Random Dictatorship and the Value in Cooperative Games with Incomplete Information☆

Andrés Salamanca Lugo

Department of Business and Economics, University of Southern Denmark

Abstract

In this paper we define a bargaining solution for cooperative games with incomplete information. Our solution concept is inspired in Myerson's [Mechanism design by an informed principal, *Econometrica*. (1983), 51, 1767-1797] theory on the informed principal problem and the random dictatorship procedure. It has the essential feature of generalizing the Maschler-Owen consistent value for non-transferable utility games. Our main results are individual rationality, incentive (second best) efficiency and existence of our cooperative solution. To obtain these results we restrict our analysis to cooperative games with stochastically independent types, private values and orthogonal coalitions.

Keywords: Cooperative games, incomplete information, virtual utility.

JEL Classification: C71, C78, D82.

1. Introduction

The value of a cooperative game is an a priori evaluation of the utility a player expects from the participation in the game. The classical concept of value was introduced by Shapley (1953) for transferable utility (TU) games. Maschler and Owen (1989, 1992) have defined a *consistent value* that extends Shapley's value to games with non-transferable utility (NTU). Their definition is based on a random order arrival procedure in which players successively enter the cooperation until the grand coalition is formed. In Hart (1994, 2005), it is shown that the consistent value may equivalently be expressed as the vector of expected marginal contributions of the players, once the marginal contributions have been appropriately defined. This alternative characterization of the consistent value can be interpreted by means of a recursive *conditional random dictatorship* procedure (see de Clippel, Peters and Zank (2004)): a player i is picked at random, with all players having equal probabilities. Then, player i is given the power of dictatorship *conditional* on giving the other players in the coalition what they would get in the value of the subgame restricted to the subcoalitions not containing player i.

In this paper, we elaborate on the conditional random dictatorship procedure to provide a generalization of the consistent value to cooperative games with incomplete information – the

☆This version: February 8, 2019. First version: April 20, 2017.

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

MO-solution –. Things become significantly more complicated when we try to extend the conditional random dictatorship procedure to games with incomplete information. Under incomplete information, agreements are generally mechanisms rather than simply utility allocations. The enforcement of a mechanism however relies on the players' claims about their private information. Yet a player may not have the incentive to truthfully reveal his information. Allowable mechanisms must therefore be self-enforcing with respect to private information, in the sense of being incentive compatible. When a player possessing private information is given all the bargaining ability to coordinate the actions inside a coalition, the choice of a particular incentive compatible mechanism may signal part of his private information to the other participants. With this new information, the members of the coalition may find new opportunities to strategically manipulate their private information or to refuse to cooperate. Myerson (1983) developed a theory of inscrutable mechanism selection by an informed individual with all the bargaining ability. We build on Myerson's approach to develop a generalization of the random dictatorship procedure.

Incentive compatible mechanisms are characterized by a system of linear inequalities. The dual variables associated to these inequalities yield shadow prices that can be used to define the *virtual utility* of the players. By allowing the players to transfer utilities in terms of these virtual utility scales, Myerson (1984a,b) generalizes the Harsanyi-Shapley fictitious-transfer procedure to games with incomplete information[1]. This approach is applied in Myerson (1984a) to extend Nash's (1950) bargaining solution to two-person bargaining problems with incomplete information. It has also been used in Myerson (1984b) to generalize Shapley's (1969) NTU value to games with incomplete information. More recently, elaborating also on the virtual utility approach, Salamanca (2016) constructs a cooperative solution extending Harsanyi's (1963) NTU value to games with incomplete information. Myerson's (1983) theory of inscrutable mechanism selection is axiomatically derived, however it can also be characterized using the virtual utilities. We exploit this characterization in order to define the MO-solution.

The MO-solution specifies a rational threat mechanism for every coalition. Each of these threats mechanisms is (interim) individually rational, incentive (second best) efficient and equitable for the coalition. This is in contrast with Myerson's (1984b) *M-solution*, which is only shown to be individually rational, incentive efficient and equitable for the grand coalition (final agreement). Unlike the M-solution, Salamanca's (2016) *H-solution* specifies threats that are equitable for every coalition. It is noteworthy that the latter two cooperative solutions are based on a notion of equity that differs from the one implied by the conditional random dictatorship procedure. In particular, the underlying principles for an equitable agreement in our solution concept takes into account the way incentive compatibility shapes the inter-type compromises for the players inside the different coalitions. These compromises are uniquely determined with respect to the grand coalition in the case of the M-solution and the H-solution. The MO-solution defines "credible" threats in the same way subgame perfection does for extensive form games. This property is called *subcoalition perfectness* by Hart and Mas-Colell (1996, p. 366). We compute and compare the above mentioned cooperative solutions in two eloquent examples proposed

[1]The reader is referred to Myerson (1992) for a detailed explanation about the Harsanyi-Shapley fictitious-transfer procedure for NTU games.

by de Clippel (2005) and Salamanca (2016). One final aspect of the MO-solution is that it coincides with Myerson's (1984a) *neutral solution* in two-person bargaining problems with incomplete information[2].

Our main results are individual rationality and existence of the MO-solution. To obtain these results we restrict our analysis to cooperative games with stochastically independent types, private values and orthogonal coalitions. Independent types is a simplifying assumption that we can make without loss of generality since the MO-solution satisfies the invariance probability axiom described in Myerson (1984a). Private values asserts that a player cares "directly" only about his private information. However, a player cares "indirectly" about the other players' private information in so far as such information affects the cooperative agreement. Finally, orthogonal coalitions means that the actions of the members of a coalition do not affect the utilities of the member of the complementary coalition. This assumption is standard in cooperative game theory and it excludes strategic externalities between the coalitions[3].

