An Egalitarian Value for Cooperative Games with Incomplete Information☆

Andrés Salamanca Lugo

Toulouse School of Economics

Abstract

A bargaining solution concept generalizing the Harsanyi NTU value is defined for cooperative games with incomplete information. Our definition of a cooperative solution implies that *all* coalitional threats are equitable when players make interpersonal utility comparisons in terms of some virtual utility scales. In contrast, Myerson's (1984b) generalization of the Shapley NTU value is only equitable for the grand coalition. When there are only two players, the two solutions are easily seen to coincide, however they may differ for general n-person games. By using the concept of virtual utility, our bargaining solution reflects the fact that players negotiate at the interim stage.

Keywords: Cooperative games, incomplete information, virtual utility.

JEL Classification: C71, C78, D82.

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].

Introducing asymmetric information in the analysis of cooperation involves two conceptual issues. First, an individual possessing non-verifiable private information may not have the incentive to truthfully reveal such information, as a consequence the final agreement may be subject to strategic manipulation. A cooperative agreement is then constrained by the necessity to provide the appropriate incentives for each party to reveal his private information. Second, when individuals negotiate at the interim stage, the bargaining process is itself a mechanism by

☆This paper makes part of my Ph.D. dissertation in Economics 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 acknowledge 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" for valuable feedback and suggestions. I would also like to thank François Salanié for helpful comments.

Email address: asalamancal.tse@gmail.com (Andrés Salamanca Lugo)

[1]The Shapley NTU value is sometimes referred as the "λ-transfer value".

which information is transmitted. Thus, individuals should take into account the information they reveal by proposing, accepting or refusing a particular agreement.

Incentive efficient mechanisms maximize the weighted sum of the players' interim utilities subject to incentive constraints. Myerson (1984a,b) proposes an approach in which the Lagrange multipliers of the incentive constraints are used to define the *virtual utility* of players. The concept of virtual utility reflects not only the signaling costs associated with incentive compatibility, but also the fact that individuals negotiate at the interim stage. By performing utility comparisons in the virtual utility scales, Myerson generalizes Shapley's (1969) fictitious-transfer procedure. Specifically, one associates to any incentive efficient mechanism a set of virtual utility scales (utility weights and Lagrange multipliers) and then considers the fictitious game in which utility is transferable in terms of such virtual scales. This approach has been used in Myerson (1984b) to extend the Shapley NTU value to an environment with incomplete information.

In this paper, we study two examples of cooperative games with incomplete information in which Myerson's (1984b) solution does not reflect well enough the game situation. Starting from the two-person bargaining problem studied in Section 10 of Myerson (1984a), 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 cooperating together. Since there is no conflict between players 1 and 2 for agreeing on an equitable and ex-post efficient allocation, it then appears that coalition $\{1, 2\}$ is more likely to form, thus leaving the informed player with a low expected payoff. According to Myerson's solution, the informed player extracts however a considerable amount of utility. Our example shares features with a complete information NTU game previously proposed by Roth (1980). In addition, we also study a three-player Bayesian cooperative game proposed by de Clippel (2005). In that example, the third player's only contribution is to partly release the other two players from the incentive constraints they face when they cooperate. One may consider this contribution important enough for rewarding the third player with a strictly positive payoff. Myerson's solution is however not sensitive to this informational contribution. De Clippel's example is an incomplete information version of a NTU game introduced by Owen (1972).

Our aim is to provide an alternative approach dealing with the "difficulties" identified in the examples above. Specifically, we construct a new solution concept for cooperative games with incomplete information that will extend the Harsanyi NTU value (cf. Theorem 1)[2]. For that, we build on Myerson's virtual utility approach to extend Myerson's (1980) balanced contributions characterization of the Harsanyi NTU value. While there might be several appealing ways to generalize the balanced contributions, here we adopt a method that is consistent with Imai's (1983) subgame value characterization of the Harsanyi NTU value. We first formulate an extended version of the subgame value condition based on Myerson's (1984b) principle for an equitable mechanism (cf. Definition 1). We then define an egalitarian criterion to be the unique extension of the balanced contributions (cf. Definition 2) that equivalently characterizes our generalized subgame value condition (cf. Proposition 1). Equipped with these equity notions, we extend Harsanyi's (1963, sec. 9) optimal threat strategies to define what we call *optimal*

[2]This actually holds only for the case of nondegenerate Harsanyi NTU values (i.e., those corresponding to strictly positive utility weights).

egalitarian threats. Specifically, we modify Myerson's (1984b) rational threats criterion by requiring coalitional threats to meet our egalitarian criterion. The idea in defining our cooperative solution is then to find an incentive efficient mechanism for which there exist virtual utility scales such that the mechanism would be equitable for the grand coalition when all intermediate coalitions commit to their optimal egalitarian threats in the fictitious game with transferable virtual utility.

Our solution concept is (interim) individually rational (cf. Theorem 2). In addition, it implies that *all* coalitional threats are equitable (cf. Proposition 2). In contrast, Myerson's solution is only equitable for the grand coalition. When there are only two players, our cooperative solution coincides with Myerson's solution. In particular, if the game corresponds to a bilateral bargaining problem both solutions coincide with Myerson's (1984a) neutral bargaining solution. However they both may differ for general n-person games. When we explicitly compute our solution concept in the examples described above, it turns out that it prescribes more intuitive and appealing outcomes than Myerson's solution.

The paper is organized as follows. Section 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 (1984b) virtual utility approach. The two motivating examples quoted in this introduction are studied in Section 3. The virtual utility approach is then used in Section 4 to introduce some equity principles for Bayesian cooperative games. In particular, we define our egalitarian criterion. In Section 5, the ideas of Section 4 are applied to define the optimal egalitarian threats. In Section 6 we introduce our cooperative solution. We then compute our solution in the examples studied in Section 3. Some additional comments about the existence of our solution concept are also discussed. Final remarks are presented in Section 7. Proofs are deferred to Section 8.

2. Formulation

2.1. Bayesian Cooperative Game

The model of a cooperative game with incomplete information is as follows. Let $N = \{1, 2, ..., n\}$ 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 coalitions[3] 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 notations[4] $t_S = (t_i)_{i\in S} \in T_S = \prod_{i\in S} T_i$, $t_{-i} = t_{N\setminus i} \in T_{-i} = T_{N\setminus i}$ 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

[3]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.

[4]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 $D_{\{i\}}$.

