac{\partial\sigma_i(r_{-i},s_{-i})}{\partial r_j}\right|+\sum_{j\neq i}\left|\frac{\partial\sigma_i(r_{-i},s_{-i})}{\partial s_j}\right|=\gamma<1.$$

For the first condition, recall that $v_i''(\cdot|\theta)\in(-\eta,-1/\eta)$, so if $\delta>\gamma\eta(n-1)$, then

$$\frac{1}{n-1}+\frac{\gamma}{\delta v_i''(\cdot|\theta)}>0.$$

Therefore, the left-hand side of (31) is equal to

$$1+\frac{\gamma(n-1)-1}{\delta v_i''(\cdot|\theta)},$$

which is less than 1 when $\gamma > 1/(n-1)$. Thus, the mechanism is contractive when $\gamma \in (1/(n-1), 1)$ and $\delta > \gamma\eta(n-1)$.

Step 2. Proving that this mechanism fully implements the Walrasian correspondence is a direct consequence of Theorem 7. Take a Nash equilibrium (r^*, s^*) for $\theta \in \Theta$. Then, from the previous step, we know $s^*_{i-1} = \gamma r^*_i$. As a result,

$$q_i(r^*_{-i}, s^*_{-i}) = \delta\left(s^*_{i-1} + \gamma \sum_{j\neq i} r^*_j\right) = \sum_{i=1}^{n} \gamma r^*_i = Q^*$$

for all $i \in \mathcal{I}$ and

$$x_i(r^*, s^*) = -Q^* y_i(r^*, s^*). \tag{32}$$

By definition of Nash equilibrium, $u_i(r^*, s^*) \geq u_i(r_i, r^*_{-i}, s^*)$ for all r_i; that is,

$$v_i(y_i(r^*, s^*)|\theta) - Q^* y_i(r^*, s^*) \geq v_i(y_i(r_i, r^*_{-i}, s^*)|\theta) - Q^* y_i(r_i, r^*_{-i}, s^*) \tag{33}$$

for all r_i. Since y_i is a surjection from $\mathbb{R}$ onto $\mathbb{R}$ (in r_i), (33) implies

$$v_i(y_i(r^*, s^*)|\theta) - Q^* y_i(r^*, s^*) \geq v_i(y_i|\theta) - Q^* y_i \tag{34}$$

for all y_i. Finally, we verify that allocation $[y_i(r^*, s^*), x_i(r^*, s^*)]_i$ is balanced. By definition,

$$\sum_i y_i(r^*, s^*) = \sum_i r^*_i - \frac{1}{n-1} \sum_i \sum_{j\neq i} r^*_j = 0.$$

Since $x^*_i(r^*, s^*) = -Q^* y_i(r^*, s^*)$, then $\sum_i x^*_i(r^*, s^*) = 0$. It follows from balancedness, (32), and (34) that allocation $[y_i(r^*, s^*), x_i(r^*, s^*)]_i$ is a Walrasian allocation. To complete the proof, start with the Walrasian allocation $[Y^*_i, X^*_i]_i$ of some environment θ. From Step 1, the mechanism is contractive, hence a Nash equilibrium exists (by Banach fixed point theorem). From the previous argument, this Nash equilibrium must correspond to a Walrasian allocation, which is necessarily $[Y^*_i, X^*_i]_i$. □

Proof of Corollary 2. For the case of $n \geq 4$, the imbalance can be done by taking the g_i function from (13) and modifying it to

$$\hat{g}_i(r, s) = g_i(r, s) + d^W_i(r_{-i}, s_{-i}),$$

where

$$d^W_i(r_{-i}, s_{-i}) = -(s_{i+1} - \gamma r_{i+2})^2 + \frac{1}{\delta}\Bigg((s_{i-2} - \gamma r_{i-1})r_{i-1} + \frac{1}{n-1}(\gamma r_{i+2} - s_{i+1})r_{i-1} + \frac{1}{n-1}\sum_{k\neq i,i-1}(\gamma r_{i+3} - s_{i+2})r_k\Bigg). \tag{35}$$

As in the proof of Corollary 1, we note that this function sums to zero at any candidate equilibrium point, does not affect equilibrium or stability since it depends only on

(r_{-i}, s_{-i}), and a bit of algebra confirms that $\sum_i d_i^W$ always equals the excess numeraire collected by the mechanism absent d_i^W. Thus, the resulting mechanism also implements Walrasian equilibrium allocations and is contractive, but now is balanced out of equilibrium as well. □

The rich type space assumptions

With multiple dimensions, it becomes overly restrictive to assume that $\nu(\theta) = \mathcal{M}$ (Assumption 4) because the dimensionality of $\nu(\Theta)$ may be strictly less than that of $\mathcal{M}$. But for any Γ, we have ruled out two types of messages that can never be Nash equilibria of Γ for any $\theta \in \Theta$ and we have defined $\mathcal{M}^*$. The analog of Assumption 4 for multiple dimensions is the following.

