THE ROLE OF THE COMMON PRIOR
IN ROBUST IMPLEMENTATION

BY

DIRK BERGEMANN and STEPHEN MORRIS

COWLES FOUNDATION PAPER NO. 1225

THE ROLE OF THE COMMON PRIOR
IN ROBUST IMPLEMENTATION

Dirk Bergemann
Yale University and CEPR

Stephen Morris
Princeton University

Abstract
We consider the role of the common prior for robust implementation in an environment with interdependent values. Specifically, we investigate a model of public good provision which allows for negative and positive informational externalities. In the corresponding direct mechanism, the agents' reporting strategies are strategic complements with negative informational externalities and strategic substitutes with positive informational externalities. We derive the necessary and sufficient conditions for robust implementation in common prior type spaces and contrast this with our earlier results without the common prior. In the case of strategic complements the necessary and sufficient conditions for robust implementation do not depend on the existence of a common prior. In contrast, with strategic substitutes, the implementation conditions are much weaker under the common prior assumption. (JEL: C79, D82)

1. Introduction

We investigate the role of the common prior assumption in robust implementation. Robust implementation requires that every equilibrium on every type space delivers outcomes that are consistent with a social choice function. In using the term "every type space" we allow for multiple copies of the same payoff type with different beliefs over the types of others; we also allow for noncommon prior–type spaces. In this paper we want to look at an intermediate notion of robustness: allowing all possible common prior type spaces.

We develop our arguments in the context of a public good model with interdependent values. We have used this specific public good model as a leading example in our previous work on ex post implementation (Bergemann and Morris forthcoming), robust implementation in direct mechanisms (Bergemann and Morris 2007a), and robust virtual implementation (Bergemann and Morris 2007b). The current objective is to analyze the importance of the common prior assumption

for the possibility of robust implementation. In particular, we identify when the necessary and sufficient conditions for robust implementation depend on whether we allow for all type spaces (as in Bergemann and Morris 2007a) or only for those type spaces with a common prior. The public good model allows for positive as well as negative informational externalities. When we consider the direct revelation mechanism, we show that the reporting strategies of the agents are strategic complements with negative informational externalities and strategic substitutes with positive informational externalities.

The analysis of the robust implementation with and without a common prior relies on epistemic results on games of incomplete information. Brandenburger and Dekel (1987) and Aumann (1987), respectively, reported formal epistemic arguments that—for complete information games—characterize the solution concepts of correlated equilibrium and rationalizability as the consequences of common knowledge of rationality with and without the common prior, respectively. In Bergemann and Morris (2007c), we report belief-free incomplete information generalizations of the solution concepts (incomplete information correlated equilibrium and incomplete information rationalizability) and their epistemic foundations; these solution concepts and results are variants or special cases of the work of Battigalli and Siniscalchi (2003) and Forges (1993), respectively. We apply these results to the direct mechanism design setting, where the strategy space is simply the payoff-type space itself. In this environment, a specific message of a payoff type is incomplete information rationalizable if and only if there exists a type space and an interim equilibrium such that the message is an equilibrium action for a type with a given payoff type in the type space. A similar result can be stated for the incomplete information correlated equilibrium—namely, a message of a specific payoff type is an element of an incomplete information correlated equilibrium if and only if there exists a type space with a common prior for which the specific message is an interim equilibrium action for a type with that payoff type. With these epistemic results in the background, we can rephrase the conditions for robust implementation with and without common prior as establishing conditions for a unique solution under incomplete information correlated equilibrium and incomplete information rationalizability, respectively. In the case of strategic complements, the necessary and sufficient conditions for robust implementation do not depend on the existence of a common prior. In other words, with strategic complements, if truthtelling is the unique incomplete information correlated equilibrium outcome, then it is also the unique incomplete information rationalizable outcome. This reflects the well-known property of supermodular environments that multiple rationalizable outcomes occur only when there are multiple equilibria (see, e.g., Milgrom and Roberts 1990). In contrast, with strategic substitutes, the common prior assumption changes the implementation conditions. In particular, as the number of participating agents in the public good model increases, the conditions

for a unique rationalizable outcome converge to requiring pure private values, whereas the conditions for robust implementation with a common prior are independent of the number of participating agents and accommodate moderate interdependence.

The public good example considered here has two notable features that facilitate the analysis. First, the willingness to pay of agent i for the public good is the weighted sum of the payoff types of all the agents. The valuation of agent i is therefore identified by an aggregator that summarizes the private information of all agents within a one-dimensional variable. Second, the cost function of the public good is quadratic, and the resulting ex post incentive compatible transfer of agent i is a quadratic function of the reports of the agents. The quadratic payoff environment leads to a linear best response property that allows us to analyze the reporting game in the direct mechanism as a potential game. The analysis of the incomplete information correlated equilibrium is then facilitated by using potential game arguments first developed by Neyman (1997) for complete information games.