$T = T_N$. 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 $i \in N$. At the interim stage each player knows his type $t_i \in T_i$, and hence, we let $p(t_{-i} \mid t_i)$ 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 $u_i : D \times T \to \mathbb{R}$. 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 $\mu_N : T \to \Delta(D)$, where $\Delta(D)$ denotes the set of probability distributions over D. The interpretation is 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

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

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

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)), 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

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

We denote as $\mathcal{M}^*_N$ 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

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

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 $i \in N$.[5] Because the set of incentive-compatible mechanisms is a compact and convex polyhedron, 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

$$\max_{\mu_N \in \mathcal{M}_N^*} \sum_{i\in N} \sum_{t_i \in T_i} \lambda_i(t_i) U_i(\mu_N \mid t_i) \tag{2.1}$$

We shall refer to this linear-programming problem as the *primal problem for* λ. Let $\alpha_i(\tau_i \mid t_i) \geq 0$ 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

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

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

$$v_i(d, t, \lambda, \alpha) = \frac{1}{p(t_i)} \left[\left(\lambda_i(t_i) + \sum_{\tau_i\in T_i} \alpha_i(\tau_i \mid t_i) \right) u_i(d, t) - \sum_{\tau_i \in T_i} \alpha_i(t_i \mid \tau_i) \frac{p(t_{-i} \mid \tau_i)}{p(t_{-i} \mid t_i)} u_i(d, (\tau_i, t_{-i})) \right] \tag{2.2}$$

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

$$\mathcal{L}(\mu_N, \lambda, \alpha) = \sum_{t\in T} p(t) \sum_{d\in D} \mu_N(d \mid t) \sum_{i\in N} v_i(d, t, \lambda, \alpha) \tag{2.3}$$

Necessary and sufficient first order conditions for the primal problem imply that an incentive compatible mechanism $\bar{\mu}_N$ is incentive efficient if and only if there exists some vectors $\lambda \geq 0$ ($\lambda \neq 0$) and $\alpha \geq 0$, such that

$$\alpha_i(\tau_i \mid t_i) \left[U_i(\bar{\mu}_N \mid t_i) - U_i(\bar{\mu}_N, \tau_i \mid t_i) \right] = 0, \quad \forall i \in N,\ \forall t_i \in T_i,\ \forall \tau_i \in T_i \tag{2.4}$$

and $\bar{\mu}_N$ maximizes the Lagrangian function over all mechanisms in $\mathcal{M}_N$, namely,

$$\sum_{d\in D} \bar{\mu}_N(d \mid t) \sum_{i\in N} v_i(d, t, \lambda, \alpha) = \max_{d\in D} \sum_{i\in N} v_i(d, t, \lambda, \alpha), \quad \forall t \in T \tag{2.5}$$

Equation (2.4) is the usual *dual complementary slackness* condition. Condition (2.5) says that any incentive efficient mechanism $\bar{\mu}_N$ must put positive probability weight only on the decisions

[5] 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.

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 $\bar{\mu}_N$ would be ex-post efficient[6]. 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 to imitate him. Myerson (1984b) 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 (2.1). This *dual problem for λ* can be written as

$$\min_{\alpha \geq 0} \sum_{t \in T} p(t) \left(\max_{d \in D} \sum_{i \in N} v_i(d, t, \lambda, \alpha) \right) \tag{2.6}$$

2.3. The Myerson Value

Using the concept of virtual utility, Myerson (1984b) generalizes Shapley's (1969) fictitious-transfer procedure in order to extend the Shapley NTU value to an environment with incomplete information. Specifically, for any incentive efficient mechanism $\bar{\mu}_N$ one associates a vector (λ, α) of virtual utility scales. These scales correspond to the utility weights λ for which $\bar{\mu}_N$ 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 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 formal definition is given in Section 4 (see Remark 1).

Myerson (1984b) defines his bargaining solution to be an incentive efficient mechanism $\bar{\mu}_N$ for which there exist virtual scales (λ, α) such that $\bar{\mu}_N$ is equitable for the grand coalition w.r.t. (λ, α). The associated interim utility allocations are called an *M-value* (short for Myerson value). A formal definition of the M-value can be deduced from our cooperative solution concept (cf. Definition 6) by removing the egalitarian restrictions from our optimal threat criterion (see Section 5). Two variants of the value can be considered depending on whether optimal threats are required to be incentive compatible or not. Each possible definition of the value can be justified according to different assumptions on the commitment structure underlying the bargaining situation (see Section 6 in Myerson (1984b) for a detailed discussion). Myerson exclusively deals with the case in which only the mechanism of the grand coalition is constrained to be incentive compatible.

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

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 $D_i = \{d_i\}$ $(i \in N)$, $D_{\{1,2\}} = \{D_1 \times D_2\} \cup \{d_{12}\} = \{[d_1, d_2], d_{12}\}$, $D_{\{i,3\}} = \{D_i \times D_3\} \cup \{d^i_{i3}, d^3_{i3}\} = \{[d_i, d_3], d^i_{i3}, d^3_{i3}\}$ $(i = 1, 2)$ 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
$[d_1, d_2, d_3]$	$(0, 0, 0)$	$(0, 0, 0)$
$[d_{12}, d_3]$	$(5, 5, 0)$	$(5, 5, 0)$
$[d^1_{13}, d_2]$	$(0, 0, 5)$	$(0, 0, 10)$
$[d^3_{13}, d_2]$	$(10, 0, -5)$	$(10, 0, 0)$
$[d^2_{23}, d_1]$	$(0, 0, 5)$	$(0, 0, 10)$
$[d^3_{23}, d_1]$	$(0, 10, -5)$	$(0, 10, 0)$

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 his type, which is unknown to the other players. If 3's type is H ("high") then the project is worth \$10 to him. However, if 3's type is L ("low") then the project is only worth \$5 to him.

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^j_{13} $(j = 1, 3)$ 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 Section 2. Assume that threats are not required to be incentive compatible. Figure 1 illustrates the set of interim efficient (and individually rational) utility allocations for this (sub)game.

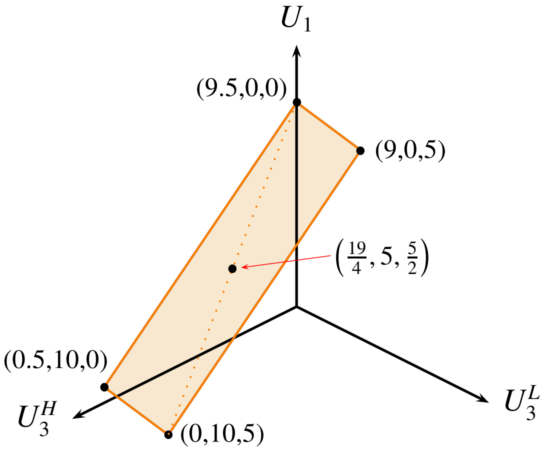


Figure 1: Efficient utility allocations for coalition $\{1, 3\}$

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 he 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^1_{13}. 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 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*[7] for player 3. On the other hand, if player 1 were a dictator, then he would demand his strong solution which implements the allocation $(19/2, 0, 0)$. Now consider a *random-dictatorship* in which each player is given equal chance of enforcing his 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 $\{1, 3\}$. Indeed, random-dictatorship together with efficiency characterize Myerson's (1984a) generalization of the Nash bargaining solution. It is then the unique M-value for this subgame[8].

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 both get 5 each, which is strictly preferred to $19/4$. When negotiating 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 player 3 less than the other players in both states.

