



A generalization of the Harsanyi NTU value to games with incomplete information

Andrés Salamanca¹ 

Accepted: 19 May 2019 / Published online: 27 May 2019
© Springer-Verlag GmbH Germany, part of Springer Nature 2019

Abstract

In this paper, we introduce a solution concept generalizing the Harsanyi non-transferable utility (NTU) value to cooperative games with incomplete information. The so-defined *S-solution* is characterized by virtual utility scales that extend the Harsanyi-Shapley fictitious weighted-utility transfer procedure. We construct a three-player cooperative game in which Myerson's (Int J Game Theory 13(2):69–96, 1984a) generalization of the Shapley NTU value does not capture some “negative” externality generated by the adverse selection. However, when we explicitly compute the *S-solution* in this game, it turns out that it prescribes a more intuitive outcome which takes into account the above mentioned informational externality.

Keywords Cooperative games · Incomplete information · Virtual utility

1 Introduction

The *value* is a central solution concept in the theory of cooperative games. Introduced by Shapley (1953) for the study of games with transferable utility (TU), the value has been extended in different ways to general games with nontransferable utility (NTU); some of the most notable NTU values are due to Harsanyi (1963) and Shapley (1969)¹.

¹ The Shapley NTU value is sometimes referred as the “ λ -transfer value”. The Harsanyi NTU value, being less tractable, has received less attention. Indeed, the Shapley NTU value was introduced as a simplification of the Harsanyi NTU value. Both values are compared in Hart (1985a) in terms of their axiomatic characterizations and in Hart (2004) by means of a simple example. The reader is referred to Peleg and Sudhölter (2007, Ch. 13) and Myerson (1991, Ch. 9) for further details and formal definitions of these two solution concepts.

Electronic supplementary material The online version of this article (<https://doi.org/10.1007/s00182-019-00686-0>) contains supplementary material, which is available to authorized users.

✉ Andrés Salamanca
salamanca@sam.sdu.dk

¹ Department of Business and Economics, University of Southern Denmark, Campusvej 55, 5230 Odense, Denmark

The value has proved to be a surprisingly useful solution concept for the analysis of cooperative outcomes in economic models under complete information (see Aumann (1994) for a partial bibliography of applications). However, many interesting economic situations are characterized by information asymmetries, such as adverse selection and moral hazard problems. Then, the question of examining the value in more realistic environments with incomplete information naturally arises.

Under incomplete information, an agreement should be seen as a mechanism (state contingent decision plan). The enforcement of any such mechanisms relies on the players' claims about their private information. As a consequence, the final agreement may be subject to strategic manipulation. A cooperative agreement must then be *incentive compatible*, in the sense that it provides the appropriate incentives for every individual to reveal honestly his private information.

Myerson (1984a, b) developed a method in which the incentive compatibility constraints are used to define the *virtual utility* of the players. Virtual utility generalizes the weighted-utility scales in the Harsanyi-Shapley fictitious transfer procedure². Elaborating on this approach, Myerson (1984a) defined a bargaining solution which extends the Shapley NTU value to games with incomplete information. The *M-solution* (short for Myerson's solution) takes into account not only the signaling costs associated to incentive compatibility, but also the fact that individuals negotiate at the interim stage. It also involves the identification of "rational threat" mechanisms for each coalition. Rational threats determine how much credit each (type of a) player can claim from the proceeds of cooperation.

In order to keep a tractable mathematical formalization allowing for general existence of the M-solution, Myerson (1984a) imposed various assumptions on the commitment structure of the underlying bargaining situation (see Sect. 6 in Myerson (1984a) for a detailed discussion). These simplifying assumptions entail, however, a reduced sensitivity of the M-solution to some informational externalities. This is evidenced by a prominent example introduced by de Clippel (2005).³ In this paper we provide another intuitive example in which the M-solution does not capture some "negative" externality generated by the adverse selection. Starting from the two-person bargaining problem studied in Sect. 10 of Myerson (1984b), we construct a three-player game in which the uninformed individuals (players 1 and 2) can overcome the potential adverse selection problem they face by ignoring the informed individual (player 3) and agreeing on an outcome that is equitable and efficient for both of them. As we will argue in Sect. 3, a reasonable outcome for this game should leave the informed player with a low expected payoff. Nevertheless, under the M-solution the informed player extracts a considerable amount of utility. Our example shares features with a complete information NTU game previously proposed by Roth (1980). In that game, the Shapley NTU value exhibits some difficulties of a similar nature to that of the M-solution in our example. Hart (1985b) showed, however, that the Harsanyi NTU value prescribes a more appealing outcome in Roth's game. Our main goal in this paper is to explore the extent to which Myerson's virtual utility approach can be used as a mathematical tool for generalizing the Harsanyi NTU value to games with incomplete information.

² Myerson (1992) provides a detailed explanation of the fictitious transfer procedure.

³ De Clippel's example is an incomplete information version of a NTU game introduced by Owen (1971).

Therewith we hope to provide an alternative compelling outcome for our three-player game.

Harsanyi (1963) introduced his NTU value using a model of bargaining in which players inside each coalition negotiate a vector of *dividends*. This dividend allocation procedure is rather intractable and difficult to extend to games with incomplete information. In this work, we shall generalize a simpler (yet equivalent) definition of the Harsanyi NTU value introduced by Myerson (1980). This definition, which dispenses with the notion of dividends, is characterized by an equity condition called *balanced contributions* (see also Myerson (1992) for a detailed explanation). While there might be several appealing ways to extend the balanced contributions to games with incomplete information, here we adopt a method that preserves a conceptual coherence with Imai's (1983) equivalent *subgame value* characterization of the Harsanyi NTU value. In Sect. 4, we build on Myerson's (1984a) virtual utility approach to formulate a "natural" extended version of the subgame value equity condition. We then define an *egalitarian criterion* to be the unique extension of the balanced contributions that is consistent with our generalized subgame value condition (cf. Proposition 2). These equity principles are then used in Sect. 5 to define optimal threat mechanisms for all coalitions. In Sect. 6 we formally define our cooperative solution concept, which we call the *S-solution*. We also exhibit its properties.

The formulation of our egalitarian criterion is inspired in the analysis of our motivating example. As a result, when we explicitly compute the S-solution in this game, it turns out that it prescribes an outcome for which there is more agreement with what we intuitively expect the outcome to be. The S-solution provides a different viewpoint when compared with the M-solution. In this sense, the S-solution illuminates the problem from another angle. Both the M-solution and the S-solution reflect different important qualitative features of our example. Therefore, one should not dispense with either one.

Our construction of the S-solution can be seen as a more sophisticated adaptation of Myerson's (1984a) theory. Indeed, the S-solution requires all threat mechanisms to be equitable, whereas the M-solution only requires equity in the case of the grand coalition. This difference between both solution concepts can be understood as a matter of "credibility" of the threats. Unfortunately, and as it might be expected, extending equity to all coalitions makes a significant difference to the analysis, and the S-solution may fail to exist. In Sect. 7 we provide a simple example of a game in which there is no S-solution. Under complete information, the same difficulty for the Harsanyi NTU value is ruled out by considering games whose characteristic function is comprehensive.⁴ This amounts to assuming free disposal of utility. The same approach does not immediately extend to games with incomplete infor-

⁴ A characteristic function game V is *comprehensive* if, for every (nonempty) coalition S , whenever $V(S) \subseteq \mathbb{R}^S$ contains an allocation u , it also contains all allocations v satisfying $v \leq u$. Further assumptions are also required for the existence of the Harsanyi NTU value: (i) if $u, v \in \partial V(N)$ (i.e., u and v are efficient for the grand coalition) and $u \geq v$, then $u = v$ (*non-levelness*); (ii) $V(N) = K + C$, where $K \subseteq \mathbb{R}^N$ is a compact set and $C \subseteq \mathbb{R}^N$ is a convex cone [see Peleg and Sudhölter (2007, Theorem 13.3.5)]. Assumption (i) excludes vanishing utility weights, while (ii) is a technical assumption guaranteeing that the set of utility weights $\lambda \in \mathbb{R}_{++}^N$ for which the "primal problem" $\max_{v \in V(N)} \lambda \cdot v$ has a finite optimum is compact and convex. It is worth noticing that these assumptions are not necessary for the definition of the Harsanyi NTU

mation. Indeed, when cooperative agreements are made at the interim stage, it is not clear how to derive an analog of the characteristic function under incomplete information.⁵ On the other hand, while in a setting with complete information free disposal of utility is usually taken as an innocuous assumption, the same cannot be done under asymmetric information. In fact, allowing players to discard utility at the interim stage may alter the incentives structure of the game, as it will be illustrated in Sect. 7. The previous difficulties prevent us from obtaining an existence result of the S-solution. This is not specific of our approach. Indeed, de Clippel (2012) encountered similar difficulties for the existence of an alternative (interim) egalitarian criterion in the context of mechanism design. The techniques and analytical tools used in the special case of complete information to achieve positive results cannot generally be extended to games with incomplete information. The reason is that incentive constraints interconnect the decisions in different states in an intricate way. This is also the case for the non-emptiness of the core of an exchange economy with incomplete information (see Forges et al. (2002) for a detailed discussion on this issue).

The virtual utility approach has already been used as a conceptual tool for understanding cooperation under incomplete information in more specific contexts. Myerson (1983) considered negotiations controlled by an informed principal and Myerson (1984b) formulated a generalized Nash bargaining solution for two-person bargaining problems. Another contribution which ought to be pointed out is Myerson (2007), where virtual utility scales were used for extending the inner core. The present paper is thus a direct continuation on this work. Indeed, our construction of the S-solution allowed us to study the significance of the virtual utility approach beyond the solution concepts mentioned above. Yet, the most important contribution from formulating our new cooperative solution may be that it led us to develop conceptual structures which have deepened our understanding of the logical issues involved in cooperation under asymmetric information. Also, it provided a way for unifying the axiomatically derived theories of Nash (1950), Harsanyi (1963) and Myerson (1984b).

As described above, the paper is organized as follows: Sect. 2 is devoted to specifying formally the model of a cooperative game with incomplete information and the notations used, including the basic assumptions on the class of games considered. We also present a summary of the facts one needs to know about Myerson's (1984a) virtual utility approach. Our motivating example is analyzed in Sect. 3. The virtual utility approach is used in Sect. 4 to define our egalitarian criterion. In Sect. 5, the ideas of Sect. 4 are applied to define optimal egalitarian threats. In Sect. 6 we introduce the S-solution. We then compute the S-solution of the example of Sect. 3. Non-existence of the S-solution is discussed in Sect. 7.

Footnote 4 continued

value. They express restrictions on the space of games for which a Harsanyi NTU value can be computed. Similar hypothesis are also required for the axiomatic characterization of the Harsanyi NTU value (see Hart (1985a)).

⁵ See Forges and Serrano (2013) for a discussion about this issue.

2 Formulation

2.1 Bayesian cooperative game

The model of a cooperative game with incomplete information is as follows. Let $N = \{1, 2, \dots, n\}$ denote the set of players. For each (non-empty) coalition $S \subseteq N$, D_S denotes the set of feasible joint actions for coalition S . We assume that the sets of joint actions are finite and *superadditive*, that is, for any two disjoint coalitions⁶ S and R ,

$$D_R \times D_S \subseteq D_{R \cup S}.$$

For any player $i \in N$, we let T_i denote the (finite) set of possible types for player i . The interpretation is that $t_i \in T_i$ denotes the private information possessed by player i . We use the notations⁷ $t_S = (t_i)_{i \in S} \in T_S = \prod_{i \in S} T_i$, $t_{-i} = t_{N \setminus i} \in T_{-i} = T_{N \setminus i}$ and $t_{-S} = t_{N \setminus S} \in T_{-S} = T_{N \setminus S}$. For simplicity, we drop the subscript N in the case of the grand coalition, so we define $D = D_N$ and $T = T_N$. We assume that players have a common prior belief p defined on T , and that all types have positive marginal probability, i.e., $p(t_i) > 0$ for all $t_i \in T_i$ and all $i \in N$.⁸ At the interim stage each player knows his type $t_i \in T_i$, and hence, we let $p(t_{-i} | t_i)$ denote the conditional probability of $t_{-i} \in T_{-i}$ that player i infers given his type t_i .

The utility function of player $i \in N$ is $u_i : D \times T \rightarrow \mathbb{R}$. As in most of the literature in cooperative game theory, we assume that coalitions are *orthogonal*, namely, when coalition $S \subseteq N$ chooses an action which is feasible for it, the payoffs to the members of S do not depend on the actions of the complementary coalition $N \setminus S$. Formally,

$$u_i((d_S, d_{N \setminus S}), t) = u_i((d_S, d'_{N \setminus S}), t),$$

for every $S \subset N$, $i \in S$, $d_S \in D_S$, $d_{N \setminus S}, d'_{N \setminus S} \in D_{N \setminus S}$ and $t \in T$. Then we can let $u_i(d_S, t)$ denote the utility for player $i \in S$ if $d_S \in D_S$ is carried out. That is, $u_i(d_S, t) = u_i((d_S, d_{N \setminus S}), t)$ for any $d_{N \setminus S} \in D_{N \setminus S}$ (recall that $D_S \times D_{N \setminus S} \subseteq D$).