The paper is organized as follows. In the rest of this section we present a summary of the facts one needs to know about the consistent value in games with complete information. Section 2 is devoted to specifying formally the model of a cooperative game with incomplete information, the assumptions on the class of games considered and the virtual utility approach. In Section 3, we introduce a simple two-person bargaining problem with incomplete information which motivates our approach. Section 4 extends Myerson's (1983) theory on the informed principal problem to obtain a generalization of the conditional random dictatorship procedure to games with incomplete information. The MO-solution is defined in Section 5. We then present the main results: characterization, existence and individual rationality. Finally, comparisons of the different value-like solutions are also found in this section. Section 6 is devoted to a summary and final comments. Proofs are relegated to Section 7.

1.1. Preliminaries

A NTU game is a pair (N, V) where $N = \{1, 2, \ldots, n\}$ is the set of players and $V(\cdot)$ is a function that assigns a subset $V(S) \subseteq \mathbb{R}^S$ to every (nonempty) coalition $S \subseteq N$. A special class of NTU games is the hyperplane games. An NTU game (N, V) is a *hyperplane game* (or *H-game*) if for each coalition S, $V(S) = \{x^S \in \mathbb{R}^S \mid \sum_{i\in S} \lambda_i^S x_i^S \leq v(S)\}$ for some real valued function $v : 2^N \to \mathbb{R}$ (with $v(\emptyset) = 0$) and a strictly positive vector $\lambda^S \in \mathbb{R}^S$. TU games are H-games with $\lambda^S = 1_S$ for every coalition S.[4]

An *order* on N is a permutation $(\pi(1), \pi(2), \ldots, \pi(n))$, where $\pi : N \to N$ is a one-to-one function. We denote $\Pi(N)$ the set of orders on N. Let $\pi \in \Pi(N)$ be an order. For an H-game (N, V), the *marginal contributions of the players in the order* π are inductively defined as follows[5]:

$$D_\pi^i(N, V) := \max\left\{ x^i \,\middle|\, \left(x^i, (D_\pi^j(N, V))_{j\in S_i(\pi)\setminus i}\right) \in V(S_i(\pi)) \right\} \tag{1.1}$$

[2]This actually holds only for the case of non-degenerate solutions, i.e., those supported by strictly positive utility weights.

[3]A more detailed discussion on the independent private values assumption is presented in Section 6.

[4]1_S denotes the 1-vector in the $|S|$-dimensional Euclidean space.

[5]We write $S \setminus i$ instead of the more cumbersome $S \setminus \{i\}$.

where $S_i(\pi) := \{j \in N \mid \pi^{-1}(j) \leq \pi^{-1}(i)\}$ is the set of players preceding (and including) i in the order π. The real number $D^i_\pi(N, V)$ is the maximal payoff that player i can get when he enters the cooperation after the players $\pi(1), \pi(2),..., \pi(\pi^{-1}(i) - 1)$ have successively entered and were paid according to $(D^j_\pi(N, V))_{j \in S_i(\pi) \setminus i}$.

Maschler and Owen (1989) define the consistent value for an H-game to be the expected marginal contribution of a player over all possible orders of N, where each one of the $n!$ orders is chosen with equal probability:

$$\varphi_i(N, V) := \frac{1}{n!} \sum_{\pi \in \Pi(N)} D^i_\pi(N, V) \tag{1.2}$$

The quantity $\varphi_i(N, v)$ thus represents the payoff allocation that player i expects in an arrival procedure in which players are successively given the power of dictatorship according to a random order (*random arrival procedure*).

Let $\partial V(S)$ denote the Pareto boundary of $V(S)$. By definition, for every order π, the payoff vector $D_\pi(N, V) = (D^i_\pi(N, V))_{i \in N}$ is efficient, i.e., $D_\pi(N, V) \in \partial V(N)$. The fact that the Pareto boundary of $V(N)$ happens to be flat guarantees that $\varphi(N, v) = (\varphi_i(N, v))_{i \in N}$ is also efficient.

We notice that the marginal contribution of a player i not only depends on the set $S_i(\pi)$, but also it may depend on the order of players inside $S_i(\pi)$. Hart (1994) defines the *marginal contribution of player i to coalition S* as

$$D^i(S, V) := \max\left\{ x^i \,\middle|\, \left(x^i, \varphi(S \setminus i, V)\right) \in V(S) \right\} \tag{1.3}$$

The idea is to summarize the payoff possibilities in the set $V(S \setminus i)$ by the value $\varphi(S \setminus i, V)$ of the subgame (S, V).[6] The payoff allocation $D^i(S, V)$ specifies the optimal choice of player $i \in S$ when he has the dictatorial power to choose for coalition S, having to guarantee the participation of the other players in $S \setminus i$, who have reservation utilities given by $\varphi(S \setminus i, V)$. The value $\varphi(S \setminus i, V)$ of the subgame $(S \setminus i, V)$ can be understood as the optimal *threat* that the members of $S \setminus i$ will enforce in case player i's proposition is not favorable to them.

Proposition 1 (Hart (1994)).
The consistent value satisfies

$$\varphi(N, V) = \frac{1}{n} \sum_{i \in N} (D^i(N, V), \varphi(N \setminus i, V))$$

The value $\varphi(N, V)$ of the game (N, V) then can be viewed as the payoff that players would get if decisions in N are taken according to the following *conditional random dictatorship procedure*: a player $i \in N$ is picked at random, with all players having equal probabilities. Then, player i is given the power of dictatorship *conditional* on giving the other players in $N \setminus i$ what they would

[6]Given a game (N, V) and a coalition $S \subseteq N$, (S, V) denotes the game whose player set is S and whose coalitional function is the restriction of V to the subsets of S.

get in the subgame $(N \setminus i, V)$. We exploit this characterization of the value in order to develop our generalization of the consistent value to games with incomplete information.

