

International Game Theory Review
© World Scientific Publishing Company

COOPERATIVE GAMES WITH INCOMPLETE INFORMATION: SOME OPEN PROBLEMS

FRANÇOISE FORGES*

*CEREMADE and LEDa, Université Paris-Dauphine,
Place du Maréchal de Lattre de Tassigny,
75775 Paris cedex 16 France.
francoise.forges@gmail.com
<http://sites.google.com/site/francoiseforges/>*

ROBERTO SERRANO†

*Department of Economics, Brown University,
Providence, RI 02912, U.S.A.
roberto_serrano@brown.edu
<http://www.econ.brown.edu/faculty/serrano/>*

Received (31 August 2011)

Revised (2 November 2011)

This is a brief survey describing some of the recent progress and open problems in the area of cooperative games with incomplete information. We discuss exchange economies, cooperative Bayesian games with orthogonal coalitions, and issues of cooperation in non-cooperative Bayesian games.

Keywords: Strategic externalities; informational externalities; exchange economies; cooperative games with orthogonal coalitions; non-cooperative Bayesian games.

Subject Classification: C71, C72, D51, D82.

1. Introduction

This is a brief survey that describes recent progress and lists some of the open problems in the area of cooperative games with incomplete information.^a Although a good number of papers have been written since the seminal work of Wilson [1978], much remains to be done. Game theory is the discipline that studies interactive strategic situations among players, in which externalities pervade. Its cooperative part uses models and solution concepts where the details of such interactions among coalitions of players, while acknowledged, are usually not described in the analysis.

*Member of Institut Universitaire de France.

†Research Associate, IMDEA-Social Sciences Institute, Madrid, Spain.

^aIn particular, our attempt here is not to be exhaustive. We refer the interested reader to the more comprehensive surveys mentioned at the beginning of Section 2.

Two kinds of externalities should be present in a cooperative game with incomplete information: *strategic externalities* and *informational externalities*. For a given coalition, the term *strategic externalities* refers to the actions of the complement set of players affecting the payoffs of the players in the coalition. In contrast, for a given coalition, *informational externalities* refers to the information held by the complement set of players affecting the payoffs of the coalition. With few exceptions, the contributions made so far have ignored strategic externalities altogether, as well as some informational externalities. This should already give the reader an idea of the vast number of open problems that lie ahead.

The survey is organized around models rather than around solution concepts. There are many conceptual issues in modeling cooperation among players with asymmetric information. Indeed, one cannot generally hope to derive an analog of the characteristic function under complete information. For instance, if the analysis is performed at the *interim* stage –that in which each player has received some private information–, for a given coalition, how much information is shared among the players in the coalition? Do they share none of the private information that each of the players has? Do they share it all? Do they share some, but not all of it? Each of these would potentially give rise to a different feasible set of payoffs for the coalition, thereby questioning the use of a single such set, as prescribed by the characteristic function. On the other hand, as is well-known, even if the derivation of a characteristic function were possible, the (possibly NTU) characteristic function does not suffice to capture some relevant aspects of the players' interaction, *e.g.*, strategic externalities across coalitions. For such cases, under complete information one can resort to the partition function. But in cooperative games with incomplete information, for the reasons just specified, one should also question the concept of the partition function.

Characteristic and partition functions alike have the advantage of abstracting from the specific details in the strategic interaction, yet we observe that addressing those details directly is important in cooperative games with asymmetric information, and perhaps that is the reason why researchers have turned to models in which such specifics of the interaction are tracked down (the consumption bundles in an exchange economy, or the actions taken in a collective decision problem or a game in strategic form). We turn to each of these models in the next pages. We start with exchange economies with asymmetric information, which do not involve any strategic interaction but allow for arbitrary forms of informational externalities. In section 3, we maintain these features in an abstract game theoretic framework, which can be viewed as the analog of the NTU characteristic function under incomplete information: Bayesian cooperative games with orthogonal coalitions. Finally, section 4 is devoted to cooperation in noncooperative Bayesian games, in which the players face both informational and strategic externalities.

2. Exchange Economies with Asymmetric Information

Cooperation under asymmetric information has been mostly studied in this model, which (1) is relevant in view of the connections between competitive equilibrium and solutions to cooperative games under complete information such as the core, and (2) is still manageable under incomplete information due to the absence of strategic externalities.^b

The basic model of an exchange economy with asymmetric information is as follows. Let T_i denote the (finite) set of agent i 's types. The interpretation is that $t_i \in T_i$ denotes the *private information* possessed by agent i . With $N = \{1, \dots, n\}$ as the finite set of agents, let $T_N = \prod_{i \in N} T_i$ denote the set of all information states. We will use the notation t_{-i} to denote $(t_j)_{j \neq i}$. Similarly $T_{-i} = \prod_{j \neq i} T_j$, and for each $S \subseteq N$, $T_S = \prod_{j \in S} T_j$ and $T_{-S} = \prod_{j \notin S} T_j$. We assume that agents have a common prior probability distribution q defined on T_N , and that no type is redundant, *i.e.*, $q(t_i) > 0$ for all $t_i \in T_i$ for all i .

There are three distinct timeframes at which one can perform the analysis, the *ex ante* stage, the *interim* stage and the *ex post* stage.^c Suppose that nature has chosen $t_N \in T_N$. At the *ex ante* stage, no agent $i \in N$ has been informed about her type. At the *interim* stage, each agent $i \in N$ knows her type, t_i , and hence, conditional probabilities will be important: for each $i \in N$ and $t_i \in T_i$, the conditional probability of $t_{-i} \in T_{-i}$, given t_i is denoted $q(t_{-i} | t_i)$. At the *ex post* stage, every agent $i \in N$ knows the information state t_N . We assume von Neumann-Morgenstern expected utility, where the probability weights are the ones just described at each of these three stages.

We assume that there are $|L| = l$, a finite number of commodities. The consumption set of agent i is $X_i = \mathbb{R}_+^l$. Agent i 's utility function in state t_N is denoted $u_i(x_i, t_N) : X_i \times T_N \mapsto \mathbb{R}$. We sometimes assume a quasilinear economy: commodity l is a numeraire, the consumption set is $X_i = \mathbb{R}_+^{l-1} \times \mathbb{R}$, and the utility function is $u_i(x_i, t_N) = v_i(x_i^{-l}, t_N) + x_i^l$. The endowment of agent i of type t_i is $\omega_i \in X_i$ (assumed to be independent of the state – with this assumption, all private information concerns agents' preferences and beliefs.)

