

## ROBUSTNESS AND APPROXIMATION IN MECHANISM DESIGN<sup>‡</sup>

### Efficient Auctions and Interdependent Types<sup>†</sup>

By DIRK BERGEMANN, STEPHEN MORRIS, AND SATORU TAKAHASHI\*

Before the work of Harsanyi (1967–1968) economists used to routinely argue that game theory could not be applied to economic settings because it required common knowledge of the environment. Following Harsanyi (1967–1968), economists accepted that rich enough “type spaces” allowed any possible lack of common knowledge to be incorporated. But very rich type spaces would be needed, and applied work remains highly sensitive to sometimes unexamined modeling choices about types. Nowhere is this more true than in mechanism design.

Two of us have written a series of papers highlighting the importance of lack of common knowledge and rich type spaces in mechanism design, now to be collected together in a book (Bergemann and Morris 2012). A new introduction examines key insights of this work using the classic example of the efficient allocation of a single good with interdependent values in a quasi linear environment to illustrate themes of the book (Bergemann and Morris 2011). In this article, we want to examine some issues about the single good problem in more detail. We discuss an approach to modeling interdependent preferences distinguishing between “payoff types” and “beliefs types.” We report a

characterization (not in the book) of when the efficient allocation can be partially Bayesian implemented on a finite type space in this setting. This characterization can be used to unify a number of sufficient conditions for efficient partial implementation in this classical setting. We also report how a canonical language for discussing interdependent types—developed in a more general setting by Bergemann, Morris, and Takahashi (2011)—applies in this setting. We note by example that this canonical language will not allow us to distinguish some types in the payoff type–belief type language. Abreu and Matsushima (1992) showed that types that cannot be distinguished in this language may always pool in some equilibrium, and, thus, the efficient allocation can only be fully implemented if it is measurable with respect to statements that can be expressed in this language.

There are  $I$  agents. Each agent  $i$  has a “payoff type”  $\theta_i$  belonging to a finite set  $\Theta_i$ . Each agent  $i$ 's monetary valuation of a good depends on the profile of payoff types  $\theta = (\theta_1, \dots, \theta_I)$  and is given by a valuation function  $v_i: \Theta_1 \times \dots \times \Theta_I \rightarrow \mathbb{R}_+$ . Agents are assumed to know their own payoff type  $\theta_i$ , but we want to allow for rich beliefs and higher order beliefs. Thus we assume each agent has a “type”  $t_i$  belonging to a finite set  $T_i$  and write  $\hat{\theta}_i: T_i \rightarrow \Theta_i$  for a function describing an agent's payoff type and  $\hat{\pi}_i: T_i \rightarrow \Delta(T_{-i})$  for his belief type. This decomposition of an agent's type into a payoff type and a belief type is a natural one in this quasi-linear setting. Two of us used this implicit description of interdependent types in our work on robust mechanism design (Bergemann and Morris 2012). We will discuss briefly below a sense in which this decomposition can be seen as without loss of generality.

We will focus on the problem of allocating the object to the agent with the highest

\*Discussants: Stephen Morris, Princeton University; Moritz Meyer-ter-Vehn, University of California-Los Angeles; Eric Maskin, Harvard University.

\*Bergemann: Department of Economics, Yale University, New Haven, CT 06520 (e-mail: [dirk.bergemann@yale.edu](mailto:dirk.bergemann@yale.edu)); Morris: Department of Economics, Princeton University, Princeton, NJ 08544 (e-mail: [smorris@princeton.edu](mailto:smorris@princeton.edu)); Takahashi: Department of Economics, Princeton University, Princeton, NJ 08544 (e-mail: [satorut@princeton.edu](mailto:satorut@princeton.edu)). We would like to thank our discussant, Eric Maskin, for his insightful comments, and Moritz Meyer-ter-Vehn and Andy Postlewaite for helpful discussions.

