# Robust Implementation: The Case of Direct Mechanisms*

Dirk Bergemann†  Stephen Morris‡

First Version: August 2005
This Version: February 2006

**Abstract**

A social choice function is robustly implementable if there is a mechanism under which the process of iteratively eliminating strictly dominated messages leads to outcomes that agree with the social choice at every type profile. In an interdependent value environment, we identify a strict contraction property on the preferences which together with strict ex post incentive compatibility and the strict single crossing property is sufficient to guarantee robust implementation in the direct mechanism.

The contraction property essentially requires that the interdependence is not too large. In a linear signal model, the contraction property is equivalent to an interdependence matrix having an eigenvalue less than one. The contraction property is also necessary for robust implementation in any mechanism.

Keywords: Mechanism Design, Implementation, Robustness, Common Knowledge, Interim Equilibrium, Iterative Deletion, Direct Mechanism.

Jel Classification: C79, D82

*This research is supported by NSF Grant #SES-0095321 and #SES-0518929. Morris is grateful for financial support from the John Simon Guggenheim Foundation and the Center for Advanced Studies in the Behavioral Sciences. We would like to thank Sandeep Baliga, Federico Echenique, Donald Goldfarb, Matt Jackson, Jim Jordan, Vijay Krishna, Tom Palfrey and Daniel Spielman for helpful discussions. We thank the seminar audiences at Pennsylvania State University, Stanford University and University of British Columbia for comments. This paper presents refined and extended versions of results we reported in Bergemann and Morris (2005b).

†Department of Economics, Yale University, 28 Hillhouse Avenue, New Haven CT 06511, dirk.bergemann@yale.edu.

‡Department of Economics, Princeton University, Prospect Street, Princeton NJ 08544, smorris@princeton.edu.

# 1 Introduction

The mechanism design literature provides a powerful characterization of which social choice functions can be achieved when the designer has incomplete information about agents' types. If we assume a commonly known common prior over the possible types of agents, the revelation principle establishes that if the social choice function can arise as an equilibrium in some mechanism, then it will arise in a truth-telling equilibrium of the direct mechanism (where each agent truthfully reports his type and the designer chooses an outcome assuming they are telling the truth). Thus the Bayesian incentive compatibility constraints characterize whether a social choice function is implementable in this sense.

But even if a truth-telling equilibrium of the direct mechanism exists, there is no guarantee that there do not exist non truth-telling equilibria that deliver unacceptable outcomes. For this reason, the literature on *full* implementation has sought to show the existence of a mechanism all of whose equilibria deliver the social choice function. A classic literature on Bayesian implementation – Postlewaite and Schmeidler (1986), Palfrey and Srivastava (1989) and Jackson (1991) - characterized when this is possible: a *Bayesian monotonicity*[1] condition is necessary for full implementation, in addition to the Bayesian incentive compatibility conditions. Bayesian monotonicity and Bayesian incentive compatibility are also "almost" sufficient for full implementation.[2]

This important literature has had a limited impact on the more applied mechanism design literature, despite the fact that the problem of multiple equilibria is real. One difficulty is that the key sufficient condition - Bayesian monotonicity - is hard to interpret. Another difficulty is that, in general, positive results rely on complicated indirect, or "augmented," mechanisms in which agents report more that their types. Such mechanisms appear impractical to many researchers. We believe that both difficulties arise because the standard formulation of the Bayesian implementation problem - assuming common knowledge of a common prior on agents' types and equilibrium as solution concept - endows the planner with more information than would be available in practise. The implementing mechanism and equilibrium then rely on that information in an implausible way.

In this paper, we characterize when a social choice function can be *robustly* implemented. We fix a social choice environment including a description of the set of possible payoff types for each agent. We ask when does there exist a mechanism with the property that every outcome consistent with common knowledge of rationality agrees with the social choice function, making no

[1] The Bayesian monotonicity condition is an incomplete information analogue of the classic "Maskin monotonicity" condition shown to be necessary and almost sufficient for complete information implementation by Maskin (1999).

[2] Jackson (1991) shows that they are sufficient in economic environments and a slight strengthening is sufficient in non-economic environments.

assumptions about agents' beliefs and higher order beliefs about other agents' payoff types. This requirement gives rise to an iterative deletion procedure: fix a mechanism and iteratively delete messages for each payoff type that are strictly dominated by another message for each payoff type profile and message profile that has survived the procedure. This notion of robust implementation is equivalent to requiring that every equilibrium on every type space corresponding to the social choice environment delivers the right outcome.

This paper identifies a class of environments where there are easily understood and tight characterizations of when robust implementation is possible. As always, there will be an incentive compatibility condition that is necessary: strict ex post incentive compatibility is necessary for robust implementation.[3] We show that if, in addition, a *contraction* property - which we explain shortly - is satisfied, robust implementation is possible in the direct mechanism. If strict ex post incentive compatibility or the contraction property fail, then robust implementation is not possible in any mechanism. Thus the "augmented" mechanisms used in the earlier complete information and Bayesian full implementation literature do not perform better than the simpler direct mechanisms. An intuition for this result is that the strong common knowledge assumptions used in the complete information and Bayesian implementation literatures can be exploited via complex augmented mechanisms. Thus an attractive feature of our approach is that the robustness requirement reduces the usefulness of complexity in mechanism design (without any ad hoc restrictions on complexity).

In the case of private values, strict ex post incentive compatibility is equivalent to strict dominant strategies incentive compatibility. Thus full implementation is obtained for free. It follows that the contraction property must have bite only if there are interdependent values. In fact, the contraction property requires exactly that there is *not too much* interdependence in players' types. The contraction property can be nicely illustrated in a class of interdependent preferences in which the private types of the agents can be linearly aggregated. If $\theta_j$ is the type of agent $j$, then agent $i$'s utility depends on $\theta_i + \gamma \sum_{j \neq i} \theta_j$. Thus if $\gamma \neq 0$, there are interdependent values - agent $j$'s type will enter agent $i$'s utility assessment - but each agent $i$ cares differently about his own type than about other agents' types. In this example, the contraction property reduces to the requirement that $|\gamma| < 1/(I-1)$, where $I$ is the number of agents. We provide characterizations of the contraction property - all equivalent to the intuition that there is not too much interdependence - in more general linear environments and when there is non-linear aggregation of agents' types.

The results of this paper apply to environments where agents' type profile can be aggregated

[3] Our earlier work on robust mechanism design, Bergemann and Morris (2005c), showed that ex post incentive compatibility was necessary and sufficient for partial robust implementation (i.e, ensuring that there exists *an* equilibrium consistent with the social choice function).

into a one dimensional sufficient statistic for each player, where preferences are single crossing with respect to that statistic. These restrictions incorporate many economic models with interdependence in the literature: we illustrate our results with a public good example with linear aggregator described above; we also apply our results to the classic problem of allocating a single private good with quasilinear utility (i.e., a single unit auction with interdependent utility). While these restrictions are strong, we provide a simply informational story that would explain environments with the properties we describe.

We focus in this paper on economically important environments and well behaved mechanisms where we get clean and tight characterizations of the robust implementation problem with direct or augmented mechanisms. An attractive feature of the methods and results is that they can be derived as applications of the rather abstract arguments in Bayesian implementation literature. Thus the contraction property is equivalent to the *robust monotonicity* condition that is necessary and almost sufficient for full implementation in general environments. Robust monotonicity is equivalent to requiring Bayesian monotonicity on all type spaces. We derive our results directly - not in this insightful but more indirect way - in this paper.[4] We discuss robust implementation in general environments in section 8.

An important paper of Chung and Ely (2001) analyzed auctions with interdependent valuations under elimination of weakly dominated strategies. In a linear and symmetric setting, they reported sufficient conditions for direct implementation that coincide with the ones derived here. We show that in the environment with linear aggregation, under strict incentive compatibility, the basic insight extends from the single unit auction model to general allocations models, with elimination of strictly dominated actions only (thus Chung and Ely (2001) require deletion of weakly dominated strategies only because incentive constraints are weak). We also prove a converse result: if there is too much interdependence, then neither the direct nor any augmented mechanism can robustly implement the social choice function (this result will also hold with deletion of weakly dominated strategies).

The ex post incentive constraints necessary for robust implementation are already strong (even without the contraction property). Jehiel, Moldovanu, Meyer-Ter-Vehn, and Zame (2005) have recently shown that in an environment with multi-dimensional signals, the ex post incentive constraints are "generically" impossible to satisfy with multi-dimensional signals. If ex post incentive compatibility fails, our positive results are moot. While this provides a natural limit for our analysis, there are many interesting applications for which ex post equilibria do exist, among them one-dimensional signal models (Dasgupta and Maskin (2000), Perry and Reny (2002), Bergemann

[4] We pursued the indirect derivation of some of the results in this paper in an earlier working paper version, Bergemann and Morris (2005b).

and Välimäki (2002)), models without allocative externalities (Bikhchandani (2005)) and other models (see the recent survey Jehiel and Moldovanu (2006)) for many positive and negative results) to which our analysis applies.

The remainder of the paper is organized as follows. Section 2 describes the formal environment and solution concepts. Section 3 considers a public good example that illustrates the main ideas and results of the paper. Section 4 establishes necessary conditions for robust implementation in the direct mechanism. Section 5 considers the preference environment with a linear aggregation of the types and obtains sharp implementation results. Section 6 reports sufficient conditions for robust implementation. Section 7 considers a single unit auction with interdependent values as a second example of robust implementation. Section 8 concludes.

# 2 Setup

## 2.1 Payoff Environment

We consider a finite set of agents, $1, 2, ..., I$. Agent $i$'s *payoff type* is $\theta_i \in \Theta_i$, where $\Theta_i$ is a compact subset of the real line. We write $\theta \in \Theta = \Theta_1 \times \cdots \times \Theta_I$. Let $Y$ be a compact set of outcomes. Let agent $i$'s utility if outcome $y$ is chosen and agents' type profile is $\theta$ be $u_i(y, \theta)$. A social choice function is a mapping $f : \Theta \to Y$.

We assume the existence of a monotonic aggregator $h_i(\theta)$ for each $i$, which allows us to rewrite the utility function of every agent $i$ as:

$$u_i(y, \theta) \equiv v_i(y, h_i(\theta)) \tag{1}$$

where $h_i : \Theta \to \mathbb{R}$ is continuous, strictly increasing in $\theta_i$ and $v_i : Y \times \mathbb{R} \to \mathbb{R}$ is continuous.[5] The content of this assumption comes from the continuity requirement and the restrictions that we will later impose on $v_i$ in section 4.1.

[5] We briefly discussed the impossibility results of ex post incentive compatibility with multi-dimensional signals, see Jehiel, Moldovanu, Meyer-Ter-Vehn, and Zame (2005) in the introduction. In this context, note that the impossibility results are obtained in a setting where the utility function is defined separately for every allocation $y$. The utility function of every agent $i$ at every allocation $y$, $u_{i,y}(\cdot)$, is then assumed to be twice differentiable in the signal. Yet, no continuity or monotonicity assumption are made across allocations. In our setting, the aggregation of private types acts independently of the particular allocation. Yet, provided the existence of an aggregating function $h_i(\theta)$, we could allow the signal space of each agent $i$ to be multi-dimensional without any further modification. Our analysis uses the single-crossing condition and hence a systematic interaction between the set of allocations and signals and provided that aggregation is possible, the dimensionality of the signal per se is not an issue.

### 2.2 Mechanisms

A planner must choose a *game form* or *mechanism* for the agents to play in order the determine the social outcome. Let $M_i$ be a compact set of messages available to agent $i$. Let $g(m)$ be the outcome chosen if action profile $m$ is chosen. Thus a mechanism is a collection:

$$\mathcal{M} = (M_1, ..., M_I, g(\cdot)),$$

where $g : M \to Y$. The *direct mechanism* has the property that $M_i = \Theta_i$ for all $i$ and $g(\theta) = f(\theta)$.