A *cooperative game with incomplete information* is defined by

$$\Gamma = \{N, (D_S)_{S \subseteq N}, (T_i, u_i)_{i \in N}, p\}.$$

A (*direct*) *mechanism* for the grand coalition N is a mapping $\mu_N : T \rightarrow \Delta(D)$, where $\Delta(D)$ denotes the set of probability distributions over D . The interpretation is

⁶ For any two sets A and B , $A \subseteq B$ denotes *weak* inclusion (i.e., possibly $A = B$), and $A \subset B$ denotes strict inclusion.

⁷ For simplicity we write $S \setminus i$, $S \cup i$ and D_i instead of the more cumbersome $S \setminus \{i\}$, $S \cup \{i\}$ and $D_{\{i\}}$.

⁸ The common prior assumption is made without loss of generality, since the solution concept developed in this paper satisfies the probability-invariance axiom described by Myerson (1984a), and so for any game with inconsistent beliefs, conditional probabilities and utilities can be jointly modified in a way that the new game satisfies the common prior assumption and both games impute probability and utility functions that are decision-theoretically equivalent [see also Myerson (1991, pp. 72–73)].

that if N forms, it makes a decision randomly as a function of its members' information. Let the set of mechanisms for N be denoted \mathcal{M}_N .

The (interim) expected utility of player i of type t_i under the mechanism μ_N when he pretends to be of type τ_i (while all other players are truthful) is

$$U_i(\mu_N, \tau_i | t_i) = \sum_{t_{-i} \in T_{-i}} p(t_{-i} | t_i) \sum_{d \in D} \mu_N(d | \tau_i, t_{-i}) u_i(d, (t_i, t_{-i})).$$

As is standard, we denote $U_i(\mu_N | t_i) = U_i(\mu_N, t_i | t_i)$.

Players can use any communication mechanism to implement a state-contingent contract. Because information is not verifiable, the only feasible contracts are those which are induced by Bayesian Nash equilibria of the corresponding communication game. By the revelation principle [see Myerson (1991, Ch. 6)], we can restrict attention to (Bayesian) incentive compatible direct mechanisms. Formally, a mechanism μ_N is *incentive compatible* (for the grand coalition) if and only if

$$U_i(\mu_N | t_i) \geq U_i(\mu_N, \tau_i | t_i), \quad \forall t_i, \tau_i \in T_i, \quad \forall i \in N.$$

We let \mathcal{M}_N^* denote the set of incentive compatible mechanisms for coalition N [“*” stands for incentive compatible as in Holmström and Myerson (1983)].

A mechanism μ_N is (interim) *individually rational* if and only if

$$U_i(\mu_N | t_i) \geq \max_{d_i \in D_i} \sum_{t_{-i} \in T_{-i}} p(t_{-i} | t_i) u_i(d_i, t), \quad \forall t_i \in T_i, \quad \forall i \in N.$$

2.2 Incentive efficiency and the virtual utility approach

Following Holmström and Myerson (1983), we say that a mechanism $\bar{\mu}_N$ for the grand coalition is (interim) *incentive efficient* if and only if $\bar{\mu}_N$ is incentive compatible and there does not exist any other incentive compatible mechanism giving a strictly higher expected utility to all types t_i of all players $i \in N$.⁹ Because the set of incentive-compatible mechanisms is a compact and convex polyhedron, (by the supporting hyperplane theorem) the mechanism $\bar{\mu}_N$ is incentive efficient if and only if there exist non-negative numbers $\lambda = (\lambda_i(t_i))_{i \in N, t_i \in T_i}$, not all zero, such that $\bar{\mu}_N$ is a solution to

$$\max_{\mu_N \in \mathcal{M}_N^*} \sum_{i \in N} \sum_{t_i \in T_i} \lambda_i(t_i) U_i(\mu_N | t_i) \quad (1)$$

We shall refer to this linear-programming problem as the *primal problem* for λ . Let $\alpha_i(\tau_i | t_i) \geq 0$ be the Lagrange multiplier (or dual variable) for the constraint that the type t_i of player i should not gain by reporting τ_i . Then the Lagrangian for this optimization problem can be written as

⁹ We have departed slightly from the formal definition of Holmström and Myerson (1983) in using strict inequalities rather than weak inequalities and one strict inequality.

$$\mathcal{L}(\mu_N, \lambda, \alpha) = \sum_{i \in N} \sum_{t_i \in T_i} \left(\lambda_i(t_i) U_i(\mu_N | t_i) + \sum_{\tau_i \in T_i} \alpha_i(\tau_i | t_i) [U_i(\mu_N | t_i) - U_i(\mu_N, \tau_i | t_i)] \right),$$

where $\mu_N \in \mathcal{M}_N$. To simplify this expression, let

$$v_i(d, t, \lambda, \alpha) = \frac{1}{p(t_i)} \left[\left(\lambda_i(t_i) + \sum_{\tau_i \in T_i} \alpha_i(\tau_i | t_i) \right) u_i(d, t) - \sum_{\tau_i \in T_i} \alpha_i(t_i | \tau_i) \frac{p(t_{-i} | \tau_i)}{p(t_{-i} | t_i)} u_i(d, (\tau_i, t_{-i})) \right] \quad (2)$$

The quantity $v_i(d, t, \lambda, \alpha)$ is called the *virtual utility* of player $i \in N$ from the joint action $d \in D$, when the type profile is $t \in T$, w.r.t. the utility weights λ and the Lagrange multipliers α . Then, the above Lagrangian can be rewritten as

$$\mathcal{L}(\mu_N, \lambda, \alpha) = \sum_{t \in T} p(t) \sum_{d \in D} \mu_N(d | t) \sum_{i \in N} v_i(d, t, \lambda, \alpha) \quad (3)$$

Necessary and sufficient first order conditions (from duality theory of linear programming) imply the following result:

Proposition 1 *An incentive compatible mechanism μ_N is incentive efficient if and only if there exist some vectors $\lambda \geq 0$ ($\lambda \neq 0$) and $\alpha \geq 0$, such that*

$$\alpha_i(\tau_i | t_i) [U_i(\mu_N | t_i) - U_i(\mu_N, \tau_i | t_i)] = 0, \quad \forall i \in N, \forall t_i \in T_i, \forall \tau_i \in T_i \quad (4)$$

and μ_N maximizes the Lagrangian in (3) over all mechanisms in \mathcal{M}_N , namely,

$$\sum_{d \in D} \mu_N(d | t) \sum_{i \in N} v_i(d, t, \lambda, \alpha) = \max_{d \in D} \sum_{i \in N} v_i(d, t, \lambda, \alpha), \quad \forall t \in T \quad (5)$$

Equation (4) is the usual *dual complementary slackness* condition. Condition (5) says that any incentive efficient mechanism μ_N must put positive probability weight only on the decisions that maximize the sum of the players' virtual utilities, on each information state. This implies that if players are given the possibility to transfer virtual utility, conditionally on every state, then μ_N would be ex-post efficient.¹⁰ Incentive compatibility forces each player to act as if he was maximizing a distorted utility, which magnifies the differences between his true type and the types that would be tempted

¹⁰ This property is specially useful for practical applications, in particular when computing value allocations.

to imitate him. Myerson (1984a) refers to this idea as the *virtual utility hypothesis*. A more detailed discussion about the meaning and significance of the virtual utility can be found in Myerson (1991, Ch. 10).

The natural vector α in this Lagrangian analysis is the vector that solves the dual problem of (1). This *dual problem* for λ can be written as

$$\min_{\alpha \geq 0} \sum_{t \in T} p(t) \left(\max_{d \in D} \sum_{i \in N} v_i(d, t, \lambda, \alpha) \right) \quad (6)$$

2.3 The M-solution

Using the concept of virtual utility, Myerson (1984a, b) generalizes the Harsanyi-Shapley fictitious transfer procedure in order to extend the Shapley NTU value to an environment with incomplete information. Specifically, for any incentive efficient mechanism μ_N one associates a vector (λ, α) of virtual utility scales. These scales correspond to the utility weights λ for which μ_N solves the primal problem and the associated Lagrange multipliers α . Then, one considers the fictitious game in which players are allowed to transfer virtual utility, conditional on every state $t \in T$, w.r.t. the scales (λ, α) . In the virtual game, each intermediate coalition $S \subset N$ commits to a *rational threat* mechanism to be carried out in case the other players refuse to cooperate with the members of S . Rational threats are the basis for computing the (virtual) worth of each coalition, and thus they determine how much credit each type of a player can claim from the proceeds of cooperation in the grand coalition. Conditionally on every state, rational threats thus define a coalitional game with transferable virtual utility. A mechanism is *equitable* for the grand coalition N if it gives each type of every player his (conditional) expected Shapley TU value of the fictitious game. A precise definition is given in Sect. 4 (see Remark 1).

Myerson (1984a) defines the *M-solution* to be an incentive efficient mechanism μ_N for which there exist virtual scales (λ, α) such that μ_N is equitable for the grand coalition. The associated interim utility allocations are called an *M-value*. A formal definition of the M-solution can be deduced from our cooperative solution concept (cf. Definition 6) by removing the egalitarian restrictions from our optimal threat criterion (see Remark 3). Two variants of the value can be considered depending on whether coalitional threats are required to be incentive compatible or not. Myerson exclusively deals with the case in which only the mechanism of the grand coalition is constrained to be (equitable and) incentive compatible. The M-solution is justified only in situations where cooperative agreements are made before a coalition structure is determined, while expecting that only the grand coalition will be forming. A detailed discussion on this issue is given in Myerson (1984a, Sect. 6).

3 Motivating examples

In this section we study two examples which motivate the introduction of our solution concept. In both examples, it is shown that the M-value exhibits some “difficulties”;

specifically, there are compelling reasons leading to an outcome not consistent with the M-value.

3.1 Example 1: a collective choice problem

We consider the following cooperative game with incomplete information. The set of players is $N = \{1, 2, 3\}$. Only player 3 has private information represented by two possible types in $T_3 = \{H, L\}$ with prior probabilities $p(H) = 1 - p(L) = 9/10$. Decision options for every coalition are $D_i = \{d_i\}$ ($i \in N$), $D_{\{1,2\}} = [D_1 \times D_2] \cup \{d_{12}\} = \{[d_1, d_2], d_{12}\}$, $D_{\{i,3\}} = [D_i \times D_3] \cup \{d_{i3}^i, d_{i3}^3\} = \{[d_i, d_3], d_{i3}^i, d_{i3}^3\}$ ($i = 1, 2$) and $D_N = [D_{\{1,2\}} \times D_3] \cup [D_{\{1,3\}} \times D_2] \cup [D_{\{2,3\}} \times D_1]$. A detailed interpretation will be given below. Finally, utility functions are as follows:

(u_1, u_2, u_3)	L	H
$[d_1, d_2, d_3]$	$(0, 0, 0)$	$(0, 0, 0)$
$[d_{12}, d_3]$	$(5, 5, 0)$	$(5, 5, 0)$
$[d_{13}^1, d_2]$	$(0, 0, 5)$	$(0, 0, 10)$
$[d_{13}^3, d_2]$	$(10, 0, -5)$	$(10, 0, 0)$
$[d_{23}^2, d_1]$	$(0, 0, 5)$	$(0, 0, 10)$
$[d_{23}^3, d_1]$	$(0, 10, -5)$	$(0, 10, 0)$

This game can be interpreted as a collective choice problem in which three individuals have the option to cooperate by investing in a work project which would benefit them. The project would cost \$10. It is commonly known that the project is worth \$10 to player 1 as well as to player 2; but its value to player 3 depends on her type, which is unknown to the other players. If 3's type is H ("high") then the project is worth \$10 to her. However, if 3's type is L ("low") then the project is only worth \$5 to her.

Decision options for all coalitions are interpreted as follows. For each player $i \in N$, d_i is the only available action for himself, which leaves him with his reservation utility normalized to \$0. If coalition $\{1, 2\}$ forms, its members may decide not to undertake the project by choosing $[d_1, d_2]$ or they can agree on the option d_{12} which carries out the project dividing the cost on equal parts. If players 1 and 3 form a coalition, decision d_{13}^j ($j = 1, 3$) denotes the option to undertake the project at j 's expense. There is no need to consider intermediate financing options, because they can be represented by randomized decisions. They may also agree on $[d_1, d_3]$ which does not implement the project. Decision options for coalition $\{2, 3\}$ are similarly interpreted. If all three form a coalition, they may use a random device to pick a two-person coalition which must then make a decision as above.