<sup>†</sup> To view additional materials, visit the article page at <http://dx.doi.org/10.1257/aer.102.3.319>.

valuation. Write  $i^*(\theta)$  for the agent who values the object the most, i.e.,  $i^*(\theta) \in \arg \max_i v_i(\theta)$ , and in this finite type setting, we can assume without loss of generality that  $i^*(\theta)$  is uniquely defined. We are interested in designing a finite mechanism involving monetary transfers which has an equilibrium where the object is always allocated to the agent who values it the most. By the revelation principle, we can restrict attention to direct mechanisms. In this context, a direct mechanism consists of a rule specifying monetary transfers to all agents,  $y: T \rightarrow \mathbb{R}^I$ . Now if agents report themselves to be of the type profile  $t = (t_1, \dots, t_I)$ , the object will be allocated to agent  $i^*(\hat{\theta}(t))$ , where we write  $\hat{\theta}(t) = (\hat{\theta}_1(t_1), \dots, \hat{\theta}_I(t_I))$ , and each agent  $i$  will receive a monetary transfer  $y_i(t)$ . This direct mechanism will be incentive compatible if

(1)

$$t_i \in \arg \max_{t'_i \in T_i} \left\{ \sum_{t_{-i} \in T_{-i}} (\mathbb{I}_i(\hat{\theta}(t'_i, t_{-i})) v_i(\hat{\theta}(t_i, t_{-i})) + y_i(t'_i, t_{-i})) \hat{\pi}_i(t_{-i} | t_i) \right\},$$

where the indicator function  $\mathbb{I}_i(\cdot)$  of agent  $i$  is one if  $i = i^*(\hat{\theta}(t))$  and zero otherwise. Thus we say efficient partial implementation is possible if and only if there exists  $y: T \rightarrow \mathbb{R}^I$  such that (1) is satisfied.

To give a characterization of the incentive compatibility of the efficient social choice function  $i^*$ , or short, the efficient partial implementation, it is useful to first identify key properties of players' incentives to report their payoff types for beliefs about others' payoff types. Fix an agent  $i$  and fix his beliefs about the payoff types of others,  $\psi_i \in \Delta(\Theta_{-i})$ . Suppose that agent  $i$  expects the object to be allocated according to  $i^*$  based on the truthful reports of other agents and his own (true or false) report. By extension of the valuation function  $v_i$ , we write  $V_i(\theta_i, \theta'_i, \psi_i)$  for the expected utility gain of agent  $i$  from a misreport  $\theta'_i$  relative to his true payoff type  $\theta_i$ , so that

$$V_i(\theta_i, \theta'_i, \psi_i) \triangleq \sum_{\theta_{-i}} (\mathbb{I}_i(\theta'_i, \theta_{-i}) - \mathbb{I}_i(\theta_i, \theta_{-i})) \psi_i(\theta_{-i}) v_i(\theta_i, \theta_{-i}).$$

Fix a subset  $\bar{\Theta}_i \subseteq \Theta_i$ ; say that  $V_i$  is cyclically monotonic on  $(\bar{\Theta}_i, \psi_i)$  if, for every sequence of types  $(\theta_i^1, \dots, \theta_i^K)$  in  $\bar{\Theta}_i$ ,

$$V_i(\theta_i^1, \theta_i^2, \psi_i) + V_i(\theta_i^2, \theta_i^3, \psi_i) + \dots + V_i(\theta_i^K, \theta_i^1, \psi_i) \leq 0.$$

Now, Theorem 1 in Rochet (1987) shows that there exists  $\bar{y}_i: \bar{\Theta}_i \rightarrow \mathbb{R}$  such that  $V_i(\theta_i, \theta'_i, \psi_i) + \bar{y}_i(\theta_i) \geq V_i(\theta_i, \theta'_i, \psi_i) + \bar{y}_i(\theta'_i)$  for all  $\theta_i, \theta'_i \in \bar{\Theta}_i$  if and only if  $V_i$  is cyclically monotonic on  $(\bar{\Theta}_i, \psi_i)$ . Thus cyclic monotonicity tells us that—ignoring belief types beyond the induced beliefs  $\psi_i$  over the payoff types—it is possible to choose transfers that give an agent an incentive to report his payoff type truthfully. To state this precisely, it will be useful to introduce the mapping  $\hat{\psi}_i: \Delta(T_{-i}) \rightarrow \Delta(\Theta_{-i})$  which describes the beliefs over payoff types induced by the agents' belief types, so that

$$\hat{\psi}_i(\pi_i) | \theta_{-i} \triangleq \pi_i(\{t_{-i}: \hat{\theta}_{-i}(t_{-i}) = \theta_{-i}\})$$

for any  $\pi_i \in \Delta(T_{-i})$ . Now the incentive compatibility condition (1) can be rewritten as:

$$(2) \quad t_i \in \arg \max_{t'_i \in T_i} \left\{ V_i(\hat{\theta}_i(t_i), \hat{\theta}_i(t'_i), \hat{\psi}_i(\hat{\pi}_i(t_i))) + \sum_{t_{-i} \in T_{-i}} y_i(t'_i, t_{-i}) \hat{\pi}_i(t_{-i} | t_i) \right\}.$$