We can now define an admissible exchange economy as

$$\mathcal{E} = \langle (u_i, X_i, \omega_i, T_i)_{i \in N}, q \rangle.$$

Note first how, as in standard Neoclassical exchange economies with complete information, strategic externalities (called consumption externalities in this context) are excluded. That is, agent i 's utility function depends on her own consumption bundle x_i only, not on anyone else's.

For coalition $S \subseteq N$, a feasible deterministic (state contingent) S -allocation, $x : T_N \mapsto \mathbb{R}^{lS}$ (where s denotes the cardinality of S), specifies a commodity bundle

^bSee Allen and Yannelis [2001] and Forges *et al.* [2002b] for previous surveys.

^cHere, we shall concentrate on the first two; see, *e.g.*, Einy *et al.* [2000a], Einy *et al.* [2000b] for results on the *ex post* stage.

for each consumer in S in each state such that $\sum_{i \in S} x_i(t_N) \leq \sum_{i \in S} \omega_i$ for all $t_N \in T_N$, and satisfying that $x(t_S, t'_{-S}) = x(t_S, t''_{-S})$ for all $t_S \in T_S$ and for all $t'_{-S}, t''_{-S} \in T_{-S}$.

Note also how the latter assumption is made to exclude one kind of information-externality across coalitions, *i.e.*, the set of feasible allocations to coalition S is independent of the information held by the complement $N \setminus S$. On the other hand, the information held by $N \setminus S$ may affect the utilities of agents in coalition S , as u_i depends on $t_N = (t_S, t_{-S})$, so informational externalities are present.

Denote by \mathcal{A}_S the set of feasible deterministic state contingent allocations of S . With confusion being avoided by the context, we also use \mathcal{A}_S to denote the set of feasible deterministic allocations in a given state: $\mathcal{A}_S = \{(x_i) \in \prod_{i \in S} X_i \mid \sum_{i \in S} x_i \leq \sum_{i \in S} \omega_i\}$. Similarly, deterministic state contingent N -allocations are simply referred to as deterministic allocations, and the set of such deterministic allocations is denoted by \mathcal{A}_N .

In environments in which the information state is not verifiable, not even at the *ex post* stage, it becomes necessary to impose incentive compatibility constraints.^d In coalitional settings such as the current one, the notion of incentive compatibility requires some discussion, particularly when it gets tangled with information transmission issues.

Let us begin again by considering deterministic allocations. Then, if agent i of type t_i pretends to be of type t'_i (while all other agents are truthful), she gets *interim* utility:

$$U_i(x, t'_i \mid t_i) = \sum_{t_{-i} \in T_{-i}} q(t_{-i} \mid t_i) u_i(x_i(t_{-i}, t'_i), (t_{-i}, t_i)).$$

As is standard, denote by $U_i(x \mid t_i) = U_i(x, t_i \mid t_i)$.

An S -allocation $x \in \mathcal{A}_S$ is *incentive compatible* (IC) if for every $i \in S$, and for every $t_i \in T_i$,

$$U_i(x \mid t_i) \geq U_i(x, t'_i \mid t_i)$$

for every $t'_i \in T_i \setminus \{t_i\}$. We denote the set of IC S -allocations by \mathcal{A}_S^* . An S -allocation x is *strictly IC* if all these inequalities are strict.

The *ex ante* stage is simpler in one important respect. With or without incorporating incentive constraints, there is a well defined characteristic function. One can specify feasible sets of *ex ante* expected utilities for each coalition. In this sense, there should be no doubt about what the core means in *ex ante* settings. An S -allocation is an *ex ante* objection to an N -allocation if every member of the coalition obtains a higher *ex ante* expected utility as a result. The *ex ante* core is the set of N -allocations that have no *ex ante* objections. The *ex ante* IC core is similarly de-

^dAs also illustrated in the next sections, with *ex post* verifiability, IC constraints are not needed, but information revelation is still an issue; other authors have relied on verifiability at the *interim* stage.

fined, but in it, both the N -allocations and the S -allocations that may qualify as *ex ante* objections are required to satisfy IC.

If one does not need to incorporate incentive constraints, the *ex ante* economy so described is simply an Arrow-Debreu economy. Under standard assumptions, the Arrow-Debreu equilibrium allocations are in the *ex ante* core, and as the economy gets replicated, the Debreu-Scarf argument applies, with just a reinterpretation of variables, to show that *ex ante* core allocations of replica economies converge to Arrow-Debreu equilibrium allocations. With incentive constraints, as shown in Forges *et al.* [2002a], the non-emptiness of the *ex ante* IC core is not so obvious. Even in quasilinear economies, while general sufficient conditions for non-emptiness can be found, there are also robust counterexamples to existence. The same goes for the core convergence question. As Forges *et al.* [2001] show, core equivalence does not generally hold, although a positive result can be established under certain conditions. One important question, still open, is what would be appropriate definitions of the *ex ante* core in contexts with strategic externalities.

At the *interim* stage we encounter the problem of the definition of the characteristic function, already described. The literature has studied the question of what should be an appropriate definition of the *interim* core for some time. The first definitions were provided by Wilson [1978]. Without relying on incentive compatible mechanisms, Wilson [1978] proposes the coarse core and the fine core. An S -allocation is a coarse objection to an N -allocation if there exists an informational event $E \subseteq T_N$ that is common knowledge among the players in coalition S over which each player in the coalition receives a higher *interim* expected utility. The coarse core is the set of N -allocations to which there are no coarse objections. An S -allocation is a fine objection to an N -allocation if there exists an informational event E such that each player in coalition S receives a higher expected utility conditional on E . The fine core is the set of N -allocations to which there are no fine objections.