ASSUMPTION 4′. *We have* $\nu(\Theta) = \mathcal{M}^*$.

This assumption is used in the next section to derive conditions on mechanisms. Here we break it into two separate (but linked) assumptions on primitives; in the case of one-dimensional mechanisms, these assumptions imply Assumption 4.

ASSUMPTION 6. *Some* $\rho \in \{2, 4, 6, \ldots\}$ *exist such that the following statements hold.*

(A) *All* ρ*th order preferences are admissible:*

$$\Theta_\rho := \left\{\theta \in \Theta : (\forall i)\ (\exists(\alpha_i, \beta_i) \in (\mathbb{R}_{++} \times \mathbb{R}))\ \text{s.t. } u_i(x_i, y_i|\theta_i) = (-\alpha_i y_i^\rho + \beta_i y_i) + x_i\right\} \subseteq \Theta.$$

(B) *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* $s_i' \in \sigma_i(r_i', r_{-i}, s_{-i})$,

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

Assumption 6(A) simply requires that all polynomial (quasilinear) preferences of order ρ be permitted in the type space. To interpret Assumption 6(B), let $\rho = 2$ and consider changes in m_i that lead to large changes in y_i. In this case, the squared term in the maximand applies, and so the assumption places quadratic upper and lower bounds on the change in x_i. For changes in m_i that lead to small changes in y_i, the upper and lower bounds are square-root bounds. In either case, the requirement is strictly weaker than requiring that χ_i be Hölder continuous of degree ρ or that χ_i be Lipschitz continuous. The bounds on χ_i imposed by this assumption are demonstrated in Figure 6. Note that as ρ increases, Assumption 6(B) becomes strictly weaker though Assumption 6(A) requires more "exotic" preferences in the economy.

Given these modified assumptions, we can now prove that Assumption 4′ (or Assumption 4) holds.

PROPOSITION 3. *Take any mechanism* $\Gamma = (\mathcal{M}_i, x_i, y_i)_{i\in\mathcal{I}}$ *and* ρ *satisfying Assumptions 1′, 2′, and 6(B), and any type space* Θ *satisfying Assumption 6(A). If* $\rho \leq 2$, *then Assumption 4′ is satisfied:* $\nu(\Theta) = \mathcal{M}^*$. *If* $\rho > 2$, *then* $\{(r, s) \in \mathcal{M}^* : y_i(r) \neq 0\ \forall i\} \subseteq \nu(\Theta)$.

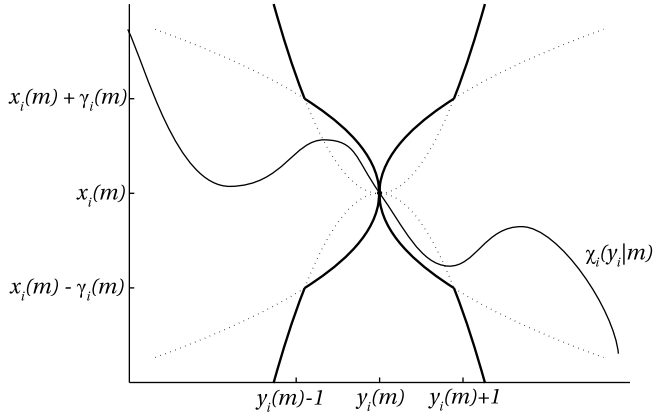

FIGURE 6. The bounds on $\chi_i(y_i|m_{-i})$ imposed by Assumption 6(B) for $\rho = 2$.

PROOF. Define

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

Note that if $\rho \in \{1,2\}$, then $\mathcal{M}^* = \mathcal{M}^{**}$ (using the convention that $0^0 = 1$). We know that $\nu(\Theta) \subseteq \mathcal{M}^*$; Proposition 3 can then be proven by showing that $\mathcal{M}^{**} \subseteq \nu(\Theta)$. This is done by constructing a mapping $\phi : \mathcal{M}^{**} \to \Theta_\rho$ such that $m \in \nu(\phi(m))$ for all $m \in \mathcal{M}^{**}$. Thus,

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

giving the result.