Now we consider what we can learn about agents' beliefs, and how we can use what we learn. The classical environment we are considering—with quasi-linear utility and no limited liability constraints—is well known to be very permissive in allowing “belief extraction.” As observed by d’Aspremont and Gerard-Varet (1979), Myerson (1981), and Cremer and McLean (1985), it is possible to elicit agents' beliefs over other types by offering them gambles. Our general characterization of efficient partial implementation will essentially state that agents' beliefs can be elicited for “free,” and what then matters is whether the set of payoff types consistent with a given belief type can be distinguished. But this reduces to a cyclic monotonicity condition. Thus, write  $\hat{\Theta}_i(\pi_i)$  for the set of payoff types of player  $i$  associated with

belief type  $\pi_i$ , so that  $\hat{\Theta}_i(\pi_i) = \{\theta_i \in \Theta_i \mid \exists t_i \text{ with } \hat{\pi}_i(t_i) = \pi_i \text{ and } \hat{\theta}_i(t_i) = \theta_i\}$ . Now we have:

**PROPOSITION 1:** *Efficient partial implementation is possible if and only if each  $V_i$  satisfies cyclic monotonicity on  $(\hat{\Theta}_i(\pi_i), \hat{\psi}_i(\pi_i))$  for each  $\pi_i$  in the range of  $\hat{\pi}_i$ .*

**PROOF:**

Suppose efficient partial implementation is possible and so (2) holds for some  $y: T \rightarrow \mathbb{R}^I$ . For any agent  $i$ ,  $\pi_i$  in the range of  $\hat{\pi}_i$  and  $\theta_i \in \hat{\Theta}_i(\pi_i)$ , write  $\hat{t}_i(\theta_i, \pi_i)$  for any type  $t_i$  with  $\hat{\pi}_i(t_i) = \pi_i$  and  $\hat{\theta}_i(t_i) = \theta_i$ . Define  $\bar{y}_i(\theta_i, \pi_i) = \sum_{t_{-i}} y_i(\hat{t}_i(\theta_i, \pi_i), t_{-i}) \pi_i(t_{-i})$ . By (2),

$$\begin{aligned} V_i(\theta_i, \theta_i, \hat{\psi}_i(\pi_i)) + \bar{y}_i(\theta_i, \pi_i) \\ \geq V_i(\theta_i, \theta'_i, \hat{\psi}_i(\pi_i)) + \bar{y}_i(\theta'_i, \pi_i) \end{aligned}$$

for all  $\theta_i, \theta'_i \in \hat{\Theta}_i(\pi_i)$ . Thus,  $V_i$  satisfies cyclic monotonicity on  $(\hat{\Theta}_i(\pi_i), \hat{\psi}_i(\pi_i))$ .

Suppose  $V_i$  satisfies cyclic monotonicity on  $(\hat{\Theta}_i(\pi_i), \hat{\psi}_i(\pi_i))$  for each  $\pi_i$  in the range of  $\hat{\pi}_i$ . Thus there exists

$$\bar{y}_i(\cdot, \pi_i): \hat{\Theta}_i(\pi_i) \rightarrow \mathbb{R},$$

such that

$$\begin{aligned} V_i(\theta_i, \theta_i, \hat{\psi}_i(\pi_i)) + \bar{y}_i(\theta_i, \pi_i) \\ \geq V_i(\theta_i, \theta'_i, \hat{\psi}_i(\pi_i)) + \bar{y}_i(\theta'_i, \pi_i) \end{aligned}$$

for all  $\theta_i, \theta'_i \in \hat{\Theta}_i(\pi_i)$ . Since  $\sum_{t_{-i}} (\log \pi_i(t_{-i}) - \log \pi'_i(t_{-i})) \pi_i(t_{-i}) > 0$  for any  $\pi'_i \neq \pi_i$ , (2) holds for  $y: T \rightarrow \mathbb{R}^I$  given by  $y_i(t) = \bar{y}_i(\theta_i(t_i), \hat{\pi}_i(t_i)) + K \log \hat{\pi}_i(t_i)[t_{-i}]$  with sufficiently large  $K$ . (If  $\hat{\pi}_i(t_i)[t_{-i}] = 0$ , then we let  $y_i(t)$  be a sufficiently small negative number.)