Wilson's core concepts are a good starting point, in which the information transmission within the coalition is exogenous (either none in the case of a coarse objection, or any kind in the case of a fine objection). Wilson [1978] proves the non-emptiness of the coarse core and provides counterexamples to the existence of the fine core. Adding incentive compatibility to the analysis, Vohra [1999] also exhibits examples to show that the IC coarse core may be empty. Since the IC coarse core includes any notion of *interim* IC core in which some information transmission takes place, Vohra's result presents a real challenge to core theory.

More recently, based on tools of mechanism design, Serrano and Vohra [2007] propose a core concept that incorporates endogenous information transmission. They answer the question of how much information gets transmitted within a coalition by requiring that the event that the objection uses coincides with the event over which an equilibrium acceptance takes place for some Bayesian equilibrium of some communication mechanism used by the coalition. This gives rise to what we refer to as the core with respect to (w.r.t.) equilibrium blocking. This core, as

6 *Françoise Forges and Roberto Serrano*

a function of the class of mechanisms one allows, reduces to the credible core of Dutta and Vohra [2005], the virtual utility core of Myerson [2007] and the randomized mediated core of Serrano and Vohra [2007]; see also Okada [2009], who adapts the Serrano-Vohra equilibrium blocking framework to a class of dynamic bargaining games where proposals carry signaling value. In general, the exploration of different negotiation procedures should be an important vehicle to tackle questions of coalition formation and stable payoff distribution. These questions, comprising a generalized Nash program, remain almost entirely unexplored in these settings. Given that we foresee this as an active area in the coming years, we proceed to details.

Information transmission concerns ruling out some states as impossible, through the identification of smaller informational events. For an event $E \subseteq T_N$ and $t_i \in T_i$, let

$$E_{-i}(t_i) = \{t_{-i} \in T_{-i} \mid (t_i, t_{-i}) \in E\}$$

and

$$E_i = \{t_i \in T_i \mid E_{-i}(t_i) \neq \emptyset\}.$$

Consider an allocation rule $x \in \mathcal{A}_N$, agent i of type t_i and an event E . Suppose $q(E_{-i}(t_i)) > 0$. Then agent i 's expected utility conditional on E can be expressed as:

$$U_i(x|t_i, E) = \sum_{t_{-i} \in E_{-i}(t_i)} \frac{q(t_{-i} \mid t_i)}{q(E_{-i}(t_i)|t_i)} u_i(x_i(t_{-i}, t_i), (t_{-i}, t_i)).$$

The corresponding expected utility (conditional on E) if type t_i pretends to be of type t'_i , while the others are truth-telling, is:

$$U_i(x, t'_i \mid t_i, E) = \sum_{t_{-i} \in E_{-i}(t_i)} \frac{q(t_{-i} \mid t_i)}{q(E_{-i}(t_i)|t_i)} u_i(x_i(t_{-i}, t'_i), (t_{-i}, t_i)).$$

Given $E \subseteq T_N$, an S -allocation $x \in \mathcal{A}_S$ is *IC over E* if for every $i \in S$ and for every $t_i, t'_i \in E_i$:

$$U_i(x \mid t_i, E) \geq U_i(x, t'_i \mid t_i, E).$$

The interpretation of the status quo $x \in \mathcal{A}_N$ is that in every state t_N , the outcome is $x(t_N)$, unless there is an agreement to change it. In particular, if there is an attempt to change it but the attempt fails, the outcome in state t_N is $x(t_N)$. As in the case of complete information, the failure of an objection makes everyone involved in the objection believe that the original status quo prevails. With incomplete information, this means that no learning takes place from the failed attempt. Of course, there are good reasons to explore alternative assumptions, and here the generalized Nash program hinted at above should play an important role.

In addition, though, a coalition may contemplate the use of random mechanisms in order to object to a status quo allocation. Thus, to potentially destabilize a status

quo, now we also consider random plans. A random coalitional plan μ consists of a collection of probability distributions over feasible allocation rules for various coalitions: $\mu(S, y^S, t_N)$, where $y^S \in \mathcal{A}_S$, denotes the probability with which coalition $S \subseteq N$ is receiving $y^S \in \mathcal{A}_S$ when the (reported) state is $t_N \in T_N$. We shall say that such a random plan is measurable with respect to coalitional information, or simply *measurable*, if for every S , for every y^S and for every t_S , $\mu(S, y^S, (t_S, t'_{-S})) = \mu(S, y^S, (t_S, t''_{-S}))$ for every $t'_{-S}, t''_{-S} \in T_{-S}$. Otherwise, we shall say that the random plan is *non-measurable*. While such non-measurabilities are in principle possible, we shall always require that, if P is the union of all coalitions in the support of μ , $\mu(S, y^S, (t_P, t'_{-P})) = \mu(S, y^S, (t_P, t''_{-P}))$ for all (S, y^S) in the support of μ , all $t_P \in T_P$ and all $t'_{-P}, t''_{-P} \in T_{-P}$.

We think of the random plan as a mediated communication mechanism used by the members of the coalition P , which follows a pre-specified IC status quo with its truthful reports behind it. The timing of the random plan mediated communication mechanism is as follows:

- Stage 1: Types are reported after players have been informed about μ . Each player is privately informed about the instances in which μ calls her to act. The type reports of this stage are used only if and when the blocking plan is implemented.
- Stage 2: players are invited to participate in the blocking plan, *i.e.*, “phone calls” are made by the blocking mediator, taking into account the probabilities $\mu(S, y^S, t_N)$, which use already the types reported within μ . Note in particular how a non-measurable plan is perfectly possible, allowing information transmission from coalition to coalition within μ . On the other hand, the plan must be measurable with respect to the information of the union of coalitions within μ .
- Stage 3: each agent in the support of μ is asked to either accept or reject the blocking plan, and he does so in private communications with the blocking mediator. An allocation proposed to coalition S is implemented if and only if every agent in S accepts the random plan; otherwise, the status quo is implemented.

An equilibrium rejection of a status quo is a Bayesian equilibrium of this communication mechanism in which the random blocking plan is accepted over an event that has positive probability. In particular, in the Bayesian equilibrium: (i) the resulting allocation rule is IC over the equilibrium rejection event; (ii) each type that accepts being part of the plan prefers accepting it to rejecting it; and (iii) each type that rejects weakly prefers rejecting it to accepting it. The core w.r.t. equilibrium blocking is the set of IC allocations against which there is no equilibrium rejection. As shown in Serrano and Vohra [2007], it turns out that, if one allows only deterministic blocking plans that have a single coalition in the support, this core is the Dutta-Vohra credible core. If one allows random measurable mechanisms, this core is the Myerson virtual utility core; and if one allows the use of unrestricted random

mechanisms, it is the Serrano-Vohra randomized mediated core.