### 2.3 Robust Implementation

In a fixed mechanism $\mathcal{M}$, we call a correspondence $S = (S_1, ...., S_I)$, with each $S_i : \Theta_i \to 2^{M_i} / \varnothing$, a message profile of the agents. We write $\mathcal{S}$ for the collection of message profiles. In the direct mechanism a message profile $S$ may be thought of as a deception $\beta = (\beta_1, ..., \beta_I)$, or

$$\beta_i : \Theta_i \to 2^{\Theta_i} / \varnothing, \text{ for all } i.$$

A deception $\beta_i(\theta_i)$ is a set of possible reports of agent $i$ to the principal regarding his true type. We shall assume without loss of generality that $\theta_i \in \beta_i(\theta_i)$ for all $i$ and all $\theta_i$. Let $\beta^*$ be the minimal deception, specifically truthtelling, with $\beta_i^*(\theta_i) = \{\theta_i\}$ for all $i$ and $\theta_i$.

Next we define the process of iterative elimination of never best responses. We denote the belief of agent $i$ over message and payoff type profiles of the remaining agents by a Borel measure $\lambda_i$:

$$\lambda_i \in \Delta(M_{-i} \times \Theta_{-i})$$

We initiate $S_i^0 = M_i$ and define inductively:

$$S_i^{k+1}(\theta_i) = \left\{ m_i \in M_i \;\middle|\; \exists \lambda_i \text{ s.th.:} \begin{array}{ll} (1) & \lambda_i\left[\left\{(m_{-i}, \theta_{-i}),\ m_j \in S_j^k(\theta_j),\ \forall\, j \neq i;\right\}\right] = 1 \\ (2) & \begin{array}{c} \int u_i(g(m_i, m_{-i}), (\theta_i, \theta_{-i}))\, d\lambda_i \ \geq \\ \int u_i(g(m_i', m_{-i}), (\theta_i, \theta_{-i}))\, d\lambda_i,\ \forall m_i' \in M_i. \end{array} \end{array} \right\}. \quad (2)$$

We observe that $S_i^k$ is (weakly) decreasing in $k$. We denote the limit set by $S^{\mathcal{M}}(\theta)$, or

$$S^{\mathcal{M}}(\theta) \triangleq \lim_{k\to\infty} S^k(\theta), \text{ for all } \theta \in \Theta.$$

By compactness of the message sets, we have

$$S_i^{\mathcal{M}}(\theta_i) \triangleq \bigcap_{k\geq 1} S_i^k(\theta_i).^6$$

[6] Because of the compactness of the message set, this procedure is equivalent, by a standard duality argument, to the iterated deletion of actions which are dominated by mixed strategies against all message type profiles that have not yet been deleted.

For brevity and for lack of a better expression, we refer to the messages $m_i \in S_i^{\mathcal{M}}(\theta_i)$ as *rationalizable messages*. We call a social choice function $f$ robust implementable if there exists a mechanism $\mathcal{M}$ under which the social choice can be recovered through a process of iterative elimination of never best responses.

**Definition 1 (Robust Implementation)**
*Social choice function $f$ is robustly implemented by mechanism $\mathcal{M}$ if $m \in S^{\mathcal{M}}(\theta) \Rightarrow g(m) = f(\theta)$.*

The set of rationalizable messages for mechanism $\mathcal{M}$ is equal to the set of messages that could be played in a Bayesian equilibrium of the game generated by the mechanism $\mathcal{M}$ and some type space. The basic logic of the argument follows the well-known argument of Brandenburger and Dekel (1987) for complete information games, showing the equivalence of correlated rationalizable actions and the set of actions that could be played in a subjective correlated equilibrium. Battigalli and Siniscalchi (2003) describe the incomplete information extension of this observation. A formal version of the equivalence is reported in Proposition 1 of our working paper, Bergemann and Morris (2005b).

## 3 A Public Good Example

We precede the formal results with an example illustrating the main insights of the paper. At the same time, the example facilitates a brief review of the key results in the implementation literature. The example involves the provision of a public good with quasilinear utility. The utility of each agent is given by:

$$u_i(\theta, y) = \left(\theta_i + \gamma \sum_{j \neq i} \theta_j\right) y_0 + y_i,$$

where $y_0$ is the level of public good provided and $y_i$ is the monetary transfer to agent $i$. The utility of agent $i$ depends on his own type $\theta_i \in [0,1]$ and the type profile of other agents, with $\gamma \geq 0$. The utility function of agent $i$ has the aggregation property with

$$h_i(\theta) = \theta_i + \gamma \sum_{j \neq i} \theta_j,$$

but we notice the aggregator function $h_i(\theta)$ depends on the agent $i$. In particular, a given type profile $\theta$ leads to a different aggregation result for $i$ and $j$, provided that $\theta_i \neq \theta_j$.

The cost of establishing the public good is given by $c(y_0) = \frac{1}{2}y_0^2$. The planner must choose $(y_0, y_1, ..., y_I) \in \mathbb{R}_+ \times \mathbb{R}^I$ to maximize social welfare, i.e., the sum of gross utilities minus the cost of the public good:

$$\left( (1+\gamma (I-1)) \sum_{i=1}^{I} \theta_i \right) y_0 - \frac{1}{2} y_0^2 .$$

The socially optimal level of the public good is therefore equal to

$$f_0 (\theta) = (1+\gamma (I-1)) \sum_{i=1}^{I} \theta_i .$$

We choose the generalized Vickrey-Groves-Clark transfers, essentially unique up to a constant, that give rise to ex post incentive compatibility:

$$f_i (\theta) = -(1+\gamma (I-1)) \left( \gamma \theta_i \sum_{j \neq i} \theta_j + \frac{1}{2} \theta_i^2 \right) . \tag{3}$$

It is useful to observe that the generalized VCG transfers given by (3) guarantee ex post incentive compatibility for any $\gamma \in \mathbb{R}$. Hence, ex post incentive compatibility per se does not impose any constraint on the interdependence parameter $\gamma$.

Now we shall argue that if $\gamma < \frac{1}{I-1}$, the social choice function $f$ is robustly implementable in the *direct mechanism* where each agent reports his payoff type $\theta_i$ and the planner chooses outcomes according to $f$ on the assumption that agents are telling the truth. Consider an iterative deletion procedure. Let $\beta^0 (\theta_i) = [0,1]$ and, for each $k = 1, 2, ...$, let $\beta^k (\theta_i)$ be the set of reports that agent $i$ might send, for some conjecture over his opponents' types and reports, with the only restriction on his conjecture being that each type $\theta_j$ of agent $j$ sends a message in $\beta^{k-1} (\theta_j)$.

Suppose that agent $i$ has payoff type $\theta_i$, but reports himself to be type $\theta_i'$ and has a point conjecture that other agents have type profile $\theta_{-i}$ and report their types to be $\theta_{-i}'$. Then his expected payoff is a constant $(1+\gamma (I-1))$ times:

$$\left( \theta_i + \gamma \sum_{j \neq i} \theta_j \right) \left( \theta_i' + \sum_{j \neq i} \theta_j' \right) - \left( \gamma \theta_i' \sum_{j \neq i} \theta_j' + \frac{1}{2} (\theta_i')^2 \right) .$$

The first order condition with respect to $\theta_i'$ is then

$$\theta_i + \gamma \sum_{j \neq i} \theta_j - \gamma \left( \sum_{j \neq i} \theta_j' \right) - \theta_i' = 0,$$

so he would wish to set

$$\theta_i' = \theta_i + \gamma \sum_{j \neq i} \left( \theta_j - \theta_j' \right) . \tag{4}$$

In other words, his best response to a misreport $\theta_{-i}'$ by the other agents is to report a type so that the aggregate type from his point of view is exactly identical to the true aggregate type generated by the true type profile $\theta$. Note that the above calculation also verifies the strict ex post incentive compatibility of $f$. The quadratic payoff / linear best response nature of this problem means that we can characterize $\beta^k(\theta_i)$ restricting attention to such point conjectures. In particular, we have

$$\beta^k(\theta_i) = \left[\underline{\beta}^k(\theta_i), \overline{\beta}^k(\theta_i)\right],$$

where

$$\begin{aligned}\overline{\beta}^k(\theta_i) &= \min\left\{1, \theta_i + \gamma \max_{\left\{\left(\theta'_{-i}, \theta_{-i}\right): \theta'_j \in \beta^k(\theta_j) \text{ for all } j \neq i\right\}} \sum_{j \neq i} \left(\theta_j - \theta'_j\right)\right\} \\ &= \min\left\{1, \theta_i + \gamma \max_{\theta_{-i}} \sum_{j \neq i} \left(\theta_j - \underline{\beta}^{k-1}(\theta_j)\right)\right\}.\end{aligned}$$

Analogously,

$$\underline{\beta}^k(\theta_i) = \max\left\{0, \theta_i - \gamma \max_{\theta_{-i}} \sum_{j \neq i} \left(\overline{\beta}^{k-1}(\theta_j) - \theta_j\right)\right\}.$$

Thus

$$\overline{\beta}^k(\theta_i) = \min\left\{1, \theta_i + (\gamma(I-1))^k\right\},$$

and

$$\underline{\beta}^k(\theta_i) = \max\left\{0, \theta_i - (\gamma(I-1))^k\right\}.$$

Thus $\theta'_i \neq \theta_i \Rightarrow \theta'_i \notin \beta^k(\theta_i)$ for sufficiently large $k$, provided that $\gamma < \frac{1}{I-1}$.

Now consider what happens when this condition fails, i.e., $\gamma > \frac{1}{I-1}$. In this case, it is possible to exploit the large amount of interdependence to construct beliefs over the opponents' types such that all types are indistinguishable. In particular, suppose that every type $\theta_i \in [0,1]$ has a degenerate belief over the types of his opponents. In particular, type $\theta_i$ is convinced that each of his opponents is of type $\theta_j$ given by:

$$\theta_j = \frac{1}{2} + \frac{1}{\gamma(I-1)}\left(\frac{1}{2} - \theta_i\right),$$

where the belief of $i$ about $j$ evidently depends on his type $\theta_i$. In this case the aggregated type profile is given by

$$\theta_i + \gamma \sum_{j \neq i} \theta_j = \frac{1}{2}(1 + \gamma(I-1)),$$

independent of $\theta_i$. Thus in *any* mechanism, for each type, we can construct beliefs so that there will be no differences across types of agent $i$ in terms of the actions which get deleted at each round of the process.

At the end of the paper we shall present an additional example, namely a single unit auction with symmetric bidders. The generalized Vickrey-Groves-Clark mechanism for the single unit auction only satisfies weak rather than strict incentive compatibility constraints. We therefore propose an $\varepsilon$-efficient allocation rule with strict ex post incentive constraints. The $\varepsilon$ efficient allocation rule can also be interpreted as the virtual implementation of the efficient rule. This rule can be robustly implemented if there is not too much interdependence among the payoff types.

# 4 Robust Implementation

## 4.1 Strict Single Crossing Environment

The following strict version of the standard single crossing property is the key economic assumption that we make about the environment in this paper:

**Definition 2 (Strict Single Crossing)**
*The environment satisfies strict single crossing (SSC) if* $v_i(y,\phi) > v_i(y',\phi)$ *and* $v_i\left(y,\phi'\right) = v_i\left(y',\phi'\right)$ *implies* $v_i\left(y,\phi''\right) < v_i\left(y',\phi''\right)$ *if either* $\phi < \phi' < \phi''$ *or* $\phi > \phi' > \phi''$*.*

The property is defined relative to the aggregation of all agents' types. So it is the combination of monotonic aggregator representation of preferences with the strict single crossing condition that drives our results. The public good model in the previous section satisfies the property and so will many environments with interdependent preferences that have been studied in the literature.

How strong is this restriction on the environment? It requires that the payoff types of the players can be aggregated into a variable that changes preferences in a monotonic way. To get some sense of the strength of this restriction, we next consider two examples. The first example involves a binary outcome space which naturally guarantees the aggregation property; the second example uses an informational foundation by means of Bayes' law to obtain the aggregation property.