In order to compute the value for the H-game (N, V), one needs to know the values of the subgames $((N \setminus i, V))_{i \in N}$. In particular, the consistent value of (N, V) is determined in exactly the same way the value of (S, V) (with $S \subseteq N$) is computed. This inductive property is denominated by Hart and Mas-Colell (1996, p. 366) "subcoalition perfectness".

Maschler and Owen (1992) extend the definition of the consistent value to general NTU games. Their approach consists in "linearizing" the boundary of each set $V(S)$ by a method similar to the Harsanyi-Shapley fictitious-transfer procedure. The reader is referred to Hart and Mas-Colell (1996) for a detailed explanation. Hart (2005) provides a characterization of the consistent value for general NTU games parallel to Proposition 1.

2. Model

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 decisions for the members of S. We assume that the sets of joint decisions are finite and *superadditive*, that is, for any two disjoint coalitions[7] 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 notation[8] $t_S = (t_i)_{i \in S} \in T_S = \prod_{i \in S} T_i$. For simplicity, we drop the subscript N in the case of the grand coalition, so we define $D := D_N$ and $T := T_N$. For each possible type $t_i \in T_i$, let $p(t_i)$ denote the probability that player i is of type t_i. We assume types are independent random variables, then we may write

$$p(t_S) = \prod_{i \in S} p(t_i), \quad \forall S \subseteq N, \ \forall t_S \in T_S$$

We also assume, without loss of generality, that all types have positive marginal probability, i.e., $p(t_i) > 0$ for all $t_i \in T_i$ and all $i \in N$.

The utility function of player $i \in N$ is defined to be $u_i : D \times T_i \to \mathbb{R}$. We notice that u_i depends only on T_i and not on $T_{N \setminus i}$. That is, we assume *private values*.

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,

[7]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.

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

$$u_i((d_S, d_{N\setminus S}), t_i) = u_i((d_S, d'_{N\setminus S}), t_i),$$

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 $t_i \in T_i$. Then we can let $u_i(d_S, t_i)$ denote the utility for player $i \in S$ if $d_S \in D_S$ is carried out. That is, $u_i(d_S, t_i) = u_i((d_S, d_{N\setminus S}), t_i)$ 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

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

For any coalition $S \subseteq N$, we denote Γ_S the game obtained by restricting Γ_N to the subcoalitions of S. Independent private values together with the orthogonal coalitions guarantee that Γ_S is well defined.

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 (1991a, sec. 6.3)), we can restrict attention to (Bayesian) incentive compatible direct mechanisms. Formally, a *(direct) mechanism* for coalition S is a mapping $\mu_S : T_S \to \Delta(D_S)$, where $\Delta(D_S)$ denotes the set of probability distributions over D_S. The interpretation is that if S forms, it makes a decision randomly as a function of its members' information. Let the set of mechanisms for S be denoted $\mathcal{M}_S$.

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

$$U_i(\mu_S, \tau_i \mid t_i) = \sum_{t_{S\setminus i}\in T_{S\setminus i}} p(t_{S\setminus i}) \sum_{d_S\in D_S} \mu_S(d_S \mid \tau_i, t_{S\setminus i}) u_i(d_S, t_i).$$

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

A mechanism μ_S is *incentive compatible* for coalition S if and only if

$$U_i(\mu_S \mid t_i) \geq U_i(\mu_S, \tau_i \mid t_i), \qquad \forall t_i, \tau_i \in T_i, \quad \forall i \in S. \tag{2.1}$$

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

A mechanism $\mu_S \in \mathcal{M}_S$ is *(interim) individually rational* for coalition S if and only if

$$U_i(\mu_S \mid t_i) \geq \max_{d_i\in D_i} u_i(d_i, t_i), \qquad \forall t_i \in T_i, \quad \forall i \in S. \tag{2.2}$$

2.2. Incentive Efficiency and The Virtual Utility

Let $S \subseteq N$ be a coalition. A mechanism $\bar{\mu}_S \in \mathcal{M}_S$ is *(interim) incentive efficient* for S if and only if $\bar{\mu}_S$ is incentive compatible for S and there does not exist any other incentive compatible mechanism for S giving a strictly higher expected utility to all types t_i of all players $i \in S$. Because the set of incentive-compatible mechanisms is a compact and convex polyhedron, the mechanism $\bar{\mu}_S$ is incentive efficient for S if and only if there exist non-negative numbers $\lambda^S = (\lambda^S_i(t_i))_{i\in S,\, t_i\in T_i}$, not all zero, such that $\bar{\mu}_S$ is a solution to

$$\max_{\mu_S\in\mathcal{M}^*_S} \sum_{i\in S} \sum_{t_i\in T_i} \lambda^S_i(t_i) U_i(\mu_S \mid t_i) \tag{2.3}$$

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

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

where $\mu_S \in \mathcal{M}_S$. To simplify this expression, let

$$v_i(d_S, t_i, \lambda^S, \alpha^S) = \frac{1}{p(t_i)} \left[\left(\lambda_i^S(t_i) + \sum_{\tau_i \in T_i} \alpha_i^S(\tau_i \mid t_i) \right) u_i(d_S, t_i) - \sum_{\tau_i \in T_i} \alpha_i^S(t_i \mid \tau_i) u_i(d_S, \tau_i) \right]. \tag{2.4}$$

The quantity $v_i(d_S, t_i, \lambda^S, \alpha^S)$ is called the *virtual utility* of player $i \in S$ from the joint decision $d_S \in D_S$, when he is type $t_i \in T_i$, w.r.t. the utility weights λ^S and the Lagrange multipliers α^S. We denote $v_i(\mu_S, t_i, \lambda^S, \alpha^S)$ the linear extension of the virtual utility over μ_S. With this definition, the above Lagrangian can be rewritten as

$$\mathcal{L}(\mu_S, \lambda^S, \alpha^S) = \sum_{t_S \in T_S} p(t_S) \sum_{i \in S} v_i(\mu_S, t_i, \lambda^S, \alpha^S) \tag{2.5}$$

The next proposition follows from duality theory of linear programming.