Note that, by building these objections on the basis of equilibrium rejections, information is endogenously transmitted within the random blocking plan. Given the equilibrium rejection event, each type that is part of the event wishes to be part of the objection, while types that are not part of the event wish to stay out and not misrepresent their information to participate in that plan.

Given the difficulties that follow the result in Vohra [1999], one needs to find alternative assumptions to establish existence of these *interim* cores. For quasilinear economies, and if one allows average feasibility in the numeraire (while exact feasibility is required for the other goods), Myerson [2007] shows non-emptiness of the virtual utility core.^e Since his argument does not rely on the measurability assumption on blocking plans, it follows that the core w.r.t. equilibrium blocking of unrestricted mechanisms is also non-empty under the same assumptions. Similarly, also in quasilinear economies, if one assumes the existence of an IC *ex post* core allocation, Dutta and Vohra [2005] show non-emptiness of the credible core, and again, their arguments can be adapted to show that the core w.r.t. equilibrium blocking of unrestricted mechanisms is also non-empty under their assumptions. Understanding the limitations that these and other conditions may have, the search for alternative interesting scenarios in which non-emptiness of the core obtains should also be part of the research agenda.

The question of core convergence at the *interim* stage leads in general to a robust negative answer. As shown in Serrano *et al.* [2001], the coarse core is “too large” and does not converge to any price equilibrium notion. This result extends to a number of other core concepts in which the information is transmitted exogenously, and it is also robust to adding IC to the feasibility requirements for coalitional allocations. The fine core, at the other extreme, is often “too small,” as it may be empty in economies in which standard price equilibrium concepts exist. In quasilinear economies, though, Kamishiro and Serrano [2011] show that the core w.r.t. equilibrium blocking of unrestricted mechanisms converges to the corresponding set of *ex post* Walrasian equilibrium allocations, while the Myerson virtual utility core does not. Thus, the measurability of the random blocking plans is a friction that prevents the full information revelation that one finds in a rational expectations equilibrium. These results are also robust to environments in which the IC constraints can be dropped. Given how important it is to understand information flows in markets, more research along these lines is definitely needed.^f

A different approach is based on the concepts of informational size and signal-based replicas, proposed in McLean and Postlewaite [2002] and McLean and Postlewaite [2005]. Although they were originally used in *ex ante* settings to establish

^eUnder standard assumptions, his proof is based on a balancedness condition, which, as he states, is met in linear exchange economies, but also in two-player Bayesian games.

^fIt is worth mentioning the results in de Clippel [2007], who based on the competitive screening framework of Rothschild-Stiglitz and in an *ex post* verifiable setting, shows convergence of his type-agent core to an equilibrium notion proposed in Wilson [1978].

existence and convergence results for the core, they have also begun to be used in *interim* economies (Kamishiro [2011], Kamishiro and Serrano [2011]). One should expect that more results should emerge using these ideas.

3. Bayesian Cooperative Games with Orthogonal Coalitions

Myerson [1984b] defines a general *cooperative game with incomplete information* as $\Gamma = \langle N, (D_S)_{S \in \mathcal{S}}, (T_i, u_i)_{i \in N}, q \rangle$, where, as in the previous section, N is the set of players, T_i is the set of types of every player $i \in N$, and q is the prior probability distribution over $T_N = \prod_{i \in N} T_i$; D_S denotes the set of feasible decisions for every coalition S , \mathcal{S} is the power set of N , and utility functions are now defined as $u_i : D_N \times T_N \rightarrow \mathbb{R}$ for every player i , $i \in N$. By assumption, all the sets D_S and T_i are finite and, for every coalitions R, S such that $R \cap S = \emptyset$, $D_R \times D_S \subseteq D_{R \cup S}$.

This model allows for both informational and strategic externalities, since the utilities of the members of a coalition S may depend on the types and the decisions of the players in $N \setminus S$. In particular, even under complete information, namely, if T_i is a singleton for every i , the model is more general than an NTU characteristic function. Following Myerson [1984b], purely informational externalities can be captured by assuming that the cooperative game Γ has *orthogonal coalitions*, namely that

$$u_i((d_S, d_{-S}), t_N) = u_i((d_S, d'_{-S}), t_N)$$

for every S , $i \in S$, $d_S \in D_S$, $d_{-S}, d'_{-S} \in D_{-S}$ and $t_N \in T_N$.

A *mechanism* for coalition S is a mapping $\mu_S : T_S \rightarrow \Delta(D_S)$, where $T_S = \prod_{i \in S} T_i$, as in section 2, and $\Delta(D_S)$ denotes the set of probability distributions over D_S . The interpretation is that if S were to form, it would make a decision in D_S randomly, as a function of its members' information.[§] In order to extend standard cooperative solution concepts, such as the core or the Shapley value, to games with incomplete information, we will typically assume that only the grand coalition's mechanism μ_N will be implemented. In this case, for every $S \neq N$, the mechanism μ_S stands as a *threat*, to be carried out only if $N \setminus S$ refuses to cooperate with S .

The issue of coalitional threats does not arise in two standard models of cooperative games with incomplete information (see Myerson [1991], chapter 10)). In a *Bayesian collective choice problem* $\Gamma^c = \langle N, D_N, (T_i, u_i)_{i \in N}, q \rangle$, a set of joint decisions, D_N , is described for the grand coalition N only. In a *Bayesian bargaining problem* $\Gamma^b = \langle N, D_N, d^*, (T_i, u_i)_{i \in N}, q \rangle$, a disagreement decision $d^* \in D_N$ is made if the players fail to agree on a joint decision in D_N . Both Γ^c and Γ^b implicitly have orthogonal coalitions.

[§]As recalled in section 2, deterministic mechanisms are common in exchange economies, even if random mechanisms have also been considered in that framework (see, *e.g.*, Forges *et al.* [2002b]). In the abstract game theoretic framework of sections 3 and 4, the tradition inherited from complete information is to allow for randomization over finitely many feasible decisions.