In a quasi-linear environment one of two allocations, $a$ or $b$, must be chosen. The outcome space can be written as $Y = [0,1] \times [-K,K]^I$, where $y_0$ is the probability of allocation $a$ (and $1-y_0$ is the probability of allocation $b$) and $y_i$ is the transfer to individual $i$. Now if $v_i^z(\theta)$ is $i$'s utility from allocation $z$ when the type profile is $\theta$, we have

$$u_i(y,\theta) = y_0 v_i^a(\theta) + (1-y_0)\, v_i^b(\theta) + y_i.$$

An equivalent representation is

$$u_i(y,\theta) = y_0\left[v_i^a(\theta) - v_i^b(\theta)\right] + y_i.$$

Clearly, we can give this a monotonic aggregator representation by setting $h_i(\theta) = v_i^a(\theta) - v_i^b(\theta)$ and $v_i(y,h(\theta)) = y_0 h(\theta) + y_i$, we have

$$u_i(y,\theta) = v_i(y,h_i(\theta)),$$

and now $v_i$ indeed satisfies the strict single crossing condition. So with quasilinear utility, the binary allocation case automatically falls in our environment.[7] But when we move beyond two allocation, this would no longer necessarily be true. For example, if player $i$'s signal was more relevant for ranking one pair of outcomes rather than another, then the aggregation property could fail.

A natural source of interdependence in preferences is informational, when an agent's payoff type corresponds to a signal which ends up being correlated with all agents' expected values of a state. In particular, suppose that each player's utility depends on the expected value of an additive random variable $\omega_0 + \omega_i$, where $\omega_0$ is a common value component and $\omega_i$ is the private value component. The random variables $\omega_0, \omega_1, \omega_2$ are assumed to independently and normally distributed with zero mean and variance $\sigma_i^2$. Let each agent observe one signal $\theta_i = \omega_0 + \omega_i + \varepsilon_i$, where each $\varepsilon_i$ is independently normally distributed with mean 0 and variance $\tau_i^2$. We are thus assuming that each agent observes only a one dimensional signal, $\theta_i$, of both the common and idiosyncratic component. Thus agent $i$ is unable to distinguish with his noisy signal $\theta_i$ between the common and the private value components. But naturally his own signal is more informative about his valuation than the others' signals because it contains his own idiosyncratic shock.

Now standard properties of the normal distribution (see DeGroot (1970)) imply that agent $i$'s expected value of $\omega_0 + \omega_i$, given the vector of signals $(\theta_i, \theta_j)$ is a constant

$$\frac{\sigma_0^2\tau_i^2 + \sigma_0^2\tau_j^2 + \sigma_0^2\sigma_i^2 + \sigma_0^2\sigma_j^2 + \tau_i^2\tau_j^2 + \tau_i^2\sigma_j^2 + \tau_j^2\sigma_i^2 + \sigma_i^2\sigma_j^2}{\sigma_0^2\tau_i^2 + \sigma_0^2\sigma_i^2 + \sigma_0^2\sigma_j^2 + \tau_j^2\sigma_i^2 + \sigma_i^2\sigma_j^2}$$

times

$$h_i(\theta) = \theta_i + \frac{\sigma_0^2\tau_i^2}{\sigma_0^2\tau_j^2 + \sigma_0^2\sigma_i^2 + \sigma_0^2\sigma_j^2 + \tau_j^2\sigma_i^2 + \sigma_i^2\sigma_j^2}\theta_j. \tag{5}$$

The calculations are reported in the appendix. Now if we assume each agent $i$'s preferences conditional on $h_i(\theta)$ satisfy strict single crossing with respect to $h_i(\theta)$, then we have an informational microfoundation for the strict single crossing environment of the paper. Moreover, in this example the aggregator takes the symmetric linear form:

$$h_i(\theta) = \theta_i + \sum_{j \neq i} \gamma_{ij}\theta_j$$

with

$$\gamma_{ij} = \frac{\sigma_0^2\tau_i^2}{\sigma_0^2\tau_j^2 + \sigma_0^2\sigma_i^2 + \sigma_0^2\sigma_j^2 + \tau_j^2\sigma_i^2 + \sigma_i^2\sigma_j^2}.$$

[7] A similar logic applies if there are two allocations and no transfers. Thus the voting example in Palfrey and Srivastava (1989) fits our framework: since the contraction property fails, robust implementation is not possible in any mechanism.

This conclusion is quite intuitive. If the variance of the common component ($\sigma_0^2$) is small or if the noise in own's own signal ($\tau_i^2$) is small, then the interdependence goes away. But a reduction in variance of one's own idiosyncratic component ($\sigma_i^2$), in one's opponent's idiosyncratic component ($\sigma_j^2$) or in one's opponent's noise ($\tau_j^2$) all tend to increase the interdependence.[8]

With this interpretation the single crossing property with respect to the aggregator reduces to assuming that there is a one dimensional parameter whose expected value effects the preferences and that there is a sufficient statistic for the vectors of signals that agents observe.[9]

## 4.2 Main Positive Result

Before we state our first positive result, we introduce the incentive compatibility condition and the contraction property as they appear in the necessary and sufficient condition for robust implementation. The standard condition for truthful implementation is:

**Definition 3 (Ex Post Incentive Compatibility)**
*The social choice function $f$ satisfies ex post incentive compatibility (EPIC) if*

$$u_i\left(f\left(\theta_i,\theta_{-i}\right),\left(\theta_i,\theta_{-i}\right)\right) \geq u_i\left(f\left(\theta_i',\theta_{-i}\right),\left(\theta_i,\theta_{-i}\right)\right)$$

*for all $i$, $\theta$ and $\theta_i'$.*

In the subsequent analysis we use the strict version of the incentive constraints.

**Definition 4 (Strict Ex Post Incentive Compatibility)**
*The social choice function $f$ satisfies strict ex post incentive compatibility (strict EPIC) if*

$$u_i\left(f\left(\theta_i,\theta_{-i}\right),\left(\theta_i,\theta_{-i}\right)\right) > u_i\left(f\left(\theta_i',\theta_{-i}\right),\left(\theta_i,\theta_{-i}\right)\right)$$

[8] The additive model with a private and a common component also appears in Hong and Shum (2003) to describe the valuation of each bidder in an ascending single unit auction. Interestingly, they prove the existence and uniqueness of an increasing bidding strategy by appealing to a dominant diagonal condition, which is implied by the contraction proprty to be defined shortly.

The example of a normal distribution fails the compact type space assumoption of our model, but we use the normal distribution here merely for its transparent updating properties.

[9] There is a possible criticism of using an informational justification for interdependent preferences like this one at the same time as insisting on a stringent robust implementation criterion, as suggested by Ilya Segal. This informational microfoundation for the environment depends on the common knowledge of the distribution of signals about the environment - among the agents and the planner. Thus there is common knowledge of a true distribution over the vectors of signals $\theta$. However, we can show that if we allowed that each agent $i$ might receive additional, conditionally independent information - not necessarily consistent with a common prior - about others' signals $\theta_{-i}$, so that the information did not change his expectation of $\zeta_i$ conditional on the vector $\theta$, then our robust implementation results would remain unchanged. Thus there is an admittedly stark story that reconciles the robust implementation environment with an informational justification of the reduced form representation of interdependent preferences.

*for all $i$, $\theta$ and $\theta_i' \neq \theta_i$.*

The key property for our analysis is the following contraction property.

**Definition 5 (Strict Contraction Property)**
*The aggregator functions $h$ satisfy the strict contraction property if, for all $\beta \neq \beta^*$, there exists $i$ and $\theta_i' \in \beta_i(\theta_i)$ with $\theta_i' \neq \theta_i$, such that*

$$\text{sign}\left(\theta_i - \theta_i'\right) = \text{sign}\left(h_i(\theta_i, \theta_{-i}) - h_i\left(\theta_i', \theta_{-i}'\right)\right),$$

*for all $\theta_{-i}$ and $\theta_{-i}' \in \beta_{-i}(\theta_{-i})$.*

The strict contraction property essentially says that for some agent $i$ the direct impact of his private signal $\theta_i$ on the aggregator $h_i(\theta)$ is always sufficiently strong such that the difference in the aggregated value between the true type profile and the reported type profile always has the same sign as the difference between the true and reported type of agent $i$ by itself. A slightly weaker version of the contraction property will emerge as a necessary condition for robust implementation.

**Definition 6 (Contraction Property)**
*The aggregator functions $h$ satisfy the contraction property if, for all $\beta \neq \beta^*$, there exists $i$ and $\theta_i' \in \beta_i(\theta_i)$ with $\theta_i' \neq \theta_i$, such that either*

$$h_i(\theta_i, \theta_{-i}) = h_i\left(\theta_i', \theta_{-i}'\right),$$

*or*

$$\text{sign}\left(\theta_i - \theta_i'\right) = \text{sign}\left(h_i(\theta_i, \theta_{-i}) - h_i\left(\theta_i', \theta_{-i}'\right)\right),$$

*for all $\theta_{-i}$ and $\theta_{-i}' \in \beta_{-i}(\theta_{-i})$.*

The strict contraction property is a very slight strengthening of the contraction property.

**Theorem 1 (Robust Implementation)**
*If strict EPIC and the strict contraction property are satisfied, then there is robust implementation in the direct mechanism.*

**Proof.** We argue by contradiction. Let $\beta = S^{\mathcal{M}}$ and suppose that $\beta \neq \beta^*$. By the strict contraction property, there exists $i$ and $\theta_i' \in \beta_i(\theta_i)$ such that

$$\text{sign}\left(\theta_i - \theta_i'\right) = \text{sign}\left(h_i(\theta_i, \theta_{-i}) - h_i\left(\theta_i', \theta_{-i}'\right)\right).$$

for all $\theta_{-i}$ and $\theta'_{-i} \in \beta_{-i}(\theta_{-i})$. Let

$$\delta \triangleq \min_{\theta_{-i},\theta'_{-i}\in\beta_{-i}(\theta_{-i})} \left|h_i(\theta_i,\theta_{-i}) - h_i\left(\theta'_i,\theta'_{-i}\right)\right|,$$

where $\delta > 0$ by the strict contraction property. Suppose (without loss of generality) that $\theta_i > \theta'_i$. Let

$$\xi(\varepsilon) \triangleq \max_{\theta'_{-i}} \left\{h_i\left(\theta'_i+\varepsilon,\theta'_{-i}\right) - h_i\left(\theta'_i,\theta'_{-i}\right)\right\}.$$

As $h_i(\cdot)$ is strictly increasing in $\theta_i$, we know that $\xi(\varepsilon)$ is increasing in $\varepsilon$ and by continuity of $h_i$ in $\theta_i$, $\xi(\varepsilon) \to 0$ as $\varepsilon \to 0$.