Specifically, consider the mapping $\phi : \mathcal{M}^{**} \to \Theta_\rho$ 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

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

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

$$\beta_i(r^*, s^*) := \alpha_i(r^*, s^*) y_i^{\rho-1}(r) + P_{ik}(r^*, s^*) \tag{36}$$

(recall that P_{ik} is the effective price function defined in (15) 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 that 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

$$-\frac{\alpha_i(m^*)}{\rho} y_i(r_i, r^*_{-i})^\rho + \beta_i(m^*) y_i(r_i, r^*_{-i}) + x_i(r_i, s_i, r^*_{-i}, s^*_{-i}). \tag{37}$$

For local optimality, the first-order conditions for each s_{ik} are already satisfied at m^* by the construction of $\mathcal{M}^{**}$ (see (21)). As for r_{ik}, agent i's first-order condition for utility maximization at (r^*, s^*) with respect to each r_{ik} is

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

But the construction of β_i ((36)) 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 are 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 (37) (after inserting the definition of $\beta_i(m^*)$ from (36)) is given by the $K_i \times K_i$ matrix

$$\mathbf{H}_i = \left[\begin{array}{c|c} -\alpha_i(m^*)(\rho-1) y_i(r^*)^{\rho-2} (\mathbf{Y}_r \cdot \mathbf{Y}_r^{\mathrm{T}}) + P_{ik}(m^*)\mathbf{Y}_{rr} + \mathbf{X}_{rr} & \mathbf{X}_{rs} \\ \hline \mathbf{X}_{rs}^{\mathrm{T}} & \mathbf{X}_{ss} \end{array}\right],$$

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 $\mathbf{d}_r = \mathbf{0}$ does 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

$$\begin{aligned}(\mathbf{d}_r, \mathbf{d}_s)^{\mathrm{T}} \cdot \mathbf{H}_i \cdot (\mathbf{d}_r, \mathbf{d}_s) &= -\alpha_i(m^*)(\rho-1) y_i(r^*)^{\rho-2} \mathbf{d}_r^{\mathrm{T}} (\mathbf{Y}_r \cdot \mathbf{Y}_r^{\mathrm{T}}) \mathbf{d}_r + K_i(m^*) \\ &= -\alpha_i(m^*)(\rho-1) y_i(r^*)^{\rho-2} (\mathbf{d}_r^{\mathrm{T}} \mathbf{Y}_r)^2 + K_i(m^*),\end{aligned}$$

where

$$K_i(m^*) = \mathbf{d}_r^{\mathrm{T}} [P_{ik}(m^*)\mathbf{Y}_{rr} + \mathbf{X}_{rr}] \mathbf{d}_r + 2\mathbf{d}_r^{\mathrm{T}} \mathbf{X}_{rs} \mathbf{d}_s + \mathbf{d}_s^{\mathrm{T}} \mathbf{X}_{ss} \mathbf{d}_s.$$

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 $y_i(r^*)^{\rho-2} \neq 0$, α_i can be chosen to be any function satisfying

$$\alpha_i(m^*) > K_i(m^*)((\rho-1) y_i(r^*)^{\rho-2})^{-1} (\mathbf{d}_r^{\mathrm{T}} \mathbf{Y}_r)^{-2}$$

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

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

which, by Assumption 6(B), is true if

$$\begin{aligned}\alpha_i'\left[\left(\frac{\rho-1}{\rho}y_i^{*\rho} + \frac{1}{\rho}y_i^{\prime\rho}\right) - (y_i^{*\rho})^{\frac{\rho-1}{\rho}}(y_i^{\prime\rho})^{\frac{1}{\rho}}\right] + P_{ik}(m^*)(y_i^* - y_i') \\ \geq \gamma_i(m^*)\rho\max\left\{|y_i^* - y_i'|^\rho, |y_i^* - y_i'|^{\frac{1}{\rho}}\right\}\end{aligned} \tag{38}$$

(the extra ρ 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 $y_i^{*\rho}$ and $y_i^{\prime\rho}$; by the inequality of arithmetic and geometric means (the AM-GM inequality) this difference is positive. Thus, there is some finite α_i' at which inequality (38) 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

$$\begin{aligned}\alpha_i(m^*)\left[\left(\frac{\rho-1}{\rho}y_i^{*\rho} + \frac{1}{\rho}y_i^{\rho}\right) - (y_i^{*\rho})^{\frac{\rho-1}{\rho}}(y_i^{\rho})^{\frac{1}{\rho}}\right] + P_{ik}(m^*)(y_i^* - y_i) \\ - \gamma_i(m^*)\rho\max\left\{|y_i^* - y_i|^\rho, |y_i^* - y_i|^{\frac{1}{\rho}}\right\} \geq 0,\end{aligned} \tag{39}$$

which is true for $y_i = y_i''$ (see inequality (38)). 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

$$\alpha_i(m^*)[y_i^{*\rho} - y_i^{*\rho-1}y_i + y_i^\rho - y_i^* y_i^{\rho-1}] + P_{ik}(m^*)(y_i^* - y_i) - \gamma_i(m^*)\rho(y_i - y_i^*)^\rho \tag{40}$$

or

$$\alpha_i(m^*)[y_i^{*\rho} - y_i^{*\rho-1}y_i + y_i^\rho - y_i^* y_i^{\rho-1}] + P_{ik}(m^*)(y_i^* - y_i) - \gamma_i(m^*)\frac{1}{\rho}(y_i - y_i^*)^{1/\rho}. \tag{41}$$

In either case, the expression is greater than the left-hand side of (39) because

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

reduces to

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

which is just the AM-GM inequality again. Thus, both (40) and (41) are positive. By continuity, (39) 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. □

Multidimensional mechanisms and the proof of Theorem 4

In the previous section, we provided two linked assumptions—one on the type space and one on the mechanism—that together imply Assumption 4′. Now we define *regular* candidate equilibrium points for which our theorem applies. Let $\mathcal{R}^*$ be the projection of $\mathcal{M}^*$ onto $\mathcal{R}$.