We note that the present result of partial implementation is stated for the (ex post) efficient allocation rule  $i^*(\theta)$ . But in fact, the above result is more generally valid for every allocation rule that is measurable with respect to the payoff type profile  $\theta$ . The only modification arises with respect to the utility gains  $V_i(\cdot)$  from misreporting the payoff type,  $\theta_i$  to  $\theta'_i$ , which have to be adapted to the specific allocation rule. In fact, a slightly more general version of this result was reported in Proposition 6.2 of Bergemann

and Morris (2003), a working paper version of Bergemann and Morris (2005), where general allocation problems were considered. In any case, the present result gives a sharp characterization of how payoff types and belief types matter for efficient partial implementation. In particular, differences in beliefs about others' types can be extracted for free, even if they are not payoff relevant. But once they are extracted, the non-payoff relevant content of the beliefs does not matter. All that matters is the implied belief over payoff types and the cyclic monotonicity condition given by that implied belief over the payoff types and the set of possible payoff types who could have had the original (perhaps payoff irrelevant) belief. Classic sufficient conditions for efficient partial implementation can now be seen as special cases of the above proposition:

(i) *Private Values.*—Under private values (i.e.,  $v_i(\theta_i, \theta_{-i})$  does not depend on  $\theta_{-i}$ ), the condition becomes vacuous as the cyclic monotonicity conditions are satisfied.

(ii) *Independent Types.*—If the agents' types are distributed independently ( $\hat{\pi}_i(t_i)$  does not depend on  $t_i$ ), then necessary and sufficient conditions for efficient partial implementation reduce to the cyclic monotonicity conditions on the entire set  $\Theta_i$  of payoff types of each agent  $i$  for his fixed beliefs over others' payoff types.

(iii) *Linear Independence and Convex Hull.*—As noted by our discussant, Eric Maskin, for any set  $\Psi_i$  of linearly independent beliefs over others' payoff types, the cyclic monotonicity conditions on the entire set  $\Theta_i$  for all beliefs in  $\Psi_i$  are sufficient for efficient partial implementation if there is common knowledge that beliefs lie in the convex hull of  $\Psi_i$ , which is less demanding than the notion of ex post equilibrium. In a similar spirit, Jehiel, Moldovanu, and Meyerter-Vehn (2012) investigate conditions for local robust incentive compatibility, and their Lemma 1 establishes necessary monotonicity conditions for partial implementation, using the above separation of payoff types and belief types.

(iv) *Belief Extraction.*—Following Neeman (2004), say that “beliefs determine preferences” (BDP) if for every  $\pi_i \in \Delta(T_{-i})$ , there exists at most one  $\theta_i \in \Theta_i$  such that  $\hat{\pi}_i(t_i) = \pi_i$  and

$\hat{\theta}_i(t_i) = \theta_i$  for some  $t_i \in T_i$ . Under BDP, the cyclic monotonicity requirements of the proposition become vacuous.<sup>1</sup>

(v) *Ex Post Incentive Compatibility*.—As in the work of Maskin (1992) and Dasgupta and Maskin (2000), suppose that a single crossing condition is satisfied with respect to the (ordered) payoff types of the agents, so that each  $v_i$  is strictly monotonic in  $\theta_i$  and  $\theta'_i > \theta_i \Rightarrow v_i(\theta'_i, \theta_{-i}) - v_i(\theta_i, \theta_{-i}) > v_j(\theta'_i, \theta_{-i}) - v_j(\theta_i, \theta_{-i})$  for all  $i, j$  with  $i \neq j$ . Then  $V_i$  satisfies cyclic monotonicity on  $(\Theta_i, \psi_i)$  for every  $\psi_i$  (which implies ex post incentive compatibility) and thus again it can be shown that the cyclic monotonicity requirements of the proposition hold.<sup>2</sup>