To analyze this game, we first consider the situation in which players 1 and 3 must reach a cooperative agreement to be implemented in case player 2 refuses to cooperate with them. In such a situation, 1 and 3 face a threat-selection subgame described by a two-person cooperative game with incomplete information that can be analyzed applying the concepts of Sect. 2. Assume that threats are not required to be incentive

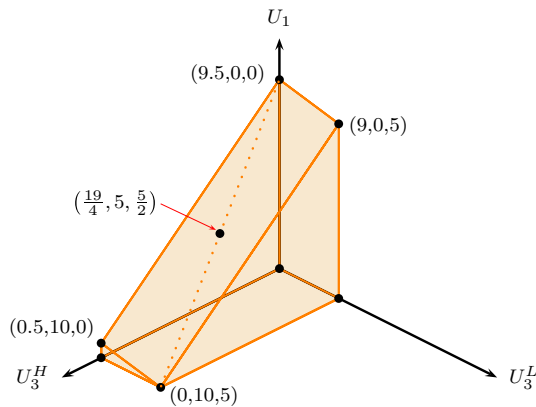


Fig. 1 Feasible allocations for $\{1, 3\}$

compatible. Figure 1 illustrates the set of feasible (i.e., individually rational) interim utility allocations for this (sub)game.

An equitable utility allocation in this game can be constructed as follows. Suppose that player 3 is given the right to act as a “dictator”, so that she may enforce any mechanism that is individually rational given the information that player 1 may infer from the selection of the mechanism. In this case, there is a clear decision that both types of player 3 would demand, namely, d_{13}^1 . This decision implements the utility allocation $(U_1, U_3^H, U_3^L) = (0, 10, 5)$ which gives *both* types of player 3 the highest expected utility they can get in the game. Moreover, it is efficient (see Fig. 1) and *safe*, i.e., it remains individually rational no matter what player 1 can infer about 3’s type from this proposal. In the terminology of Myerson (1983), it is a *strong solution*¹¹ for player 3. On the other hand, if player 1 were a dictator, then he would clearly demand the mechanism implementing the allocation $(19/2, 0, 0)$, which yields the largest possible expected utility he can get, while leaving both types of player 3 with their individual rationality levels (see Fig. 1). Now consider a *random-dictatorship* in which each player is given equal chance of enforcing his/her strong solution. Then, the interim efficient allocation $(19/4, 5, 5/2) = \frac{1}{2}(0, 10, 5) + \frac{1}{2}(19/2, 0, 0)$ is equitable for $\{1, 3\}$.¹² Indeed, random-dictatorship together with efficiency characterize Myerson’s (1984b) generalization of the Nash bargaining solution. Thus, this allocation is also the unique M-value for this (sub)game.

The value of a player is an index based on his ability to guarantee high payoffs to all members of the coalitions to which he belongs (marginal contribution). From that perspective, player 3 should be considered as a weak player. By agreeing to cooperate with player 3, player 1 cannot expect to get more than $19/4$ in an equitable allocation. Because players 1 and 2 are symmetric, the same reasoning is also true for a negotiation between players 2 and 3. Hence, both players 1 and 2 are better off in coalition $\{1, 2\}$, in which case they get 5 each, which is strictly preferred to $19/4$. When negotiating

¹¹ A strong solution may not exist, but if so it is unique up to equivalence in utility.

¹² This allocation is implemented by the mechanism $\mu_{\{1,3\}}(d_{13}^1 | L) = 1 - \mu_{\{1,3\}}(d_{13}^3 | L) = 3/4$, $\mu_{\{1,3\}}(d_{13}^1 | H) = \mu_{\{1,3\}}(d_{13}^3 | H) = 1/2$.

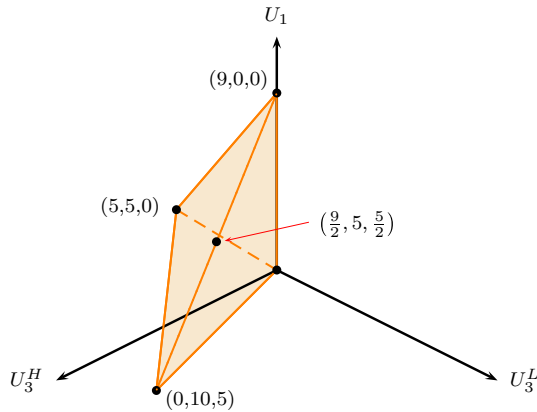


Fig. 2 Incentive feasible allocations for $\{1, 3\}$

with player 3, 1 and 2 are adversely affected by the likely presence of 3's "bad" low type. However, by acting together players 1 and 2 face no uncertainty at all. Indeed, it is commonly known that the project is equally worth to each of them. A value allocation for our three player game should thus reward both types of player 3 less than the other two players¹³.

Let us suppose now that threats are required to be incentive compatible. Figure 2 depicts the set of incentive feasible (i.e., incentive compatible and individually rational) interim utility allocations for the subgame faced by coalition $\{1, 3\}$. For this modified threat-selection game, the strong solution for player 3 implements again the utility allocation $(0, 10, 5)$.¹⁴ However, the strong solution for player 1 now implements the allocation $(9, 0, 0)$. Proceeding as before, random-dictatorship prescribes the value allocation $(9/2, 5, 5/2)$.¹⁵ We notice that both types of player 3 get the same expected utility in an equitable allocation regardless of whether incentive constraints are relevant or not. In contrast, 1's expected utility is reduced in the presence of incentive constraints. Incentive compatibility leads to efficiency losses that are mainly beared by the uninformed party, hence increasing the incentives for 1 and 2 to form a coalition, and thus reducing 3's bargaining ability. Therefore, we argue that 3's expected payoff from a value allocation in the whole game should be further reduced when coalitional threats are required to be incentive compatible.

¹³ Although we have tried to make our arguments as compelling as possible, this sort of discussion may leave room for disagreement.

¹⁴ When incentive constraints matter, a safe mechanism for player 3 is one which would be incentive compatible and individually rational if player 1 knew 3's type.

¹⁵ This allocation is implemented by the mechanism $\mu_{\{1,3\}}(d_{13}^1 | L) = \mu_{\{1,3\}}([d_1, d_3] | L) = 1/2$, $\mu_{\{1,3\}}(d_{13}^1 | H) = \mu_{\{1,3\}}(d_{13}^3 | H) = 1/2$.

The unique M-value of our three-player game is the utility allocation¹⁶

$$(U_1, U_2, U_3^H, U_3^L) = \left(\frac{10}{3}, \frac{10}{3}, \frac{10}{3}, \frac{5}{3}\right). \quad (7)$$

The M-value rewards players proportionally to their valuations of the project, as if the likely presence of 3's low type did not adversely affect players 1 and 2. This is so even when threats are required to be incentive compatible. In this example, the M-value is insensitive to the negative externality that adverse selection exerts on 3's bargaining position.

One possible interpretation for the counterintuitive behavior of the M-value here can be obtained by applying the random-dictatorship procedure to the grand coalition: the strong solution for player 3 in N implements the allocation $(U_1, U_2, U_3^H, U_3^L) = (0, 0, 10, 5)$. The strong solution for player 1 (resp. 2) in N implements the allocation $(19/2, 1/2, 0, 0)$ (resp. $(1/2, 19/2, 0, 0)$). Averaging these utility vectors we obtain (7). It is worth emphasizing that this procedure does not generally characterize the M-value. Yet for our example, it exhibits why both types of player 3 extract a considerable amount of utility. The random dictatorship procedure applied to N ignores the possibilities of cooperation among subsets of players, hence it is only acceptable when coalitions are treated symmetrically. Indeed, Myerson's rational threats criterion cares only about the joint overall gains that can be allocated inside a coalition, but not about the way in which they are distributed. Because all coalitions can achieve the total surplus from the project, the M-value treats all coalitions symmetrically. For instance, the mechanism that implements d_{j3}^j ($j = 1, 2$) in both states is a rational threat for coalition $\{j, 3\}$. This mechanism however gives the whole surplus of cooperation to player 3, which is manifestly not equitable. Such a threat can be considered as being not "credible", in the sense that player $i \notin \{j, 3\}$ could not believe that player j would agree to implement d_{j3}^j in case cooperation in N breaks down.

It seems then that, if we want our considerations to be well reflected in a value allocation for this game, we require our cooperative solution to take into account the equity restrictions that coalitions face when sharing the proceeds of cooperation. This intuition will guide our formulation of the equity principles introduced in the next section. But, before proceeding to Sect. 4, and for the sake of completeness, we shall briefly analyze an additional example introduced by de Clippel (2005).

3.2 Example 2: a bilateral trade problem

Let us consider the following cooperative game with incomplete information. $N = \{1, 2, 3\}$, $T_1 = \{H, L\}$, $p(H) = 1 - p(L) = 4/5$, $D_i = \{d_i\}$ ($i = 1, 2, 3$),

¹⁶ The incentive efficient mechanism $\mu_N([d_{12}, d_3] | t) = \frac{2}{3}$, $\mu_N([d_{23}^2, d_1] | t) = \mu_N([d_{13}^1, d_2] | t) = \frac{1}{6}$ for all $t \in T_3$ is an M-solution. The value is supported by the utility weights $(\lambda_1, \lambda_2, \lambda_3^H, \lambda_3^L) = (1, 1, 9/10, 1/5)$ and the Lagrange multipliers $(\alpha_1(L | H), \alpha_1(H | L)) = (0, 0)$. We focus only on non-degenerated values, i.e., those which are supported by strictly positive utility weights λ . Utility weights are determined up to a positive scalar multiplication. We then normalize utility weights so that virtual utilities of the uninformed players coincide with their real utilities. This is possible since 1 and 2 are symmetric. Explicit computations are given in the Online supplementary material.

$D_{\{1,2\}} = \{[d_1, d_2], d_{12}^1, d_{12}^2\}$, $D_{\{1,3\}} = \{[d_1, d_3]\}$, $D_{\{2,3\}} = \{[d_2, d_3]\}$, $D_N = \{[d_1, d_2, d_3], [d_{12}^1, d_3], [d_{12}^2, d_3], d_{23}, d_{32}\}$ and

(u_1, u_2, u_3)	$[d_1, d_2, d_3]$	$[d_{12}^1, d_3]$	$[d_{12}^2, d_3]$	d_{23}	d_{32}
H	$(0, 0, 0)$	$(90, 0, 0)$	$(0, 90, 0)$	$(0, 90, 0)$	$(0, 0, 90)$
L	$(0, 0, 0)$	$(30, 0, 0)$	$(-60, 90, 0)$	$(0, 30, 0)$	$(0, 0, 30)$

The game can be interpreted as follows. Player 2 is the seller of a single good that has no value for himself. Player 1 is the only potential buyer and he has a valuation of the good that can be low (30\$), with probability $1/5$, or high (90\$), with probability $4/5$. Decision $[d_1, d_2]$ represents the no-exchange alternative. Decision d_{12}^1 (resp. d_{12}^2) represents the situation where player 1 receives the good from player 2 for free (resp. in exchange of 90\$). Any other transfer of money from player 1 to player 2 (between 0\$ and 90\$) can be represented by a lottery defined on $\{d_{12}^1, d_{12}^2\}$. Because of the necessity to give player 1 an incentive to participate honestly, both players are limited in their abilities to share the gains from trade. Indeed, the mechanism that gives the entire surplus to player 2 in both states, is not incentive compatible. On the other hand, player 3 acts as a pure intermediary (broker), which does not generate any additional surplus from the trade. Yet, his participation partly releases players 1 and 2 from the incentive constraints they face when they cooperate. Indeed, when she joins coalition $\{1, 2\}$ (so that the grand coalition forms), decisions d_{23} and d_{32} are added to $D_{\{1,2\}} \times D_{\{3\}}$. Decision d_{23} (resp. d_{32}) gives the whole surplus to player 2 (resp. 3) in both states¹⁷.

As it is shown by de Clippel (2005), the unique M-value of this game is the interim utility allocation

$$(U_1^H, U_1^L, U_2, U_3) = (45, 15, 39, 0). \quad (8)$$

We observe that player 3 is considered a null player. Even though player 3 does not create any additional surplus, it would be fair to give her some positive payoff, as players 1 and 2 have to rely on her in order to weaken the incentive constraints they face. As in the previous example, requiring optimal threats to be incentive compatible does not change the M-value allocation. We conclude that the M-value is not sensitive to the informational contribution of player 3.