Definition 3. A candidate equilibrium $(r^*, s^*) \in \mathcal{M}^*$ is *regular* if $\sigma(r)$ is locally threaded by some differentiable function $\varsigma = \bigtimes_i \varsigma_i$ for each i; formally, (r^*, s^*) is regular if for each i, there is some open set $\mathcal{R}_i^0 \subset \mathcal{R}_i^*$ containing r_i^* and a differentiable function $\varsigma_i : \mathcal{R} \times \mathcal{S}_{-i} \to \mathcal{S}_i$ such that $\varsigma_i(r^*, s_{-i}^*) = s_i^*$ and $(r_i', r_{-i}^*, \varsigma_i(r_i', s_{-i}^*), s_{-i}^*) \in \mathcal{M}^*$ for all $r_i' \in \mathcal{R}_i^0$.[23]

A candidate equilibrium (r^*, s^*) is regular if differential deviations in r_i can always be accompanied by a differential change in s_i such that the joint deviation does not lead to a strategy profile outside of $\mathcal{M}^*$. We refer to ς_i as i's *adjustment function*. An example of a nonregular equilibrium is one for which a differential change in r_i leads to a point at which $\sigma(r)$ is empty or has a jump discontinuity.

We now prove higher-dimensional analogs to Theorem 4 and Corollary 3.

Theorem 8 (Necessity). *Suppose 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 4′. Writing the mechanism as*

$$x_i(r, s) = -q_i(r, s)y_i(r, s_{-i}) - g_i(r, s),$$

it must be the case that for every regular point $(r^*, s^*) \in \mathcal{M}^*$ with adjustment functions $(\varsigma_i)_i$,

$$\frac{dq_i(r^*, \varsigma_i(r^*, s_{-i}^*), s_{-i}^*)}{dr_{ik}} = 0 \quad \forall i \in \mathcal{I}, k \in \{1, \ldots, J_i\}$$

[23]The locally threaded condition rules out space-filling Peano functions, for example. Mount and Reiter (1974, 1977) describe the communication requirements of implementation via a "message process" $\nu : \Theta \twoheadrightarrow \mathcal{M}$ that may or may not be an equilibrium correspondence. They assume ν is locally threaded to rule out pathological cases where Peano functions are used to economize on message space dimensions.

and

$$g_i(r^*, s^*) = 0.$$

PROOF. 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 regular equilibrium point $m^* = (r^*, s^*)$ in $\mathcal{M}^*$ 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 $i \in \mathcal{I}$.

(A) Because m^* is a Nash equilibrium for some $\theta \in \Theta$, the first-order condition

$$\frac{\partial u_i(x_i^*, y_i^*|\theta_i)}{\partial y_i} \frac{\partial y_i(r_i^*)}{\partial r_{ik}} = \frac{\partial u_i(x_i^*, y_i^*|\theta_i)}{\partial x_i} \left[-\frac{\partial x_i(r^*, s^*)}{\partial r_{ik}} \right] \tag{42}$$

is satisfied for each $k \in \{1, \ldots, J_i\}$.

(B) 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 equal the transfers of the numeraire required by the Walrasian or Lindahl equilibrium:

$$x_i(r^*, s^*) = -p_i(\phi(r^*, s^*)) y_i(r^*). \tag{43}$$

(C) 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:

$$\frac{\partial u_i(x_i^*, y_i^*|\theta_i)/\partial y_i}{\partial u_i(x_i^*, y_i^*|\theta_i)/\partial x_i} = p_i(\phi(r^*, s^*)). \tag{44}$$

Dividing both sides of (42) by $\partial u_i/\partial x_i$, inserting (44), and rearranging gives

$$\frac{\partial x_i(r^*, s^*)}{\partial r_{ik}} = -p_i(\phi(r^*, s^*)) \frac{\partial y_i(r^*)}{\partial r_{ik}} \tag{45}$$

for each i and k.