(vi) *Combining Belief Extraction and Private Values*.—McLean and Postlewaite (2004) analyze efficient auctions with interdependent values and multidimensional types. More precisely, each agent has a two-dimensional payoff type  $\theta_i = (\theta_i^i, \theta_i^c)$ , with an idiosyncratic and a common component,  $\theta_i^i$  and  $\theta_i^c$ , respectively, as reflected by the valuation function  $v_i(\theta) = v_i^i(\theta_i^i) + v_i^c(\theta^c)$ . The common part of the valuation,  $v_i^c(\theta^c)$ , depends on the entire profile of common components  $\theta^c = (\theta_1^c, \dots, \theta_j^c)$ . Now, in terms of our language, their environment in Section 4 can be described by a type space  $T_i = \Theta_i$ , the payoff types can be described by the identity mapping  $\theta_i$ , and the belief types can be described by  $\pi_i[\theta_i^i | ((\theta_j)_{j \neq i})] = \pi_i^c[\theta_i^c | ((\theta_j^c)_{j \neq i})] \cdot \pi_i^i((\theta_j^i)_{j \neq i})$ . Finally, their condition of positive informational variability implies that the common component  $\theta_i^c$  can be extracted from every agent  $i$ , and the residual private information  $\theta_i^i$  is independent and pertains to the private value; and hence, again, our necessary and sufficient conditions are satisfied.

Importantly, Proposition 1 establishes only partial implementation, i.e., that the efficient

allocation arises in some equilibrium, but there may exist other equilibria with inefficient allocations. Suppose there are three agents,  $I = 3$ , and each agent has two possible types,  $\Theta_1 = \Theta_2 = \Theta_3 = \{0, 1\}$ . Each agent's valuation of the object is given by

$$(3) \quad v_i(\theta) = \theta_i + \frac{2}{3} \sum_{j \neq i} \theta_j.$$

Each agent has only two types, and thus a single belief type for each payoff type. Suppose, in particular, that if a type has valuation  $\theta_i$ , he assigns independent probability  $\frac{1}{8}$  to each of the other agents having valuation  $\theta_j = \theta_i$ , and the remaining probability  $\frac{7}{8}$  to  $\theta_j = 1 - \theta_i$ .

Efficient partial implementation is possible in this example (by both a belief extraction argument, i.e., sufficient condition (iv) above, or an ex post incentive compatibility argument, e.g., sufficient condition (v) above). But observe that each type of each agent has an expected value of  $\frac{7}{6}$  for the object. To see why, note that if agent 1, say, has  $\theta_1 = 1$ , his expectation of  $\theta_2$  (or  $\theta_3$ ) is  $\frac{1}{8}$ , and, thus, his expectation of  $\theta_1 + \frac{2}{3}\theta_2 + \frac{2}{3}\theta_3 = 1 + \frac{1}{12} + \frac{1}{12} = \frac{7}{6}$ . But if agent 1 has  $\theta_1 = 0$ , his expectation of  $\theta_2$  (or  $\theta_3$ ) is  $\frac{7}{8}$ , and, thus, his expectation of  $\theta_1 + \frac{2}{3}\theta_2 + \frac{2}{3}\theta_3 = 0 + \frac{7}{12} + \frac{7}{12} = \frac{7}{6}$ . Thus, there will always be an equilibrium where both types of each player behave in the same way, which cannot give rise to the efficient allocation, as shown in a general payoff environment in Bergemann and Morris (2009).

Following Bergemann, Morris, and Takahashi (2011), we propose a natural description of agents' types in this setting. An agent's (unconditional) willingness to pay for the object does not depend on other agents' types and, thus, has a natural meaning. If it makes sense to speak of all agents' willingness to pay for the object, it makes sense to talk about an agent's beliefs about other agents' willingness to pay for the object, and his willingness to pay conditional on others' willingness to pay for the object. Call the agent's unconditional willingness to pay his first-order type. Call his beliefs over others' willingness to pay and conditional willingness to pay his second-order type. Now we can inductively define an agent's  $(n + 1)$ th order type to be his beliefs over others'  $n$ th order types and his willingness to pay conditional on others'  $n$ th order types. We can thus identify an agent with

<sup>1</sup> Arguments going back at least to Cremer and McLean (1985) say BDP holds generically on finite type spaces, since if we fix a finite set of types and perturb beliefs, then they will all be different and BDP will hold. Recent contributions, notably Heifetz and Neeman (2006), examine when BDP holds on infinite type spaces, and in particular if it can be said to hold "generically" on infinite type spaces.