Proposition 2 (Incentive Efficiency).
Let $S \subseteq N$ be a coalition. A mechanism $\bar{\mu}_S \in \mathcal{M}_S$ is incentive efficient for S if and only it is incentive compatible for S and there exist vectors $\lambda^S \geq 0$ ($\lambda^S \neq 0$) and $\alpha^S \geq 0$, such that

$$\alpha_i^S(\tau_i \mid t_i) \left[U_i(\bar{\mu}_S \mid t_i) - U_i(\bar{\mu}_S, \tau_i \mid t_i) \right] = 0, \quad \forall i \in S, \ \forall t_i \in T_i, \ \forall \tau_i \in T_i \tag{2.6}$$

and $\bar{\mu}_S$ maximizes the Lagrangian in (2.5) over all mechanisms in $\mathcal{M}_S$, namely,

$$\sum_{i \in S} v_i(\bar{\mu}_S, t_i, \lambda^S, \alpha^S) = \max_{d_S \in D_S} \sum_{i \in S} v_i(d_S, t_i, \lambda^S, \alpha^S), \quad \forall t_S \in T_S. \tag{2.7}$$

Equation (2.6) is the usual *dual complementary slackness* condition. Condition (2.7) says that if players are given the possibility to transfer virtual utility (as if virtual payoffs were in money), then $\bar{\mu}_S$ would be ex-post efficient for S. 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 (1991a, ch. 10) refers to this idea as the *virtual utility hypothesis*. The utility weights λ^S give the utility scales in which players make interpersonal utility comparisons, while the Lagrange multipliers α^S determine the signaling costs associated to incentive compatibility.

A vector α^S satisfying (2.6) for λ^S is a vector that solves the dual problem of (2.3). This *dual problem for S w.r.t. λ^S* can be written as

$$\min_{\alpha^S \geq 0} \sum_{t_S \in T_S} p(t_S) \left(\max_{d_S \in D_S} \sum_{i \in S} v_i(d_S, t_i, \lambda^S, \alpha^S) \right) \tag{2.8}$$

3. Motivating Example

The following example, which we call Example 1, was originally proposed by Myerson (1991b). It motivates our approach. Also, it illustrates the difficulties of extending the conditional random dictatorship procedure to games with incomplete information.

The set of players is $N = \{1, 2\}$. Each player has private information represented by two possible types: $T_1 = \{1_H, 1_L\}$ and $T_2 = \{2_H, 2_L\}$. Prior probabilities are $p(1_H) = 1 - p(1_L) = 1/5$ and $p(2_H) = 1 - p(2_L) = 4/5$. Feasible decisions for each coalition are: $D_i = \{d_i^0\}$ $(i \in N)$ and $D_N = \{[d_1^0, d_2^0], d_1, d_2\}$. Finally, utility functions are given by[9]:

(u_1, u_2)	$[d_1^0, d_2^0]$	d_1	d_2
H	$(0, 0)$	$(20, 0)$	$(-80, 100)$
L	$(0, 0)$	$(100, -80)$	$(0, 20)$

This game can be interpreted as follows. Player 1 is the seller of a single and indivisible good. Player 2 is the only potential buyer. The good may be worth either \$0 (type L) or \$80 (type H) to the seller, and it may be worth \$20 (type L) or \$100 (type H) to the buyer. Each individual only knows his/her own valuation of the good, but thinks that the other individual's values are likely to be either of the two possible numbers. Utilities from no-trade are normalized to be zero. Decision $d_0 := [d_1^0, d_2^0]$ denotes the no-exchange alternative. Decision d_1 (resp. d_2) represents the situation in which player 2 receives the good from player 1 in exchange of \$100 (resp. for free). Any other transfer of money from player 2 to player 1 between \$0 and \$100 can be represented by a lottery defined on $\{d_1, d_2\}$.

The maximal expected utility that any type of a player can get by himself is 0. Let us proceed as in the conditional random dictatorship procedure for games with complete information. Assume that player 1 is picked to be a dictator in the grand coalition. Then, player 1 faces the problem to select an incentive compatible mechanism for N giving both types of player 2 a non-negative expected utility. Among all feasible mechanisms for player 1, the following maximizes simultaneously the interim utility of her both types[10]:

$$\bar{\mu}_N(d_1 \mid H, H) = \bar{\mu}_N(d_0 \mid H, L) = \bar{\mu}_N(d_1 \mid L, H) = \bar{\mu}_N(d_0 \mid L, L) = 1$$

In this mechanism selection problem, it is clear that player 1 should choose $\bar{\mu}_N$. According to this mechanism trade only occurs when player 2 claims to be type H, and in this case, player 1 offers the take-it-or-leave-it price \$100. Thus, the expected payoffs each type of every player gets under $\bar{\mu}_N$ are $(\bar{U}_1^H, \bar{U}_1^L, \bar{U}_2^H, \bar{U}_2^L) = (16, 80, 0, 0)$.

Likewise, player 2 faces a similar mechanism selection problem when he is given the power to choose for coalition N. The following mechanism maximizes simultaneously the expected utility of his both types among all incentive compatible mechanism for N giving both types of

[9]When looking at the payoffs of player $i = 1, 2$, H (resp. L) denotes the payoffs for type i_H (resp. i_L). Player i's payoffs are independent of j's $(j \neq i)$ type. Thus, values are private.