4 Equity principles for Bayesian cooperative games

The Harsanyi NTU value can be characterized using two different fair allocation rules. The first of these two equity notions, introduced by Myerson (1980) under the name of

¹⁷ It can be shown that when player 3 drops out of the game and coalition $\{1, 2\}$ forms, the constraint asserting that type 1_H has no incentive to report to be type 1_L is binding in any incentive efficient mechanism for this coalition.

balanced contributions, requires that, for any two members of a coalition, the amount that each player would gain by the other's participation should be equal when utility comparisons are made in some weighted utility scale. The second equity principle, denominated *subgame value equity* by Imai (1983), says that, for every coalition $S \subseteq N$, each player in S should obtain his Shapley TU value from the game restricted to the subcoalitions of S when utility has been made comparable in some weighted utility scale. These two equity notions are in dual relationship: for fixed utility scales both allocation rules are equivalent (*equity equivalence*). In this Section, we extend this result to the case of incomplete information.

Given a vector of utility weights λ and a vector of Lagrange multipliers α , let us consider the fictitious game in which players make interpersonal utility comparisons in the virtual utility scales (λ, α) . In such a virtual game, each player's payoffs are represented in the virtual utility scales and virtual payoffs are transferable among the players (conditionally on every state). We assume that, as a threat during the bargaining process within the grand coalition N , each coalition $S \subset N$ commits to some mechanism $\mu_S : T_S \rightarrow \Delta(D_S)$.¹⁸ We denote by \mathcal{M}_S the set of mechanisms for S . Let $\mathcal{M} = \prod_{S \subseteq N} \mathcal{M}_S$ denote the set of possible profiles of mechanisms that all various coalitions might select.

Let $v_i(\mu_S, t, \lambda, \alpha)$ denote the linear extension of $v_i(\cdot, t, \lambda, \alpha)$ (as defined in (2)) over μ_S . We define $W_S(\mu_S, t, \lambda, \alpha)$ as the sum of virtual utilities that the members of $S \subseteq N$ would expect in state t when they select the mechanism μ_S , that is

$$W_S(\mu_S, t, \lambda, \alpha) = \sum_{i \in S} v_i(\mu_S, t, \lambda, \alpha). \quad (9)$$

Let $W(\eta, t, \lambda, \alpha) = (W_S(\mu_S, t, \lambda, \alpha))_{S \subseteq N}$ denote the characteristic function game when the vector of threats $\eta = (\mu_S)_{S \subseteq N} \in \mathcal{M}$ is selected by the various coalitions¹⁹ in the virtual game. For any vector $\eta \in \mathcal{M}$, let $\eta_S = (\mu_R)_{R \subseteq S}$ denote its restriction to the subcoalitions of S . We define $W|_S(\eta_S, t, \lambda, \alpha)$ as the subgame of $W(\eta, t, \lambda, \alpha)$ obtained by restricting the domain of $W(\eta, t, \lambda, \alpha)$ to the subsets of S . Let ϕ be the Shapley TU value operator; for $i \in S \subseteq N$, $\phi_i(S, W|_S(\eta_S, t, \lambda, \alpha))$ will thus denote the Shapley TU value of player i in the subgame restricted to S when the vector of threats η_S is selected in the virtual game.

We denote $V_i(\mu_S | t_i, \lambda, \alpha)$ the expected virtual utility of type t_i of player $i \in S$ when the members of S agree on μ_S , i.e.,

$$V_i(\mu_S | t_i, \lambda, \alpha) := \sum_{t_{-i} \in T_{-i}} p(t_{-i} | t_i) v_i(\mu_S, t, \lambda, \alpha). \quad (10)$$

¹⁸ When a coalition S forms, it cannot rely on the information possessed by the players outside S . In other words, a communication mechanism for a coalition must be measurable with respect to the private information of its members. This is equivalent to define a mechanism as $\mu_S : T \rightarrow \Delta(D_S)$ with $\mu_S(t) = \mu_S(t')$ for every $t, t' \in T$ such that $t_S = t'_S$.

¹⁹ Strictly speaking, the component $\mu_N \in \mathcal{M}_N$ of η is not a threat, since there is no coalition to threaten. However, we keep this terminology in order to simplify the exposition.

Definition 1 (*Equitable mechanism*) For any coalition $S \subseteq N$, the mechanism μ_S is *equitable* for S w.r.t. η_S , λ and α if

$$V_i(\mu_S | t_i, \lambda, \alpha) = \sum_{t_{-i} \in T_{-i}} p(t_{-i} | t_i) \phi_i(S, W|_S(\eta_S, t, \lambda, \alpha)), \quad \forall t_i \in T_i, \quad \forall i \in S. \quad (11)$$

If for all coalitions $R \subseteq S$, μ_R is equitable for R w.r.t. η_R , λ and α , then the vector of threats $\eta_S = (\mu_R)_{R \subseteq S}$ is called *equitable* w.r.t. λ and α .

Then, a mechanism for coalition S is said to be equitable for S if it gives every type of a player in S his (conditionally) expected Shapley TU value from the virtual subgame obtained by restricting $W(\eta, t, \lambda, \alpha)$ to the subcoalitions of S . This equity notion extends Imai's subgame value equity condition.²⁰

Remark 1 When $S = N$, the equality in (11) reduces to Myerson's (1984a) principle for equitable compromises.

Definition 2 (*Egalitarian mechanism*) For any coalition $S \subseteq N$, the mechanism μ_S is *egalitarian* for S w.r.t. $(\mu_{S \setminus i})_{i \in S}$, λ and α if

$$\begin{aligned} & \sum_{t_{-i} \in T_{-i}} p(t_{-i} | t_i) \sum_{j \in S \setminus i} [v_i(\mu_S, t, \lambda, \alpha) - v_i(\mu_{S \setminus j}, t, \lambda, \alpha)] \\ &= \sum_{t_{-i} \in T_{-i}} p(t_{-i} | t_i) \sum_{j \in S \setminus i} [v_j(\mu_S, t, \lambda, \alpha) - v_j(\mu_{S \setminus i}, t, \lambda, \alpha)], \quad \forall t_i \in T_i, \quad \forall i \in S. \end{aligned} \quad (12)$$

If for all coalitions $R \subseteq S$, μ_R is egalitarian for R w.r.t. $(\mu_{R \setminus i})_{i \in R}$, λ and α , then the vector of threats $\eta_S = (\mu_R)_{R \subseteq S}$ is called *egalitarian* w.r.t. λ and α .

Equation (12) says that the expected average virtual contribution of the different players in S to player i equals the expected average virtual contribution of player i to the different players in S as assessed by his type t_i . This egalitarian criterion generalizes Myerson's balanced contributions condition²¹. Indeed, when information is complete (i.e., T_i is a singleton for every $i \in N$, so that we can set $\alpha = 0$), condition (12) implies that the j -th terms on both sides are equal: the marginal contribution of j to i , measured by $v_i(\mu_S, \lambda) - v_i(\mu_{S \setminus j}, \lambda)$, equals the marginal contribution of i to j , symmetrically measured by $v_j(\mu_S, \lambda) - v_j(\mu_{S \setminus i}, \lambda)$. The same implication cannot be expected to generally hold in the case of asymmetric information. The reason is that, since negotiations take place at the interim stage, the individual probability assessments of the different types of the various players need not be the same. Then,

²⁰ When information is complete, so that T_i is a singleton for every $i \in N$, (11) reduces to the first condition in Proposition 6 of Imai (1983).

²¹ It also extends the "preservation of average differences" principle introduced by Hart and Mas-Colell (1996).

i 's personal evaluation of j 's gains may not coincide with j 's evaluation of her own gains.

For given arbitrary vectors $(\mu_R)_{R \subseteq S}$, λ and α , equity and egalitarianism are in general two different notions of "fairness" for coalition $S \subseteq N$. In particular, notice that while an egalitarian mechanism μ_S depends only on the mechanisms $(\mu_{S \setminus i})_{i \in S}$, an equitable mechanism depends on the whole profile of threats $(\mu_R)_{R \subseteq S}$. However, it turns out that if the *whole* profile η_S is egalitarian, then it is also equitable, and viceversa.

Proposition 2 (*Equity equivalence*) *For any coalition $S \subseteq N$, the vector of threats $\eta_S = (\mu_R)_{R \subseteq S}$ is equitable (w.r.t. λ and α) if and only if it is egalitarian (w.r.t. λ and α).*

This result is significant, first, in establishing a dual relationship between equity (as defined by the Shapley TU value) and the balanced contributions in environments with incomplete information. Second, and most important, Proposition 2 helps us to justify why our egalitarian criterion is (probably) the most appropriate generalization of the balanced contributions condition.

When information is asymmetric, so that the probability assessments of the various types of distinct players are different, Proposition 2 cannot be deduced from the equity equivalence under complete information simply by taking (conditional) expectations. Instead we use a "consistency property" of the Shapley TU value: the value of a player is the average of his marginal contribution to the grand coalition $W_N - W_{N \setminus i}$ and his values $\phi_i(N \setminus j, W|_{N \setminus j})$ in the subgames with $|N| - 1$ players [see Hart (2004, p. 39)]. Apart from this clarification, the proof of Proposition 2 is straightforward.

Proof We start proving the "only if" part. Let $\eta \in \mathcal{M}$ be a vector of equitable threats (w.r.t. λ and α). Let $S \subseteq N$ and $i \in S$ be fixed. Then, for any $j \in S \setminus i$, $\mu_{S \setminus j}$ is equitable for $S \setminus j$ (w.r.t. $\eta_{S \setminus j}$, λ and α). Thus, for all $t_i \in T_i$,

$$\begin{aligned} & \sum_{t_{-i} \in T_{-i}} p(t_{-i} | t_i) \sum_{j \in S \setminus i} \phi_i(S \setminus j, W|_{S \setminus j}(\eta_{S \setminus j}, t, \lambda, \alpha)) \\ &= \sum_{t_{-i} \in T_{-i}} p(t_{-i} | t_i) \sum_{j \in S \setminus i} v_i(\mu_{S \setminus j}, t, \lambda, \alpha). \end{aligned} \quad (13)$$

On the other hand, because μ_S is equitable for S (w.r.t. η_S , λ and α), we have that for all $t_i \in T_i$,

$$\begin{aligned} & \sum_{t_{-i} \in T_{-i}} p(t_{-i} | t_i) v_i(\mu_S, t, \lambda, \alpha) \\ &= \sum_{t_{-i} \in T_{-i}} p(t_{-i} | t_i) \phi_i(S, W|_S(\eta_S, t, \lambda, \alpha)) \\ &= \sum_{t_{-i} \in T_{-i}} p(t_{-i} | t_i) \frac{1}{|S|} \left[W_S(\mu_S, t, \lambda, \alpha) - W_{S \setminus i}(\mu_{S \setminus i}, t, \lambda, \alpha) \right] \end{aligned}$$

$$\begin{aligned}
& + \sum_{j \in S \setminus i} \phi_i(S \setminus j, W|_{S \setminus j}(\eta_{S \setminus j}, t, \lambda, \alpha)) \Big] \\
& = \frac{1}{|S|} \sum_{t_{-i} \in T_{-i}} p(t_{-i} | t_i) \left[v_i(\mu_S, t, \lambda, \alpha) + \sum_{j \in S \setminus i} v_i(\mu_{S \setminus j}, t, \lambda, \alpha) \right. \\
& \quad \left. + \sum_{j \in S \setminus i} (v_j(\mu_S, t, \lambda, \alpha) - v_j(\mu_{S \setminus i}, t, \lambda, \alpha)) \right], \quad (14)
\end{aligned}$$

where (13) has been used in the last equality. Finally, rearranging terms in (14) we get (12).

Consider now the “if” part. Let $\eta \in \mathcal{M}$ be a vector of egalitarian threats (w.r.t. λ and α). For any coalition $S \subseteq N$ and any player $i \in S$ of type t_i , the (conditionally) expected marginal contribution of player i to coalition S is

$$\begin{aligned}
& \sum_{t_{-i} \in T_{-i}} p(t_{-i} | t_i) [W_S(\mu_S, t, \lambda, \alpha) - W_{S \setminus i}(\mu_{S \setminus i}, t, \lambda, \alpha)] \\
& = \sum_{t_{-i} \in T_{-i}} p(t_{-i} | t_i) v_i(\mu_S, t, \lambda, \alpha) \\
& \quad + \sum_{t_{-i} \in T_{-i}} p(t_{-i} | t_i) \sum_{j \in S \setminus i} [v_j(\mu_S, t, \lambda, \alpha) - v_j(\mu_{S \setminus i}, t, \lambda, \alpha)] \\
& = \sum_{t_{-i} \in T_{-i}} p(t_{-i} | t_i) v_i(\mu_S, t, \lambda, \alpha) \\
& \quad + \sum_{t_{-i} \in T_{-i}} p(t_{-i} | t_i) \sum_{j \in S \setminus i} [v_i(\mu_S, t, \lambda, \alpha) - v_i(\mu_{S \setminus j}, t, \lambda, \alpha)] \\
& = \sum_{t_{-i} \in T_{-i}} p(t_{-i} | t_i) \left[|S| v_i(\mu_S, t, \lambda, \alpha) - \sum_{j \in S \setminus i} v_i(\mu_{S \setminus j}, t, \lambda, \alpha) \right],
\end{aligned}$$

where the second equality is due to the fact that μ_S is egalitarian for S w.r.t. $(\mu_{S \setminus j})_{j \in S}$, λ and α . Therefore,

$$\sum_{t_{-i} \in T_{-i}} p(t_{-i} | t_i) \phi_i(S, W|_S(\eta_S, t, \lambda, \alpha)) = \sum_{t_{-i} \in T_{-i}} p(t_{-i} | t_i) v_i(\mu_S, t, \lambda, \alpha).$$

□

We conclude this section with a convenient characterization of an equitable mechanism for the grand coalition. It will allow us to identify the real interim utilities corresponding to an equitable allocation in the virtual game.