Thus we have

$$h_i(\theta_i,\theta_{-i}) - h_i\left(\theta'_i,\theta'_{-i}\right) \geq \delta, \tag{6}$$

for all $\theta_{-i}$ and $\theta'_{-i} \in \beta_i(\theta_{-i})$; and

$$h_i\left(\theta'_i,\theta'_{-i}\right) \geq h_i\left(\theta'_i+\varepsilon,\theta'_{-i}\right) - \xi(\varepsilon), \tag{7}$$

for all $\theta'_{-i}$. By strict EPIC,

$$v_i\left(f\left(\theta'_i,\theta'_{-i}\right), h_i\left(\theta'_i,\theta'_{-i}\right)\right) > v_i\left(f\left(\theta'_i+\varepsilon,\theta'_{-i}\right), h_i\left(\theta'_i,\theta'_{-i}\right)\right), \tag{8}$$

for all $\varepsilon > 0$ and

$$v_i\left(f\left(\theta'_i+\varepsilon,\theta'_{-i}\right), h_i\left(\theta'_i+\varepsilon,\theta'_{-i}\right)\right) > v_i\left(f\left(\theta'_i,\theta'_{-i}\right), h_i\left(\theta'_i+\varepsilon,\theta'_{-i}\right)\right), \tag{9}$$

for all $\varepsilon > 0$. Now continuity of $u_i$ with respect to $\theta$ implies that for each $\varepsilon > 0$ and $\theta'_{-i}$, there exists

$$\phi^*\left(\varepsilon,\theta'_{-i}\right) \leq h_i\left(\theta'_i+\varepsilon,\theta'_{-i}\right), \tag{10}$$

such that

$$v_i\left(f\left(\theta'_i,\theta'_{-i}\right), \phi^*\left(\varepsilon,\theta'_{-i}\right)\right) = v_i\left(f\left(\theta'_i+\varepsilon,\theta'_{-i}\right), \phi^*\left(\varepsilon,\theta'_{-i}\right)\right);$$

and SSC implies that

$$v_i\left(f\left(\theta'_i,\theta'_{-i}\right), \phi\right) < v_i\left(f\left(\theta'_i+\varepsilon,\theta'_{-i}\right), \phi\right),$$

for all $\phi > \phi^*\left(\varepsilon,\theta'_{-i}\right)$. Now fix any $\varepsilon$ with

$$\xi(\varepsilon) < \delta. \tag{11}$$

Now for all $\theta'_{-i} \in \beta_{-i}(\theta_{-i})$,

$$\begin{aligned}
h_i(\theta_i,\theta_{-i}) &\geq h_i\left(\theta'_i,\theta'_{-i}\right) + \delta, \text{ by (6)} \\
&\geq h_i\left(\theta'_i+\varepsilon,\theta'_{-i}\right) - \xi(\varepsilon) + \delta, \text{ by (7)} \\
&> h_i\left(\theta'_i+\varepsilon,\theta'_{-i}\right), \text{ by (11)} \\
&\geq \phi^*\left(\varepsilon,\theta'_{-i}\right), \text{ by (10).}
\end{aligned}$$

So

$$v_i\left(f\left(\theta_i'+\varepsilon,\theta_{-i}'\right),h_i\left(\theta_i,\theta_{-i}\right)\right)>v_i\left(f\left(\theta_i',\theta_{-i}'\right),h_i\left(\theta_i,\theta_{-i}\right)\right),$$

for every $\theta_{-i}$ and $\theta_{-i}'\in\beta_{-i}\left(\theta_{-i}\right)$. This contradicts our assumption that $\beta=S^{\mathcal{M}}$. ■

The surprising element in this result is that we do not need to impose any conditions on how the social choice function varies with the type profile. In particular, it does not have to respond to the reported profile $\theta$ in a manner similar to the response of any of the aggregators $h_i$. Merely, the strong single crossing condition is sufficient to make full use of the contraction property. In contrast to the classic results in Nash and Bayesian Nash implementation we do not have to impose a condition on the number of agents, such as $I>2$.

The argument is centered around the true type profile $\theta=\left(\theta_i,\theta_{-i}\right)$ and a reported profile $\theta'=\left(\theta_i',\theta_{-i}'\right)$. Without loss of generality we may assume that $\theta_i>\theta_i'$. We use the contraction property to establish a positive lower bound on the difference $h\left(\theta_i,\theta_{-i}\right)-h\left(\theta_i',\theta_{-i}'\right)$ for all $\theta_{-i}$ and $\theta_{-i}'\in\beta_{-i}\left(\theta_{-i}\right)$. With this positive lower bound, we then show that agent $i$ is strictly better off to move his misreport $\theta_i'$ marginally upwards in the direction of $\theta_i$, in other words to report $\theta_i'+\varepsilon$. This is achieved by showing that there is an intermediate value $\phi^*$ for the aggregator, with $h_i\left(\theta_i',\theta_{-i}'\right)<\phi^*<h_i\left(\theta_i'+\varepsilon,\theta_{-i}'\right)$, such that agent $i$ with the utility profile corresponding to the aggregator value $\phi^*$ would be indifferent between the social allocations $f\left(\theta_i',\theta_{-i}'\right)$ and $f\left(\theta_i'+\varepsilon,\theta_{-i}'\right)$. By choosing $\varepsilon$ sufficiently small, we know that $h\left(\theta_i,\theta_{-i}\right)>\phi^*$ and strict single crossing then allows us to assert that an agent with a true preference profile $\theta=\left(\theta_i,\theta_{-i}\right)$ would also prefer to obtain $f\left(\theta_i'+\varepsilon,\theta_{-i}'\right)$ rather than $f\left(\theta_i',\theta_{-i}'\right)$. But this yields the contradiction to $\theta_i'\in\beta_i\left(\theta_i\right)$ being part of the fixed point of the iterative elimination. Consequently we show that the misreport $\theta_i'$, which established the same sign on the difference between private type profiles and aggregated public profiles can be eliminated as a best response to the set of misreports of the remaining agents.

# 5 The Linear Model

In this section, we consider the special case in which the preference aggregator $h_i\left(\theta\right)$ is linear for each $i$ and given by:

$$h_i\left(\theta\right)=\sum_{j=1}^{I}\gamma_{ij}\theta_j,$$

with $\gamma_{ij}\in\mathbb{R}$ for all $i$, $j$ and $\gamma_{ii}>0$ for all $i$. Without loss of generality, we set $\gamma_{ii}=1$ for all $i$:

$$h_i\left(\theta\right)=\theta_i+\sum_{j\neq i}\gamma_{ij}\theta_j.$$

The parameters $\gamma_{ij}$ represent the influence of the signal of agent $j$ on the value of agent $i$. With the exception of $\gamma_{ii} > 0$ for all $i$, we do not impose any further a priori sign restrictions on $\gamma_{ij}$. We denote the square matrix generated by the absolute values of $\gamma_{ij}$, namely $\left|\gamma_{ij}\right|$, for all $i, j$ with $i \neq j$ and zero entries on the diagonal by $\Gamma$:

$$\Gamma \triangleq \begin{bmatrix} 0 & |\gamma_{12}| & \cdots & |\gamma_{1I}| \\ |\gamma_{21}| & 0 & & \\ \vdots & & \ddots & \\ |\gamma_{I1}| & & & 0 \end{bmatrix}.$$

We refer to the matrix $\Gamma$ as the *interdependence matrix*. The matrix $\Gamma = 0$ then constitutes the case of pure private values.

## 5.1 Contraction Property

We shall first give necessary and sufficient conditions for the matrix $\Gamma$ to satisfy the (strict) contraction property. We then use duality theory to give a dual characterization of the contraction property, which is very useful to finally obtain necessary and sufficient conditions for the contraction property in terms of the eigenvalue of the matrix $\Gamma$.

**Lemma 1 (Linear Aggregator)**
*Linear aggregator functions $h$ satisfy the strict contraction property if and only if, for all $c \in \mathbb{R}_+^I$ with $c \neq \mathbf{0}$, there exists $i$ such that*

$$c_i > \sum_{j \neq i} \left|\gamma_{ij}\right| c_j. \tag{12}$$

**Proof.** We proof the contrapositive. Thus suppose there exists $c \in \mathbb{R}_+^I$ with $c \neq \mathbf{0}$, such that for all $i$:

$$c_i \leq \sum_{j \neq i} \left|\gamma_{ij}\right| c_j.$$

We now show that this implies that the strict contraction property fails. Choose $\varepsilon > 0$ such that $2c_i\varepsilon < \overline{\theta}_i - \underline{\theta}_i$ for all $i$. Now consider deceptions of the form:

$$\beta_i(\theta_i) = [\theta_i - \varepsilon c_i, \theta_i + \varepsilon c_i] \cap \Theta_i, \tag{13}$$

for all $i$. Then for all $i$ and all $j \neq i$, let $\theta_j = \frac{1}{2}\left(\underline{\theta}_j + \overline{\theta}_j\right)$ and let $\theta_j' = \theta_j - \varepsilon c_i$ if $\gamma_{ij} \geq 0$ and $\theta_j' = \theta_j + \varepsilon c_i$ if $\gamma_{ij} < 0$. By (13), we have $\theta_j' \in \beta_j(\theta_j)$ for each $j \neq i$. Also observe that $\gamma_{ij}\left(\theta_j - \theta_j'\right) = \varepsilon \left|\gamma_{ij}\right| c_j$. Thus

$$\sum_{j \neq i} \gamma_{ij}\left(\theta_j - \theta_j'\right) = \varepsilon \sum_{j \neq i} \left|\gamma_{ij}\right| c_j \geq \varepsilon c_i.$$

Now if $\theta_i' = \theta_i + \varepsilon c_i$, $\theta_i - \theta_i'$ is strictly negative but

$$\theta_i - \theta_i' + \sum_{j \neq i} \gamma_{ij} \left(\theta_j - \theta_j'\right),$$

is non-negative. A symmetric argument works if $\theta_i > \theta_i'$. So the strict contraction property, which says that for all $\beta \neq \beta^*$, there exists $i$ and $\theta_i' \in \beta_i(\theta_i)$ with $\theta_i' \neq \theta_i$, such that

$$\text{sign}\left(\theta_i - \theta_i'\right) = \text{sign}\left(h_i\left(\theta_i, \theta_{-i}\right) - h_i\left(\theta_i', \theta_{-i}'\right)\right) = \text{sign}\left(\theta_i - \theta_i' + \sum_{j \neq i} \gamma_{ij}\left(\theta_j - \theta_j'\right)\right), \qquad (14)$$

for all $\theta_{-i}$ and $\theta_{-i}' \in \beta_{-i}(\theta_{-i})$ fails. This proves the necessity of condition (12) of Lemma 1.**??**.

($\Leftarrow$) To show sufficiency, suppose that condition (12) of the lemma holds. Fix any deception $\beta$. For all $j$, let:

$$c_j = \max_{\theta_j' \in \beta_j(\theta_j)} \left|\theta_j' - \theta_j\right|.$$

By hypothesis, there exists $i$ such that $c_i > \sum_{j \neq i} \left|\gamma_{ij}\right| c_j$. Let

$$\left|\theta_i - \theta_i'\right| = c_i,$$

and suppose without loss of generality that $\theta_i > \theta_i'$. Observe that for all $\theta_{-i}$ and $\theta_{-i}' \in \beta_{-i}(\theta_{-i})$, $\gamma_{ij}\left(\theta_j - \theta_j'\right) \leq \left|\gamma_{ij}\right| c_j$ and thus

$$\sum_{j \neq i} \gamma_{ij}\left(\theta_j - \theta_j'\right) \leq \sum_{j \neq i} \left|\gamma_{ij}\right| c_j;$$

so

$$\begin{aligned} \left(\theta_i - \theta_i'\right) - \sum_{j \neq i} \gamma_{ij}\left(\theta_j - \theta_j'\right) &= c_i - \sum_{j \neq i} \gamma_{ij}\left(\theta_j - \theta_j'\right) \\ &\geq c_i - \sum_{j \neq i} \left|\gamma_{ij}\right| c_j > 0, \end{aligned}$$

and hence the strict contraction property, or (14), is satisfied. ■

The corresponding characterization for the contraction property is given by replacing the strict inequality in (12) by a weak inequality. The absolute values of the matrix $\Gamma$ are required to guarantee that the linear inequality (12) implies the strict contraction property. We observe that the condition (12) is only required to hold for a single agent $i$. In fact, for $c \gg 0$, the condition (12) could hold for all $i$ only in the case of pure private values, or $\Gamma = \mathbf{0}$.

The proof of the contraction property is constructive. We identify for each player $i$ an initial deception of the form $\beta_i(\theta_i) = [\theta_i - c_i\varepsilon, \theta_i + c_i\varepsilon]$ for some $\varepsilon > 0$, common across all agents. The size of $c_i$ is therefore proportional to the size of the set of candidate reports by agent $i$. It can be thought of as the set of rationalizable strategies at an arbitrary stage $k$. The inequality of the contraction property then says that for any arbitrary set of deceptions, characterized by the vector $c$, there is always an agent $i$ whose set of deceptions is too large (in the sense of being rationalizable) relative to the set of deceptions by the remaining agents. It then follows that the set of deceptions for this agent can be chosen smaller than $c_i$, allowing us to reduce the set of possible reports for a given agent $i$ with a given type $\theta_i$. The inequality (12) asserts that for any given set of deceptions, there is always at least one agent $i$ whose deception $\beta_i$ represents a set too large to be rationalizable. Moreover, if the set of deceptions by $i$ is too large, then there is an "overhang" which can be "nipped and tucked". In the appendix, we report a dual interpretation of the condition (12) which leads us from the idea of the overhang directly to the contraction property. We use this dual interpretation to derive the following simple test of the contraction property:

**Theorem 2 (Contraction Property via Eigenvalue)**
*The matrix $\Gamma$ has the contraction property if and only if its largest eigenvalue $\lambda < 1$.*

**Proof.** See appendix. ∎

### 5.2 Examples

By linking the contraction property to the eigenvalue of the matrix $\Gamma$, we can immediately obtain necessary and sufficient condition for robust implementation for different classes of preference environments.