Since (r^*, s^*) is a regular equilibrium point, differential changes in r_i accompanied by the requisite change in s_i lead to other points in $\mathcal{M}^*$ at which the above equations hold. Now take the total derivative of (43) with respect to r_i (allowing for the adjustment in s_i); since s_i maximizes x_i, the envelope theorem guarantees that $dx_i/dr_{ik} = \partial x_i/\partial r_{ik}$ and so the total derivative is

$$\frac{\partial x_i(r^*, s^*)}{\partial r_{ik}} = -p_i(\phi(r^*, s^*)) \frac{\partial y_i(r^*)}{\partial r_{ik}} - \frac{dp_i(\phi(r^*, s^*))}{dr_{ik}} y_i(r^*). \tag{46}$$

Comparing (45) and (46), it must be that either $y_i(r^*) = 0$ or $dp_i(\phi(r^*, s^*))/dr_{ik} = 0$ for all k.

If $y_i(r^*) \neq 0$ but $dp_i(\phi(r^*, s^*))/dr_{ik} = 0$ for all k, then, by (43),

$$g_i(r^*, s^*) = \big[p_i(\phi(r^*, s^*)) - q_i(r^*, s^*)\big] y_i(r^*)$$

and so $g_i(r^*, s^*)$ can be expressed as $h_i(r^*, s^*) y_i(r^*)$ for some function h_i such that $dh_i/dr_{ik} = 0$ for all k. But then $x_i(r^*, s^*)$ can be rewritten as

$$x_i(r^*, s^*) = -[q_i(r^*, s^*) + h_i(r^*, s^*)] y_i(r^*).$$

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

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

with $d\tilde{q}/dr_{ik} = 0$, giving the result.

If $y_i(r^*) = 0$, then by (43) we have $g_i(r^*, s^*) = 0$. It remains to show that $dq_i(r^*, s^*)/dr_{ik} = 0$. Since dy_i/dr_{ik} is bounded away from zero, any perturbation of r_{ik} leads to $y_i \neq 0$; by regularity of (r^*, s^*), small perturbations lead to other regular equilibria with $y_i \neq 0$ at which $dq_i/dr_{ik} = 0$. Since q_i is continuously differentiable, it must be that $dq_i(r^*, s^*)/dr_{ik} = 0$ as well. □

Proof of Theorems 5 and 7. Consider the case of Lindahl equilibrium. Under the maintained assumptions, an allocation $(x_1^*, \dots, x_n^*, y^*)$ is Lindahl equilibrium allocation at θ if there exists some $(p_i^*)_i$ such that the following conditions hold.

(A) For each i, $(x_i^*, y^*) \in \arg\max_{x_i, y} u_i(x_i, y|\theta_i)$ subject to $x_i = -p_i^* y$.

(B) We have $\sum_i p_i^* = \kappa$.

For the first part of the theorem, fix a Nash equilibrium $m^* = (r^*, s^*)$ of Γ at θ and let $p_i^* = q_i(m_{-i}^*)$ for each i. Then condition B is satisfied by hypothesis. Condition A can be rewritten as

$$y^* \in \arg\max_y u_i(-q_i(m_{-i}^*) y, y|\theta_i). \tag{47}$$

Since y is bijective in r_i for each m_{-i}, this is equivalent to

$$(r_i^*, s_i^*) \in \arg\max_{(r_i, s_i)} u_i(-q_i(r_{-i}^*, s_{-i}^*) y(r_i, r_{-i}^*, s_{-i}^*), y(r_i, r_{-i}^*, s_{-i}^*)|\theta_i).$$

Because $g_i \geq 0$ and $g_i = 0$ at any equilibrium point, and u_i is increasing in the first argument, condition A is also equivalent to

$$(r_i^*, s_i^*) \in \arg\max_{(r_i, s_i)} u_i(-q_i(r_{-i}^*, s_{-i}^*) y(r_i, r_{-i}^*, s_{-i}^*) - g_i(r_i, r_{-i}^*, s_i, s_{-i}^*), y(r_i, r_{-i}^*, s_{-i}^*)|\theta_i).$$

But this is clearly satisfied since (r_i^*, s_i^*) is a best response to (r_{-i}^*, s_{-i}^*). Thus,

$$(x_1(m^*), \dots, x_n(m^*), y(m^*))$$

is a Lindahl equilibrium allocation at θ with prices $(p_i(m_{-i}^*))_i$.

For the second part of the theorem, fix a Lindahl allocation (x^*, y^*) with prices $(p_i^*)_i$ such that message $m' = (r', s')$ maps to (x^*, y^*) and $q_i(m'_{-i}) = p_i^*$ for each i; Assumption 5 guarantees that at least one such m' exists.

Condition A for Lindahl equilibria is equivalent to (47); since $s' \in \sigma(r')$, this is equivalent to

$$(r'_i, s'_i) \in \arg\max_{(r_i, s_i)} u_i(-q_i(r'_{-i}, s'_{-i})y(r_i, r'_{-i}, s'_{-i}) - g_i(r_i, r'_{-i}, s_i, s'_{-i}), y(r_i, r'_{-i}, s'_{-i})|\theta_i).$$

But this implies that (r'_i, s'_i) is a best response for each i, so (r', s') is a Nash equilibrium of Γ at θ.