<sup>2</sup> Such sufficient conditions do not naturally arise in multidimensional payoff type space  $\Theta_i$ , except for some special cases, as noted by Maskin (1992) and shown in much more general environments by Jehiel et al. (2006).

a hierarchy of statements about preferences and conditional preferences.

In the simple setting of this note, such hierarchies can be defined formally as follows. An  $n$ th-level type is pair  $t_n = (b_n, v_n)$  consisting of a belief component and valuation component. Write  $T_n$  for the set of  $n$ th-level types that can arise in a finite type space. We describe the sets  $T_n$  inductively. Let  $T_1 = \{\emptyset\} \times \mathbb{R}_+$  with a typical first-level type  $t_1 = (\emptyset, v_1)$  consisting of a degenerate belief type and an unconditional valuation of the object. Now an  $(n + 1)$ th-level type  $t_{n+1} = (b_{n+1}, v_{n+1})$  consists of a simple (i.e., finite support) probability distribution  $b_{n+1} \in \Delta(T_n)$  with support  $\text{supp}(b_{n+1})$  and a valuation function  $v_{n+1} : \text{supp}(b_{n+1}) \rightarrow \mathbb{R}_+$ . Now a hierarchy of types is a sequence of  $n$ th-order types  $(t_1, t_2, \dots, t_n, \dots) \in T_1 \times T_2 \times \dots \times T_n \times \dots$ . A sequence of types is coherent if each  $(n + 1)$ th-type  $t_{n+1}$  induces beliefs over other agents'  $(n - 1)$ th-level types and willingness to pay conditional on other agents'  $(n - 1)$ th-level types that are consistent with those of  $t_n$ . (We omit the formal statement of this condition). Now if we write  $T_f^*$  for the set of all coherent hierarchies of higher-order types that can arise in finite type spaces, we have a natural language in which to discuss agents' types.

In Bergemann, Morris, and Takahashi (2011), we discuss the extension to infinite types and construct a universal space of higher-order preferences  $T^*$  from hierarchies. We can identify those hierarchies with beliefs over others' types and valuations conditional on their hierarchies. Thus, we identify a type  $t^* \in T^*$  with a probability measure over the types of others  $b^* \in \Delta((T^*)^{I-1})$  and an equivalence class of  $b^*$ -integrable valuation functions  $v^* : (T^*)^{I-1} \rightarrow \mathbb{R}$  (where we identify two functions if they agree  $b^*$ -almost surely). In this sense, we have a canonical way of representing type spaces with a decomposition of types into "payoff types" and "belief types," as we did earlier. However, it is important to realize that these payoff types do not specify valuations on 0-probability events.

This language is closely related to full implementation. Abreu and Matsushima (1992) showed that a necessary condition for Bayesian full implementation of a social choice function using a finite mechanism on a finite type space was that two types which had the same hierarchies of preferences received the same allocation. While they expressed this "measurability"

condition as a property of the fixed finite type space, it can be expressed without reference to a particular finite type space as we described above. Thus, if full implementation is required, we can at most achieve constrained efficiency, with the object allocated to the agent  $i^{**}(t_1^*, \dots, t_i^*)$  that maximizes  $v_i^*[t_i^*](t_{-i}^*)$ , where  $v_i^*[t_i^*]$  denotes the payoff type component of  $t_i^*$ , and Proposition 1 must hold on the coarser type space where types with the same hierarchy of preferences are merged.