Whatever the basic model - Γ , Γ^c or Γ^b - a mechanism for the grand coalition N is a mapping $\mu_N : T_N \rightarrow \Delta(D_N)$. In this section, we assume that types are not verifiable and that the players negotiate at the *interim* stage. Hence, we require that μ_N be *incentive compatible*, which can be defined as in section 2. μ_N is *incentive efficient* if and only if μ_N is IC and there does not exist any other IC mechanism giving a higher *interim* expected utility to all types t_i of all players $i \in N$. As explained in detail in Holmström and Myerson [1983], this extension of *Pareto efficiency* to cooperative games with incomplete information makes sense if the players must choose a mechanism μ_N after having learnt their type.^h Incentive efficient mechanisms are the optimal solutions of a linear programming problem and can be usefully characterized through the *dual* of this problem, which leads to *virtual utilities* for the players (see Myerson [1991], section 10.5). As already suggested by Myerson [2007]'s core concept (which is introduced for exchange economies in section 2 but applies to any Bayesian cooperative game with orthogonal coalitions) and as will be further illustrated below, virtual utilities can be a useful tool to define and characterize cooperative solution concepts under incomplete information.

In a Bayesian collective choice problem Γ^c , it is understood that participation is compulsory, but in Γ and Γ^b , *interim individual rationality* is a property to consider, in addition to incentive efficiency. If Γ has orthogonal coalitions, the definition is straightforward. In particular, in Γ^b , μ_N is *interim* individually rational if and only if μ_N gives at least as high *interim* expected utility as the disagreement outcome d^* to all types t_i of all players $i \in N$. We defer a possible general definition of *interim* individual rationality to the next section (see (1)).

Harsanyi and Selten [1972] proposed an extension of the Nash solution to Bayesian bargaining problems, which was later modified by Myerson [1979]. However, Myerson himself acknowledges that this solution concept has several drawbacks (see Myerson [1984b]). Building on Myerson [1983], which assumes that one of the players has the whole bargaining power (see also Myerson [1991], section 10.7), Myerson [1984a] proposes a new generalization of the Nash bargaining solution for two player games with incomplete information and equal bargaining power.

Myerson [1984a] identifies two axioms for incentive efficient mechanisms: *random dictatorship* and *extension* (see also Myerson [1991], section 10.8). These axioms turn out to be satisfied by many solution correspondences \mathcal{B} , which associate incentive efficient mechanisms with every two-person Bayesian bargaining problem. A *neutral bargaining solution* to Γ^b is defined as any (incentive efficient) mechanism μ such that, for every correspondence \mathcal{B} that satisfies the axioms, $\mu \in \mathcal{B}(\Gamma^b)$. Myerson [1984a] proves that (i) neutral bargaining solutions satisfy the two axioms; (ii) the set of neutral bargaining solutions is nonempty for every two-person Bayesian bargaining problem; (iii) if Γ^b is a bargaining problem with complete information, μ is a neutral bargaining solution of Γ^b if and only if μ is the Nash bargaining solution

^hAs an illustration, N -allocations in the IC coarse core (see section 2) are incentive efficient.

of Γ^b . Myerson [1984a] also characterizes neutral bargaining solutions in terms of the players' virtual utilities. Myerson [1985] solves two examples in full details.

Myerson [1984b] explains that, in the presence of incentive constraints, transferable utility at the *interim* stage cannot be modelled without referring to an appropriate NTU framework (see also Forges *et al.* [2002a]). He proposes a solution concept that extends Myerson [1984a]'s bargaining solution and Shapley [1969]'s NTU value (the latter being defined for cooperative games with complete information).¹ Myerson's value is not derived from axioms, but can be computed by performing the fictitious transfers behind Shapley [1969]'s value in terms of the players' virtual utilities. As the Shapley NTU value, Myerson's solution concept is not single valued. The solutions consist of incentive compatible mechanisms μ_N for which there exist virtual utility scales such that μ_N would be equitable and efficient if interpersonal comparisons could be made in terms of these virtual utility scales. In the same way as the solution concepts developed in Holmström and Myerson [1983], Myerson [1983] and Myerson [1984a], Myerson [1984b]'s value reflects the fact that players negotiate at the *interim* stage, so that they should take account of the information that they reveal by proposing, accepting or refusing a particular IC mechanism.

Myerson's value involves the identification of a rational threat mechanism μ_S for every coalition S . Any cooperative game Γ with incomplete information has a value, which is *interim* individually rational (in the sense of (1), see below). Myerson [1984b] further conjectures that his solution set should be generically finite. However, a satisfactory justification of the rational threats, in terms of blocking plans as in section 2, is only available when Γ has orthogonal coalitions. Even in this case, it is understood that every coalition S commits to a threat mechanism μ_S in view of full cooperation, before knowing which coalition structure will effectively prevail, so that μ_S need not be incentive compatible or equitable for $S \neq N$.

de Clippel [2005] explicitly computes Myerson [1984b]'s value in a three-person example in which player 1 is the only informed player and player 3's only contribution is to facilitate the fulfilment of player 1's incentive constraints in the grand coalition. Contrary to what intuition would suggest, Myerson's value treats player 3 as a dummy. Nonetheless, de Clippel [2005] shows that player 3 gets a positive payoff in a procedure of random arrival, which can be viewed as a generalization of Myerson [1984b]'s random dictatorship to games with more than two players. de Clippel formalizes the procedure in a game in extensive form, which accounts for the players' negotiation possibilities. As a topic for further research, Myerson [1984b]'s value should be further investigated and challenged.