Definition 3 (*Warranted claims*) Let (λ, α) be a vector of virtual scales and $\eta \in \mathcal{M}$ a vector of threats. The interim allocation $\omega \in \prod_{i \in N} \mathbb{R}^{T_i}$ is *warranted* by λ, α and η if

$$\begin{aligned} & \left(\lambda_i(t_i) + \sum_{\tau_i \in T_i} \alpha_i(\tau_i \mid t_i) \right) \omega_i(t_i) - \sum_{\tau_i \in T_i} \alpha_i(t_i \mid \tau_i) \omega_i(\tau_i) \\ &= \sum_{t_{-i} \in T_{-i}} p(t) \phi_i(N, W(\eta, t, \lambda, \alpha)), \quad \forall t_i \in T_i, \quad \forall i \in N. \end{aligned} \quad (15)$$

The quantity $\omega_i(t_i)$ is called the *warranted claim* of type t_i of player i .

Remark 2 By Lemma 1 in Myerson (1983), the warrant equations have a unique solution in the vector of warranted claims of player i , provided that $\lambda > 0$. Furthermore, the solution (weakly) increases (in the vector sense) as the right-hand side is increased.

The following result follows from the equalities (3.10) and (3.11) in Myerson (1984a).

Lemma 1 Let (λ, α) be a vector of virtual scales such that α is a solution of the dual for λ . Let $\eta \in \mathcal{M}$ be a vector of threats such that μ_N is a solution of the primal for λ . The mechanism μ_N is equitable for N w.r.t. η, λ and α if and only if the vector of interim utilities $U(\mu_N) := (U_i(\mu_N \mid t_i))_{i \in N, t_i \in T_i}$ is warranted by λ, α and η .

We can thus interpret the warrant equations: they implicitly define ω to be the real utility allocation which would give every type of each player (in the grand coalition) his expected Shapley TU value in the virtual game.

5 Optimal threats

In this section we use the equity principles previously developed in order to extend Harsanyi's (1963, Sect. 9) optimal threat strategies. Specifically, we modify Myerson's (1984a) rational threats by requiring coalitional threats to meet our egalitarian criterion.

Definition 4 (*Optimal egalitarian threats*) The mechanism $\bar{\mu}_S \in \mathcal{M}_S$ is an *optimal egalitarian threat* for $S \subseteq N$ w.r.t. $(\mu_{S \setminus i})_{i \in S}, \lambda$ and α if and only if $\bar{\mu}_S$ is a solution to

$$\begin{aligned} & \max_{\mu_S \in \mathcal{M}_S} \sum_{t \in T} p(t) W_S(\mu_S, t, \lambda, \alpha) \\ & \text{s.t. (12)} \end{aligned} \quad (16)$$

The optimal threats criterion in (16) postulates that each coalition should maximize the ex-ante expected total virtual utility that its members would earn when coalitions commit to a vector of egalitarian threats. In view of Proposition 2, we could also have defined an optimal threat replacing the egalitarian constraints (12) in (16) by the equity conditions in (11). However, this alternative definition is less tractable since

threats of one coalition cannot be determined without knowledge of threats of *all* its subcoalitions²².

We notice that the maximization in (16) is carried out over all mechanisms in \mathcal{M}_S . However, we can alternatively require threats to be incentive compatible. A mechanism μ_S is *incentive compatible for coalition* $S \subseteq N$ if and only if²³

$$\begin{aligned} & \sum_{t_{-i} \in T_{-i}} p(t_{-i} | t_i) \sum_{d_S \in D_S} \mu_S(d_S | t_S) u_i(d_S, t) \\ & \geq \sum_{t_{-i} \in T_{-i}} p(t_{-i} | t_i) \sum_{d_S \in D_S} \mu_S(d_S | \tau_i, t_{S \setminus i}) u_i(d_S, t), \quad \forall i \in S, \quad \forall t_i, \tau_i \in T_i. \end{aligned}$$

We denote as \mathcal{M}_S^* the set of incentive-compatible mechanisms for coalition S .

Definition 5 (*Incentive compatible optimal egalitarian threats*) A mechanism $\bar{\mu}_S \in \mathcal{M}_S$ is an *incentive compatible optimal egalitarian threat* for $S \subseteq N$ w.r.t. $(\mu_{S \setminus i})_{i \in S}$, λ and α if and only if it solves (16) over all mechanisms in \mathcal{M}_S^* .

Given some virtual scales (λ, α) , (incentive compatible) optimal egalitarian threats must be recursively constructed: for each S , given the threats $(\bar{\mu}_{S \setminus i})_{i \in S}$, $\bar{\mu}_S$ is determined solving (16). This recursion leads to a profile of threats $\bar{\eta} = (\bar{\mu}_S)_{S \subseteq N}$ which we call an (*coalitionally incentive compatible*) *egalitarian solution* w.r.t. λ and α .²⁴

Myerson (1984a, Sect. 6) argues that maximizing the ex-ante expected virtual worth of a coalition is appropriate in games where only the mechanism chosen by the grand coalition will be implemented. In such a situation, the final payoffs are granted by the grand coalition and therefore the mechanisms $(\mu_S)_{S \subseteq N}$ need not be either equitable or incentive compatible. Thus, Myerson's (1984a) rational threats maximize the objective function in (16) constrained only by the feasibility of the mechanisms, i.e., $\mu_S \in \mathcal{M}_S$. Even if we agree with this reasoning, the examples in Sect. 3 illustrate situations in which some relevant aspects of the intermediate coalitions are ignored by Myerson's rational threat criterion. In contrast, we think that for a mechanism μ_S to constitute an appropriate measure of the strength of coalition S , it must be equitable regardless of whether it is expected to be implemented or not. It should be clear that a vector $\eta = (\mu_S)_{S \subseteq N}$ of (inductively constructed) optimal egalitarian threats is egalitarian. Therefore, by Proposition 2, η is also equitable. This reasoning is summarized in the following proposition.

Proposition 3 *Let $\eta = (\mu_S)_{S \subseteq N}$ be an (coalitionally incentive compatible) egalitarian solution w.r.t. λ and α . Then, for each coalition $S \subseteq N$, η_S is equitable w.r.t. λ and α .*

²² A definition like that would be consistent with Imai's (1983) characterization of the Harsanyi NTU value.

²³ Clearly, when $S = N$ this definition coincides with the one introduced in Sect. 2.

²⁴ Egalitarian solutions generalize the monotonic solutions introduced by Kalai and Samet (1985) to games with incomplete information.

6 The S-solution

In this section we apply the ideas developed in the preceding sections to construct an egalitarian-based cooperative solution.

Definition 6 (*S-solution*) A mechanism $\bar{\mu}_N \in \mathcal{M}_N$ is an *S-solution* if and only if there exist vectors $\lambda > 0$, $\alpha \geq 0$ and $\eta = (\mu_S)_{S \subseteq N} \in \mathcal{M}$ with $\mu_N = \bar{\mu}_N$ such that

- (i) μ_N is a solution of the primal problem for λ .
- (ii) α is a solution of the dual problem for λ .
- (iii) For each coalition $S \subset N$, μ_S is an optimal egalitarian threat for S w.r.t. $(\mu_{S \setminus i})_{i \in S}$, λ and α .
- (iv) μ_N is an egalitarian threat for N w.r.t. $(\mu_{N \setminus i})_{i \in N}$, λ and α .

The vector of interim utilities $U(\bar{\mu}_N)$ is called an *S-value*.

Alternatively, a bargaining solution can be defined replacing condition (iii) by

- (iii') For each coalition $S \subset N$, μ_S is an incentive compatible optimal egalitarian threat for S w.r.t. $(\mu_{S \setminus i})_{i \in S}$, λ and α .

In that case an S-solution is called *coalitionally incentive compatible*.

Conditions (i)-(iv) in our definition of an S-value have natural interpretations: (i) generalizes the λ -weighted utilitarian criterion, (ii) says that α is the vector of Lagrange multipliers associated with (i), and (iii) extends Harsanyi's (1963, sec. 9) optimal threats criterion to games with incomplete information. It follows from (i) that μ_N maximizes the Lagrangian in (3). Hence, (i) and (iv) imply that μ_N is also an optimal egalitarian threat for N w.r.t. $(\mu_{N \setminus i})_{i \in N}$, λ and α (i.e., μ_N solves (16) for N). Therefore, the whole vector of threats $\eta = (\mu_S)_{S \subseteq N}$ is an egalitarian solution.

Remark 3 By Lemma 1, we can equivalently define an S-solution replacing condition (iv) by

- (iv') $U(\mu_N)$ is warranted by λ , α and η .

This equivalent definition puts the S-solution in a form that makes it easily comparable with the M-solution. Indeed, with (iv'), the S-solution differentiates itself from the M-solution only in that the latter does not require optimal threats to meet our egalitarian criterion. Then, it follows that when $n = 2$ both solution concepts coincide²⁵.

Theorem 1 (Two-player games) *Let Γ be a two-player cooperative game with incomplete information. Then, any (non-degenerated) M-solution is an S-solution and viceversa. Moreover, if Γ is a two-person bargaining problem,²⁶ both solution concepts coincide with Myerson's (1984b) generalization of the Nash bargaining solution.*

Theorem 2 (Individual rationality) *Both variants of the S-solution are interim individually rational.*

²⁵ Singleton coalitions are not constrained by the egalitarian restrictions in (12)

²⁶ A two-person bargaining problem is a cooperative game satisfying: $n = 2$, $D_i = \{d_i\}$ for all $i \in N$ and $u_i(d^*, t) = 0$ for all $i \in N$ and $t \in T$, where $d^* := [d_i, d_j]$ is the disagreement outcome.

Proof Let μ_N be an S-solution supported by η, λ and α . For each $i \in N$, let $\hat{\mu}_i \in \mathcal{M}_i$ be defined by

$$\begin{aligned} & \sum_{t_{-i} \in T_{-i}} p(t_{-i} \mid t_i) \sum_{d_i \in D_i} \hat{\mu}_i(d_i \mid t_i) u_i(d_i, t) \\ &= \max_{d_i \in D_i} \sum_{t_{-i} \in T_{-i}} p(t_{-i} \mid t_i) u_i(d_i, t), \quad \forall t_i \in T_i, \end{aligned} \quad (17)$$

For each $t \in T$, the TU game $W(\eta, t, \lambda, \alpha)$ is weakly superadditive²⁷ (since decision sets are superadditive). Then, $\phi_i(N, W(\eta, t, \lambda, \alpha)) \geq v_i(\mu_i, t, \lambda, \alpha)$ for every $t \in T$. Also, for all $i \in N$, $\sum_{t_{-i} \in T_{-i}} p(t_{-i} \mid t_i) v_i(\mu_i, t, \lambda, \alpha) \geq \sum_{t_{-i} \in T_{-i}} p(t_{-i} \mid t_i) v_i(\hat{\mu}_i, t, \lambda, \alpha)$ for all $t_i \in T_i$, since μ_i is an optimal egalitarian threat for i . Then, we have that for each $i \in N$ and $t_i \in T_i$,

$$\begin{aligned} & \left(\lambda_i(t_i) + \sum_{\tau_i \in T_i} \alpha_i(\tau_i \mid t_i) \right) U_i(\mu_N \mid t_i) - \sum_{\tau_i \in T_i} \alpha_i(t_i \mid \tau_i) U_i(\mu_N \mid \tau_i) \\ &= \sum_{t_{-i} \in T_{-i}} p(t) \phi_i(N, W(\eta, t, \lambda, \alpha)) \\ &\geq \sum_{t_{-i} \in T_{-i}} p(t) v_i(\hat{\mu}_i, t, \lambda, \alpha) \\ &\geq \left(\lambda_i(t_i) + \sum_{\tau_i \in T_i} \alpha_i(\tau_i \mid t_i) \right) \max_{d_i \in D_i} \sum_{t_{-i} \in T_{-i}} p(t_{-i} \mid t_i) u_i(d_i, t) \\ &\quad - \sum_{\tau_i \in T_i} \alpha_i(t_i \mid \tau_i) \max_{d_i \in D_i} \sum_{t_{-i} \in T_{-i}} p(t_{-i} \mid \tau_i) u_i(d_i, (\tau_i, t_{-i})), \end{aligned} \quad (18)$$

where the first line follows from the fact that $U(\mu_N)$ is warranted by η, λ and α (cf. condition (iv')); the second line follows from the first part of the proof; and finally, the last inequality uses (17) applied to τ_i . The desired conclusion is obtained from (18) together with Remark 2. \square

The following result follows directly from the definitions [cf. Harsanyi (1963, Sect. 10)].