**Symmetric Preferences** In the symmetric model, the parameters for interdependent values are given by

$$\gamma_{ij} = \begin{cases} 1, & \text{if} \quad j = i, \\ \gamma, & \text{if} \quad j \neq i. \end{cases}$$

The eigenvalue $\lambda$ of the resulting matrix satisfies:

$$1 + \lambda = 1 + \gamma (I - 1),$$

and hence from Theorem 2, we immediately obtain the necessary and sufficient condition:

$$\gamma < \frac{1}{I - 1}.$$

**Cyclic Preferences** A weaker form of symmetry is incorporated in the following model of cyclic preferences. Here, the interdependence matrix is determined by the distance between $i$ and $j$ (modulo $I$), or

$$\gamma_{ij} = \gamma_{(i-j)_{\text{mod } I}}.$$

In this case, the positive eigenvalue is given by:

$$1 + \lambda = 1 + \sum_{j \neq i} \gamma_{(i-j)},$$

and consequently a necessary and sufficient condition for robust implementation is given by:

$$\sum_{j \neq i} \gamma_{(i-j)} < 1.$$

**Two Bidders** With two bidders, the matrix of interdependence, $\Gamma$, is given by

$$\Gamma = \begin{bmatrix} 1 & \gamma_{12} \\ \gamma_{21} & 1 \end{bmatrix}.$$

The eigenvalue of the matrix $\Gamma$ can again be immediately computed by requiring that

$$1 + \lambda = 1 + \sqrt{\gamma_{12}\gamma_{21}},$$

or

$$\gamma_{12}\gamma_{21} < 1.$$

**Central Bidder** Finally, we may consider a model in which each bidder only cares about his own type and the type of bidder 1, the central or informed bidder. The matrix of interdependence is then given by

$$\gamma_{ij} = \begin{cases} 1 & \text{if} \quad j = i, \\ \gamma & \text{if} \quad i \neq 1 \text{ and } j = 1, \\ 0 & \text{if} \quad \text{otherwise.} \end{cases}$$

In this case, the eigenvalue is given by:

$$1 + \lambda = 1 + 0,$$

and hence robust monotonicity holds vacuously for all $\gamma$, independent of $I$. The intuition in this case is that bidder 1 has a private value utility model. In conjunction with the strict ex post incentive constraints, this essentially means that agent 1 will always have a strict incentive to tell the truth. But as the utility of all the other agents depends only their own utility and the utility of agent 1, and agent 1 is known to tell the truth, all other agents will also want to report truthfully.

The linear model has the obvious advantage that the local conditions for contraction agree with the global conditions for contraction as the derivatives of the mapping $h_i(\theta)$ are constant and independent of $\theta$. In the appendix, we extend the idea behind the linear aggregator function to a general nonlinear and differentiable aggregator function $h_i(\theta)$, but with a gap between necessary and sufficient conditions.

# 6 Necessity of Contraction Property

The contraction property appears to be a natural condition in the context of robust implementation. In fact, we now show that the contraction property is necessary for robust implementation. In particular, the necessity of the contraction property allows us to give a sharp impossibility result in the context of the linear model just discussed. The idea behind the necessity argument is to show that the hypothesis of robust implementation leads inevitably to a conflict with a deception profile $\beta$ which fails to satisfy the contraction property.

**Theorem 3 (Necessity)**
*If $f$ is robustly implementable, then $f$ satisfies strict EPIC and the contraction property.*

**Proof.** The restriction to compact mechanisms ensures that $S^{\mathcal{M}}$ is non-empty. It follows that if mechanism $\mathcal{M}$ robustly implements $f$, then, for each $i$, there exists $m_i^*: \Theta_i \to M_i$ such that

$$g(m^*(\theta)) = f(\theta) \text{ and } m^*(\theta) \in S^{\mathcal{M}}(\theta),$$

we can simply let $m_i^*(\theta_i)$ be any element of $S_i^{\mathcal{M}}(\theta_i)$.

We first establish strict EPIC. Suppose strict EPIC fails. Then there exists $i$, $\theta$ and $\theta_i'$ such that $f(\theta_i', \theta_{-i}) \neq f(\theta_i, \theta_{-i})$ and

$$u_i\left(f\left(\theta_i', \theta_{-i}\right), \theta\right) \geq u_i(f(\theta), \theta).$$

Now, for any message $m_i$ with

$$m_i \in \underset{m_i'}{\arg\max}\, u_i\left(g\left(m_i', m_{-i}^*(\theta_{-i})\right), (\theta_i, \theta_{-i})\right),$$

since $m_{-i}^*(\theta_{-i}) \in S_i^{\infty}(\theta_{-i})$, we must have $m_i \in S_i^{\infty}(\theta_i)$ and thus $g\left(m_i, m_{-i}^*\left(\theta_{-i}'\right)\right) = f\left(\theta_i, \theta_{-i}'\right)$ for all $\theta_{-i}'$. Thus let

$$m_i^*\left(\theta_i'\right) \in \underset{m_i'}{\arg\max}\, u_i\left(g\left(m_i', m_{-i}^*(\theta_{-i})\right), (\theta_i, \theta_{-i})\right),$$

and $f\left(\theta_i', \theta_{-i}'\right) = g\left(m_i^*\left(\theta_i'\right), m_{-i}^*\left(\theta_{-i}'\right)\right) = f\left(\theta_i, \theta_{-i}'\right)$ for all $\theta_{-i}'$, a contradiction.

Now we establish the contraction property. The proof is by contradiction and we show that if $f$ is robustly implementable, then every $\beta$, with $\beta \neq \beta^*$, must satisfy the contraction property. To this end, suppose $\beta \neq \beta^*$ does not satisfy the contraction property. Then for all $i$, $\theta_i' \in \beta_i(\theta_i)$ with $\theta_i' \neq \theta_i$, there exists $\theta_{-i}$ and $\theta_{-i}' \in \beta_{-i}(\theta_{-i})$ such that:

$$\text{sign}\left(\theta_i - \theta_i'\right) = -\,\text{sign}\left(h_i(\theta_i, \theta_{-i}) - h_i\left(\theta_i', \theta_{-i}'\right)\right).$$

Thus for $\theta_i > \theta_i'$, there exists $\theta_{-i}$ and $\theta_{-i}' \in \beta_{-i}(\theta_{-i})$ such that

$$h_i\left(\theta_i, \theta_{-i}'\right) > h_i\left(\theta_i', \theta_{-i}'\right) > h_i(\theta_i, \theta_{-i}).$$

Now, by single crossing, if

$$v_i(y, h_i(\theta_i, \theta_{-i})) > v_i\left(f\left(\theta_i', \theta_{-i}'\right), h_i(\theta_i, \theta_{-i})\right) \tag{15}$$

and

$$v_i\left(y, h_i\left(\theta_i', \theta_{-i}'\right)\right) \leq v_i\left(f\left(\theta_i', \theta_{-i}'\right), h_i\left(\theta_i', \theta_{-i}'\right)\right),$$

then

$$v_i\left(y, h_i\left(\theta_i, \theta_{-i}'\right)\right) < v_i\left(f\left(\theta_i', \theta_{-i}'\right), h_i\left(\theta_i, \theta_{-i}'\right)\right). \tag{16}$$

In other words, there does not exist $y$ such that:

$$v_i\left(y, h_i\left(\theta_i', \theta_{-i}'\right)\right) \leq v_i\left(f\left(\theta_i', \theta_{-i}'\right), h_i\left(\theta_i', \theta_{-i}'\right)\right) \tag{17}$$

and

$$v_i\left(y, h_i\left(\theta_i, \widetilde{\theta}_{-i}\right)\right) > v_i\left(f\left(\theta_i', \theta_{-i}'\right), h_i\left(\theta_i, \widetilde{\theta}_{-i}\right)\right), \tag{18}$$

for all $\widetilde{\theta}_{-i}$ such that $\widetilde{\theta}_{-i} \in \beta_{-i}(\theta_{-i})$, as by hypothesis, $\theta_{-i}, \theta_{-i}' \in \beta_{-i}(\theta_{-i})$ and hence both (15) and (16) apply.

We now show that (15) and (16) are in conflict with the hypothesis of robust implementation. Consider an arbitrary deception $\beta \neq \beta^*$. Let $\widehat{k}$ be the largest $k$ such that for every $i$, $\theta_i$ and $\theta_i' \in \beta_i(\theta_i)$:

$$S_i^\infty\left(\theta_i'\right) \subseteq S_i^k(\theta_i).$$

We know that such a $\widehat{k}$ exists because $S_i^0(\theta_i) \cap S_i^\infty\left(\theta_i'\right) = S_i^\infty\left(\theta_i'\right)$ ,and since $\mathcal{M}$ robustly implements $f$, we must have $S_i^\infty(\theta_i) \cap S_i^\infty\left(\theta_i'\right) = \varnothing$.

Now we know that there exists $i$ and $\theta_i' \in \beta_i(\theta_i)$ such that

$$S_i^{\widehat{k}+1}(\theta_i) \cap S_i^\infty\left(\theta_i'\right) \neq S_i^\infty\left(\theta_i'\right).$$

Let

$$\widehat{m}_i \in S_i^{\widehat{k}}(\theta_i) \cap S_i^{\infty}(\theta_i'),$$

and

$$\widehat{m}_i \notin S_i^{\widehat{k}+1}(\theta_i) \cap S_i^{\infty}(\theta_i').$$

Since message $\widehat{m}_i$ gets deleted for $\theta_i$ at round $\widehat{k}+1$, we know that for every $\lambda_i \in \Delta(M_{-i} \times \Theta_{-i})$ such that

$$\lambda_i(m_{-i}, \theta_{-i}) > 0 \Rightarrow m_j \in S_j^{\widehat{k}}(\theta_j) \text{ for all } j \neq i,$$

there exists $m_i^*$ such that

$$\sum_{m_{-i},\theta_{-i}} \lambda_i(m_{-i}, \theta_{-i}) u_i(g(m_i^*, m_{-i}), (\theta_i, \theta_{-i})) > \sum_{m_{-i},\theta_{-i}} \lambda_i(m_{-i}, \theta_{-i}) u_i(g(\widehat{m}_i, m_{-i}), (\theta_i, \theta_{-i})).$$

Let

$$\widehat{m}_j \in S_j^{\infty}(\theta_j')$$

for all $j \neq i$. Now the above claim remains true if we restrict attention to distributions $\lambda_i$ putting probability 1 on $\widehat{m}_{-i}$. Thus for every $\psi_i \in \Delta(\Theta_{-i})$ such that

$$\psi_i(\theta_{-i}) > 0 \Rightarrow \widehat{m}_j \in S_j^{\widehat{k}}(\theta_j) \text{ for all } j \neq i,$$

there exists $m_i^*$ such that

$$\sum_{\theta_{-i}} \psi_i(\theta_{-i}) u_i(g(m_i^*, \widehat{m}_{-i}), (\theta_i, \theta_{-i})) > \sum_{\theta_{-i}} \psi_i(\theta_{-i}) u_i(g(\widehat{m}_i, \widehat{m}_{-i}), (\theta_i, \theta_{-i})).$$