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

[8]This allocation is implemented by the mechanism $\mu_{\{1,3\}}(d^1_{13} \mid L) = 1 - \mu_{\{1,3\}}(d^3_{13} \mid L) = 3/4$, $\mu_{\{1,3\}}(d^1_{13} \mid H) = \mu_{\{1,3\}}(d^3_{13} \mid H) = 1/2$.

Let us suppose now that threats are required to be incentive compatible. Figure 2 depicts the set of interim incentive efficient (and individually rational) 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 $(0,10,5)$. However, the strong solution for player 1 now implements the allocation $(9,0,0)$. Proceeding as before, random-dictatorship prescribes the value allocation $(9/2,5,5/2)$.[9] 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 the bargaining position of player 3. Therefore, 3's payoff from a value allocation should be further reduced when coalitional incentive constraints matter.

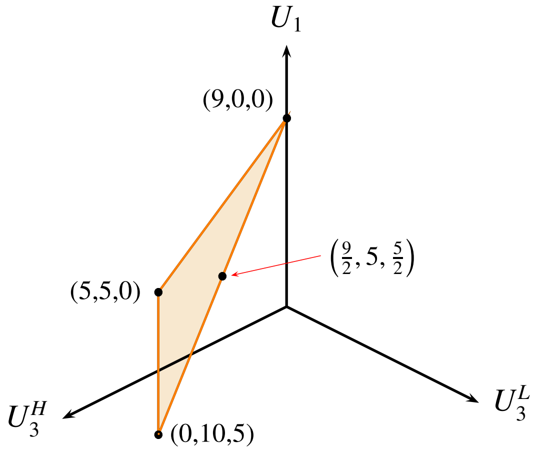


Figure 2: Incentive efficient utility allocations for coalition $\{1,3\}$

The unique (non-degenerated) M-value of our three player game is the utility allocation

$$\left(U_1, U_2, U_3^H, U_3^L\right) = \left(\tfrac{10}{3}, \tfrac{10}{3}, \tfrac{10}{3}, \tfrac{5}{3}\right). \tag{3.1}$$

For instance, the incentive efficient mechanism $\mu_N([d_{12}, d_3] \mid t) = \frac{2}{3}$, $\mu_N([d_{23}^2, d_1] \mid t) = \mu_N([d_{13}^1, d_2] \mid t) = \frac{1}{6}$ for all $t \in T_3$ is an M-solution. The value is supported by the utility weights[10] $(\lambda_1, \lambda_2, \lambda_3^H, \lambda_3^L) = (1, 1, 9/10, 1/5)$ and the Lagrange multipliers $(\alpha_1(L \mid H), \alpha_1(H \mid L)) = (0,0)$.[11]

An easy way to compute the M-solution in this game is simply to apply 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

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

[10]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.

[11]Explicit computations are given in the Appendix A.

allocation $(19/2, 1/2, 0, 0)$ (resp. $(1/2, 19/2, 0, 0)$). Averaging these utility vectors we obtain (3.1). Is is worth emphasizing that this procedure does not generally characterize the M-value. Yet for our example, it exhibits the reason why both types of player 3 extract a considerable amount of utility; namely, players are treated symmetrically in both states. Indeed, the random dictatorship procedure applied to N assumes coalitions are symmetric, so that threats can be disregarded. 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. Since all coalitions can achieve the maximal gains from the project in both states of the transferable virtual utility game, the M-value treats all coalitions symmetrically. This is so even when threats are required to be incentive compatible. For instance, the mechanism that implements d^j_{j3} ($j = 1, 2$) 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 not equitable). Moreover, such a threat is not "credible" since player $i \notin \{j, 3\}$ could not believe that player j would agree to implement d^j_{j3} in case cooperation in N breaks down.

The following feature of the game also explains why player 3 obtains a significantly high expected payoff. Suppose that player 2 has *definitely* dropped out of the game. Then, type L of player 3 becomes "surprisingly strong", since he has very little to lose by not agreeing to undertake the project and, at the same time, player 1 cannot go to close a deal with player 2. When player 3 is in a such surprisingly strong position, the outcome of the Myerson's (1984a) bargaining solution tends to be similar to what would have been the outcome if player 3 had become a dictator. Myerson (1991, p. 523) calls this property of his bargaining solution *arrogance of strength*. Because 3's type is not verifiable by player 1, type H also benefits from L's arrogance. A similar reasoning applies when player 1 leaves the game. Consider again the value allocations in the different subgames obtained when one player drops out of the game[12]. It will be convenient to write these allocations as a 4-dimensional vector with "−" for players outside coalition S:

$$\begin{aligned}(U_1, U_2, U_3^H, U_3^L) &= \left(\tfrac{9}{2}, -, 5, \tfrac{5}{2}\right)\\ (U_1, U_2, U_3^H, U_3^L) &= \left(-, \tfrac{9}{2}, 5, \tfrac{5}{2}\right)\\ (U_1, U_2, U_3^H, U_3^L) &= (5, 5, -, -)\end{aligned}$$

Now, for each allocation, we add a payoff for the missing player (in boldface) so that the resulting payoff vector is incentive efficient for N:

$$\begin{aligned}(U_1, \mathbf{U_2}, U_3^H, U_3^L) &= \left(\tfrac{9}{2}, \boldsymbol{\tfrac{1}{2}}, 5, \tfrac{5}{2}\right)\\ (\mathbf{U_1}, U_2, U_3^H, U_3^L) &= \left(\boldsymbol{\tfrac{1}{2}}, \tfrac{9}{2}, 5, \tfrac{5}{2}\right)\\ (U_1, U_2, \mathbf{U_3^H}, \mathbf{U_3^L}) &= (5, 5, \mathbf{0}, \mathbf{0})\end{aligned}$$

These assigned payoffs can be interpreted as a compensation that the forming two-person coalition grants to the remaining player for agreeing to leave the grand coalition. Averaging the

[12]For this analysis we assume that threats are required to be incentive compatible. The same conclusions can be obtained when threats do not need to be incentive compatible.

above three utility allocations we obtain (3.1). The underlying assumption in this computation is that any two-person coalition is equally likely to form. We argue however that the probability that $\{1,3\}$ or $\{2,3\}$ be the forming coalitions is rather small.

In this example, the M-value is insensitive to the negative externality created by 3's private information. The challenge for our solution concept is to better reflect the game situation.

3.2. Example 2: A Bilateral Trade Problem

Let us consider the following cooperative game with incomplete information originally proposed by de Clippel (2005). $N = \{1,2,3\}$, $T_1 = \{H,L\}$, $p(H) = 1 - p(L) = 4/5$, $D_i = \{d_i\}$ $(i = 1,2,3)$, $D_{\{1,2\}} = \{[d_1,d_2], d^1_{12}, d^2_{12}\}$, $D_{\{1,3\}} = \{[d_1,d_3]\}$, $D_{\{2,3\}} = \{[d_2,d_3]\}$, $D_N = \{[d_1,d_2,d_3], [d^1_{12},d_3], [d^2_{12},d_3], d_{23}, d_{32}\}$ and

(u_1,u_2,u_3)	$[d_1,d_2,d_3]$	$[d^1_{12},d_3]$	$[d^2_{12},d_3]$	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^1_{12} (resp. d^2_{12}) 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 $\{d^1_{12}, d^2_{12}\}$. 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. Player 3 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 he 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 states[13].