2. Setup

There are I agents. Player i has a payoff type $\theta_i \in \Theta_i$, where each $\Theta_i = [0, 1]$ is a compact interval of the real line. Each agent gets utility from a social choice $x \in X$, where X is a compact set, and transfers $t_i \in \mathbb{R}$; the agent's utility is given by $u_i(x, \theta) - t_i$. A direct mechanism specifies the social choice as a function of the profile of types, $f : \Theta \to X$, and a transfer rule for each agent, $t_i : \Theta \to \mathbb{R}$. Each agent sends an announcement of his type in the form of a message, $m_i \in \Theta_i$. This mechanism $(f, (t_i)_{i=1}^{I})$ is ex post incentive compatible if, for all i, θ, m_i:

$$u_i(f(\theta_i, \theta_{-i}), (\theta_i, \theta_{-i})) - t_i(\theta_i, \theta_{-i}) \geq u_i(f(m_i, \theta_{-i}), (\theta_i, \theta_{-i})) - t_i(m_i, \theta_{-i}).$$

A number of papers have described single crossing characterizations under which ex post incentive compatible transfers exist (e.g., Dasgupta and Maskin 2000; Bergemann and Valimaki 2002). In this article we focus on the case where they exist. If the true type profile is θ and the announced profile is m, then agent i's payoff is

$$u_i^+(m, \theta) \triangleq u_i(f(m), \theta) - t_i(m).$$

By construction, truthtelling is an ex post equilibrium in the direct mechanism. We consider two solution concepts for this setting. The first notion is incomplete information rationalizability.

DEFINITION 1. (Incomplete Information Rationalizability)   *The incomplete information rationalizable correspondences are* $R = (R_i)_{i=1}^{I}$, *with each* $R_i : \Theta_i \to$

$2^{\Theta_i}/\varnothing$, *being defined recursively. Let* $R_i^0(\theta_i) = \Theta_i$; *then*

$$R_i^{k+1}(\theta_i) = \left\{ m_i \in R_i^k(\theta_i) \;\middle|\; \begin{array}{l} \exists \mu_i \in \Delta(\Theta_{-i} \times \Theta_{-i}) \text{ such that} \\ \mu_i\big[\{(m_{-i}, \theta_{-i}) : m_j \in R_j^k(\theta_j) \forall j \neq i\}\big] = 1 \text{ and} \\ m_i \in \arg\max_{m_i'} \int_{m_{-i},\theta_{-i}} u_i^+\big((m_i', m_{-i}), (\theta_i, \theta_{-i})\big) d\mu_i \end{array} \right\}$$

for each $k = 1, 2, \ldots,$ *and*

$$R_i(\theta_i) = \bigcap_{k \geq 0} R_i^k(\theta_i).$$

The second notion is the incomplete information version of the correlated equilibrium.

DEFINITION 2. (Incomplete Information Correlated Equilibrium)   *A probability distribution* $\mu \in \Delta(\Theta \times \Theta)$ *is an incomplete information correlated equilibrium (ICE) of the direct mechanism if, for each i and each measurable* $\phi_i : \Theta_i \times \Theta_i \to \Theta_i$,

$$\int_{m,\theta} u_i^+((m_i, m_{-i}), \theta) d\mu \geq \int_{m,\theta} u_i^+((\phi_i(m_i, \theta_i), m_{-i}), \theta) d\mu.$$

We define $C_i(\theta_i)$ as the set of messages that can be played by type θ_i in an incomplete information correlated equilibrium of the direct mechanism. We will say that $m_i^* \in C_i(\theta_i^*)$ if, for each $\varepsilon > 0$, there exists an ICE μ with

$$\mu\big[\{(m, \theta)|m_i \in [\theta_i^* - \varepsilon, \theta_i^* + \varepsilon] \text{ and } \theta_i \in [\theta_i^* - \varepsilon, \theta_i^* + \varepsilon]\}\big] > 0.$$

In Bergemann and Morris (2007c) we report on these solution concepts in a general game theoretic environment. We observe that the solution concepts $R_i(\theta_i)$ and $C_i(\theta_i)$ are "belief-free" solution concepts in the sense that they do not refer to a specific common prior or to specific higher order beliefs of the agents. Rather, they represent the sets of actions that can be observed as rationalizable or correlated equilibrium actions for some beliefs of agent i given his payoff type θ_i. In Bergemann and Morris (2007c) we show that incomplete information rationalizability and correlated equilibrium share the same epistemic properties as their complete information equivalents, as outlined in the Introduction. With these belief-free notions in place, there is no further need to refer to beliefs and higher order beliefs of agent i. As a consequence, we shall hereafter refer to the payoff type θ_i simply as the type θ_i of agent i.