The proof for Walrasian equilibria is identical, setting $\kappa = 0$. □

Proof of Theorem 6. We know that no one-dimensional mechanism can exist that Nash-implements the Walrasian correspondence—contractive or not. For the public-goods setting, suppose, to the contrary, that the mechanism $(y, (q_i, g_i)_{i=1}^n)$ Nash-implements the Lindahl correspondence and is contractive. By Theorem 8, we know that $g_i \equiv 0$, and so for any quasilinear environment

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

with $v_i'' < 0$, we have

$$U_i(r_i, r_{-i}) = v_i(y(r)|\theta_i) - q_i(r_{-i})y(r).$$

Agent i's best response is given by $\rho_i(r_{-i})$ and satisfies the first-order condition

$$v_i'(y(\rho_i, r_{-i})|\theta_i) = q_i(r_{-i})$$

for all r_{-i}. Take any r^* and θ for which r^* is a Nash equilibrium at θ. By the implicit function theorem, the slope of ρ_i at r^* with respect to each r_j is

$$\frac{\partial \rho_i}{\partial r_j} = \frac{\partial q_i/\partial r_j - v_i''(y|\theta_i)\partial y/\partial r_j}{v_i''(y|\theta_i)\partial y/\partial r_i}.$$

For the mechanism to be contractive, it is necessary (though not sufficient) that for all i and $j \neq i$,

$$\left| \frac{\partial y/\partial r_j}{\partial y/\partial r_i} - \frac{\partial q_i/\partial r_j}{v_i''(y|\theta_i)\,\partial y/\partial r_i} \right| < 1. \tag{48}$$

Now select the agent j^* such that $|\partial y(r^*)/\partial r_{j^*}| \geq |\partial y(r^*)/\partial r_i|$ for all i. To satisfy (48), it must be that $\partial y/\partial r_{j^*}$ and $\partial q_i/\partial r_{j^*}$ have the opposite signs for all i and that $\partial q_i/\partial r_{j^*} \neq 0$ for all i. Therefore, each $\partial q_i/\partial r_{j^*}$ has the same sign for all $i \neq j^*$ (and $\partial q_{j^*}/\partial r_{j^*} = 0$), so that $\sum_i \partial q_i/\partial r_{j^*} \neq 0$. But since all r are Nash equilibria for some θ and the mechanism implements Lindahl allocations, it must be that $\sum_i q_i(r_{-i}) = \kappa$ for all r and, therefore, that $\sum_i \partial q_i/\partial r_{j^*} = 0$; this is a contradiction. □

References

Abreu, Dilip and Hitoshi Matsushima (1992), "Virtual implementation in iteratively undominated strategies: Complete information." *Econometrica*, 60, 993–1008. [615]

Arifovic, Jasmina and John O. Ledyard (2011), "A behavioral model for mechanism design: Individual evolutionary learning." *Journal of Economic Behavior and Organization*, 78, 374–395. [623, 624]

Arrow, Kenneth J. (1951), *Social Choice and Individual Values*. Wiley, New York. [611]

Bergemann, Dirk and Stephen Morris (2011), "Rationalizable implementation." *Journal of Economic Theory*, 146, 1253–1274. [615]

Bergstrom, Theodore, Carl P. Simon, and Charles J. Titus (1983), "Counting Groves–Ledyard equilibria via degree theory." *Journal of Mathematical Economics*, 12, 167–184. [613]

Brock, William A. (1980), "The design of mechanisms for efficient allocation of public goods." In *Quantitative Economics and Development* (Lawrence R. Klein, Marc Nerlove, and S. C. Tsiang, eds.), 45–80, Academic Press, New York. [611]

Cabrales, Antonio (1999), "Adaptive dynamics and the implementation problem with complete information." *Journal of Economic Theory*, 86, 159–184. [615]

Cabrales, Antonio and Roberto Serrano (2011), "Implementation in adaptive better-response dynamics: Towards a general theory of bounded rationality in mechanisms." *Games and Economic Behavior*, 73, 360–374. [615]

Cason, Tim, Tatsuyoshi Saijo, Tomas Sjöström, and Takehiko Yamato (2006), "Secure implementation experiments: Do strategy-proof mechanisms really work?" *Games and Economic Behavior*, 57, 206–235. [615]

Chen, Yan (2002), "A family of supermodular Nash mechanisms implementing Lindahl allocations." *Economic Theory*, 19, 773–790. [612, 614, 622, 623, 627, 638]

Chen, Yan (2008), "Incentive compatible mechanisms for pure public goods: A survey of experimental research." In *The Handbook of Experimental Economics Results*, Vol. 1 (Charles R. Plott and Vernon L. Smith, eds.), 625–643, North-Holland, Amsterdam. [622]