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

$$\left(U^H_1, U^L_1, U_2, U_3\right) = (45, 15, 39, 0)\,. \tag{3.2}$$

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 him some positive payoff, as players 1 and 2 have to rely on him 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. Thus, the M-value is not sensitive to the informational contribution of player 3.

[13] 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.

4. Equity Principles for Bayesian Cooperative Games

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 a condition called *balanced contributions* (see also Myerson (1992) for a more detailed explanation). According to this condition, for any two members of every coalition, the amount that each player would gain by the other's participation should be equal. In this section we build on Myerson's virtual utility approach to define equity principles for cooperative games with incomplete information generalizing the balanced contributions condition.

While there might be several appealing ways to generalize the balanced contributions condition, here we adopt a method that preserves a conceptual coherence with the equity principles developed by Myerson (1984b) in his M-solution. A well known property of the balanced contributions condition is that it can be equivalently characterized through a *subgame value* equity condition (see for instance Imai (1983)). The most important consequence of this dual relationship is that it reveals a reasonable way to extend the balanced contributions. We proceed first to formulate a "natural" extended version of the subgame value condition based on Myerson's (1984b) principle for an equitable mechanism (cf. Definition 1). We then construct an egalitarian criterion to be the unique extension of the balanced contributions (cf. Definition 2) that is consistent with our generalized subgame value condition (cf. Proposition 1).

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).

For the virtual game, we would like to identify mechanisms that are equitable in some well defined sense. For that, we assume that, as a threat during the bargaining process within the grand coalition N, each coalition $S \subset N$ commits to some mechanism $\mu_S : T_S \to \Delta(D_S)$.[14] 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 (2.2)) over μ_S. The expected virtual utility of type t_i of player $i \in S$ when the members of S agree on μ_S is

$$V_i(\mu_S \mid t_i, \lambda, \alpha) := \sum_{t_{-i} \in T_{-i}} p(t_{-i} \mid t_i) v_i(\mu_S, t, \lambda, \alpha). \tag{4.1}$$

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

$$W_S(\mu_S, t, \lambda, \alpha) = \sum_{i \in S} v_i(\mu_S, t, \lambda, \alpha). \tag{4.2}$$

[14]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.

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 coalitions[15] and virtual utility is conditionally transferable in state t w.r.t (λ, α). 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 and virtual utility is conditionally transferable in state t w.r.t. (λ, α).

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

$$V_i(\mu_S \mid t_i, \lambda, \alpha) = \sum_{t_{-i}\in T_{-i}} p(t_{-i} \mid t_i)\, \phi_i(S, W|_S(\eta_S, t, \lambda, \alpha)), \quad \forall t_i \in T_i,\ \forall i \in S. \tag{4.3}$$

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.

Remark 1. When $S = N$, the equality in (4.3) reduces to Myerson's (1984b) principle for equitable compromises.

Based on the principle of "equal gains", according to which cooperating players within a coalition should have equal compensation for their cooperation, we define the following egalitarian criterion.

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

$$\sum_{t_{-i}\in T_{-i}} p(t_{-i} \mid t_i) \sum_{j\in S\setminus i} \Big[v_i(\mu_S, t, \lambda, \alpha) - v_i(\mu_{S\setminus j}, t, \lambda, \alpha)\Big] = \\ \sum_{t_{-i}\in T_{-i}} p(t_{-i} \mid t_i) \sum_{j\in S\setminus i} \Big[v_j(\mu_S, t, \lambda, \alpha) - v_j(\mu_{S\setminus i}, t, \lambda, \alpha)\Big], \quad \forall t_i \in T_i,\ \forall i \in S. \tag{4.4}$$

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 (4.4) 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 the balanced contribution condition[16]. 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 (4.4) 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

[15] 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.

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

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

For given arbitrary mechanisms $(\mu_R)_{R\subset S}$, 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 $(\mu_R)_{R\subset S}$. However, it turns out that if the *whole* profile η_S is egalitarian, then it is also equitable, and viceversa.

Proposition 1 (**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 principle of equal gains in environments with incomplete information. Second, and most important, Proposition 1 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, equity equivalence cannot be established simply by taking expectations over the Myerson's (1980) balanced contributions characterization of the Shapley TU value. 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 and his TU values in the subgames with $|N| - 1$ players[17]. Apart from this clarification, the proof of Proposition 1 is straightforward. A detailed reasoning is presented in Section 8.

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

$$\left(\lambda_i(t_i) + \sum_{\tau_i\in T_i} \alpha_i(\tau_i \mid t_i)\right)\omega_i(t_i) - \sum_{\tau_i\in T_i} \alpha_i(t_i \mid \tau_i)\omega_i(\tau_i) = \sum_{t_{-i}\in T_{-i}} p(t)\phi_i(N, W(\eta, t, \lambda, \alpha)), \quad \forall t_i \in T_i, \quad \forall i \in N. \quad (4.5)$$

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

[17] This property can also be used to characterize Maschler-Owen's (1992) consistent NTU value (see Hart and Mas-Colell (1996)).

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 (1984b).

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

A cooperative solution for the virtual game will take into account not only the equity compromises among the different types of the various members of a coalition, but also the efficiency of the selected threat. Because in the virtual game payoffs are transferable, a natural efficiency criterion is given by the maximization of the total expected virtual worth of the coalition. Thus, considerations of equity and efficiency in the virtual game lead us to the following optimal threat criterion:

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

$$\max_{\mu_S \in \mathcal{M}_S} \sum_{t \in T} p(t) W_S(\mu_S, t, \lambda, \alpha) \quad \text{s.t. (4.4)} \tag{5.1}$$

The optimal threats criterion in (5.1) 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 1, we could also have defined an optimal threat replacing the egalitarian constraints (4.4) in (5.1) by the equity conditions in (4.3). However, this alternative definition is less tractable since threats of one coalition cannot be determined without knowledge of threats of *all* its subcoalitions[18]. It is easy to see that (5.1) generalizes Harsanyi's (1963, sec. 9) optimal threats criterion[19].

We can alternatively require threats to be incentive compatible. A mechanism μ_S is *incentive*

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

[19] Myerson (1992) has a formula for Harsanyi's optimal threats that immediately implies (5.1) when information is complete.

compatible for coalition $S \subset N$ if and only if

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

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* $S \subset N$ *w.r.t.* $(\mu_{S \setminus i})_{i \in S}$, λ *and* α *if and only if it solves* (5.1) *over all mechanisms in* $\mathcal{M}_S^*$.

Myerson (1984b, sec. 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 $(\mu_S)_{S \subset N}$ need not be either equitable or incentive compatible. Thus, Myerson's (1984b) rational-threat criterion maximizes the objective function in (5.1) 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 Section 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 a "credible threat", regardless of whether it is expected to be implemented or not. Then, at the very least, an optimal threat must be equitable.