[10]A mechanism depends on the reported types of both players: the first component denotes the reported type of player 1, while the second component denotes the type reported by player 2.

player 1 a nonnegative expected utility:

$$\hat{\mu}_N(d_0 \mid H, H) = \hat{\mu}_N(d_0 \mid H, L) = \hat{\mu}_N(d_2 \mid L, H) = \hat{\mu}_N(d_2 \mid L, L) = 1$$

Under this mechanism, player 2 makes a take-it-or-leave-it offer equal to \$0, provided that player 1 claims to be type L. Otherwise, trade does not occur. This mechanism gives expected payoffs $(\hat{U}_1^H, \hat{U}_1^L, \hat{U}_2^H, \hat{U}_2^L) = (0, 0, 80, 16)$.

Since $n = 2$, each player has 1/2 probability to be chosen as a dictator for N. Then, the following interim allocation can be regarded as an equitable outcome for the grand coalition:

$$(U_1^H, U_1^L, U_2^H, U_2^L) = (8, 40, 40, 8) = \tfrac{1}{2}(16, 80, 0, 0) + \tfrac{1}{2}(0, 0, 80, 16)$$

This allocation is incentive efficient for N, thus it may be called the value of the game.

Let us assume now that the prior probabilities are changed to be $p(1_H) = p(1_L) = 1/2$ and $p(2_H) = p(2_L) = 1/2$. The analysis becomes more complex in this modified example. The problem is that the best feasible mechanism for each player now depends on what his type is. Consider again the situation in which player 1 is chosen to be the dictator. The best incentive feasible mechanism for type H is $\bar{\mu}_N$. On the other hand, the best incentive feasible mechanism for type L is

$$\tilde{\mu}_N(d_1 \mid H, H) = 1 - \tilde{\mu}_N(d_2 \mid H, H) = \tfrac{4}{5}, \quad \tilde{\mu}_N(d_0 \mid H, L) = 1$$
$$\tilde{\mu}_N(d_1 \mid L, H) = 1 - \tilde{\mu}_N(d_2 \mid L, H) = \tfrac{34}{35}$$
$$\tilde{\mu}_N(d_0 \mid L, L) = \tfrac{5}{7}, \quad \tilde{\mu}_N(d_1 \mid L, L) = \tfrac{2}{35}, \quad \tilde{\mu}_N(d_2 \mid L, L) = \tfrac{8}{35}$$

This mechanism gives expected payoffs $(\tilde{U}_1^H, \tilde{U}_1^L, \tilde{U}_2^H, \tilde{U}_2^L) = (0, 51.4, 11.4, 0)$.

A simplistic analysis may suggest that player 1 should select $\bar{\mu}_N$ when he is type H and $\tilde{\mu}_N$ when he is type L. However, by acting like that, player 2 will infer the type of player 1 from her choice of the mechanism. Hence, when $\tilde{\mu}_N$ is selected, player 2 anticipates that by claiming to be type H, he will end up paying a price \$97.14, while if he declares to be type L, he will only pay \$20 in case of trade. Therefore, type H of player 2 will pretend to be type L. By the *inscrutability principle* (see Myerson (1983)), player 1 can equivalently select the mechanism that coincides with $\bar{\mu}_N$ if he is type H and with $\tilde{\mu}_N$ if he is type L, so that his actual choice of the mechanism conveys no information to player 2. This inscrutable mechanism however does not satisfy incentive compatibility for type H of player 2.

A similar analysis reveals analogous difficulties in the case player 2 is given the power of dictatorship.

Regardless of his actual type, any player cannot implement the mechanism that is the best for him, that is, he must select a feasible mechanism other than the one he prefers. Therefore, a player must use a bargaining strategy that achieves a balance between the objectives of his various types. Myerson (1983) developed a theory of inscrutable mechanism selection by an informed individual with all the bargaining ability. We build on Myerson's theory to provide a generalization of the conditional random dictatorship procedure to cooperative games with incomplete information. This will allow us to define a bargaining solution extending the consistent value.

4. Conditional Random Dictatorship and the Neutral Optima

Let $S \subseteq N$ be a coalition (with $|S| \geq 2$) and $i \in S$ be a fixed player. Assume that the members of coalition $S \setminus i$ have agreed on a mechanism $\mu_{S\setminus i} \in \mathcal{M}_{S\setminus i}$ before the arrival of player i into S. Suppose that once i joins $S \setminus i$, he is chosen to be a dictator in S, that is, he is given all the bargaining ability to determine a coordination mechanism for S. However, to maintain his dictatorship, he must use a mechanism that offers each type t_j of every player $j \in S \setminus i$ an expected payoff larger than $U_j(\mu_{S\setminus i} \mid t_j)$, guaranteeing that they all cannot be worse off than if they refused to cooperate. We denote this mechanism selection problem by $\Gamma^i_S(\mu_{S\setminus i})$.

Definition 1 (Feasibility).
A mechanism $\mu_S \in \mathcal{M}_S$ is feasible in $\Gamma^i_S(\mu_{S\setminus i})$ if it is incentive compatible for S and, for every $j \in S \setminus j$, $U_j(\mu_S \mid t_j) \geq U_j(\mu_{S\setminus i} \mid t_j)$ for all $t_j \in T_j$.

By the inscrutability principle of Myerson (1983), there is no loss of generality in assuming that all types of player i select the same feasible mechanism, so that his actual choice of the mechanism conveys no information. Any revelation of private information can be equivalently postponed to the implementation of the mechanism.