Theorem 3 (Generalization of the Harsanyi NTU value) *Let Γ be a cooperative game with complete information, i.e., T_i is a singleton for every $i \in N$. If $\bar{\mu}_N$ is an S-solution of Γ , then the utility allocation $U(\bar{\mu}_N)$ is a Harsanyi NTU value of Γ . Conversely, if the utility allocation $\bar{U} = (\bar{U}_i)_{i \in N}$ is a (non-degenerated) Harsanyi NTU value of Γ , then there exists an S-solution of Γ , $\bar{\mu}_N$, such that $\bar{U} = U(\bar{\mu}_N)$.*

We are now ready to compute our bargaining solution in the examples introduced in Sect. 3.

²⁷ A TU game (N, W) is weakly superadditive if and only if for each player $i \in N$, $W(S \setminus i) + W(\{i\}) \leq W(S)$ for all coalitions $S \subseteq N$ containing i . Clearly, by definition of the Shapley TU value, weak superadditivity implies that $\phi_i(N, W) \geq W(\{i\})$ for every $i \in N$.

6.1 Example 1

Let us consider the vector of utility weights $\bar{\lambda} = (\bar{\lambda}_1, \bar{\lambda}_2, \bar{\lambda}_3^H, \bar{\lambda}_3^L) = (1, 1, 9/10, 1/5)$. First, we notice that for any feasible mechanism $\mu_N \in \mathcal{M}_N$ we have that

$$U(\mu_N, \bar{\lambda}) := U_1(\mu_N) + U_2(\mu_N) + \frac{9}{10}U_3(\mu_N | H) + \frac{1}{5}U_3(\mu_N | L) \leq 10, \quad (19)$$

Consider now the problem of finding the best incentive compatible and individually rational utility allocation for each possible type of every player. Straightforward computations yield that the best allocation for player 1 is $(U_1, U_2, U_3^H, U_3^L) = (19/2, 1/2, 0, 0)$. By symmetry, the best allocation for player 2 is $(1/2, 19/2, 0, 0)$. Finally, $(0, 0, 10, 5)$ is simultaneously the best allocation for *both* types of player 3. These three allocations are incentive efficient, and they lie on the hyperplane $U(\mu_N, \bar{\lambda}) = 10$. Then, by convexity of \mathcal{M}_N^* , any individually rational and incentive efficient mechanism μ_N must satisfy $U(\mu_N, \bar{\lambda}) \geq 10$. Thus, (19) implies that the incentive efficient frontier coincides with the hyperplane $U(\mu_N, \bar{\lambda}) = 10$ on the individually rational zone. Therefore, in view of Theorem 2, condition (i) implies that a value allocation can only be supported by the utility weights $\bar{\lambda}$.²⁸

The utility weights $\bar{\lambda}$ reflect the optimal inter-type compromise between both types of player 3. To conceal her type, player 3 must achieve a balance that puts extra weight on the payoff maximization goals of type L (*inscrutability principle*). This is what explains that $\bar{\lambda}_3^L$ differs from the prior probability $p(L)$ by scaling up the actual utility of type L . On the other hand, the optimal value of the dual variables in the dual problem for $\bar{\lambda}$ is $(\bar{\alpha}_3(L | H), \bar{\alpha}_3(H | L)) = (0, 0)$.

Given these virtual scales, it can be easily verified that the only S-value of this game is²⁹

$$(U_1, U_2, U_3^H, U_3^L) = \left(\frac{61}{18}, \frac{61}{18}, \frac{60}{18}, \frac{20}{18}\right). \quad (20)$$

According to (20), the S-value rewards *both* types of player 3 less than players 1 and 2. This is in contrast to the M-value, which rewards players proportionally to their valuations of the project, as if there were no adverse selection problem. This feature of the S-value is due to the fact that, by requiring optimal threats to satisfy our egalitarian criterion, coalitions $\{1, 3\}$ and $\{2, 3\}$ cannot agree to fully distribute the total gains of cooperation. Indeed, because players in coalition $\{i, 3\}$ ($i = 1, 2$) are constrained to choose a feasible allocation giving them equal gains (in the virtual utility scales), then they have to settle for a sum of payoffs of at most $\$20/3 (< \$10)$ in state L . This implies that, in a two-person coalition with 3, players 1 and 2 cannot expect to get more than $\$29/6 (< \$5)$ each. That is, the expected “marginal contribution” of player 3 in a two-person coalition is strictly lower than what 1 and 2 can get in coalition $\{1, 2\}$. Consequently, 3 is perceived to have a weak bargaining position. It then appears that the S-value reflects the game situation better than the M-value.

²⁸ The same utility weights support the unique M-value (see Sect. 3).

²⁹ Detailed computations are provided in the Online supplementary material.

The asymmetry reflected in the allocation (20) comes uniquely from the fact that players 1 and 2 are adversely affected by 3's low type. None of the inefficiencies created by the incentive compatibility is taken into account: on one hand, incentive constraints are not essential for the grand coalition (i.e., incentive constraints do not impose any restriction for achieving an ex-post efficient allocation) and, on the other hand, optimal egalitarian threats are not required to be incentive compatible. The unique coalitionally incentive compatible S-value of this game is

$$(U_1, U_2, U_3^H, U_3^L) = \left(\frac{41}{12}, \frac{41}{12}, \frac{40}{12}, \frac{10}{12}\right). \quad (21)$$

When we take account of the incentive constraints that coalitions $\{1, 3\}$ and $\{2, 3\}$ face, our bargaining solution gives much less to player 3 in both states compared to the situation in which incentive constraints are only imposed for the grand coalition [compare (20) and (21)]. In fact, when coalition $\{i, 3\}$ (with $i = 1, 2$) is required to choose a mechanism that is incentive compatible, its members cannot agree on a virtual utility allocation giving them equal gains without an efficiency loss. Thus player 3's bargaining ability is further lowered by the necessity for players to trust each other.

It seems that, in this particular game, our solution concept provides much more agreement with what we expect the outcome to be.

6.2 Example 2

Proceeding as in Example 1, it can be shown that any incentive compatible and individually rational mechanism is incentive efficient if and only if it satisfies

$$\frac{4}{3}U_1(\mu_N | H) + \frac{1}{3}U_1(\mu_N | L) + U_2(\mu_N) + U_3(\mu_N) = 78, \quad (22)$$

The natural vector of utility weights is thus $\bar{\lambda} = (\bar{\lambda}_1^H, \bar{\lambda}_1^L, \bar{\lambda}_2, \bar{\lambda}_3) = (4/5, 1/5, 1, 1)$. For these utility weights, the corresponding dual variables are $(\bar{\alpha}_1(L | H), \bar{\alpha}_1(H | L)) = (0, 0)$. Then, we conclude that incentive constraints do not matter for the grand coalition. As it was previously discussed in Sect. 3.2, the participation of player 3 in the grand coalition releases players 1 and 2 from the incentive constraints they face in coalition $\{1, 2\}$. Unlike Example 1, here utility weights and prior probabilities coincide. This is so because player 3 allows 1 and 2 to fully distribute the gains from trade. Types are then essentially verifiable, as any transfer of utility can be implemented by a utility equivalent incentive compatible mechanism.

Given these virtual scales, it can be checked that the interim allocation in (8) is also the unique S-value of this game. Both the M-value and the S-value coincide because the virtual value of coalition $\{1, 2\}$ is computed while using the vector (λ, α) as specified for the grand coalition. By doing so, we act as if incentive constraints do not matter for coalition $\{1, 2\}$, although they do. To remedy this, we impose incentive constraints for all intermediate coalitions. The unique coalitionally incentive compatible S-value of this game is the allocation

$$(U_1, U_2, U_3^H, U_3^L) = (45, 13, 38.6, 0.8). \quad (23)$$

The S-value generates an interesting alternative to the M-value in de Clippel's example. This game however also puts in evidence some "difficulties" with our bargaining solution. First, notice that while it is the case that the coalitionally incentive compatible S-value rewards player 3, it is as if *both* players 1 and 2 pay \$0.8 to player 3 in exchange of her service. This may be considered as not reasonable since only player 2 needs the help of player 3 in order to extract the whole cooperative surplus. Second, the virtual worths of all coalitions in our bargaining solution are computed using the vector (λ, α) specified for the grand coalition. As a consequence, the efficiency losses due to the incentive compatibility at the level of all subcoalitions are not taken into account, unless incentive constraints are explicitly required.

All in all, it turns out that both examples presented in this paper are similar in nature, and that the S-solution prescribes intuitively appealing outcomes in each case.

7 Some comments about the (Non-)existence of the S-solution

The S-solution is characterized by strong equity conditions that may lead to its non-existence in some cases. In this section we shall exhibit an example of a 4-player cooperative game with *complete* information in which there is no S-solution. The following hinders the existence of the S-solution in this example: first, optimal egalitarian threats do not exist for some utility weights; second, optimal egalitarian threats vary discontinuously with the utility weights, which makes impossible the consistency of conditions (i) and (iv) in the definition of the S-solution. This example can be used to construct a game with incomplete information satisfying the same properties. The method is outlined in footnote 30 below. We study instead the game with complete information, this being however easier to analyze. Finally, we discuss the reasons why the methods and techniques used to obtain existence results of the Harsanyi NTU value cannot be well adapted to games with incomplete information.

7.1 Example 3: non-existence of the S-solution

Let Γ_C be the following cooperative game (with complete information): the set of players is $N = \{1, 2, 3, 4\}$. Decision options for every coalition are $D_i = \{d_i\}$ ($i \in N$), $D_{\{1,2\}} = [D_1 \times D_2] \cup \{d_{12}\} = \{[d_1, d_2], d_{12}\}$, $D_{\{1,3\}} = D_1 \times D_3 = \{[d_1, d_3]\}$, $D_{\{2,3\}} = [D_2 \times D_3] \cup \{d_{23}\} = \{[d_2, d_3], d_{23}\}$, $D_{\{1,2,3\}} = [D_1 \times D_{\{2,3\}}] \cup [D_{\{1,2\}} \times D_3]$, $D_{S \cup \{4\}} = D_S \times D_4$ ($S \subset N \setminus \{4\}$) and $D_N = [D_{\{1,2,3\}} \times D_4] \cup \{d_N^1, d_N^2, d_N^3, d_N^4\}$. Finally, utility functions are as follows:

d_N	(u_1, u_2, u_3, u_4)
$[d_1, d_2, d_3, d_4]$	$(0, 0, 0, 0)$
$[d_{12}, d_3, d_4]$	$(2, 2, 0, 0)$
$[d_1, d_{23}, d_4]$	$(0, 1, 1, 0)$
d_N^1	$(-1, 3, 3, 3)$
d_N^2	$(3, -1, 3, 3)$
d_N^3	$(3, 3, -1, 3)$
d_N^4	$(3, 3, 3, -1)$

We notice that decisions $[d_1, d_2, d_3, d_4]$, $[d_{12}, d_3, d_4]$ and $[d_1, d_{23}, d_4]$ are strictly Pareto dominated. Hence, the Pareto frontier (of the grand coalition) is a bounded surface contained in the hyperplane

$$u_1 + u_2 + u_3 + u_4 = 8.$$

Natural utility weights supporting an S-solution are $\bar{\lambda}_i = 1$ for all $i \in N$. However, no egalitarian solution can be constructed for $\bar{\lambda}$. Let us see this. For coalitions consisting of a single player i ($i \in N$), the unique optimal egalitarian threat is the (deterministic) mechanism d_i . Similarly, coalition $\{1, 2\}$ has a unique optimal egalitarian threat given by the (deterministic) mechanism d_{12} . For coalition $\{2, 3\}$ (resp. $\{1, 3\}$) we have that the unique optimal egalitarian threat is d_{23} (resp. $[d_1, d_3]$). Then, the egalitarian restrictions in (12) for coalition $S = \{1, 2, 3\}$ reduce to:

$$u_1(\mu_S) - u_3(\mu_S) = 1 \quad (24)$$

$$u_2(\mu_S) - u_1(\mu_S) = 1 \quad (25)$$

Condition (25) implies that $\mu_S([d_{12}, d_3]) = 1 - \mu_S([d_1, d_{23}]) = 0$. However, (24) requires that $\mu_S([d_{12}, d_3]) > 0$, which is a contradiction. We conclude that no egalitarian solution exists w.r.t. $\bar{\lambda}$.