By its definition, a profile $(\bar{\mu}_S)_{S \subset N}$ of optimal egalitarian threats must be recursively constructed. Start with all coalitions of the form $\{i\}$ $(i \in N)$. Then, (5.1) amounts to maximizing i's ex-ante expected virtual utility on the set $\mathcal{M}_i$. This problem is always feasible and its solution $\bar{\mu}_i$ is such that, for all $t_i \in T_i$, $\bar{\mu}_i(d_i \mid t_i) > 0$ only if d_i maximizes $\sum_{t_{-i} \in T_{-i}} p(t_{-i} \mid t_i) v_i(d_i, t, \lambda, \alpha)$. Now, for any coalition S with $1 < |S| < n$, given the threats $(\bar{\mu}_R)_{R \subset S}$ (already defined by induction), an optimal threat $\bar{\mu}_S$ is determined solving (5.1).

Unfortunately, the recursion above may be unfeasible. The reason is that mechanisms satisfying our egalitarian criterion may not always exist. When information is complete, this possibility is ruled out by the assumption that the NTU game is comprehensive ("free disposal" assumption). 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 not only the incentive structure of the game, but also the efficient frontier: free disposal can be used for signaling purposes, i.e., for weakening incentive compatibility.

Another alternative is to consider a more general class of mechanisms allowing players in any coalition $S \subseteq N$ to agree to discard utility. More formally, we have in mind mechanisms of the form (μ_S, x_S), where $\mu_S \in \mathcal{M}_S$ is a random joint plan and $x_S : T_S \to \mathbb{R}_-^S$ is a type-contingent deterministic vector of utility decrements. For one unit of decreased (real) utility, type t_i of player i will experience a reduction of

$$\beta_i(t_i, \lambda, \alpha) := \left[\lambda_i(t_i) + \sum_{\tau_i \in T_i} \alpha_i(\tau_i \mid t_i) - \sum_{\tau_i \in T_i} \alpha_i(t_i \mid \tau_i) \right] \frac{1}{p(t_i)}$$

units of his virtual utility. It can be shown that if α solves the dual for λ, all coefficients $\beta_i(t_i, \lambda, \alpha)$ are non-negative. However, we cannot prevent some of the coefficients to vanish. This is so even when the utility weights λ are all strictly positive. Hence, there may be players for which disposing of (real) utility does not result in virtual utility decrements, thus removing free disposal from the virtual utility game.

All previous difficulties are not specific of our egalitarian criterion, they also arise for other notions of egalitarianism in the context of mechanism design (see de Clippel (2012)). An extensive discussion about these issues is presented in the Appendix B.

6. The General Bargaining Solution

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

Definition 6 (H-bargaining solution).
A mechanism $\bar{\mu}_N \in \mathcal{M}_N$ is an H-bargaining 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) *$U(\mu_N)$ is warranted by λ, α and η.*

The vector of interim utilities $U(\bar{\mu}_N)$ is called an H-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 H-bargaining solution is called coalitionally incentive compatible.

Conditions $(i) - (iv)$ in our definition of an H-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) optimal threats criterion to games with incomplete information. Condition (iv) asserts that, to be a bargaining solution, the final agreement $\bar{\mu}_N$ must give every type of each player his expected Shapley TU value in the (λ, α)-virtual game.

By Lemma 1, conditions (i), (ii) and (iv) imply that an H-bargaining solution μ_N is equitable for N w.r.t. η, λ and α. On the other hand, for each coalition $S \subset N$, condition (iii) implies that μ_S is egalitarian for S w.r.t. $(\mu_{S \setminus i})_{i \in S}$, λ and α. Therefore, by Proposition 1, for each $S \subset N$, μ_S is equitable for S w.r.t. η_S, λ and α. This reasoning is summarized in the following proposition.

Proposition 2.
Let $\eta = (\mu_S)_{S\subseteq N}$ be part of an H-bargaining solution supported by the virtual scales (λ, α). Then, for each coalition $S \subseteq N$, η_S is equitable w.r.t. λ and α.

Unlike the M-solution, for which only the final agreement μ_N is equitable, our cooperative solution concept implies that coalitional threats $(\mu_S)_{S\subset N}$ are also equitable. Proposition 1 then implies that we can equivalently define an H-bargaining solution replacing condition (iv) by

(iv') μ_N is egalitarian for N w.r.t $(\mu_{N\setminus i})_{i\in N}$, λ and α.

Indeed, under this alternative formulation, it becomes evident that the H-value generalizes Harsanyi's (1963) definition of his NTU value.

Theorem 1 (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 H-bargaining solution, 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 H-bargaining solution of Γ, $\bar{\mu}_N$, such that $\bar{U} = U(\bar{\mu}_N)$.

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

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

6.1. Example 1

Based on the coalitional analysis presented in Section 3, we argue that a reasonable outcome for this example should satisfy the following three properties. First, it should reward both types of player 3 strictly less than players 1 and 2: despite the fact that in state H all players are symmetric, 1 and 2 are adversely affected by the likely presence of 3's "low" type. On the other hand, in state L, 3 has a weak bargaining position. In addition, incentive compatibility forces 1 and 2 to accept an efficiency loss in a bilateral bargaining with 3. Therefore, 3's payoff should be further reduced when coalitional incentive constraints matter. Second, because the bargaining position of type H is more favorable than that of type L, player 3 should be rewarded more in state H than in state L. Finally, since players 1 and 2 are symmetric, they both should get the same expected utility.

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

$$U(\mu_N, \bar{\lambda}) := U_1(\mu_N) + U_2(\mu_N) + \tfrac{9}{10} U_3(\mu_N \mid H) + \tfrac{1}{5} U_3(\mu_N \mid L) \leq 10, \qquad (6.1)$$

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, (6.1) 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 $\bar{\lambda}$.[20]

The utility weights $\bar{\lambda}$ reflect the optimal inter-type compromise between both types of player 3. To conceal his 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 $(\bar{\alpha}_3(L \mid H), \bar{\alpha}_3(H \mid L)) = (0, 0)$.

Given these virtual scales, it can be easily verified that the only H-value of this game is[21]

$$\left(U_1, U_2, U_3^H, U_3^L\right) = \left(\tfrac{61}{18}, \tfrac{61}{18}, \tfrac{60}{18}, \tfrac{20}{18}\right). \tag{6.2}$$

The value allocation gives less to player 3 in *both* states. This 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 in state L. 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. Hence, the expected "marginal contribution" of player 3 to the other players in a two-person coalition with him is strictly lower than what 1 and 2 can get by acting together. Consequently, 3 is perceived to have a weak bargaining position. It then appears that the H-value reflects the game situation better that the M-value.

The asymmetry reflected in the allocation (6.2) 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: incentive constraints are not essential for the grand coalition and optimal egalitarian threats are not required to be incentive compatible. The unique coalitionally incentive compatible H-value of this game is

$$\left(U_1, U_2, U_3^H, U_3^L\right) = \left(\tfrac{41}{12}, \tfrac{41}{12}, \tfrac{40}{12}, \tfrac{10}{12}\right). \tag{6.3}$$

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 only matter for the grand coalition (compare (6.2) and (6.3)). 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.