Exact full implementation is generally not possible in the allocation of a single good: even with private values, while bidding one's true value is a dominant strategy, it is often only a weak best response, and it is easy to construct inefficient equilibria in dominated strategies. For sufficient conditions for full implementation with finite mechanisms, the objective must be weakened to virtual implementation, so that  $i^{**}(t_1^*, \dots, t_i^*)$  is allocated the object with arbitrarily high probability. Abreu and Matsushima (1992) showed that virtual Bayesian full implementation is possible under Bayesian incentive compatibility and their measurability condition. In this context, this implies that the present constrained efficient allocation of the good is possible if the condition of Proposition 1 holds on the coarsened type space that is expressible in our language.<sup>3</sup>

## REFERENCES

- Abreu, Dilip, and Hitoshi Matsushima. 1992. "Virtual Implementation in Iteratively Undominated Strategies: Incomplete Information." Unpublished.
- Bergemann, Dirk, and Stephen Morris. 2003. "Robust Mechanism Design." Yale University, Cowles Foundation Discussion Paper 1421R.
- Bergemann, Dirk, and Stephen Morris. 2005. "Robust Mechanism Design." *Econometrica* 73 (6): 1771–813.

<sup>3</sup> In Bergemann and Morris (2009), we examined the efficient allocation rule (with the single unit auction as a special case) and gave conditions under which full implementation can be achieved by a simpler mechanism, namely the direct mechanism, than the mechanism of Abreu and Matsushima (1992). It is an open question when a simpler mechanism can achieve full implementation more generally.



- Bergemann, Dirk, and Stephen Morris.** 2009. "Robust Implementation in Direct Mechanisms." *Review of Economic Studies* 76 (4): 1175–204.
- Bergemann, Dirk, and Stephen Morris.** 2011. "Robust Mechanism Design: An Introduction." Yale University, Cowles Foundation Discussion Paper 1818.
- Bergemann, Dirk, and Stephen Morris.** 2012. *Robust Mechanism Design: The Role of Private Information and Higher Order Beliefs*. Singapore: World Scientific Publishing.
- Bergemann, Dirk, Stephen Morris, and Satoru Takahashi.** 2011. "Interdependent Preferences and Strategic Distinguishability." Yale University, Cowles Foundation Discussion Paper 1772R.
- Cremer, Jacques, and Richard P. McLean.** 1985. "Optimal Selling Strategies under Uncertainty for a Discriminating Monopolist When Demands Are Interdependent." *Econometrica* 53 (2): 345–61.
- Dasgupta, Partha, and Eric S. Maskin.** 2000. "Efficient Auctions." *Quarterly Journal of Economics* 115 (2): 341–88.
- d'Aspremont, Claude, and Louis-Andre Gerard-Varet.** 1979. "Incentives and Incomplete Information." *Journal of Public Economics* 11 (1): 25–45.
- Harsanyi, John.** 1967. "Games with Incomplete Information Played by 'Bayesian' Players, Part I. The Basic Model." *Management Science* 14 (3): 159–82.
- Harsanyi, John.** 1968. "Games with Incomplete Information Played by 'Bayesian' Players, Part II. Bayesian Equilibrium Points." *Management Science* 14 (5): 320–34.
- Harsanyi, John.** 1968. "Games with Incomplete Information Played by 'Bayesian' Players, Part III. The Basic Probability Distribution of the Game." *Management Science* 14 (7): 486–502.
- Heifetz, Aviad, and Zvika Neeman.** 2006. "On the Generic (Im)possibility of Full Surplus Extraction in Mechanism Design." *Econometrica* 74 (1): 213–33.
- Jehiel, Philippe, Benny Moldovanu, and Moritz Meyer-ter-Vehn.** 2012. "Locally Robust Implementation and Its Limits." Unpublished.
- Jehiel, Philippe, Moritz Meyer-ter-Vehn, Benny Moldovanu, and William R. Zame.** 2006. "The Limits of Ex Post Implementation." *Econometrica* 74 (3): 585–610.
- Maskin, Eric S.** 1992. "Auctions and Privatization." In *Privatization: Symposium in Honor of Herbert Giersch*, edited by Horst Siebert, 115–36. Tübingen: Mohr (Siebeck).
- McLean, Richard, and Andrew Postlewaite.** 2004. "Informational Size and Efficient Auctions." *Review of Economic Studies* 71 (3): 809–27.
- Myerson, Roger.** 1981. "Optimal Auction Design." *Mathematics of Operations Research* 6: 53–73.
- Neeman, Zvika.** 2004. "The Relevance of Private Information in Mechanism Design." *Journal of Economic Theory* 117 (1): 55–77.
- Rochet, Jean-Charles.** 1987. "A Necessary and Sufficient Condition for Rationalizability in a Quasi-linear Context." *Journal of Mathematical Economics* 16 (2): 191–200.