But $\widehat{m} \in S^{\infty}(\theta')$, so (since $\mathcal{M}$ robustly implements $f$), $g(\widehat{m}_i, \widehat{m}_{-i}) = f(\theta')$. Also observe that if $\theta_{-i}' \in \beta_{-i}(\theta_{-i})$, then $\widehat{m}_{-i} \in S_{-i}^{\widehat{k}}(\theta_{-i})$. Thus for every $\psi_i \in \Delta\left(\beta_{-i}^{-1}(\theta_{-i}')\right)$, there exists $m_i^*$ such that

$$\sum_{\theta_{-i}} \psi_i(\theta_{-i}) u_i(g(m_i^*, \widehat{m}_{-i}), (\theta_i, \theta_{-i})) > \sum_{\theta_{-i}} \psi_i(\theta_{-i}) u_i\left(f(\theta'), (\theta_i, \theta_{-i})\right). \tag{19}$$

We also observe that the message profile $m_i^*$ (and the associated allocation $g(m_i^*, \widehat{m}_{-i})$) which eliminates $\widehat{\theta}_i$ must be strictly dominated by the social choice function at the type profiles $\left(\widetilde{\theta}_i, \theta_{-i}'\right)$ for all $\widetilde{\theta}_i \in \Theta_i$, or

$$u_i\left(f\left(\widetilde{\theta}_i, \theta_{-i}'\right), \left(\widetilde{\theta}_i, \theta_{-i}'\right)\right) > u_i\left(g(m_i^*, \widehat{m}_{-i}), \left(\widetilde{\theta}_i, \theta_{-i}'\right)\right), \ \forall \widetilde{\theta}_i \in \Theta_i. \tag{20}$$

The argument here is by contradiction. Suppose the ex post incentive inequalities, (20), are not satisfied strictly, and hence:

$$u_i\left(g(m_i^*, \widehat{m}_{-i}), \left(\widetilde{\theta}_i, \theta_{-i}'\right)\right) \geq u_i\left(f\left(\widetilde{\theta}_i, \theta_{-i}'\right), \left(\widetilde{\theta}_i, \theta_{-i}'\right)\right),$$

for some $\widetilde{\theta}_i \in \Theta_i$. Now, for any

$$m_i \in \arg\max_{m_i'} u_i \left( g\left(m_i', \widehat{m}_{-i}\right), \left(\widetilde{\theta}_i, \theta_{-i}'\right)\right), \tag{21}$$

since $\widehat{m}_{-i} \in S_i^{\infty}\left(\theta_{-i}'\right)$, we must have $m_i \in S_i^{\infty}\left(\widetilde{\theta}_i\right)$ and thus $g\left(m_i, \widehat{m}_{-i}\right) = f\left(\theta_i, \theta_{-i}'\right)$. Thus from (21) we also know that $m_i^*$ achieves the maximum:

$$m_i^* \in \arg\max_{m_i'} u_i \left( g\left(m_i', \widehat{m}_{-i}\right), \left(\widetilde{\theta}_i, \theta_{-i}'\right)\right)$$

and, for all $\widetilde{\theta}_i$, if

$$u_i\left(g\left(m_i^*, \widehat{m}_{-i}\right), \left(\widetilde{\theta}_i, \theta_{-i}'\right)\right) \geq u_i\left(f\left(\widetilde{\theta}_i, \theta_{-i}'\right), \left(\widetilde{\theta}_i, \theta_{-i}'\right)\right),$$

then $g\left(m_i^*, \widehat{m}_{-i}\right) = f\left(\widetilde{\theta}_i, \theta_{-i}'\right)$.

Now setting $y \equiv g\left(m_i^*, \widehat{m}_{-i}\right)$, we have established that for each $\theta_{-i}' \in \beta_{-i}\left(\theta_{-i}\right)$ and $\psi_i \in \Delta\left(\beta_{-i}^{-1}\left(\theta_{-i}'\right)\right)$, there exists $y$ such that

$$u_i\left(f\left(\widetilde{\theta}_i, \theta_{-i}'\right), \left(\widetilde{\theta}_i, \theta_{-i}'\right)\right) > u_i\left(y, \left(\widetilde{\theta}_i, \theta_{-i}'\right)\right), \ \forall \widetilde{\theta}_i \in \Theta_i, \tag{22}$$

and

$$\sum_{\theta_{-i}} \psi_i\left(\theta_{-i}\right) u_i\left(y, \left(\theta_i, \theta_{-i}\right)\right) > \sum_{\theta_{-i}} \psi_i\left(\theta_{-i}\right) u_i\left(f\left(\theta'\right), \left(\theta_i, \theta_{-i}\right)\right). \tag{23}$$

But now we arrive at the desired contradiction as the inequalities (22)-(23), coming from robust implementation, and the inequalities (17)-(18), coming from the failure of the contraction property cannot be true simultaneously. A symmetric argument works if $\theta_i < \theta_i'$. ■

We briefly sketch the idea of the necessity part of the proof. The proof is by contradiction. We start with the hypothesis of robust implementation and consider a deception $\beta$ for which the contraction property fails to be satisfied. The failure of the contraction property in the single crossing environment is then shown to imply that for all $i$, $\theta_i' \in \beta_i\left(\theta_i\right)$ and some $\theta_{-i}, \theta_{-i}'$ with $\theta_i' \in \beta_i\left(\theta_i\right)$ we *cannot* find an allocation $y$ such that it weakly dominated if $\theta'$ is the true type profile, or

$$u_i\left(f\left(\theta'\right), \theta'\right) \geq u_i\left(y, f\left(\theta'\right)\right), \tag{24}$$

but is preferred over $f\left(\theta'\right)$ if agent $i$ is of type $\theta_i$:

$$u_i\left(y, \left(\theta_i, \widetilde{\theta}_{-i}\right)\right) > u_i\left(f\left(\theta'\right), \left(\theta_i, \widetilde{\theta}_{-i}\right)\right), \tag{25}$$

for all $\widetilde{\theta}_{-i} \in \beta_{-i}\left(\theta_{-i}\right)$.

Yet, if we start the iterative process, then for $f$ to be robustly implementable, it must be that there is a first time, denoted by stage $\widehat{k}$, such that for some $i$ and some $\theta_i$:

$$\theta'_i \in S_i^{\widehat{k}}(\theta_i)$$

but then in stage $\widehat{k}+1$:

$$\theta'_i \notin S_i^{\widehat{k}+1}(\theta_i).$$

The question then arises what is a necessary condition for $\theta'_i$ to be eliminated if all misreports feasible by $\beta(\theta)$ are still possible candidate strategies at stage $\widehat{k}$. The answer is simply that there must be some allocation $y$, induced by a report $\theta_i^*$ of agent $i$, or $y = f\left(\theta_i^*, \theta'_{-i}\right)$ such that

$$u_i\left(y, \left(\theta_i, \widetilde{\theta}_{-i}\right)\right) > u_i\left(f\left(\theta'\right), \left(\theta_i, \widetilde{\theta}_{-i}\right)\right),$$

but of course $f\left(\theta_i^*, \theta'_{-i}\right)$ would still have to satisfy incentive compatibility relative to a truthful report of $\theta'_i$ (if and when the true type is indeed $\theta'_i$), or

$$u_i\left(f\left(\theta'\right), \theta'\right) \geq u_i\left(y, f\left(\theta'\right)\right).$$

But now we observe that the necessary condition for $\theta'_i$ to be eliminated from $S_i^k(\theta_i)$ in stage $\widehat{k}+1$, is precisely the condition which fails to hold if the contraction property fails to hold, leading to the desired conclusion.

The above inequalities, (24) and (25) in fact describe a condition, termed robust monotonicity, in Bergemann and Morris (2005b). There it shown that robust monotonicity is a necessary and almost sufficient condition if we want to guarantee Bayesian equilibrium implementation for all possible priors. In Bergemann and Morris (2005b), the notion of Bayesian equilibrium implementation for all possible priors allows the use of complicated augmented mechanism. In contrast, here we focus on robust implementation in the direct mechanism, yet as the argument shows the robust monotonicity condition emerges again as necessary condition for implementation.

For the linear model discussed in the previous section, with

$$h_i(\theta) = \sum_j \gamma_{ij}\theta_j,$$

we have an impossibility result as an immediate consequence of Theorem 3.

**Corollary 1 (Impossibility of Robust Implementation)**
*If the contraction property fails, i.e. there exists $c \in \mathbb{R}_+ \setminus \{0\}$ such that for all $i$:*

$$c_i < \sum_{j \neq i} \left|\gamma_{ij}\right| c_j,$$

*then robust implementation fails.*

# 7 Single Unit Auction

We conclude our analysis with a second example, namely a single unit auction with symmetric bidders. The model has $I$ agents and agent $i$'s payoff type is $\theta_i \in [0,1]$. If the type profile is $\theta$, agent $i$'s valuation of the object is

$$\theta_i + \gamma \sum_{j \neq i} \theta_j, \tag{26}$$

where $0 \leq \gamma \leq 1$.

An allocation rule in this context is a function $q : \Theta \to [0,1]^I$, where $q_i(\theta)$ is the probability that agent $i$ gets the object and so $\sum_i q_i(\theta) \leq 1$. The symmetric efficient allocation rule is given by:

$$q_i^*(\theta) = \begin{cases} \frac{1}{\#\{j:\theta_j \geq \theta_k \text{ for all } k\}}, & \text{if } \quad \theta_i \geq \theta_k \text{ for all } k, \\ 0, & \text{if } \quad \text{otherwise.} \end{cases}$$

Maskin (1992) and Cremer and McLean (1985) have shown that the efficient allocation can be truthfully implemented in a generalized Vickrey-Clark-Groves mechanism, according to which the monetary transfer of the winning agent $i$ is given by

$$y_i(\theta) = \max_{j \neq i} \theta_j + \gamma \sum_{j \neq i} \theta_j.$$

We observe that the winning probability $q_i(\theta)$ and the monetary transfer are piecewise constant. The generalized VCG mechanism therefore does not satisfy the strict ex post incentive compatibility conditions which we assumed as part of our analysis. We therefore modify the generalized VCG mechanism to a symmetric $\varepsilon$-efficient allocation rule given by:

$$q_i^{**}(\theta) = \varepsilon \frac{\theta_i}{I} + (1-\varepsilon) q_i^*(\theta).$$

Under this allocation rule, the object is not allocated with probability $\frac{\varepsilon}{2}$.[10] We then argue that the symmetric $\varepsilon-$efficient allocation rule can be robustly implemented if $\gamma < \frac{1}{I-1}$. Alternatively, we can say that the generalized VCG mechanism itself is virtually and robustly implementable if $\gamma < \frac{1}{I-1}$.

It is easy to verify that the resulting generalized VCG transfers satisfy strict ex post incentive compatibility and show that this $\varepsilon$-efficient allocation is robustly implementable. The corresponding

[10] At the cost of some additional algebra, we could modify the allocation rule $q_i^{**}(\theta)$ to be:

$$\varepsilon \frac{\theta_i}{\sum_j \theta_j} + (1-\varepsilon) q_i^*(\theta),$$

which in turn allocates the object with probability 1.

essentially unique ex post transfer rule is:

$$y_i(\theta) = \frac{\varepsilon}{2I}(\theta_i)^2 + \frac{\varepsilon\gamma}{I}\left(\sum_{j\neq i}\theta_j\right)\theta_i + (1-\varepsilon)\left(\max_{j\neq i}\left\{\theta_j + \gamma\sum_{j\neq i}\theta_j\right\}\right)q_i^*(\theta).$$

The first two components of the transfers guarantee incentive compatibility with the respect to the linear probability assignment and the third component with respect to the efficient allocation rule. The best response of agent $i$ for misreport $\theta'_{-i}$ of the remaining agents at a true type profile $\theta$ is given as the public good example by:

$$\theta'_i = \theta_i + \gamma\sum_{j\neq i}\left(\theta_j - \theta'_j\right).$$

We can therefore exactly repeat our earlier argument in the context of the public good and get robust implementation in the direct mechanism if $\gamma < \frac{1}{I-1}$.