Definition 2 (Undominated mechanisms).
A mechanism $\mu_S \in \mathcal{M}_S$ is undominated for player i in $\Gamma^i_S(\mu_{S\setminus i})$ if it is feasible and there does not exist any other feasible mechanism giving a strictly higher expected utility to all types $t_i \in T_i$ of player i.

Player i should never be expected to select a mechanism that is dominated for him in $\Gamma^i_S(\mu_{S\setminus i})$. Because the set of feasible mechanisms is convex, a mechanism $\bar{\mu}_S \in \mathcal{M}_S$ is undominated for i in $\Gamma^i_S(\mu_{S\setminus i})$ if and only if there exist non-negative numbers $\lambda^S_i = (\lambda^S_i(t_i))_{t_i \in T_i}$, not all zero, such that $\bar{\mu}_S$ is a solution to

$$\begin{array}{ll} \max\limits_{\mu_S \in \mathcal{M}^*_S} & \sum\limits_{t_i \in T_i} \lambda^S_i(t_i) U_i(\mu_S \mid t_i) \\ \text{s.t.} & U_j(\mu_S \mid t_j) \geq U_j(\mu_{S\setminus i} \mid t_j), \quad \forall j \in S \setminus i,\ t_j \in T_j. \end{array} \tag{4.1}$$

This linear programming problem will be called the *primal problem for i in $\Gamma^i_S(\mu_{S\setminus i})$ w.r.t. λ^S_i.*

Remark 1. The optimization problem in (4.1) is feasible provided that $\mu_{S\setminus i}$ is incentive compatible for $S \setminus i$. Indeed, let $\hat{\mu}_i \in \mathcal{M}_i$ be the mechanism defined by:

$$U_i(\hat{\mu}_i \mid t_i) = \max_{d_i \in D_i} u_i(d_i, t_i), \quad \forall t_i \in T_i. \tag{4.2}$$

Define the mechanism $(\hat{\mu}_i, \mu_{S\setminus i}) \in \mathcal{M}_S$ by

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

It can be easily checked that $(\hat{\mu}_i, \mu_{S\setminus i})$ is feasible for i in $\Gamma^i_S(\mu_{S\setminus i})$ whenever $\mu_{S\setminus i}$ is incentive compatible for $S \setminus i$.

As in Section 2, let $\alpha_i^S(\tau_i \mid t_i) \geq 0$ be the Lagrange multiplier for the constraint that the type t_i of player i should not gain by reporting τ_i. Let also $\lambda_j^S(t_j) \geq 0$ denote the Lagrange multiplier for the constraint that μ_S must give at least $U_j(\mu_{S\setminus i} \mid t_j)$ to type t_j of player $j \in S \setminus i$. Then, using the concept of virtual utility, the Lagrangian of the above optimization problem can be written as

$$\mathcal{L}(\mu_S, \mu_{S\setminus i}, \lambda^S, \alpha^S) = \sum_{t_S \in T_S} p(t_S) \sum_{j \in S} v_j(\mu_S, t_j, \lambda^S, \alpha^S) - \sum_{j \in S\setminus i} \sum_{t_j \in T_j} \lambda_j^S(t_j) U_j(\mu_{S\setminus i} \mid t_j) \tag{4.3}$$

Assume that the mechanism $\mu_{S\setminus i}$ is incentive efficient for $S \setminus i$. Then, by Proposition 2, there exists a vector $\alpha^{S\setminus i} \geq 0$ such that

$$\alpha_j^{S\setminus i}(\tau_j \mid t_j)\left[U_j(\mu_{S\setminus i} \mid t_j) - U_j(\mu_{S\setminus i}, \tau_j \mid t_j)\right] = 0, \quad \forall j \in S \setminus i,\ \forall t_j, \tau_j \in T_j. \tag{4.4}$$

Hence, the following chain of equalities holds:

$$\begin{aligned}
&\sum_{j \in S\setminus i} \sum_{t_j \in T_j} \lambda_j^S(t_j) U_j(\mu_{S\setminus i} \mid t_j) \\
&\quad = \sum_{j \in S\setminus i} \sum_{t_j \in T_j} \lambda_j^S(t_j) U_j(\mu_{S\setminus i} \mid t_j) + \sum_{j \in S\setminus i} \sum_{t_j \in T_j} \sum_{\tau_j \in T_j} \alpha_j^{S\setminus i}(\tau_j \mid t_j)\left[U_j(\mu_{S\setminus i} \mid t_j) - U_j(\mu_{S\setminus i}, \tau_j \mid t_j)\right] \\
&\quad = \sum_{t_{S\setminus i} \in T_{S\setminus i}} p(t_{S\setminus i}) \sum_{j \in S\setminus i} v_j(\mu_{S\setminus i}, t_j, \lambda^S, \alpha^{S\setminus i})
\end{aligned} \tag{4.5}$$

Therefore, given some vector $\alpha^{S\setminus i}$ satisfying (4.4), the Lagrangian in (4.3) can alternatively be written as:

$$\mathcal{L}(\mu_S, \mu_{S\setminus i}, \lambda^S, \alpha^S, \alpha^{S\setminus i}) = \sum_{t_S \in T_S} p(t_S) \left(\sum_{j \in S} v_j(\mu_S, t_j, \lambda^S, \alpha^S) - \sum_{j \in S\setminus i} v_j(\mu_{S\setminus i}, t_j, \lambda^S, \alpha^{S\setminus i}) \right) \tag{4.6}$$

When implementing the mechanism $\mu_{S\setminus i}$, the members of coalition $S \setminus i$ experience an efficiency loss due to the incentive compatibility. Those inefficiencies have an indirect impact on the participation constraints that player i face in coalition S. Equation (4.6) help us to understand how the signaling costs associated to incentive compatibility in coalition $S \setminus i$ affect the surplus that player i is able to extract in coalition S.