Proposition 4 *The game Γ_C has no S-solution. This holds even if we allow some (but not all) utility weights to vanish.*³⁰

A detailed proof is given in the Appendix. A further difficulty prevents the existence of an S-solution in this game, namely, the optimal solution correspondence of (16) may not be upper-hemicontinuous in the utility weights. Consider, for instance, coalition $\{1, 2\}$. The set of feasible expected utility allocations for $\{1, 2\}$ is given by the line segment ow illustrated in Fig. 3.

³⁰ The game Γ_C can be used to construct a game with incomplete information for which there is no S-solution. Let N and $(D_S)_{S \subseteq N}$ be defined as in Γ_C . For each $i = 1, 2, 3$, let T_i be a singleton. Player 4 has private information in the form of two possible types $T_4 = \{A, B\}$ with prior probabilities $q(A) = 1 - q(B) > 0$. Utility functions are defined as follows: $w_i(d_N, A) = u_i(d_N)$ and $w_i(d_N, B) = \beta u_i(d_N)$ (with $\beta > 0$), where $(u_i)_{i \in N}$ is defined as in Γ_C . Then, the game $\Gamma_I = \{N, (D_S)_{S \subseteq N}, (w_i, T_i)_{i \in N}, q\}$ has no S-solution. Indeed, because the incentives of player 4 are fully aligned in both states, incentive constraints are not essential. Thus, we can set the Lagrange multipliers to be $\alpha_4(A | B) = \alpha_4(B | A) = 0$. Virtual utilities then reduce to λ -weighted utilities. The rest of the analysis follows, *mutatis mutandis*, the same reasoning as for Γ_C .

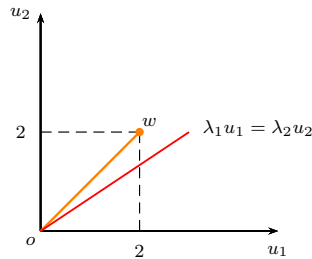


Fig. 3 Feasible allocations for $\{1, 2\}$

For any vector $\lambda > 0$ such that $\lambda_1/\lambda_2 \neq 1$, the unique optimal egalitarian threat is $\bar{\mu}_{\{1,2\}}(d_{12}) = 1 - \bar{\mu}_{\{1,2\}}([d_1, d_2]) = 0$, this being the unique feasible mechanism satisfying the egalitarian constraint $\lambda_1 u_1 = \lambda_2 u_2$. The corresponding utility allocation is o . However, when $\lambda_1/\lambda_2 = 1$, the unique optimal egalitarian threat is $\tilde{\mu}_{\{1,2\}}(d_{12}) = 1 - \tilde{\mu}_{\{1,2\}}([d_1, d_2]) = 1$, achieving the utility allocation w . We conclude that the optimal solutions correspondence of (16) for $S = \{1, 2\}$, viewed as a function of λ_1/λ_2 , is discontinuous. As stated earlier, this lack of continuity impedes conditions (i) and (iv) to be simultaneously satisfied as the utility weights λ accommodate. This issue can only be appreciated while exhaustively analyzing conditions (i) and (iv) for all values of λ (see proof of Proposition 4).

7.2 Free disposal and the structure of incentives

When information is complete, the above difficulties are ruled out by considering games whose characteristic function is comprehensive (free disposal of utility). Then, one is tempted to accommodate free disposal activities by introducing decisions in each D_S specifying how much utility a player may discard. This has no significant consequence when information is complete, however under asymmetric information, adding new decisions may change the incentive structure of the game: free disposal can be used for signaling purposes, i.e., for weakening incentive compatibility. As a result, for any interim utility allocation on the interim incentive efficient frontier (of the grand coalition), we cannot *generally* extend the original game by introducing additional decisions allowing players to discard utility (conditional on every state), while leaving the original utility allocation efficient in the expanded problem³¹. In order to illustrate this issue, consider again the (sub)game faced by players 1 and 3 in Example 1. Assume now that player 3 is allowed to dispose of her utility in state H . Specifically, let \tilde{d} be such that $u_3(\tilde{d}, H) = 0$, $u_3(\tilde{d}, L) = 5$ and $u_1(\tilde{d}, H) = u_1(\tilde{d}, L) = 0$. Decision \tilde{d} is equivalent to implement decision d_{13}^1 first, but then player 3 agrees to discard 10 units

³¹ Clearly, this issue is not present in Γ_C , this being a game with complete information. Neither is it in Γ_I , as incentive constraints are not essential in this game. Nevertheless, in more general games in which incentives constraints are binding, the same difficulties are also encountered. In those cases, in addition to the utility weights λ , also the dual variables α have to be taken into consideration. Exemplify such situations is, however, more difficult due to the endogenous nature of the dual variables.

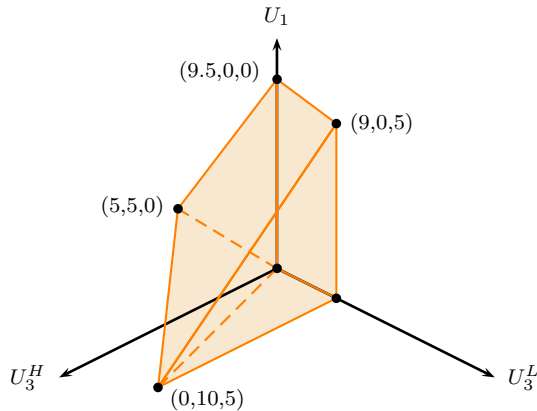


Fig. 4 Incentive feasible allocations for $\{1, 3\}$ in the expanded problem

of her utility in state H . Now consider the expanded problem with decision set $\tilde{D}_{\{1,3\}} = D_{\{1,3\}} \cup \{\tilde{d}\}$. The new set of incentive feasible interim utility allocations is depicted in Fig. 4.

When comparing Figs. 2 and 4, we observe that the game has substantially changed after \tilde{d} was introduced. As required, type H of player 3 can now achieve all the allocations in which he discards any nonnegative amount of utility. However, permitting free disposal facilitated also the fulfillment of incentive constraints, thus allowing both players to achieve higher interim utilities with respect to the original problem.³² In particular, any incentive efficient allocation in the expanded game is ex-post efficient, which is not the case in the original game (cf. Fig. 1). This implies that incentive constraints are not essential in the enlarged problem. In addition, the supporting utility weights to the Pareto frontier of the expanded bargaining problem are no longer the same as in the original problem.³³

Acknowledgements This paper makes up part of my Ph.D. dissertation written at Toulouse School of Economics. I am very much indebted to Françoise Forges for her insights, her continuous guidance and for innumerable discussions. I am also grateful to Thomas Mariotti, François Salanié, Peter Sudhölter and two anonymous referees of this journal for their comments and remarks that helped me to improve the paper. I acknowledge valuable feedback and suggestions from participants at “Dynamic Approach to Game and Economic Theory: Conference Celebrating the 65th Birthday of Sergiu Hart” and “10th BiGSEM Doctoral Workshop on Economic Theory”.

³² In this game, only type H has incentives to impersonate type L . In an effort to distinguish himself from type H , type L may agree on a mechanism that discards an appropriate amount of utility in state H . Clearly, this “commitment strategy” does not affect type L , but is harmful for type H . Therefore, type H will never accept any such mechanism. In this way, types become essentially verifiable.

³³ It is worth noticing that in two-person games the S-solution always exists. This follows from the existence of the M-solution, since by Theorem 1 both solutions coincide whenever $n = 2$. The issue illustrated in the previous example does not bring any difficulty for the existence of the S-solution in two-person games. The reason is that, equity imposes no restrictions for singleton coalitions. Hence, in this case, allowing for free disposal is no longer necessary for guaranteeing existence and (upper-hemi)continuity of the optimal egalitarian threats.

8 Appendix

8.1 Proof of Proposition 4

Let μ_N be an S-solution of Γ_C supported by λ and $\eta = (\mu_S)_{S \subseteq N}$. We verify recursively conditions (i)-(iv). Because Γ_C has complete information, there are no incentive constraints, which is equivalent to set $\alpha = 0$, so that virtual utility reduces to λ -weighted utility and the egalitarian criterion in (12) becomes

$$\lambda_i (u_i(\mu_S) - u_i(\mu_{S \setminus j})) = \lambda_j (u_j(\mu_S) - u_j(\mu_{S \setminus i})), \quad \forall i, j \in S. \quad (26)$$

For coalitions consisting of a single player i , it is clear that $u_i(\mu_i) = 0$. For all two-person coalitions containing player 4, $D_{\{i,4\}} = \{[d_i, d_4]\}$. Then, it follows immediately that $u_i(\mu_{\{i,4\}}) = u_4(\mu_{\{i,4\}}) = 0$ for all $i \in N \setminus 4$. Similarly, $D_{\{1,3\}} = \{[d_1, d_3]\}$, thus $u_1(\mu_{\{1,3\}}) = u_3(\mu_{\{1,3\}}) = 0$. Consider now coalition $\{1, 2\}$. It can be easily verified that an optimal egalitarian threat for this coalition satisfies:

$$(u_1(\mu_{\{1,2\}}), u_2(\mu_{\{1,2\}})) = \begin{cases} (2, 2), & \text{if } \lambda_1 = \lambda_2 > 0 \\ (y, y), \ y \in [0, 2], & \text{if } \lambda_1 = \lambda_2 = 0 \\ (0, 0), & \text{if } \lambda_1 \neq \lambda_2 \end{cases}$$

Similarly,

$$(u_2(\mu_{\{2,3\}}), u_3(\mu_{\{2,3\}})) = \begin{cases} (1, 1), & \text{if } \lambda_2 = \lambda_3 > 0 \\ (y, y), \ y \in [0, 1], & \text{if } \lambda_2 = \lambda_3 = 0 \\ (0, 0), & \text{if } \lambda_2 \neq \lambda_3 \end{cases}$$

We proceed now by cases.

Case 1: $\lambda_1 = \lambda_2 = \lambda_3 > 0$ Condition (26) applied to $S = \{1, 2, 3\}$ leads to Eqs. (24) and (25). We have already shown in Sect. 7 that these two equations are incompatible.

Case 2: $\lambda_1 = \lambda_2 > 0, \lambda_2 \neq \lambda_3$ Without loss of generality, we can set $\lambda_1 = \lambda_2 = 1$. It can be easily verified that, (26) implies: $u_1(\mu_{\{1,2,3\}}) = u_2(\mu_{\{1,2,3\}}) = u_1(\mu_{\{1,2,4\}}) = u_2(\mu_{\{1,2,4\}}) = 2$ and $u_3(\mu_{\{1,2,3\}}) = u_4(\mu_{\{1,2,4\}}) = u_2(\mu_{\{2,3,4\}}) = u_3(\mu_{\{2,3,4\}}) = u_4(\mu_{\{2,3,4\}}) = 0$. Then, condition (26) applied to N reduces to

$$u_1(\mu_N) = u_2(\mu_N) \quad (27a)$$

$$u_1(\mu_N) = \lambda_3 u_3(\mu_N) + 2 \quad (27b)$$

$$u_1(\mu_N) = \lambda_4 u_4(\mu_N) + 2 \quad (27c)$$

On the other hand, we have that

$$\begin{aligned} & u_1(d_N) + u_2(d_N) + \lambda_3 u_3(d_N) + \lambda_4 u_4(d_N) \\ &= \begin{cases} 2 + 3(\lambda_3 + \lambda_4), & \text{if } d_N = d_N^1, d_N^2 \\ 6 - \lambda_3 + 3\lambda_4, & \text{if } d_N = d_N^3 \\ 6 + 3\lambda_3 - \lambda_4, & \text{if } d_N = d_N^4 \end{cases} \end{aligned} \quad (27d)$$

Subcase 2.1: $\lambda_3 > 1, \lambda_4 > 1$ Condition (i) implies that $\mu_N(d_N^3) = \mu_N(d_N^4) = 0$. But then, (27a) requires that $\mu_N(d_N^1) = \mu_N(d_N^2) = 1/2$. Hence, $u_1(\mu_N) = u_2(\mu_N) = 1$ and $u_3(\mu_N) = u_4(\mu_N) = 3$. With this, (27b) reduces to $\lambda_3 = -1/3$, which is a contradiction.

Subcase 2.2: $\lambda_3 > 1, \lambda_4 = 1$ Condition (i) implies that $\mu_N(d_N^3) = 0$. But then, $u_1(\mu_N) < \lambda_3 u_3(\mu_N) + 2$, which contradicts (27b).

Subcase 2.3: $\lambda_3 > 1, \lambda_4 < 1$ Condition (i) implies that $\mu_N(d_N^4) = 1$. But then, (27c) reduces to $\lambda_4 = -1$.

Subcase 2.4: $\lambda_3 < 1, \lambda_4 \geq 1$ Condition (i) implies that $\mu_N(d_N^3) = 1$. But then, $u_1(\mu_N) > \lambda_3 u_3(\mu_N) + 2$, which contradicts (27b).