The implementation conditions here are substantially different from the average crossing condition of Krishna (2003), the generalized single crossing conditions of Birulin and Izmalkov (2003) and the dominant effect property identified by Echenique and Manelli (2004). The average crossing condition provides sufficient while the generalized single crossing conditions provide necessary and sufficient conditions for the existence of an efficient equilibrium in an English auction for a single object.[11] In our linear and symmetric environment, see (26), their necessary and sufficient condition reduces to the condition of $\gamma < 1$, independent of the number of agents, $I$. Echenique and Manelli (2004) present a dominant effect property which guarantees the existence of an efficient ex post equilibrium without further continuity and differentiability conditions. However in our linear and symmetric environment, their condition again reduces to $\gamma < 1$, independent of the number of agents, $I$.

# 8 Discussion

## 8.1 Relation to Partial and Ex Post Implementation

The results in this paper concern full implementation. An earlier paper of ours, Bergemann and Morris (2005c), addresses the analogous questions of robustness to rich type spaces, but looking at the question of *truthtelling* in the *direct* mechanism. In the literature, this is frequently referred to as

[11] Their conditions generalize earlier results by Maskin (1992) for two bidders. The novel issue with many bidders is that a marginal change in the signal profile should favor one of the currently winning bidders over a currently loosing bidder. With many bidders, a pairwise condition comparing the effect of signal $\theta_i$ on $i$ and $j$ is not sufficient anymore.

*partial* implementation. The notion of partial implementation asks whether there exist a mechanism such that *some* equilibrium under that mechanism implements the social choice function. By the revelation principle, it is then sufficient to look at truthtelling in the direct mechanism. In Bergemann and Morris (2005c), we showed that a social choice function robustly satisfies the interim incentive constraints, i.e. satisfies the interim incentive constraints for any type space, if and only if the ex post incentive constraints are satisfied.

It is important to note, however, that robust implementation is not equivalent to full ex post implementation, i.e., the requirement that every ex post equilibrium delivers the right outcome. Often ex post implementation will be possible - because there are no bad ex post equilibria - even though there exist type spaces and interim equilibria that deliver bad outcomes. In Bergemann and Morris (2005a), we identify the ex post monotonicity that is necessary and sufficient for full ex post implementation. It is much weaker than robust monotonicity and the contraction property.

### 8.2 Robust and Virtual Implementation in General Environments

The existing Bayesian implementation literature - Postlewaite and Schmeidler (1986), Palfrey and Srivastava (1989) and Jackson (1991) - has shown that on a fixed type space with a common knowledge common prior, Bayesian incentive compatibility and a Bayesian monotonicity condition are necessary and almost sufficient for full implementation. The proof of the sufficiency part of the result relies on complex augmented mechanisms.

In a working paper version of this paper, Bergemann and Morris (2005b), we developed the results in this paper as a special case of a general approach to robust implementation. The results reported in this section appear in that working paper.

Our robust implementation notion is equivalent to requiring Bayesian implementation on all type spaces. Ex post incentive compatibility is equivalent to Bayesian incentive compatibility on all type spaces. It is possible to define a notion of robust monotonicity which is equivalent to Bayesian monotonicity on all type spaces. Ex post incentive compatibility and robust monotonicity are thus necessary and almost sufficient for full implementation. However, this result relies on allowing complex augmented mechanisms including integer games. If we restrict attention to well-behaved mechanisms - with the compact message space assumption of this paper - then strict ex post incentive compatibility is also necessary.

The contraction property is an implication of robust monotonicity in the environment studied in this paper. The robust monotonicity condition requires the existence of allocations that can be used to reward individuals for reporting deceptions from desirable equilibria. In the environment of this paper, we are able to show that we can always use rewards from misreports in the direct mechanism.

In our treatment of the single good auction example, we noted that since the efficient allocation failed ex post incentive compatibility, robust implementation of the efficient allocation would surely not be possible. However, we were able to show that virtual implementation was possible. However, this begs the question of how much can be achieved in general with virtual implementation.[12] However, one can show that when the contraction property fails, robust virtual implementation is not possible in any mechanism either.

### 8.3 Interdependent Valuations

In this paper we considered implementation in an environment with interdependent valuations. We provided conditions for full implementation which did not depend on the prior or posterior belief of the agents. More precisely, we provided conditions under which the social choice function can be implemented in the direct mechanism by iteratively eliminating strictly dominated reports.

In contrast to much of the recent literature on implementation which relies heavily on complicated augmented mechanisms to achieve full implementation, here we pursued implementation in the direct mechanism without relying on augmented mechanisms. The resulting sufficient and almost necessary condition for robust implementation, the contraction property, was shown to essentially require that there is not too much interdependence in the valuation of each agent across signals received by the agents. In the important case of the linear model in signals, the contraction property was shown to reduce to a single condition on the eigenvalue of the interdependence matrix.

The nature of the contraction property also highlighted that robust implementation is considerably more demanding than ex post truthful implementation. Finally, the example of the efficient single unit auction suggested that the dividing line between positive and negative robust implementation results might also be the exact dividing line for the more permissive notion of virtual implementation.

### 8.4 Contraction Property

The robust implementation argument rested essentially on the single crossing property and the contraction property. The single crossing is essentially symmetric in allocation and type. It therefore would have been possible to impose the contraction property on the outcome function rather than on the preference aggregator. In fact, given that the misreports can only alter the outcome function, but certainly not the preferences, one might have thought it would be more natural to

[12] Abreu and Matsushima (1992a) and Abreu and Matsushima (1992b) show very permissive results about virtual implementation in complete and incomplete environments, respectively.

impose the contraction property on the outcome function rather than on the preference aggregator.[13] The advantage of using the contraction property on the aggregator function arises from the single crossing condition. The true type $\theta$ and the misreported types $\theta'$ can potentially be very far from each other. Consequently, the preferences at the type profiles $\theta$ and $\theta'$ over a pair allocations, in particular $f\left(\theta_i', \theta_{-i}'\right)$ and $f\left(\theta_i' + \varepsilon, \theta_{-i}\right)$, can be very different. With the contraction property on the preference aggregator, it suffices to compare the allocations, $f\left(\theta_i', \theta_{-i}'\right)$ and $f\left(\theta_i' + \varepsilon, \theta_{-i}\right)$ for a type profile near $\theta' = \left(\theta_i', \theta_{-i}'\right)$ and then extend the ranking to be valid for $\theta$ through the existence of an aggregator $h_i$ and the single crossing property. Without the aggregator $h_i$ but a contraction property on the social choice function, we would be forced to rank the allocations $f\left(\theta_i', \theta_{-i}'\right)$ and $f\left(\theta_i' + \varepsilon, \theta_{-i}\right)$ for some preferences near the true type profile $\theta = (\theta_i, \theta_{-i})$. In particular, in order to be able to use the single crossing condition fruitfully, it would have to be the case that the allocations $f\left(\theta_i', \theta_{-i}'\right)$ and $f\left(\theta_i' + \varepsilon, \theta_{-i}\right)$ would also arise as the equilibrium allocation for some reports $\theta_i^*, \theta_i^{**}$ of agent $i$ given the truthful report $\theta_{-i}$ of the remaining agents. But such a "full support" requirement is rather strong. In particular, it will rarely be satisfied in models with quasilinear utilities, where each agent has preferences over a two-dimensional object, the allocation and the monetary transfer.

### 8.5 The Common Prior Assumption and Strategic Substitutes/Complements

The definition of robust implementation in this paper is equivalent to requiring that every equilibrium on every type space delivers outcomes consistent with the social choice function. By "very type space", we are allowing for multiple copies of the same payoff type with different beliefs over the types of others. And we are allowing for non common prior type spaces. An interesting question is what happens when we look at an intermediate notion of robustness: allowing all possible common prior type spaces. This interesting question goes beyond the scope of this paper but we can use our leading example to illustrate why it is interesting.

Consider the public good example in the case where there is negative interdependence in valuations, i.e., $\gamma < 0$. Recall the ex post best response function in that example: if type $\theta_i$ is sure that his opponents have type profile $\theta_{-i}$ and is sure that they will report themselves to be type profile $\theta_{-i}'$, his best response is to report himself to be type

$$\theta_i' = \theta_i + \gamma \sum_{j \neq i} \left(\theta_j - \theta_j'\right).$$

We see that there are strategic complements in misreporting strategies (if others misreport upwards,

[13]In fact, the contraction property has been employed successfully in games with complete information and linear best responses to prove the uniqueness of the Nash equilibrium, see Luenberger (1978) and Gabay and Moulin (1980).

I have an incentive to misreport upwards). This means that when we carry out the iterated deletion procedure, the profile of largest and smallest misreports that survive must constitute an ex post equilibrium of the game (Milgrom and Roberts (1990)). Thus a failure of robust implementation also implies that there exists a bad equilibrium on any common prior type space.

On the other hand, in the standard case with positive interdependence, i.e., $\gamma > 0$, there is strategic substitutability in misreports and this argument does not go through. In fact, one can show in the example that even when the contraction property fails (i.e., $\gamma > \frac{1}{I-1}$), every equilibrium on any common prior type space delivers the right outcome.

# 9 Appendix

The appendix contains the arguments and proofs missing in the main text.

**Informational Foundation for Interdependence** The vector of the random variables

$$\begin{pmatrix} \omega_0+\omega_1 \\ \theta_1 \\ \theta_2 \end{pmatrix}$$

is normally distributed with mean zero and variance matrix

$$\begin{pmatrix} \sigma_0^2+\sigma_1^2 & \sigma_0^2+\sigma_1^2 & \sigma_0^2 \\ \sigma_0^2+\sigma_1^2 & \sigma_0^2+\sigma_1^2+\tau_1^2 & \sigma_0^2 \\ \sigma_0^2 & \sigma_0^2 & \sigma^2+\sigma_2^2+\tau_2^2 \end{pmatrix}$$

By a standard property of the multivariate normal distribution, see DeGroot (1970), this implies that the expectation of $\omega_0+\omega_1$ conditional on $\theta_1$ and $\theta_2$ is given by:

$$\begin{pmatrix} \sigma_0^2+\sigma_1^2 & \sigma_0^2 \end{pmatrix} \begin{pmatrix} \sigma_0^2+\tau_1^2+\sigma_1^2 & \sigma_0^2 \\ \sigma_0^2 & \sigma_0^2+\sigma_2^2+\tau_2^2 \end{pmatrix}^{-1} \begin{pmatrix} \theta_1 \\ \theta_2 \end{pmatrix},$$

which equals

$$\frac{\left(\sigma_0^2\tau_2^2+\sigma_0^2\sigma_1^2+\sigma_0^2\sigma_2^2+\tau_2^2\sigma_1^2+\sigma_1^2\sigma_2^2\right)\theta_1+\sigma^2\tau_1^2\theta_2}{\sigma_0^2\tau_1^2+\sigma_0^2\tau_2^2+\sigma_0^2\sigma_1^2+\sigma_0^2\sigma_2^2+\tau_1^2\tau_2^2+\tau_1^2\sigma_2^2+\tau_2^2\sigma_1^2+\sigma_1^2\sigma_2^2}.$$

If we multiply the above expression by the constant

$$\frac{\sigma_0^2\tau_1^2+\sigma_0^2\tau_2^2+\sigma_0^2\sigma_1^2+\sigma_0^2\sigma_2^2+\tau_1^2\tau_2^2+\tau_1^2\sigma_2^2+\tau_2^2\sigma_1^2+\sigma_1^2\sigma_2^2}{\sigma_0^2\tau_2^2+\sigma_0^2\sigma_1^2+\sigma_0^2\sigma_2^2+\tau_2^2\sigma_1^2+\sigma_1^2\sigma_2^2},$$

we obtain:

$$\theta_1+\frac{\sigma^2\tau_1^2}{\sigma_0^2\tau_2^2+\sigma_0^2\sigma_1^2+\sigma_0^2\sigma_2^2+\tau_2^2\sigma_1^2+\sigma_1^2\sigma_2^2}\theta_2,$$

as reported in (5).

**Dual Characterization of the Contraction Property** The following lemma gives a dual representation of the strict contraction property for the linear case. In turn, it allows us to characterize the contraction property in terms of the eigenvalue of the interdependence matrix $\Gamma$.