Optimality conditions from duality theory imply the following result.

Proposition 3 (Characterizing undominated mechanisms).
A mechanism $\bar{\mu}_S \in \mathcal{M}_S$ is undominated for i in $\Gamma_S^i(\mu_{S\setminus i})$ if and only it is feasible for i in $\Gamma_S^i(\mu_{S\setminus i})$ and there exist vectors $\lambda^S = (\lambda_j^S(t_j))_{t_j \in T_j\, j \in S} \geq 0$ ($\lambda_i^S \neq 0$) and $\alpha^S \geq 0$, such that

$$\alpha_j^S(\tau_j \mid t_j)\left[U_j(\bar{\mu}_S \mid t_j) - U_j(\bar{\mu}_S, \tau_j \mid t_j)\right] = 0, \quad \forall j \in S,\ \forall t_j, \tau_j \in T_j, \tag{4.7}$$

$$\lambda_j^S(t_j)\left[U_j(\bar{\mu}_S \mid t_j) - U_j(\mu_{S\setminus i} \mid t_j)\right] = 0, \quad \forall j \in S \setminus i,\ \forall t_j \in T_j, \tag{4.8}$$

and $\bar{\mu}_S$ maximizes the Lagrangian in (4.3) over all mechanisms in $\mathcal{M}_S$, namely,

$$\sum_{j \in S} v_j(\bar{\mu}_S, t_j, \lambda^S, \alpha^S) = \max_{d_S \in D_S} \sum_{j \in S} v_j(d_S, t_j, \lambda^S, \alpha^S), \quad \forall t_S \in T_S \tag{4.9}$$

Remark 2. Propositions 2 and 3 imply that any undominated mechanism for i in $\Gamma_S^i(\mu_{S\setminus i})$ is incentive efficient for S.

In order to identify a solution among the many undominated mechanisms for player i, it is necessary to define some principles for an inscrutable compromises among the different types of player i. Myerson (1983) defined a solution concept, called the *neutral optimum*, which predicts which mechanisms an informed individual with all the bargaining ability might select. We elaborate on his theory to extend the definition of neutral optima to the game $\Gamma_S^i(\mu_{S\setminus i})$.[11]

We say that a vector $\omega_i^S = (\omega_i^S(t_i))_{t_i\in T_i}$ of interim utilities for player i is *warranted* by λ^S, α^S, $\alpha^{S\setminus i}$ and $\mu_{S\setminus i}$ if

$$\frac{1}{p(t_i)}\left[\left(\lambda_i^S(t_i) + \sum_{\tau_i\in T_i}\alpha_i^S(\tau_i \mid t_i)\right)\omega_i^S(t_i) - \sum_{\tau_i\in T_i}\alpha_i^S(t_i\mid\tau_i)\omega_i^S(\tau_i)\right]$$
$$= \sum_{t_{S\setminus i}\in T_{S\setminus i}} p(t_{S\setminus i})\left(\max_{d_S\in D_S}\sum_{j\in S} v_j(d_S, t_j, \lambda^S, \alpha^S) - \sum_{j\in S\setminus i} v_j(\mu_{S\setminus i}, t_j, \lambda^S, \alpha^{S\setminus i})\right), \quad \forall t_i\in T_i. \quad (4.10)$$

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

Lemma 1 (Myerson (1983)).
The warrant equations in (4.10) *have a unique solution in the warranted claims provided that* $\lambda_i^S > 0$. *Furthermore, the solution is weakly increasing in the vector of right hand side independent terms.*

Lemma 2.
Let $\lambda_i^S > 0$. *Suppose that* ω_i^S *is warranted by* λ^S, α^S, $\alpha^{S\setminus i}$ *and* $\mu_{S\setminus i}$, *where the mechanism* $\mu_{S\setminus i}$ *is incentive compatible for* $S\setminus i$ *and* $\alpha^{S\setminus i}$ *satisfies* (4.4). *Then,*

$$\omega_i^S(t_i) \geq \max_{d_i\in D_i} u_i(d_i, t_i), \qquad \forall t_i\in T_i, \quad \forall i\in S. \qquad (4.11)$$

According to Lemma 2, by demanding his warranted claims, a player can never be worse than when he refuses to cooperate (individual rationality). It is worth emphasizing that this lemma crucially hinges on the assumption of independent private values. This point is discussed in Section 6.

Definition 3 (Neutral optimum (Myerson (1983))).
Let $\mu_{S\setminus i}$ *be an incentive efficient mechanism for* $S\setminus i$. *A feasible mechanism* μ_S *in* $\Gamma_S^i(\mu_{S\setminus i})$ *is a neutral optimum for player i if there exist vectors* $\lambda^S \geq 0$, $\alpha^S \geq 0$ *and* $\alpha^{S\setminus i}\geq 0$ *such that:*

(i) $\lambda_i^S(t_i) > 0, \quad \forall t_i\in T_i$.

[11] Myerson's theory does not allow for participation constraints with arbitrary outside options. Indeed, the neutral optimum is only defined for participation constraints with reservations utilities normalized to zero. This apparently innocuous modification entails technical complications. For instance, in order to establish Lemma 2 below, which is analogous to Lemma 3 in Myerson (1983), we require the independent private values hypothesis.

(ii) $U_i(\mu_S) := (U_i(\mu_S \mid t_i))_{t_i \in T_i}$ *is warranted by* λ^S, α^S, $\alpha^{S\setminus i}$ *and* $\mu_{S\setminus i}$.