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

[20] The same utility weights support the unique M-value (see Section 3).

[21] Detailed computations are provided in the Appendix A.

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

$$\tfrac{4}{5}U_1(\mu_N \mid H) + \tfrac{1}{5}U_1(\mu_N \mid L) + U_2(\mu_N) + U_3(\mu_N) = 78, \tag{6.4}$$

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 $(\bar{\alpha}_1(L \mid H), \bar{\alpha}_1(H \mid L)) = (0, 0)$. Then, we conclude that incentive constraints do not matter for the grand coalition. As it was previously discussed in Section 3, 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 (3.2) is also the unique H-value of this game. Both the M-value and the H-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.

By imposing incentive constraints for all intermediate coalitions, we have that the unique coalitionally incentive compatible H-value of this game is the allocation

$$\left(U_1, U_2, U_3^H, U_3^L\right) = (45, 13, 38.6, 0.8). \tag{6.5}$$

The H-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 value allocation rewards player 3, it is as if *both* players 1 and 2 pay \$0.8 to player 3 in exchange of his 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 in both states. 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.

It turns out that both examples presented in this paper are similar in nature, and that our solution concept prescribes intuitively appealing outcomes in each case.

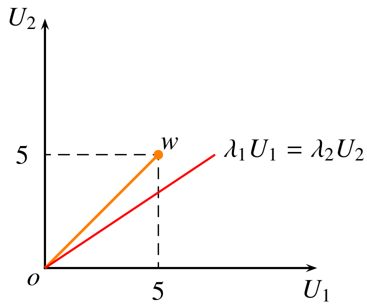
6.3. Some Comments about the Existence of the H-solutions

Two difficulties are imposed for proving an existence result of our cooperative solution concept. First, as it was already established in Section 5, the optimization problem (5.1) may not be feasible for some combination of parameters $(\lambda, \alpha, (\mu_{S\setminus i})_{i\in S})$. Second, even if (5.1) has a solution for any $(\lambda, \alpha, (\mu_{S\setminus i})_{i\in S})$ in a subset $\Theta \subseteq \prod_{i\in N} \mathbb{R}_+^{T_i} \times \prod_{i\in N} \mathbb{R}_+^{T_i\times T_i} \times \prod_{i\in S} \mathcal{M}_{S\setminus i}$, the corresponding optimal solutions correspondence is not generally upper-hemicontinuous on Θ. This would be required in case Kakutani's fixed point theorem were to be employed.

When there are only two players, an H-solution always exists. This follows from the existence theorem in Myerson (1984b), since both solutions coincide whenever $N = 2$. In the three-player case, the feasibility of (5.1) is not an issue. Indeed, for any two-person coalition $\{i, j\}$ $(i, j \in N)$, (5.1) has an optimal solution for all combination of parameters $(\lambda, \alpha, \bar{\mu}_i, \bar{\mu}_j)$. To see this, define the mechanism $\bar{\mu}_{\{i,j\}} \in \mathcal{M}_{\{i,j\}}$ by

$$\begin{aligned} \bar{\mu}_{\{i,j\}}((d_i, d_j) \mid (t_i, t_j)) &= \bar{\mu}_i(d_i \mid t_i)\bar{\mu}_j(d_j \mid t_j), \quad \text{if } (d_i, d_j) \in D_i \times D_j \subseteq D_{\{i,j\}}. \\ \bar{\mu}_{\{i,j\}}(d_{\{i,j\}} \mid (t_i, t_j)) &= 0, \quad \text{if } d_{\{i,j\}} \in D_{\{i,j\}} \setminus D_i \times D_j. \end{aligned}$$

It can be easily checked that $\bar{\mu}_{\{i,j\}}$ is incentive compatible and equitable for $\{i, j\}$ (w.r.t. λ, α and $(\bar{\mu}_i, \bar{\mu}_j)$). In fact, $v_i(\bar{\mu}_{\{i,j\}}, t, \lambda, \alpha) = v_i(\bar{\mu}_i, t, \lambda, \alpha)$ for every $i, j \in N$ and $t \in T$. This argument cannot be extended to more general games with $n > 3$ players. The difficulty in the three-player case is, however, that the optimal solutions correspondence may not be upper-hemicontinuous. Consider for instance problem (5.1) for coalition $\{1, 2\}$ in Example 1. A mechanism $\bar{\mu}_{\{1,2\}} \in \Delta(D_{\{1,2\}})$ is an optimal egalitarian threat for $\{1, 2\}$ if and only if it maximizes $\lambda_1 U_1(\mu_{\{1,2\}}) + \lambda_2 U_2(\mu_{\{1,2\}})$ subject to $\lambda_1 U_1(\mu_{\{1,2\}}) = \lambda_2 U_2(\mu_{\{1,2\}})$.[22] The set of feasible expected utility allocations for coalition $\{1, 2\}$ is given by the line segment $\vec{ow}$ as in the following figure:



For any vector $\lambda > 0$ such that $\lambda_1 \neq \lambda_2$, the unique solution of (5.1) is then $\bar{\mu}_{\{1,2\}}(d_{12}) = 1 - \bar{\mu}_{\{1,2\}}([d_1, d_2]) = 0$, since it is the unique feasible mechanism satisfying the egalitarian constraints. The corresponding utility allocation is o. However, when $\lambda_1 = \lambda_2$, the unique solution is $\tilde{\mu}_{\{1,2\}}(d_{12}) = 1 - \tilde{\mu}_{\{1,2\}}([d_1, d_2]) = 1$, achieving the utility allocation w. Hence the optimal solutions correspondence, viewed as a function of (λ_1, λ_2), is discontinuous.

When information is complete, the above difficulties are overcome by the free disposal assumption. As it was previously discussed in Section 5, free disposal is however more difficult to accommodate within the virtual utility approach. General existence of our cooperative solution remains an open problem for future research.

7. Concluding Remarks

The contribution of this paper is twofold. On one hand, we develop equity principles for cooperative games with incomplete information preserving a conceptual coherence with Myerson's (1984b) virtual utility approach. In particular, we obtain generalizations of Imai's (1983) subgame value equity condition and Myerson's (1980) balanced contributions condition. We also show that these two generalized notions of equity are in a dual relationship. On the other hand, we extend Harsanyi's (1963) NTU value to games with incomplete information. As most of the

[22]Notice that the virtual utility of players 1 and 2 do not depend on the dual variables α.

literature on cooperative games, our analysis is restricted to games with orthogonal coalitions, which rules out the possibility of strategic externalities.

The relevance of our solution concept is evidenced in two eloquent examples in which our value allocation is computed and contrasted with Myerson's (1984b) extension of the Shapley NTU value. It is shown that in both cases our solution concept provides much more agreement with what we expect the outcome to be. The fact that our considerations are better reflected by our bargaining solution must not be taken to be controversial. In the same way, the Shapley NTU value was introduced as a simplification of the Harsanyi NTU value, our cooperative solution is constructed to be a more sophisticated adaptation of Myerson's (1984b) theory. The proposed games are not intended to claim that the Myerson value must be abandoned. The value is an index and, to that extent, different values reflect different qualitative features of a same game.