**Lemma 2 (Duality)**

*The following two properties of $\Gamma$ are equivalent:*

1. *for all* $c \in \mathbb{R}_+^I$ *with* $c \neq \mathbf{0}$, *there exists* $i$ *such that:*

$$c_i > \sum_{j \neq i} \left| \gamma_{ij} \right| c_j; \tag{27}$$

2. *there exists* $d \in \mathbb{R}_+^I$ *such that:*

$$d_i > \sum_{j \neq i} \left| \gamma_{ji} \right| d_j, \tag{28}$$

   *for all* $i$.

**Proof.** Consider the following contrapositive restatement of condition (27): there does not exist $c \in \mathbb{R}_+^I$ such that

$$\sum_{i=1}^{I} c_i > 0, \tag{a}$$

and

$$\sum_{j \neq i} \left| \gamma_{ij} \right| c_j - c_i \geq 0 \text{ for each } i. \tag{b}$$

Writing $\mu$ for the multiplier of constraint (a) and $d_i$ for the $i$ multiplier of constraint (b), Farkas' lemma states that such a $c$ does not exist if and only if there exist $d \in \mathbb{R}_+^I$ and $\mu \in \mathbb{R}_+$ such that

$$\mu - d_i + \sum_{j \neq i} \left| \gamma_{ji} \right| d_j = 0 \text{ for all } i, \tag{a'}$$

and

$$\mu > 0. \tag{b'}$$

But this is true if and only if condition (28) of the lemma holds. ■

An analogous exercise leads to the duality result for the contraction property, where the strict inequalities in (27) and (28) are simply replaced by weak inequalities.

**Proof of Theorem 2.** If we try to find a solution for the strict inequalities (28):

$$d_i > \sum_{j \neq i} \left| \gamma_{ji} \right| d_j, \quad \text{for all } i \tag{29}$$

with the assistance of a contraction constant $\lambda < 1$, or

$$d_i \lambda = \sum_{j \neq i} \left| \gamma_{ji} \right| d_j,$$

then by the Froebenius-Perron Theorem for nonnegative matrices (see Minc (1988), Theorem 1.4.2), there exists a positive right and a left eigenvector, both with the same positive eigenvalue $\lambda$. The associated eigenvector is positive as well. We can use the above dual property to establish that clearly a $(\lambda, d)$ solution exists for:

$$\lambda d_i = \sum_{j \neq i} |\gamma_{ji}| \, d_j,$$

but from the duality relationship (28), we know that for every $d > 0$,

$$d_i > \sum_{j \neq i} |\gamma_{ji}| \, d_j,$$

so it follows that $\lambda < 1$. ■

**Nonlinear Conditions** The linear model has the obvious advantage that the local conditions for contraction agree with the global conditions for contraction as the derivatives of the mapping $h_i(\theta)$ are constant and independent of $\theta$. Conversely, with a nonlinear model, we can present weak local conditions for every $\theta$ and stronger global conditions. With this we can extend the idea behind the linear aggregator function to a general nonlinear and differentiable aggregator function $h_i(\theta)$ as follows.

**Definition 7 (Local and Global Contraction Property)**

1. *The aggregator function $h_i$ satisfies the local contraction property if for all $c \in \mathbb{R}^I_+$ and $\theta \in \operatorname{int}(\Theta)$, there exists $i$ such that*

$$c_i \frac{\partial h_i(\theta)}{\partial \theta_i} > \sum_{j \neq i} c_j \left| \frac{\partial h_i(\theta)}{\partial \theta_j} \right|. \tag{30}$$

2. *The aggregator function $h_i$ satisfies the global contraction property if for all $c \in \mathbb{R}^I_+$, there exists $i$ such that,*

$$c_i \frac{\partial h_i(\theta)}{\partial \theta_i} > \sum_{j \neq i} c_j \left| \frac{\partial h_i(\theta)}{\partial \theta_j} \right| \tag{31}$$

*for all $\theta$.*

**Proposition 1 (Local and Global Contraction Property)**

1. *If $h_i$ satisfies the strict contraction property, then it satisfies the local contraction property.*

2. *If $h_i$ satisfies the global contraction property, then it satisfies the strict contraction property.*

**Proof.** (**1.**) The proof is by contradiction. The strict contraction property requires that if, for all $\beta \neq \beta^*$, there exists $i$ and $\theta_i' \in \beta_i(\theta_i)$ with $\theta_i' \neq \theta_i$, such that

$$\operatorname{sign}\left(\theta_i - \theta_i'\right) = \operatorname{sign}\left(h_i(\theta_i, \theta_{-i}) - h_i\left(\theta_i', \theta_{-i}'\right)\right),$$

for all $\theta_{-i}$ and $\theta_{-i}' \in \beta_{-i}(\theta_{-i})$. Fix any $c \in \mathbb{R}_+^I$ and choose small $\varepsilon > 0$. Now consider deceptions of the form

$$\beta_i(\theta_i) = [\theta_i - \varepsilon c_i, \theta_i + \varepsilon c_i] \cap \Theta_i.$$

If for some $\theta \in \operatorname{int}(\Theta)$,

$$c_i \frac{\partial h_i(\theta)}{\partial \theta_i} \leq \sum_{j \neq i} c_j \frac{\partial h_i(\theta)}{\partial \theta_j}.$$

for all $i$, then if $\theta_i' \in \beta_i(\theta_i)$ and (wlog) $\theta_i' > \theta_i$, then $\theta_i - \theta_i'$ is negative. Now choose $\theta_{-i}'$ such that $\theta_j' = \theta_j - \varepsilon c_j$. Now

$$\frac{h_i(\theta_i, \theta_{-i}) - h_i\left(\theta_i', \theta_{-i}'\right)}{\varepsilon} \to -c_i \frac{\partial h_i(\theta)}{\partial \theta_i} + \sum_{j \neq i} c_j \frac{\partial h_i(\theta)}{\partial \theta_j} \geq 0,$$

as $\varepsilon \to 0$. This contradicts the strict contraction property.

(**2.**) Fix any deception. Let

$$c_j = \max_{\theta_j' \in \beta_j(\theta_j)} \left|\theta_j' - \theta_j\right|.$$

There exists $i$

$$c_i \frac{\partial h_i(\theta)}{\partial \theta_i} > \sum_{j \neq i} c_j \left|\frac{\partial h_i(\theta)}{\partial \theta_j}\right|,$$

for all $\theta$. Let

$$\left|\theta_i - \theta_i'\right| = c_i$$

and suppose wlog that $\theta_i > \theta_i'$. Now fix any $\theta_{-i}' \in \beta_{-i}(\theta_{-i})$, we can then write the difference $h_i(\theta_i, \theta_{-i}) - h_i\left(\theta_i', \theta_{-i}'\right)$ as:

$$\begin{aligned}
& \int_{t=0}^{1} \sum_{j=1}^{I} \frac{\partial h_i\left(t\theta + (1-t)\theta'\right)}{\partial \theta_j} \left(\theta_j - \theta_j'\right) dt \\
= & \int_{t=0}^{1} \frac{\partial h_i\left(t\theta + (1-t)\theta'\right)}{\partial \theta_i} \left(\theta_i - \theta_i'\right) dt + \int_{t=0}^{1} \sum_{j \neq i} \frac{\partial h_i\left(t\theta + (1-t)\theta'\right)}{\partial \theta_j} \left(\theta_j - \theta_j'\right) dt \\
\geq & \int_{t=0}^{1} \frac{\partial h_i\left(t\theta + (1-t)\theta'\right)}{\partial \theta_i} c_i dt - \int_{t=0}^{1} \sum_{j \neq i} \left|\frac{\partial h_i\left(t\theta + (1-t)\theta'\right)}{\partial \theta_j}\right| c_j dt \\
> & \ 0,
\end{aligned}$$

where the last inequality comes from the hypothesis of the global contraction property. This establishes the claim. ■

## References

Abreu, D., and H. Matsushima (1992a): "Virtual Implementation in Iterative Undominated Strategies: Complete Information," *Econometrica*, 60, 993–1008.

——— (1992b): "Virtual Implementation In Iteratively Undominated Strategies: Incomplete Information," Discussion paper, Princeton University and University of Tokyo.

Battigalli, P., and M. Siniscalchi (2003): "Rationalization and Incomplete Information," *Advances in Theoretical Economics*, 3, Article 3.

Bergemann, D., and S. Morris (2005a): "Ex Post Implementation," Discussion Paper 1502, Cowles Foundation for Research in Economics, Yale University.

——— (2005b): "Robust Implementation: The Role of Large Type Spaces," Discussion Paper 1519, Cowles Foundation, Yale University.

——— (2005c): "Robust Mechanism Design," *Econometrica*, 73, 1771–1813.

Bergemann, D., and J. Välimäki (2002): "Information Acquisition and Efficient Mechanism Design," *Econometrica*, 70, 1007–1033.

Bikhchandani, S. (2005): "The Limits of Ex Post Implementation Revisited," Discussion paper, University of California at Los Angeles.

Birulin, O., and S. Izmalkov (2003): "On Efficiency of the English Auction," Discussion paper, Penn State University and MIT.

Brandenburger, A., and E. Dekel (1987): "Rationalizability and Correlated Equilibria," *Econometrica*, 55, 1391–1402.

Chung, K.-S., and J. C. Ely (2001): "Efficient and Dominance Solvable Auctions with Interdependent Valuations," Discussion paper, Northwestern University.

Cremer, J., and R. McLean (1985): "Optimal Selling Strategies Under Uncertainty for a Discriminating Monopolist When Demands are Interdependent," *Econometrica*, 53, 345–361.

Dasgupta, P., and E. Maskin (2000): "Efficient Auctions," *Quarterly Journal of Economics*, 115, 341–388.

DeGroot, M. (1970): *Optimal Statistical Decisions.*

Echenique, F., and A. Manelli (2004): "Comparative Statics, English Auctions, and the Stolper-Samuelson Theorem," Discussion paper, California Institute of Technology and Arizona State University.

Gabay, D., and H. Moulin (1980): "On the Uniqueness and Stability of Nash Equilibria in Noncooperative Games," in *Applied Stochastic Control in Econometrics and Management Science*, ed. by A. Bensoussan, pp. 271–293. North-Holland, Amsterdam.

Hong, H., and M. Shum (2003): "Econometric Models of Asymmetric Ascending Auctions," *Journal of Econometrics*, 112, 327–358.

Jackson, M. (1991): "Bayesian Implementation," *Econometrica*, 59, 461–477.

Jehiel, P., and B. Moldovanu (2006): "Allocative and Informational Externalities in Auction and Mechanism Design," in *Advances in Economics and Econometrics*, ed. by R. Blundell, W. Newey, and T. Persson. Cambridge University Press, Cambridge.

Jehiel, P., B. Moldovanu, M. Meyer-Ter-Vehn, and B. Zame (2005): "The Limits of Ex Post Implementation," *Econometrica*, forthcoming.

Krishna, V. (2003): "Asymmetric English Auctions," *Journal of Economic Theory*, 112, 261–288.

Luenberger, D. G. (1978): "Complete Stability of Noncooperative Games," *Journal of Optimization Theory and Applications*, 25, 485–505.

Maskin, E. (1992): "Auctions and Privatization," in *Privatization: Symposium in Honor of Herbert Giersch*, ed. by H. Siebert, pp. 115–136. J.C.B. Mohr, Tuebingen.

——— (1999): "Nash Equilibrium and Welfare Optimality," *Review of Economic Studies*, 66, 23–38.

Milgrom, P., and J. Roberts (1990): "Rationalizability, Learning and Equilibrium in Games with Strategic Complementarities," *Econometrica*, 58, 1255–1277.

Minc, H. (1988): *Nonnegative Matrices*. Wiley and Sons, New York.

Palfrey, T., and S. Srivastava (1989): "Mechanism Design with Incomplete Information: A Solution to the Implementation Problem," *Journal of Political Economy*, 97, 668–691.

Perry, M., and P. Reny (2002): "An Ex Post Efficient Auction," *Econometrica*, 70, 1199–1212.

Postlewaite, A., and D. Schmeidler (1986): "Implementation in Differential Information Economies," *Journal of Economic Theory*, 39, 14–33.