3. A Public Good Example

We discuss the role and the importance of the common prior for robust implementation in a public good example with interdependent values. In the final section, we discuss which special properties of the environment are used in establishing our results. We consider the provision of a public good $x \in \mathbb{R}_+$. The utility of each agent i for the public good is given by $u_i(x, \theta) = h_i(\theta) \cdot x$, where each

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

aggregates the agents' payoff types. Thus the utility of agent i depends on his own type with weight 1 and the types of the other agents with weight $\gamma \in \mathbb{R}$. The weight γ represents the preference interdependence among the agents. If $\gamma = 0$ we have a private values model, whereas $\gamma < 0$ represents negative informational externalities and $\gamma > 0$ represents positive informational externalities. The cost of establishing the public good is given by $c(x) = (1/2)x^2$. The planner must choose x to maximize social welfare-that is, the sum of gross utilities minus the cost of the public good. The socially optimal level of the public good is therefore equal to

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

The generalized Vickrey–Clarke–Groves (VCG) transfers are given by

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

The transfers $\{t_i(\theta)\}_{i=1}^{I}$ of the generalized VCG mechanism guarantee that truthtelling is an ex post incentive compatible strategy as long as $\gamma \geq -1/(I-1)$. If the (negative) externality γ falls below this threshold, then the single crossing condition ceases to hold.

Within the generalized VCG mechanism, we can define for every agent i an ex post best response as a mapping from the true payoff types of all agents and the reported types of all agents except i into a report of agent i: $b_i : \Theta \times \Theta_{-i} \to \Theta_i$. Given that agent i has type θ_i but reports himself to be of type m_i, and given that he has a point conjecture that other agents have type profile θ_{-i} and report their types to be m_{-i}, it follows that the net payoff of agent i is a constant $(1+\gamma(I-1))$ times:

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

The ex post best response of agent i is simply the solution to the first-order condition of the payoff function (3) with respect to m_i:

$$b_i(\theta, m_{-i}) \triangleq \theta_i + \gamma \sum_{j \neq i}(\theta_j - m_j). \qquad (4)$$

In other words, the best response by i to a (mis)report m_{-i} by the other agents is to report a type such that the aggregate type from his point of view is exactly identical to the true aggregate type (under the aggregation function $h_i(\theta)$ of agent i) generated by the true type profile θ. We note that the foregoing calculation also verifies the strict ex post incentive compatibility of f, because setting $m_i = \theta_i$ is the unique best response if $m_j = \theta_j$ for all $j \neq i$.

Whether there are strategic complements or substitutes plays an important role in determining the role of the common prior in implementation. We say that the strategies of i and j are *strategic complements* if $\partial b_i(\theta, m_j, m_{-ij})/\partial m_j > 0$, and they are *strategic substitutes* if $\partial b_i(\theta, m_j, m_{-ij})/\partial m_j < 0$, for all i, j, θ, m. Given the best response (4), it follows that the reports of the agents are strategic complements if there are negative informational externalities ($\gamma < 0$) but are strategic substitutes if there are positive informational externalities ($\gamma > 0$).

Bergemann and Morris (2007a) showed that if interdependence is small with $\gamma \in (-1/(I - 1), +1/(I - 1))$, then truthtelling is the unique rationalizable outcome and for all i and θ_i, $R_i(\theta_i) = \{\theta_i\}$, but if the interdependence is large with $\gamma \notin (-1/(I-1), +1/(I-1))$ then any message is rationalizable and for all i and θ_i, $R_i(\theta_i) = [0, 1]$. We refer the reader to Bergemann and Morris (2007a) for a formal statement and the derivation of this result. There we present necessary and sufficient conditions for robust implementation in environments where, for each agent i, the payoff types of all agents can be aggregated within a one-dimensional variable. The environment is general in the sense that neither the aggregator nor the utility function of each agent i must be linear as in the current example. We show that robust implementation is possible in any mechanism if and only if it is possible in the direct mechanism; and we show that robust implementation is possible if and only if the aggregator function satisfies a contraction property that reduces to the condition of a small γ with $\gamma \in (-1/(I - 1), +1/(I - 1))$.