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. [610, 614, 624]

Chen, Yan and Charles R. Plott (1996), "The Groves–Ledyard mechanism: An experimental study of institutional design." *Journal of Public Economics*, 59, 335–364. [610, 613, 614]

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. [610, 613, 614, 622, 623]

Conlisk, John (1973), "Quick stability checks and matrix norms." *Economica*, 40, 402–409. [618]

de Trenqualye, Pierre (1989), "Stable implementation of Lindahl allocations." *Economic Letters*, 29, 291–294. [611, 613]

Gibbard, Allan (1973), "Manipulation of voting schemes: A general result." *Econometrica*, 41, 587–601. [609]

Glazer, Jacob and Robert W. Rosenthal (1992), "A note on Abreu–Matsushima mechanisms." *Econometrica*, 60, 1435–1438. [615]

Green, Jerry and Jean-Jacques Laffont (1977), "Characterization of satisfactory mechanisms for the revelation of preferences for public goods." *Econometrica*, 45, 427–438. [615]

Groves, Theodore and John O. Ledyard (1977), "Optimal allocation of public goods: A solution to the 'free-rider' problem." *Econometrica*, 45, 783–809. [609, 613, 616, 623, 633]

Groves, Theodore and John O. Ledyard (1980), "The existence of efficient and incentive compatible equilibria with public goods." *Econometrica*, 48, 1487–1506. [634]

Groves, Theodore and John O. Ledyard (1987), "Incentive compatibility since 1972." In *Information, Incentives, and Economic Mechanisms* (Theodore Groves, Roy Radner, and Stanley Reiter, eds.), 48–114, University of Minnesota Press, Minneapolis. [611]

Harstad, Ronald M. and Michael Marrese (1982), "Behavioral explanations of efficient public good allocations." *Journal of Public Economics*, 19, 367–383. [613]

Healy, Paul J. (2006), "Learning dynamics for mechanism design: An experimental comparison of public goods mechanisms." *Journal of Economic Theory*, 129, 114–149. [610, 613, 614, 615, 619, 622, 634]

Hirota, Masayoshi (1985), "Global stability in a class of markets with three commodities and three consumers." *Journal of Economic Theory*, 36, 186–192. [614]

Hurwicz, Leonid (1979a), "On allocations attainable through Nash equilibria." *Journal of Economic Theory*, 21, 140–165. [613, 616]

Hurwicz, Leonid (1979b), "Outcome functions yielding Walrasian and Lindahl allocations at Nash equilibrium points." *Review of Economic Studies*, 46, 217–225. [609, 613]

Hurwicz, Leonid and Stanley Reiter (2006), *Designing Economic Mechanisms*. Cambridge University Press, Cambridge. [611]

Hurwicz, Leonid and Mark Walker (1990), "On the generic nonoptimality of dominant-strategy allocation mechanisms: A general theorem that includes pure exchange economies." *Econometrica*, 58, 683–704. [609]

Jackson, Matthew O. (2001), "A crash course in implementation theory." *Social Choice and Welfare*, 18, 655–708. [616]

Jordan, James S. (1986), "Instability in the implementation of Walrasian allocations." *Journal of Economic Theory*, 39, 301–328. [610, 613, 625, 639]

Kim, Taesung (1987), *Stability Problems in the Implementation of Lindahl Allocations*. Ph.D. thesis, University of Minnesota. [610, 613, 625, 639]

Kim, Taesung (1993), "A stable Nash mechanism implementing Lindahl allocations for quasi-linear environments." *Journal of Mathematical Economics*, 22, 359–371. [611, 613, 614]

Kim, Taesung (1996), "A stable Nash mechanism for quasi-additive public-good environments." *Japanese Economic Review*, 47, 144–156. [611, 613]

Ledyard, John O. (1995), "Public goods: A survey of experimental research." In *Handbook of Experimental Economics* (John H. Kagel and Alvin E. Roth, eds.), 111–194, Princeton University Press, Princeton. [614]

Maskin, Eric (1999), "Nash equilibrium and welfare optimality." *Review of Economic Studies*, 66, 23–38. [611, 615]

Mathevet, Laurent A. (2010), "Supermodular mechanism design." *Theoretical Economics*, 5, 403–443. [612, 622]

McKelvey, Richard D. and Thomas R. Palfrey (1992), "An experimental study of the centipede game." *Econometrica*, 60, 803–836. [615]

Milgrom, Paul and John Roberts (1990), "Rationalizability, learning, and equilibrium in games with strategic complementarities." *Econometrica*, 58, 1255–1277. [610, 612, 614, 620, 621]

Mount, Kenneth and Stanley Reiter (1974), "The informational size of message spaces." *Journal of Economic Theory*, 8, 161–192. [653]