The previous paragraphs illustrate that, even in the absence of strategic externalities, the axiomatic approach to cooperation, which was so fruitful under complete information, to date renders much less clear conclusions when negotiation takes

¹Under complete information, several extensions of the Shapley [1953] value have been proposed in the NTU case: Harsanyi [1959], Harsanyi [1963], Maschler and Owen [1992], Shapley [1969].

place between privately informed players. As recalled above, Myerson [1984a] proposes a partial axiomatization of a bargaining solution in the case of two equally powerful players. Focusing on the issue of information revelation at the negotiation stage, de Clippel and Minelli [2004] pursue this analysis. They provide cooperative and noncooperative characterizations of Myerson [1983] and Myerson [1984a]'s solutions under the additional assumption that types become verifiable at the stage where a mechanism is used to make decisions. de Clippel and Minelli [2004] propose in particular a refinement of Wilson [1978]'s coarse core. An obvious open problem is the extension of these results when types remain unverifiable, even at the decision stage. In any case, this contribution, as other recent ones, *e.g.*, Serrano and Vohra [2007], detailed in section 2, and de Clippel [2005], indicates that the analysis of simple, explicit negotiation procedures is a promising approach given the state of the art.

4. Cooperation in Noncooperative Bayesian Games

Under complete information, the most natural way to take strategic externalities into account is to start with a strategic form game. As recalled in Myerson [1991] (section 8.5), one can represent the effects of binding agreements in a two-person strategic form game by various bargaining problems, leading for instance to Nash [1953]'s bargaining solution with variable threats. Harsanyi [1959] and Harsanyi [1963] extended the latter solution to n -person games. Many other approaches are conceivable (see, *e.g.*, Ray and Vohra [1997], Ray [2007] and Hart and Mas-Colell [2010]). Under incomplete information, the analog of a strategic form game is a *noncooperative Bayesian game* $G = \langle N, (T_i, D_i, u_i)_{i \in N}, q \rangle$, where the parameters have the same interpretation as in the previous section, but decision sets D_i are now described for individual players $i \in N$ only and D_N is defined as $D_N = \prod_{i \in N} D_i$ in the utility functions u_i , $i \in N$.^j

As a first step, let us think of a possible characterization of the feasible set when the players can cooperate in G . As in the previous sections, a mechanism for the set N of all players is a mapping $\mu_N : T_N \rightarrow \Delta(D_N)$. As in section 2, we denote as $U_i(\mu_N | t_i)$ the *interim* expected payoff of player i of type t_i when decisions are made according to μ_N , namely

$$U_i(\mu_N | t_i) = \sum_{t_{-i} \in T_{-i}} q(t_{-i} | t_i) \sum_{d_N \in D_N} \mu_N(d_N | t_i, t_{-i}) u_i(d_N, (t_i, t_{-i})).$$

Following Myerson [1991] (section 6.6), μ_N is *interim individually rational* if and only if, for every player $i \in N$, there exists a mechanism $\nu_{-i} : T_{-i} \rightarrow \Delta(D_{-i})$ for

^jThe current model is thus a particular case of the general cooperative game $\Gamma = \langle N, (D_S)_{S \in \mathcal{S}}, (T_i, u_i)_{i \in N}, q \rangle$ of the previous section, with $D_S = \prod_{i \in S} D_i$, for every $S \subseteq N$. As pointed out above, Myerson [1984b]'s value applies to this model but its relevance is disputable when coalitions are not orthogonal.

the players in $N \setminus i$ such that, for every $t_i \in T_i$,

$$U_i(\mu_N | t_i) \geq \max_{d_i \in D_i} \sum_{t_{-i} \in T_{-i}} q(t_{-i} | t_i) \sum_{d_{-i} \in D_{-i}} \nu_{-i}(d_{-i} | t_{-i}) u_i((d_i, d_{-i}), (t_i, t_{-i})) \quad (1)$$

Observe that, under complete information, in the case of two players, the corresponding individual rationality level of player i is the standard, fixed threats, minmax level, namely $\min_{\nu_{-i} \in \Delta(D_{-i})} \max_{d_i \in D_i} \sum_{d_{-i} \in D_{-i}} \nu_{-i}(d_{-i}) u_i(d_i, d_{-i})$.

Let $\mathcal{F}(G)$ be the set of all *interim* expected payoffs that can be achieved by means of an incentive compatible, *interim* individually rational mechanism μ_N . The idea is that the players are asked to reveal their types once and for all, in view of the cooperative agreement μ_N . Hence, μ_N , but none of the ν_{-i} 's, is required to be IC. The set $\mathcal{F}(G)$ can be interpreted as the set of all *interim* expected payoffs that the players can achieve by signing voluntary, binding agreements at the *interim* stage of the Bayesian game G (see Myerson [1991], sections 6.1 and 6.6, and Forges [2011]). A formal statement of this property involves some subtleties, even under complete information (see Kalai *et al.* [2010]). It seems natural to define a cooperative solution for G as an element of $\mathcal{F}(G)$. For instance, as pointed out above, Myerson [1984b]'s value is *interim* individually rational in the sense of (1).^k

Kalai, A. and Kalai, E. [2011] propose a new solution, the cooperative-competitive (“coco”) solution, for two-person Bayesian games with transferable utility. The coco solution is defined by an explicit, simple formula, in terms of the players' *ex ante* expected payoffs. It is based on a decomposition of the game G into two components, summarizing (i) what the players can get as a team and (ii) their respective payoff advantages.^l The coco solution is *ex post* Pareto (namely, “first best”) efficient and is monotonic with respect to both individual decision sets and individual information sets. By identifying some further axioms, Kalai, A. and Kalai, E. [2011] provide a complete characterization of their solution.

In the spirit of the Nash program, Kalai, A. and Kalai, E. [2011] also investigate to which extent the coco solution can be implemented at the *ex ante* or the *interim* stage of an appropriately designed game extending G . *Interim* implementation is particularly delicate, since we know, from Myerson and Satterthwaite [1983], that there is a tension between *ex post* Pareto efficiency, incentive compatibility and *interim* individual rationality. In order to show that coco can be implemented at the *ex ante* (resp., *interim*) stage, Kalai, A. and Kalai, E. [2011] assume that G takes place in an environment of “weakly (resp., strongly) revealed payoffs”, in which, at

^kTo see this, let μ_S be the rational threat of coalition S associated with the underlying value, take $\nu_{-i} = \sum_{S \ni i} \frac{(|S|-1)! (|N|-|S|)!}{|N|!} (\mu_{S \setminus i}, \mu_{N \setminus S})$ and use lemma 2 in Myerson [1984b].