In this article, we contrast the uniqueness result with incomplete information rationalizability with the incomplete information correlated equilibrium case. We use results regarding the uniqueness of the incomplete information correlated equilibrium in potential games derived in Bergemann and Morris (2007c). We say that a belief-free incomplete information game

$$\Gamma = \left\{ I, \{A_i\}_{i=1}^{I}, \{\Theta_i\}_{i=1}^{I}, \{u_i(a, \theta)\}_{i=1}^{I} \right\}$$

has a *weighted potential* $v : A \times \Theta \to \mathbb{R}$ if there exists a $w \in \mathbb{R}_{++}^{I}$ such that

$$u_i((a_i, a_{-i}), \theta) - u_i((a_i', a_{-i}), \theta) = w_i\left[v((a_i, a_{-i}), \theta) - v((a_i', a_{-i}), \theta) \right]$$

for all i and all $a_i, a'_i \in A_i$, $a_{-i} \in A_{-i}$ and $\theta \in \Theta$. This is an incomplete information generalization of the definition of a weighted potential in Monderer and Shapley (1996); in particular, it is equivalent to requiring that each complete information game $(u_i(\cdot, \theta))^I_{i=1}$ be a weighted potential game in the sense of Monderer and Shapley, using the same weights for each $\theta \in \Theta$. We say that v is a *strictly concave potential* if $v(\cdot, \theta)$ is a strictly concave function of a for all $\theta \in \Theta$. In Bergemann and Morris (2007c) we show that, if Γ has a strictly concave smooth potential function *and* an ex post equilibrium s^*, then $s^*_i(\theta_i) = C_i(\theta_i)$ for all i and θ_i. In the direct mechanism, the set of actions is the set of types. We argued earlier that the direct mechanism has truthtelling as an ex post equilibrium provided that $\gamma \geq -1/(I-1)$. By the result in Bergemann and Morris (2007c), the sufficiency condition for a unique correlated equilibrium can then be established by verifying the existence of a potential of the direct mechanism that is strictly concave.

PROPOSITION 1. (Incomplete Information Correlated Equilibrium)   *The set of incomplete information correlated equilibrium actions for all i and θ_i is $C_i(\theta_i) = \{\theta_i\}$ if and only if $\gamma \in (-1/(I-1), 1)$.*

Proof. We first establish the sufficiency result. We consider the following function $v(m, \theta)$ as a potential function for the direct mechanism:

$$v(m, \theta) = -\sum_{i=1}^{I}(m_i - \theta_i)\left[(m_i - \theta_i) + \gamma\sum_{j \neq i}(m_j - \theta_j)\right].$$

The partial derivative of the function $v(m, \theta)$ with respect to m_i is

$$\frac{\partial v}{\partial m_i}(m, \theta) = -2(m_i - \theta_i) - 2\gamma\sum_{j \neq i}(m_j - \theta_j), \tag{5}$$

and the cross-derivatives are given by

$$\frac{\partial^2 v}{\partial m_i \partial m_j}(m, \theta) = \begin{cases} -\frac{1}{2} & \text{if } j = i, \\ -\frac{1}{2}\gamma & \text{if } j \neq i. \end{cases} \tag{6}$$

It follows from (6) that $v(m, \theta)$ is a potential function, and it follows from (5) and the ex post best response (4) that v is a potential for the belief free incomplete information game Γ with type- and person-independent weights $w_i = 1/2$. Finally, because $v(m, \theta)$ is a quadratic function, the (constant) Hessian H is given by definition (6). With elementary linear algebra, we can now verify that H is negative semi-definite if and only if $-1/(I-1) \leq \gamma \leq 1$ and that H is negative definite if $-1/(I-1) < \gamma < 1$. Now the potential function is strictly concave if and only if its Hessian H is negative definite, which establishes sufficiency.

The necessity result follows from the best-response function (4). It suffices to show that, for $\gamma \geq 1$, there exist complete information correlated equilibria that do not involve truthtelling. Consider a payoff type profile $\hat{\theta}$ given by $\hat{\theta} = (1/2, 1/2, \hat{\theta}_3, \ldots, \hat{\theta}_I)$ for some $\hat{\theta}_{-12} = (\hat{\theta}_3, \ldots, \hat{\theta}_I)$. Consider the correlated equilibrium μ with $\mu((0, 1, \hat{\theta}_3, .., \hat{\theta}_I), (1/2, 1/2, \hat{\theta}_3, \ldots, \hat{\theta}_I)) = 1$. It is easy to verify that the equilibrium conditions (4) will be satisfied at μ. We observe that, for $\gamma > 1$, the equilibrium condition (4) will be a corner solution and hence there will be strict inequalities for $i = 1, 2$.   □

The results in Bergemann and Morris (2007a) together with Proposition 1 show how the common prior assumption has a considerable impact on the possibility of robust implementation with positive interdependence (and thus strategic substitutes) but no impact with negative interdependence (and thus strategic complementarities). The following corollary is then an immediate consequence of Bergemann and Morris (2007a) and Proposition 1.