A clear drawback of our solution concept is the difficulty to get a general existence result. This is not only due to technical difficulties, but more importantly it is connected to the way incentive constraints modify the shape of the feasible interim utility sets. Incentive compatibility has an impact on the signaling opportunities for the players that makes arguments significantly more complicated than in the special case of complete information. Despite the identified difficulties, we see our cooperative solution (and *a fortiori* our egalitarian criterion) as the most appealing way to extend the Harsanyi NTU value to games with incomplete information. Existence is clearly an issue that remains to be investigated.

8. Proofs

8.1. Proof of Proposition 1

It suffices to establish Proposition 1 only for the case of coalition N. For any other subcoalition $S \subset N$, the same arguments apply by replacing N by S. 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 player $j \in S \setminus i$, $\mu_{S\setminus j}$ is equitable for $S \setminus j$ (w.r.t. $\eta_{S\setminus j}$, λ and α). Thus

$$\sum_{t_{-i}\in T_{-i}} p(t_{-i} \mid t_i) \sum_{j\in S\setminus i} \phi_i\left(S \setminus j, W|_{S\setminus j}(\eta_{S\setminus j}, t, \lambda, \alpha)\right) = \sum_{t_{-i}\in T_{-i}} p(t_{-i} \mid t_i) \sum_{j\in S\setminus i} v_i(\mu_{S\setminus j}, t, \lambda, \alpha), \tag{8.1}$$

for all $t_i \in T_i$. Because μ_S is equitable for S (w.r.t. η_S, λ and α), we have that for any type t_i of a player $i \in S$,

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

$$= \frac{1}{|S|} \sum_{t_{-i} \in T_{-i}} p(t_{-i} \mid t_i) \Bigg[v_i(\mu_S, t, \lambda, \alpha) + \sum_{j \in S \setminus i} v_i(\mu_{S \setminus j}, t, \lambda, \alpha)$$
$$+ \sum_{j \in S \setminus i} \Big(v_j(\mu_S, t, \lambda, \alpha) - v_j(\mu_{S \setminus i}, t, \lambda, \alpha) \Big) \Bigg], \tag{8.2}$$

where in the second line we used the fact that the Shapley value of player i is the average of his marginal contribution to the grand coalition $W_S - W_{S \setminus i}$, and his values in the subgames with $|S| - 1$ players (see Hart (2004, p. 39)); and in the third line we used the definition of W_S together with (8.1). Finally, rearranging terms in (8.2) we get (4.4).

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 we have

$$\sum_{t_{-i} \in T_{-i}} p(t_{-i} \mid t_i) \left[W_S(\mu_S, t, \lambda, \alpha) - W_{S \setminus i}(\mu_{S \setminus i}, t, \lambda, \alpha) \right]$$
$$= \sum_{t_{-i} \in T_{-i}} p(t_{-i} \mid t_i) \left[\sum_{j \in S} v_j(\mu_S, t, \lambda, \alpha) - \sum_{j \in S \setminus i} v_j(\mu_{S \setminus i}, t, \lambda, \alpha) \right]$$
$$= \sum_{t_{-i} \in T_{-i}} p(t_{-i} \mid t_i) v_i(\mu_S, t, \lambda, \alpha) + \sum_{t_{-i} \in T_{-i}} p(t_{-i} \mid t_i) \sum_{j \in S \setminus i} \left[v_j(\mu_S, t, \lambda, \alpha) - v_j(\mu_{S \setminus i}, t, \lambda, \alpha) \right]$$
$$= \sum_{t_{-i} \in T_{-i}} p(t_{-i} \mid t_i) v_i(\mu_S, t, \lambda, \alpha) + \sum_{t_{-i} \in T_{-i}} p(t_{-i} \mid t_i) \sum_{j \in S \setminus i} \left[v_i(\mu_S, t, \lambda, \alpha) - v_i(\mu_{S \setminus j}, t, \lambda, \alpha) \right]$$
$$= \sum_{t_{-i} \in T_{-i}} p(t_{-i} \mid t_i) \left[|S| v_i(\mu_S, t, \lambda, \alpha) - \sum_{j \in S \setminus i} v_i(\mu_{S \setminus j}, t, \lambda, \alpha) \right], \tag{8.3}$$

where the first equality comes from the definition of W_S and the third 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,

$$\sum_{t_{-i} \in T_{-i}} p(t_{-i} \mid t_i) \phi_i(S, W|_S(\eta_S, t, \lambda, \alpha))$$
$$= \sum_{\substack{R \subseteq S \\ i \in R}} \frac{(|R| - 1)!(|S| - |R|)!}{|S|!} \sum_{t_{-i} \in T_{-i}} p(t_{-i} \mid t_i) \left[W_R(\mu_R, t, \lambda, \alpha) - W_{R \setminus i}(\mu_{R \setminus i}, t, \lambda, \alpha) \right]$$
$$= \sum_{t_{-i} \in T_{-i}} p(t_{-i} \mid t_i) \sum_{\substack{R \subseteq S \\ i \in R}} \frac{(|R| - 1)!(|S| - |R|)!}{|S|!} \left[|R| v_i(\mu_R, t, \lambda, \alpha) - \sum_{j \in R \setminus i} v_i(\mu_{R \setminus j}, t, \lambda, \alpha) \right]$$
$$= \sum_{t_{-i} \in T_{-i}} p(t_{-i} \mid t_i) \Bigg[\sum_{\substack{R \subseteq S \\ i \in R}} \frac{|R|!(|S| - |R|)!}{|S|!} v_i(\mu_R, t, \lambda, \alpha)$$
$$- \sum_{\substack{R \subseteq S \\ i \in R}} \sum_{j \in R \setminus i} \frac{(|R| - 1)!(|S| - |R|)!}{|S|!} v_i(\mu_{R \setminus j}, t, \lambda, \alpha) \Bigg]$$

$$= \sum_{t_{-i} \in T_{-i}} p(t_{-i} \mid t_i) \left[\sum_{\substack{R \subseteq S \\ i \in R}} \frac{|R|!(|S| - |R|)!}{|S|!} v_i(\mu_R, t, \lambda, \alpha) \right.$$

$$\left. - \sum_{\substack{R \subset S \\ i \in R}} \frac{|R|!(|S| - |R| - 1)!}{|S|!} |S \setminus R| v_i(\mu_R, t, \lambda, \alpha) \right]$$

$$= \sum_{t_{-i} \in T_{-i}} p(t_{-i} \mid t_i) v_i(\mu_S, t, \lambda, \alpha),$$

where the first equality follows from the definition of the Shapley TU value and the second equality uses (8.3). This completes the proof.