^lThe coco solution for player i is easily computed from the parameters of the Bayesian game G , with $N = \{1, 2\}$, as $\frac{1}{2} \sum_t q(t) \max_d [u_1(d, t) + u_2(d, t)] + \frac{1}{2} \text{val}_i(G^{ad})$ where $t = (t_1, t_2)$ and $d = (d_1, d_2)$, G^{ad} is the zero-sum Bayesian game defined as G except that player i 's utility function is $u_i - u_{-i}$ and $\text{val}_i(G^{ad})$ denotes the *ex ante* expected value of G^{ad} for player i . In particular, under complete information, the coco solution is equal to the TU bargaining solution with variable threats.

the end of the game, types become verifiable to some extent.^m The latter assumption is similar to the one made in de Clippel and Minelli [2004], but as seen above, the latter paper builds on Myerson [1984a], in which, unlike in Kalai, A. and Kalai, E. [2011], utilities are nontransferable and the disagreement decision is fixed. On the other hand, de Clippel and Minelli [2004] address the issue of the revelation of information at the negotiation stage, while Kalai, A. and Kalai, E. [2011]'s *interim* implementation procedure does not allow the players to choose among mechanisms.

Myerson [1984b]'s value applies in particular to two-person Bayesian games and extends Nash's bargaining solution with variable threats. Myerson's value deeply differs from the coco solution since it reflects bargaining over mechanisms at the *interim* stage and is thus "second best" efficient by nature. Nevertheless, a precise comparison of the treatment of threats in the two papers could be helpful.

Kalai, A. and Kalai, E. [2011] point out the following open problems: extension of the coco value in the TU case for more than two players and in the NTU case for two players or more, and construction of a second best analog of the coco value for environments that do not satisfy *ex post* observability conditions.

Unlike several contributions analyzed in section 2, the papers surveyed up to now in sections 3 and 4 do not impose any IC requirement on the mechanism μ_S to be chosen by a coalition $S \neq N$, either because they focus on the two-person case or because they implicitly assume a specific timing in the coalitions' commitment process. The formulation of meaningful incentive compatibility conditions for coalitions S (such that $2 < |S| < |N|$) in a Bayesian noncooperative game can be a real challenge. In the absence of an *ad hoc* scenario, we face a circular problem since incentives inside a coalition should take into account the commitments to threats, and the ability of a coalition to commit to threats depends on its members' information.

Biran and Forges [2011] show that the previous problem can be avoided when informational externalities are limited, *i.e.*, in a Bayesian game G with independent private values, namely in which $q = \bigotimes_{i \in N} q_i$ for some $q_i \in \Delta(T_i)$, $i \in N$, and $u_i(d_N, (t_i, t_{-i})) = u_i(d_N, t_i)$ for every $i \in N$, $d_N \in D_N$, $(t_i, t_{-i}) \in T_N$. Biran and Forges [2011] assume that utilities are transferable and that every coalition S commits to a mechanism μ_S at the *ex ante* stage (of course, μ_S is used at the *interim* stage).ⁿ For every partition Π of the players, they define a *coalitional equilibrium*^o relative to Π as a Bayesian Nash equilibrium $(\mu_S)_{S \in \Pi}$ of the auxiliary game in which the players are the coalitions S in the partition Π , with private information in T_S .

^mThe coco solution is formally defined only at the *ex ante* stage so that its *interim* individual rationality cannot be assessed directly, that is, without specifying an *interim* protocol. To guarantee *interim* participation in a procedure leading to the coco value, the key is to design transfers that depend on true types (as opposed to reported types in Myerson and Satterthwaite [1983]).

ⁿGiven the TU framework, μ_S determines *ex post* balanced transfers in addition to a decision in D_S .

^oRay and Vohra [1997] and Ray [2007] introduced this solution concept for games with complete information.

A coalitional equilibrium $(\mu_S)_{S \in \Pi}$ is incentive compatible if and only if, for every $S \in \Pi$, given that the players in $N \setminus S$ make decisions according to $(\mu_R)_{R \in \Pi, R \neq S}$, μ_S is IC in the usual sense. Observe that, if coalition S looks for a “credible” threat given the coalition structure Π , it must take the behavior of the players in $N \setminus S$ into account, even if S is not subjected to incentive constraints (e.g, under complete information, see Ray and Vohra [1997], or if types are verifiable as early as at the *interim* stage). The previous definition illustrates that incentive compatibility for coalitions S such that $2 < |S| < |N|$ is easy to formulate in a context where coalitions play best responses to each other.

By relying on the techniques of d’Aspremont and Gérard-Varet [1979] and d’Aspremont and Gérard-Varet [1982], Biran and Forges [2011] prove that *every* coalitional equilibrium can be made incentive compatible by means of an appropriate transfer scheme.^P For every partition Π of the players, every coalitional equilibrium relative to Π generates a cooperative game in partition form. As in section 2 for exchange economies, once the *ex ante* cooperative game is well-defined (here, as a partition form game), it can be solved as under complete information. Proceeding in this way, Biran and Forges [2011] propose a core solution concept which, recalling the core concepts introduced in section 2, can be interpreted as an *ex ante* IC, sequentially rational, α -core, the sequential rationality being inherited from the absence of incredible threats in a coalitional equilibrium. This core is shown to be nonempty in standard first price and second price auctions. Independent private values in the basic game G and *ex ante* commitment of coalitions play a crucial role in the previous results and relaxation of these assumptions is an obvious topic for future research. Furthermore, even if the previous core concept captures some form of sequentially rational threats, it assumes, as the traditional α -core, that (blocking) coalitions are cautious. It thus leaves open the question of the “equilibrium” partition formed by the players outside a blocking coalition.

Acknowledgments

We thank Geoffroy de Clippel, Yusuke Kamishiro and two anonymous referees for helpful comments.

References

- Allen, B. and Yannelis, N. C. [2001] Differential information economies: introduction, *Economic Theory* **18**, 263–273.
- d’Aspremont, C. and Gérard-Varet, L.-A. [1979] Incentives and incomplete information, *Journal of Public Economics* **11**, 25–45.
- d’Aspremont, C. and Gérard-Varet, L.-A. [1982] Bayesian incentive compatible beliefs, *Journal of Mathematical Economics* **10**, 83–103.

^PForges *et al.* [2002a] already proved a similar result in the absence of strategic externalities. As in Kalai, A. and Kalai, E. [2011], first best efficiency is achieved.