Mount, Kenneth and Stanley Reiter (1977), "Economic environments for which there are Pareto satisfactory mechanisms." *Econometrica*, 45, 821–842. [653]

Muench, Thomas and Mark Walker (1983), "Are Groves–Ledyard equilibria attainable?" *Review of Economic Studies*, 50, 393–396. [613]

Nagel, Rosemarie C. (1995), "Unraveling in guessing games: An experimental study." *American Economic Review*, 85, 1313–1326. [615]

Novshek, William (1985), "On the existence of Cournot equilibrium." *Review of Economic Studies*, 52, 85–98. [630]

Page, Scott E. and Troy Tassier (2004), "Equilibrium selection and stability for the Groves Ledyard mechanism." *Journal of Public Economic Theory*, 6, 311–335. [613]

Reichelstein, Stefan and Stanley Reiter (1988), "Game forms with minimal message spaces." *Econometrica*, 56, 661–692. [611, 633]

Roberts, Kevin (1979), "The characterization of implementable social choice rules." In *Aggregation and Revelation of Preferences* (Jean-Jacques Laffont, ed.), 321–348, North-Holland, Amsterdam. [615]

Saijo, Tatsuyoshi, Tomas Sjöström, and Takehiko Yamato (2007), "Secure implementation." *Theoretical Economics*, 2, 203–229. [615]

Sandholm, William H. (2002), "Evolutionary implementation and congestion pricing." *Review of Economic Studies*, 69, 667–689. [615]

Sandholm, William H. (2005), "Negative externalities and evolutionary implementation." *Review of Economic Studies*, 72, 885–915. [615]

Sandholm, William H. (2007), "Pigouvian pricing and stochastic evolutionary stability." *Journal of Economic Theory*, 132, 367–382. [615]

Satterthwaite, Mark A. (1975), "Strategy-proofness and Arrow's conditions: Existence and correspondence theorems for voting procedures and social welfare functions." *Journal of Economic Theory*, 10, 187–217. [609]

Scarf, Herbert (1960), "Some examples of global instability of the competitive equilibrium." *International Economic Review*, 1, 157–172. [614]

Scherr, Bruce A. and Emerson M. Babb (1975), "Pricing public goods: An experiment with two proposed pricing systems." *Public Choice*, 23, 35–48. [613]

Sefton, Martin and Abdullah Yavaş (1996), "Abreu–Matsushima mechanisms: Experimental evidence." *Games and Economic Behavior*, 16, 280–302. [615]

Smith, Vernon L. (1979), "Incentive compatible experimental processes for the provision of public goods." In *Research in Experimental Economics* (Vernon L. Smith, ed.), 392–455, JAI Press, Greenwich, Connecticut. [613]

Stahl, Dale O. and Paul W. Wilson (1995), "On players' models of other players: Theory and experimental evidence." *Games and Economic Behavior*, 10, 218–254. [615]

Tian, Guoqiang (1990), "Completely feasible and continuous implementation of the Lindahl correspondence with a message space of minimal dimension." *Journal of Economic Theory*, 51, 443–452. [633]

Tideman, T. Nicolaus (1983), "An experiment in the demand-revealing process." *Public Choice*, 41, 387–401. [613]

Van Essen, Matthew (2009), "A note on the stability of Chen's Lindahl mechanism." *Social Choice and Welfare*, 38, 365–370. [622, 623, 627]

Van Essen, Matthew (2010), "Information complexity, punishment, and stability in two Nash efficient Lindahl mechanisms." *Review of Economic Design*, 16, 15–40. [614]

Van Essen, Matthew (forthcoming), "A simple supermodular mechanism that implements Lindahl allocations." *Journal of Public Economic Theory*. [614, 623, 627, 638]

Van Essen, Matthew, Natalia Lazzati, and Mark Walker (2012), "Out-of-equilibrium performance of three Lindahl mechanisms: Experimental evidence." *Games and Economic Behavior*, 74, 366–381. [614]

Varian, Hal R. (1994), "A solution to the problem of externalities when agents are well-informed." *American Economic Review*, 84, 1278–1293. [624]

Vega-Redondo, Fernando (1989), "Implementation of Lindahl equilibrium: An integration of static and dynamic approaches." *Mathematical Social Sciences*, 18, 211–228. [611, 613]

Walker, Mark (1981), "A simple incentive compatible scheme for attaining Lindahl allocations." *Econometrica*, 49, 65–71. [609, 613, 622, 633]

Walker, Mark (1984), "A simple auctioneerless mechanism with Walrasian properties." *Journal of Economic Theory*, 32, 111–127. [615]

Zhou, Lin (1991), "Impossibility of strategy-proof mechanisms in economies with public goods." *Review of Economic Studies*, 58, 107–119. [609]

Submitted 2010-11-7. Final version accepted 2011-8-16. Available online 2011-8-16.