COROLLARY 1. (Robust Implementation)

1. *If the reports are strategic complements, then robust implementation with common prior implies robust implementation without common prior.*
2. *If the reports are strategic substitutes, then robust implementation with common prior fails to imply robust implementation without common prior.*

The public good example shows how large the gap between robust implementation (with or without common prior) can be. In particular, as the number of agents increases, essentially only the private value model with $\gamma = 0$ can be robustly implemented without a common prior. In contrast, the interdependent model can be robustly implemented with a common prior for all $\gamma < 1$. This shows that the role of the common prior is critical in many mechanism design environments and for our understanding of robust implementation.

4. Discussion

Strategic complements and strategic substitutes. In the linear best-response environment of the public good problem, the notions of strategic complement and strategic substitute emerged directly from the best response. In general environments with differentiable mechanisms, the link between the information externality and the strategic properties of the reports have yet to be shown. For preciseness if we assume that the supermodularity conditions: $\partial f / \partial \theta_i > 0$ and $\partial^2 u_i / \partial x \partial \theta_i > 0$ support ex post incentive compatibility, then it follows that, locally at truthful reporting, the strategies of the agents are strategic substitutes if $\partial^2 u_i / \partial x \partial \theta_j > 0$ and strategic complements if $\partial^2 u_i / \partial x \partial \theta_j < 0$.

Potential and mechanism design. Our proof that there is a unique incomplete information correlated equilibrium action for each payoff type if $\gamma \in$

$(-1/(I-1), 1)$ used the fact that we could construct a potential function for the direct mechanism. This turns out to be a very strong property. In a differentiable environment, it is straightforward to show that a sufficient condition for Bayesian potential games is that the cross-derivatives of agent i and j equalize at every true and reported type profile. Although the cross-derivative is equal to zero in the current linear quadratic environment (because f is linear), we lose that property if we replace the quadratic cost function with a general concave cost function. It remains an open question to identify a larger class of environments where the potential argument goes through.

Jehiel, Meyer-Ter-Vehn, and Moldovanu (forthcoming) have used potential arguments to characterize when ex post incentive compatible transfers exist. Their definition of the potential function implicitly assumes that the agents are telling the truth and so—as the authors note—this is a much weaker requirement than the one that the direct mechanism be a potential game.

References

Aumann, Robert (1987). "Correlated Equilibrium as an Expression of Bayesian Rationality," *Econometrica*, 55, 1–18.

Battigalli, Pierpaolo, and Marciano Siniscalchi (2003). "Rationalization and Incomplete Information." *Advances in Theoretical Economics*, 3, Article 3.

Bergemann, Dirk, and Stephen Morris (forthcoming). "Ex Post Implementation." *Games and Economic Behavior*.

Bergemann, Dirk, and Stephen Morris (2007a). "Robust Implementation: The Case of Direct Mechanisms." Discussion Paper No. 1561R, Cowles Foundation for Research in Economics, Yale University.

Bergemann, Dirk, and Stephen Morris (2007b). "Strategic Distinguishability with an Application to Robust Virtual Implementation." Discussion Paper No. 1609, Cowles Foundation, Yale University.

Bergemann, Dirk, and Stephen Morris (2007c). "Belief Free Incomplete Information Games." Discussion Paper No. 1629, Cowles Foundation for Research in Economics, Yale University.

Bergemann, Dirk, and Juuso Valimaki (2002). "Information Acquisition and Efficient Mechanism Design." *Econometrica*, 70, 1007–1033.

Brandenburger, Adam, and Eddie Dekel (1987). "Rationalizability and Correlated Equilibria." *Econometrica*, 55, 1391–1402.

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

Forges, F. (1993). "Five Legitimate Definitions of Correlated Equilibrium in Games with Incomplete Information." *Theory and Decision*, 35, 277–310.

Jehiel, Philippe, Moritz Meyer-Ter-Vehn, and Benny Moldovanu (forthcoming). "Ex-Post Implementation and Preference Aggregation Via Potentials." *Economic Theory*.

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

Monderer, Dov, and Lloyd Shapley (1996). "Potential Games." *Games and Economic Behavior*, 14, 124–143.

Neyman, Abraham (1997). "Correlated Equilibrium and Potential Games." *International Journal of Game Theory*, 26, 223–227.