- Biran, O. and Forges, F. [2011] Core-stable rings in auctions with independent private values, *Games and Economic Behavior* **73**, 52–64.
- de Clippel, G. [2005] Values for cooperative games with incomplete information: an eloquent example, *Games and Economic Behavior* **53**, 73–82.
- de Clippel, G. [2007] The type-agent core of exchange economies with asymmetric information, *Journal of Economic Theory* **135**, 144–158.
- de Clippel, G. and Minelli, E. [2004] Two-person bargaining with verifiable types, *Journal of Mathematical Economics* **40**, 799–813.
- Dutta, B. and Vohra, R. [2005] Incomplete information, credibility and the core, *Mathematical Social Sciences* **50**, 148–165.
- Einy, E., Moreno, D. and Shitovitz, B., [2000a] Rational expectations equilibria and the ex-post core of an economy with asymmetric information, *Journal of Mathematical Economics* **34**, 527–535.
- Einy, E., Moreno, D. and Shitovitz, B., [2000b] On the core of an economy with differential information, *Journal of Economic Theory* **94**, 262–270.
- Forges, F., [2011] Bayesian games with contracts, *mimeo*, Université Paris-Dauphine.
- Forges, F., Heifetz, A. and Minelli, E., [2001] Incentive compatible core and competitive equilibria in differential information economies, *Economic Theory* **18**, 349–365.
- Forges, F., Mertens, J.-F. and Vohra, R. [2002] The *ex ante* incentive compatible core in the absence of wealth effects, *Econometrica* **70**, 1865–1892.
- Forges, F., Minelli, E. and Vohra, R. [2002] Incentives and the core of an exchange economy: a survey, *Journal of Mathematical Economics* **38**, 1–41.
- Harsanyi, J. C. [1959] A bargaining model for the cooperative n-person game in Tucker, A. and D. Luce (eds), *Contributions to the Theory of Games IV*, Princeton University Press.
- Harsanyi, J. C. [1963] A simplified bargaining model for the n-person cooperative game, *International Economic Review* **4**, 194–220.
- Harsanyi, J. C. and Selten, R. [1972] A generalized Nash solution for two-person bargaining games with incomplete information, *Management Science* **18**, 80–106.
- Hart, S. and Mas-Colell, A. [2010] Bargaining and cooperation in strategic form games, *Journal of the European Economic Association* **8**, 7–33.
- Holmström, B. and Myerson, R. B. [1983] Efficient and durable decision rules with incomplete information, *Econometrica* **51**, 1799–1819.
- Kalai, A. and Kalai, E. [2011] On semi-cooperative games with incomplete information, *mimeo*, Northwestern University.
- Kalai, A., Kalai, E., Lehrer, E. and Samet, D. [2010] A commitment Folk theorem, *Games and Economic Behavior* **69**, 127–137.
- Kamishiro, Y. [2011] Informational size and the incentive compatible coarse core in quasilinear economies, *Games and Economic Behavior* **71**, 513–520.
- Kamishiro, Y. and Serrano, R. [2011] Equilibrium blocking in large quasilinear economies, *Mathematics of Operations Research* **36**, 552–567.
- Maschler, M. and Owen, G. [1992] The consistent Shapley value for games without side-payments, in: Selten, R. (ed.), *Rational Interaction*, Springer-Verlag, 5–11.
- McLean, R. P. and Postlewaite, A. [2002] Informational size and incentive compatibility, *Econometrica* **70**, 2421–2453.
- McLean, R. and Postlewaite, A. [2005] Core convergence with asymmetric information, *Games and Economic Behavior* **50**, 58–78.
- Myerson, R. B. [1979] Incentive compatibility and the bargaining problem, *Econometrica* **47**, 61–73.
- Myerson, R. B. [1983] Mechanism design by an informed principal, *Econometrica* **51**,

- 1767–1797.
- Myerson, R. B. [1984a] Two-person bargaining problems with incomplete information, *Econometrica* **52**, 461–487.
- Myerson, R. B. [1984b] Cooperative games with incomplete information, *International Journal of Game Theory* **13**, 69–86.
- Myerson, R. B. [1985] Analysis of two bargaining problems with incomplete information, in Roth, A. (ed.), *The Shapley Value*, Cambridge University Press, 115–147.
- Myerson, R. B. [1991] Game theory: analysis of conflict, Harvard University Press.
- Myerson, R. B. [2007] Virtual utility and the core for games with incomplete information, *Journal of Economic Theory* **136**, 260–285.
- Myerson, R. and Satterthwaite, M. [1983] Efficient mechanisms for bilateral trading, *Journal of Economic Theory* **29**, 265–281.
- Nash, J. F. [1950] The bargaining problem, *Econometrica* **18**, 155–162.
- Nash, J. F. [1953] Two-person cooperative games, *Econometrica* **21**, 128–140.
- Okada, A. [2009] Non-cooperative bargaining and the incomplete information core, *mimeo*, Hitotsubashi University.
- Ray, D. [2007] A game-theoretic perspective on coalition formation, Oxford University Press.
- Ray, D. and Vohra, R. [1997] Equilibrium binding agreements, *Journal of Economic Theory* **73**, 30–78.
- Serrano, R. and Vohra, R. [2007] Information transmission in coalitional voting games, *Journal of Economic Theory* **134**, 117–137.
- Serrano, R., Vohra, R. and Volij, O. [2001] On the failure of core convergence in economies with asymmetric information, *Econometrica* **69**, 1685–1696.
- Shapley, L. S. [1953] A value for n-person games, in Kuhn, H. and A. Tucker (eds), *Contributions to the Theory of Games II*, Princeton University Press, 307–317.
- Shapley, L. S. [1969] Utility comparisons and the theory of games, in: Guilbaud, G. Th. (ed.), *La Décision*, CNRS, 251–263. Reprinted in Roth, A. (ed.), *The Shapley value*, Cambridge University Press, 307–319.
- Vohra, R. [1999] Incomplete information, incentive compatibility and the core, *Journal of Economic Theory* **86**, 123–147.
- Wilson, R. [1978] Information, efficiency and the core of an economy, *Econometrica* **46**, 807–816.