8.2. Proof of Theorem 2

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

$$\sum_{t_{-i} \in T_{-i}} p(t_{-i} \mid t_i) \sum_{d_i \in D_i} \hat{\mu}_i(d_i \mid t_i) u_i(d_i, t) = \max_{d_i \in D_i} \sum_{t_{-i} \in T_{-i}} p(t_{-i} \mid t_i) u_i(d_i, t), \quad \forall t_i \in T_i, \tag{8.4}$$

For each $t \in T$, the TU game $W(\eta, t, \lambda, \alpha)$ is weakly superadditive[23] (this is a consequence of Lemma 2 below). 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 $t_i \in T_i$,

$$\left(\lambda_i(t_i) + \sum_{\tau_i \in T_i} \alpha_i(\tau_i \mid t_i) \right) U_i(\mu_N \mid t_i) - \sum_{\tau_i \in T_i} \alpha_i(t_i \mid \tau_i) U_i(\mu_N \mid \tau_i)$$
$$= \sum_{t_{-i} \in T_{-i}} p(t) \phi_i(N, W(\eta, t, \lambda, \alpha))$$
$$\geq \sum_{t_{-i} \in T_{-i}} p(t) v_i(\hat{\mu}_i, t, \lambda, \alpha)$$
$$\geq \left(\lambda_i(t_i) + \sum_{\tau_i \in T_i} \alpha_i(\tau_i \mid t_i) \right) \max_{d_i \in D_i} \sum_{t_{-i} \in T_{-i}} p(t_{-i} \mid t_i) u_i(d_i, t)$$
$$- \sum_{\tau_i \in T_i} \alpha_i(t_i \mid \tau_i) \max_{d_i \in D_i} \sum_{t_{-i} \in T_{-i}} p(t_{-i} \mid \tau_i) u_i(d_i, (\tau_i, t_{-i})). \tag{8.5}$$

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 (8.4) applied to τ_i. The desired conclusion is obtained from (8.5) together with Remark 2.

In the proof we have used the following lemma.

[23] 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 $i \in N$.

Lemma 2.
Let $\eta \in \mathcal{M}$ be a vector of egalitarian threats (w.r.t. λ and α). Let $i \in N$ be a fixed player and take $\hat{\mu}_i \in \mathcal{M}_i$. Define $\hat{\eta} \in \mathcal{M}$ by

$$\hat{\mu}_S = \begin{cases} (\hat{\mu}_i, \mu_{S\setminus i}), & \text{if } i \in S \\ \mu_S, & \text{if } i \notin S \end{cases}$$

where $(\hat{\mu}_i, \mu_{S\setminus i}) \in \mathcal{M}_S$ is defined by

$$\begin{aligned} (\hat{\mu}_i, \mu_{S\setminus i})((d_i, d_{S\setminus i}) \mid t_S) &= \hat{\mu}_i(d_i \mid t_i)\mu_{S\setminus i}(d_{S\setminus i} \mid t_{S\setminus i}), \quad \text{if } (d_i, d_{S\setminus i}) \in D_i \times D_{S\setminus i} \subseteq D_S \\ (\hat{\mu}_i, \mu_S)(d_S \mid t_S) &= 0, \quad \text{if } d_S \in D_S \setminus D_i \times D_{S\setminus i}. \end{aligned}$$

Then, $\hat{\eta}$ is egalitarian (w.r.t. λ and α). Moreover, if μ_S is incentive compatible for S, so is $\hat{\mu}_S$.

Proof. Let $i \in N$ be a fixed player. Let $S \subseteq N$ be such that $i \notin S$. Then, $\hat{\mu}_S = \mu_S$ and $\hat{\mu}_{S\setminus j} = \mu_{S\setminus j}$ for all $j \in S$. Then, clearly $\hat{\mu}_S$ is egalitarian for S w.r.t. $(\hat{\mu}_{S\setminus j})_{j\in S}$, λ and α. Also, if μ_S is incentive compatible for S, so is $\hat{\mu}_S$. Now, let $S \subseteq N$ be such that $i \in S$. Then, for all $t \in T$ we have that

$$v_i(\hat{\mu}_S, t, \lambda, \alpha) - v_i(\hat{\mu}_{S\setminus j}, t, \lambda, \alpha) = 0 = v_j(\hat{\mu}_S, t, \lambda, \alpha) - v_j(\hat{\mu}_{S\setminus i}, t, \lambda, \alpha), \quad \forall j \in S \setminus i.$$

Then, it follows immediately that $\hat{\mu}_S$ is egalitarian for S w.r.t. $(\hat{\mu}_{S\setminus j})_{j\in S}$, λ and α. It is straightforward to check that if μ_S is incentive compatible for S, so is $\hat{\mu}_S$.

□

9. References

1. de Clippel, G. (2005). "Values for cooperative games with incomplete information: An eloquent example". *Games and Economic Behavior*, 53 pp. 73-82.
2. de Clippel, G. (2012). "Egalitarianism in mechanism design". Working paper, Department of Economics, Brown University.
3. Harsanyi, J. (1963). "A simplified bargaining model for the *n*-person cooperative game". *International Economic Review*, 4, pp. 194-220.
4. Hart, S. (2004). "A comparison of non-transferable utility values". *Theory and Decisions*, 56, pp 35-46.
5. Hart, S. and Mas-Colell, A. (1996). "Bargaining and the value". *Econometrica*, 64, pp. 357-380.
6. Holmström, B. and Myerson, R. (1983). "Efficient and durable decision rules with incomplete information". *Econometrica*, 51, pp 1779-1819.
7. Imai, H. (1983). "On Harsanyi's solution". *International Journal of Game Theory*, 12, pp. 161-179.
8. Maschler, M. and Owen, G. (1992). The consistent Shapley value for games without sidepayments. In: *Rational Interaction*. R. Selten (Ed.). New York: Springer-Verlag, pp. 5-12.
9. Myerson, R. (1980). "Conference structures and fair allocation rules". *International Journal of Game Theory*, 9, pp. 169-182.

10. Myerson, R. (1983). “Mechanism design by an informed principal”. *Econometrica*, 51, pp. 1767-1797.
11. Myerson, R. (1984a). “Two-person bargaining problems with incomplete information”. *Econometrica*, 52, pp. 461-488.
12. Myerson, R. (1984b). “Cooperative games with incomplete information”. *International Journal of Game Theory*, 13, pp. 69-96.
13. Myerson, R. (1991). *Game Theory: Analysis of Conflict*, Harvard University Press.
14. Myerson, R. (1992). Fictitious-transfers solutions in cooperative game theory. In: *Rational Interaction*. R. Selten (Ed.). New York: Springer-Verlag, pp. 13-33.
15. Owen, G. (1972). “A value for games without side payments”. *International Journal of Game Theory*, 1, pp. 95-109.
16. Roth, A. (1980). “Values for games without side-payments: Some difficulties with current concepts”. *Econometrica*, 48, pp. 457-465.
17. Shapley, L. (1953). A Value for *n*-person Games. In: *Contributions to the Theory of Games II*. H. W. Kuhn and A. W. Tucker (Eds.). Princeton: Princeton University Press, pp. 307-317.
18. Shapley, L. (1969). Utility comparisons and the theory of games. In: *La Decision*. Paris: Editions du CNRS, pp. 251-263.