Subcase 2.5: $\lambda_4 < \lambda_3 < 1$ Condition (i) implies that $\mu_N(d_N^4) = 1$. The same conclusion as in the case 2.3 is obtained.

Subcase 2.6: $\lambda_3 < \lambda_4 < 1$ Condition (i) implies that $\mu_N(d_N^3) = 1$. The same conclusion as in the case 2.4 is obtained.

Subcase 2.7: $0 < \lambda_3 = \lambda_4 < 1$ Condition (i) implies that $\mu_N(d_N^3) = 1 - \mu_N(d_N^4) = \beta$ with $\beta \in [0, 1]$. Condition (27b) implies

$$\beta[1 + \lambda_3] + (1 - \beta)[1 - 3\lambda_3] = 0 \Rightarrow \beta = \frac{3\lambda_3 - 1}{4\lambda_3}$$

Similarly, (27c) implies

$$\beta[1 - 3\lambda_3] + (1 - \beta)[1 + \lambda_3] = 0 \Rightarrow \beta = \frac{1 + \lambda_3}{4\lambda_3}$$

Therefore, $3\lambda_3 - 1 = 1 + \lambda_3$, or equivalently, $\lambda_3 = 1$, which is a contradiction.

Subcase 2.8: $\lambda_3 = \lambda_4 = 0$ As in the previous case, condition (i) implies that $\mu_N(d_N^3) = 1 - \mu_N(d_N^4) = \beta$ with $\beta \in [0, 1]$. However, (27a) and (27b) imply that $u_1(\mu_N) = u_2(\mu_N) = 2$, which cannot be satisfied by any such mechanism.

Case 3: $\lambda_1 = \lambda_2 = \lambda_3 = 0$ From (26) with $S = N, i = 2$ and $j = 4$, we get that $\lambda_4 u_4(\mu_N) = 0$ (since $u_4(\mu_{\{1,3,4\}}) = 0$). Hence, $\sum_{i \in N} \lambda_i u_i(\mu_N) = 0$. However, condition (i) implies that $\sum_{i \in N} \lambda_i u_i(\mu_N) = 3$, which is a contradiction.

Case 4: $\lambda_1 = \lambda_2 = 0, \lambda_2 \neq \lambda_3$ Condition (26) applied to $S = N$ with $i = 2$ and $j = 3, 4$ gives $\lambda_3 u_3(\mu_N) = \lambda_4 u_4(\mu_N) = 0$ (since $u_3(\mu_{\{1,3,4\}}) = u_4(\mu_{\{1,3,4\}}) = 0$). Hence, $\sum_{i \in N} \lambda_i u_i(\mu_N) = 0$. However, condition (i) implies that $\sum_{i \in N} \lambda_i u_i(\mu_N) > 0$, which is a contradiction.

Case 5: $\lambda_1 \neq \lambda_2, \lambda_2 = \lambda_3 > 0$ Without loss of generality, we can set $\lambda_2 = \lambda_3 = 1$. It can be easily verified that, (26) implies: $u_2(\mu_{\{1,2,3\}}) = u_3(\mu_{\{1,2,3\}}) = u_2(\mu_{\{2,3,4\}}) = u_3(\mu_{\{2,3,4\}}) = 1$ and $u_1(\mu_{\{1,2,3\}}) = u_1(\mu_{\{1,2,4\}}) = u_2(\mu_{\{1,2,4\}}) = u_4(\mu_{\{1,2,4\}}) = u_4(\mu_{\{2,3,4\}}) = 0$. Then, condition (26) applied to N reduces to

$$\lambda_1 u_1(\mu_N) = u_2(\mu_N) - 1 \quad (28a)$$

$$\lambda_1 u_1(\mu_N) = u_3(\mu_N) - 1 \quad (28b)$$

$$\lambda_1 u_1(\mu_N) = \lambda_4 u_4(\mu_N) \quad (28c)$$

On the other hand, we have that

$$\begin{aligned} & \lambda_1 u_1(d_N) + u_2(d_N) + u_3(d_N) + \lambda_4 u_4(d_N) \\ &= \begin{cases} 2 + 3(\lambda_1 + \lambda_4), & \text{if } d_N = d_N^2, d_N^3 \\ 6 - \lambda_1 + 3\lambda_4, & \text{if } d_N = d_N^1 \\ 6 + 3\lambda_1 - \lambda_4, & \text{if } d_N = d_N^4 \end{cases} \end{aligned} \quad (28d)$$

Subcase 5.1: $\lambda_1 > 1, \lambda_4 > 1$ Condition (i) implies that $\mu_N(d_N^2) = 1 - \mu_N(d_N^3) = \beta$ with $\beta \in [0, 1]$. But then, since $u_2(\mu_N) = u_3(\mu_N)$ (by (28a) and (28b)), we must necessarily have that $\beta = 1/2$. Therefore, $u_2(\mu_N) = u_3(\mu_N) = 1$ and $u_1(\mu_N) = u_4(\mu_N) = 3$. However, this together with (28a) imply that $\lambda_1 = 0$, which is a contradiction.

Subcase 5.2: $\lambda_1 > 1, \lambda_4 = 1$ Condition (i) implies that $\mu_N(d_N^1) = 0$. Hence, (since $\lambda_1 > 2/3$) $\lambda_1 u_1(\mu_N) > u_2(\mu_N) - 1$, which contradicts (28a).

Subcase 5.3: $\lambda_1 > 1, \lambda_4 < 1$ Condition (i) implies that $\mu_N(d_N^4) = 1$. With this, (28a) implies that $\lambda_1 = 2/3 < 1$, which is a contradiction.

Subcase 5.4: $\lambda_1 < 1, \lambda_4 \geq 1$ Condition (i) implies that $\mu_N(d_N^1) = 1$. This together with (28a) imply that $\lambda_1 = -2$.

Subcase 5.5: $\lambda_1 < \lambda_4 < 1$ Condition (i) implies that $\mu_N(d_N^1) = 1$. The same conclusion as in case 5.4 is obtained.

Subcase 5.6: $\lambda_4 < \lambda_1 < 1$ Condition (i) implies that $\mu_N(d_N^4) = 1$. Hence, (28c) implies that $3\lambda_1 = -\lambda_4$. Therefore, (since $\lambda_1, \lambda_2 \geq 0$) $\lambda_1 = \lambda_2 = 0$, which is a contradiction.

Subcase 5.7: $0 < \lambda_4 = \lambda_1 < 1$ Condition (i) implies that $\mu_N(d_N^1) = 1 - \mu_N(d_N^4) = \beta$, with $\beta \in [0, 1]$. On the other hand, (28c) implies that $u_1(\mu_N) = u_4(\mu_N)$. Hence, we must have that $\beta = 1/2$. But then, (28a) implies that $\lambda_1 = 2$, which is a contradiction.

Subcase 5.8: $0 = \lambda_4 = \lambda_1$ As in the previous case, condition (i) implies that $\mu_N(d_N^1) = 1 - \mu_N(d_N^4) = \beta$, with $\beta \in [0, 1]$. Hence, $u_2(\mu_N) = u_3(\mu_N) = 3$. However, by (28a) and (28b), $u_2(\mu_N) = u_3(\mu_N) = 1$, which contradicts the previous fact.

Case 6: $\lambda_1 \neq \lambda_2, \lambda_2 = \lambda_3 = 0$ Condition (26) applied to $S = N$ with $i = 2$ and $j = 1, 4$ gives $\lambda_1 u_1(\mu_N) = \lambda_4 u_4(\mu_N) = 0$ (since $u_1(\mu_{\{1,3,4\}}) = u_4(\mu_{\{1,3,4\}}) = 0$). Hence, $\sum_{i \in N} \lambda_i u_i(\mu_N) = 0$. However, condition (i) implies that $\sum_{i \in N} \lambda_i u_i(\mu_N) = 3(1 + \lambda_4) > 0$, which is a contradiction.

Case 7: $0 = \lambda_2 \neq \lambda_1, \lambda_2 \neq \lambda_3$ Condition (26) applied to $S = N$ with $i = 2$ and $j = 1, 3, 4$ gives $\lambda_1 u_1(\mu_N) = \lambda_3 u_3(\mu_N) = \lambda_4 u_4(\mu_N) = 0$ (since $u_1(\mu_{\{1,3,4\}}) = u_3(\mu_{\{1,3,4\}}) = u_4(\mu_{\{1,3,4\}}) = 0$). Hence, $\sum_{i \in N} \lambda_i u_i(\mu_N) = 0$. However, condition (i) implies that $\sum_{i \in N} \lambda_i u_i(\mu_N) = 3(\lambda_1 + \lambda_3 + \lambda_4) > 0$ (since we must have $\lambda_1 > 0$ and $\lambda_3 > 0$). But this is a contradiction.

Case 8: $\lambda_2 > \lambda_1, \lambda_2 \neq \lambda_3$ Condition (26) applied to $S = N$ with $i = 1$ and $j = 2$ gives $\lambda_1 u_1(\mu_N) = \lambda_2 u_2(\mu_N)$ (since $u_1(\mu_{\{1,3,4\}}) = u_2(\mu_{\{2,3,4\}}) = 0$). On the other hand, condition (i) implies that $\mu_N(d_N^2) = 0$. But then, this implies that $\lambda_1 u_1(\mu_N) < \lambda_2 u_2(\mu_N)$, which is a contradiction.

Case 9: $\lambda_2 < \lambda_1, \lambda_2 \neq \lambda_3$ As in the previous case, we have that $\lambda_1 u_1(\mu_N) = \lambda_2 u_2(\mu_N)$. We also have that condition (i) implies $\mu_N(d_N^1) = 0$. But then, this implies that $\lambda_1 u_1(\mu_N) > \lambda_2 u_2(\mu_N)$, which is a contradiction.

We conclude that Γ_C has no S-solution. This completes the proof.

References

- Aumann RJJ (1994) Economic applications of the shapley value. Springer, Dordrecht, pp 121–133
- de Clippel G (2005) Values for cooperative games with incomplete information: an eloquent example. *Games Econ Behav* 53(1):73–82
- de Clippel G (2012) Egalitarianism in mechanism design. Working paper, Department of Economics, Brown University
- Forges F, Serrano R (2013) Cooperative games with incomplete information: some open problems. *Int Game Theory Rev* 15(02):1340009
- Forges F, Minelli E, Vohra R (2002) Incentives and the core of an exchange economy: a survey. *J Math Econ* 38(1):1–41 (Special Issue on Equilibrium with Asymmetric Information)
- Harsanyi JC (1963) A simplified bargaining model for the n-person cooperative game. *Int Econ Rev* 4(2):194–220
- Hart S (1985a) An axiomatization of harsanyi's nontransferable utility solution. *Econometrica* 53(6):1295–1313
- Hart S (1985b) Nontransferable utility games and markets: some examples and the harsanyi solution. *Econometrica* 53(6):1445–1450
- Hart S (2004) A comparison of non-transferable utility values. *Theory Decis* 56(1):35–46
- Hart S, Mas-Colell A (1996) Bargaining and value. *Econometrica* 64(2):357–380
- Holmström B, Myerson RB (1983) Efficient and durable decision rules with incomplete information. *Econometrica* 51(6):1799–1819
- Imai H (1983) On harsanyi's solution. *Int J Game Theory* 12(3):161–179
- Kalai E, Samet D (1985) Monotonic solutions to general cooperative games. *Econometrica* 53(2):307–327
- Myerson RB (1980) Conference structures and fair allocation rules. *Int J Game Theory* 9(3):169–182
- Myerson RB (1983) Mechanism design by an informed principal. *Econometrica* 51(6):1767–1797
- Myerson RB (1984a) Cooperative games with imcomplete information. *Int J Game Theory* 13(2):69–96
- Myerson RB (1984b) Two-person bargaining problems with incomplete information. *Econometrica* 52(2):461–487
- Myerson RB (1991) *Game Theory: analysis of conflict*. Harvard University Press, Harvard
- Myerson RB (1992) Fictitious-transfer solutions in cooperative game theory. Springer, Berlin, pp 13–33
- Myerson RB (2007) Virtual utility and the core for games with incomplete information. *J Econ Theory* 136(1):260–285
- Nash JF (1950) The bargaining problem. *Econometrica* 18(2):155–162
- Owen G (1971) Values of games without side payments. *Int J Game Theory* 1(1):95–109
- Peleg B, Sudhölter P (2007) *Introduction to the theory of cooperative games*. Springer, Berlin
- Roth AE (1980) Values for games without sidepayments: some difficulties with current concepts. *Econometrica* 48(2):457–465
- Shapley L (1953) *A value for n-person Games*. Princeton University Press, Princeton, pp 307–317
- Shapley L (1969) Utility comparisons and the theory of games. *Editions du CNRS*, Paris, pp 251–263

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.