Remark 3. Definition 3 involves strictly positive utility weights λ_i^S, that is, we only consider non-degenerate solutions. In general, this greatly complicates matters for obtaining existence results of the neutral optimum. On the other hand, the warrant equations are only know to be solvable when $\lambda_i^S > 0$ (see Lemma 1), thus allowing for zero weights also creates difficulties. Myerson (1983) solves this dilemma by slightly enlarging the solution set to include utility allocations that are reasonable as emerging from a closure argument. Our objective is to present the main insights of Myerson's (1983) theory in the framework of the game $\Gamma_S^i(\mu_{S\setminus i})$ and how it relates to our cooperative solution. Thus, we restrict attention to non-degenerate neutral optima.

Although the neutral optima were axiomatically derived, they can be interpreted in terms of a fictitious-transfer procedure. Assume that $\mu_{S\setminus i}$ is incentive efficient for $S \setminus i$ and let $\alpha^{S\setminus i}$ be such that (4.4) is satisfied for $\mu_{S\setminus i}$. For any vector (λ^S, α^S), with $\lambda_i^S > 0$, let us consider the *virtual mechanism selection problem* that differs from $\Gamma_S^i(\mu_{S\setminus i})$ in the following. First, according to the virtual utility hypothesis, each player's payoffs are in terms of the virtual utility scales, where the utility weights λ^S are used for interpersonal utility comparisons and the signaling costs are given by α^S or $\alpha^{S\setminus i}$, depending on whether incentives are evaluated inside coalition S or $S \setminus i$, respectively. Second, virtual utility is assumed to be transferable between the players. Because players are paid in transferable units of virtual utility, this virtual problem has a clear solution for player i: he extracts the total expected virtual surplus that coalition S can share in every state and then he rewards the participation of the other players according to their virtual reservation utilities from $\mu_{S\setminus i}$. The expected virtual utility for type t_i of player i in this allocation would be

$$v_i(t_i) = \sum_{t_{S\setminus i} \in T_{S\setminus i}} p(t_{S\setminus i}) \left(\max_{d_S \in D_S} \sum_{j \in S} v_j(d_S, t_j, \lambda^S, \alpha^S) - \sum_{j \in S\setminus i} v_j(\mu_{S\setminus i}, t_j, \lambda^S, \alpha^{S\setminus i}) \right) \tag{4.12}$$

Let μ_S be an undominated mechanism satisfying Proposition 3 for (λ^S, α^S) such that $U_i(\mu_S)$ is warranted by λ^S, α^S, $\alpha^{S\setminus i}$ and $\mu_{S\setminus i}$. Then we have

$$\begin{aligned} v_i(t_i) &= \frac{1}{p(t_i)} \left[\left(\lambda_i^S(t_i) + \sum_{\tau_i \in T_i} \alpha_i^S(\tau_i \mid t_i) \right) U_i(\mu_S \mid t_i) - \sum_{\tau_i \in T_i} \alpha_i^S(t_i \mid \tau_i) U_i(\mu_S \mid \tau_i) \right] \\ &= \frac{1}{p(t_i)} \left[\left(\lambda_i^S(t_i) + \sum_{\tau_i \in T_i} \alpha_i^S(\tau_i \mid t_i) \right) U_i(\mu_S \mid t_i) - \sum_{\tau_i \in T_i} \alpha_i^S(t_i \mid \tau_i) U_i(\mu_S, t_i \mid \tau_i) \right] \\ &= \sum_{t_{S\setminus i} \in T_{S\setminus i}} p(t_{S\setminus i}) v_i(\mu_S, t_i, \lambda^S, \alpha^S), \end{aligned} \tag{4.13}$$

where the second equality follows from the complementary slackness condition in (4.7).

Definition 3 thus asserts that μ_S is a neutral optimum for i in $\Gamma_S^i(\mu_{S\setminus i})$ if $U_i(\mu_S)$ is the vector of interim utilities that the various types of player i would obtain in a mechanism that is a solution of the virtual problem.

Proposition 4.
Let μ_S be a neutral optimum for player i in $\Gamma_S^i(\mu_{S\setminus i})$ supported by the scales λ^S, α^S and $\alpha^{S\setminus i}$. Then, μ_S is undominated for player i in $\Gamma_S^i(\mu_{S\setminus i})$. Moreover, it satisfies Proposition 3 for (λ^S, α^S).

The following result is an immediate consequence of Lemma 2.

Proposition 5.
A neutral optimum μ_S for player i in $\Gamma^i_S(\mu_{S\setminus i})$ is individually rational for this player, i.e., it satisfies (2.2) for all $t_i \in T_i$.

Example 1. We study again the motivating example of Section 3 with prior probabilities $p(1_H) = p(1_L) = 1/2$ and $p(2_H) = p(2_L) = 1/2$. Let $\mu^0_i \in \mathcal{M}_i$ be the mechanism that always implements d^0_i with probability 1. Consider the case in which player 1 becomes a dictator for coalition N. The reservation utilities of both types of player 2 are $U_2(\mu^0_2 \mid H) = U_2(\mu^0_2 \mid L) = 0$. The problem for player 1 is thus to select an incentive compatible mechanism for N giving both types of player 2 a nonnegative expected utility. As it was previously shown in Section 3, player 1 must establish an inscrutable compromise between the payoff maximization goals of her both types. In this game, there are compelling reasons to think that the conflict of interests between 1's types must be resolved in favor or type H. Indeed, among all undominated mechanisms for player 1, only $\bar{\mu}_N$ remains incentive compatible when player 2 learns that player 1 is of type L. In this sense, the mechanism $\bar{\mu}_N$ is "safe". Therefore, giving the impression of being type H all the time, even when she is actually type L, is an inscrutable strategy for player 1. In the terminology of Myerson (1983), $\bar{\mu}_N$ is a "strong solution" for player 1. Let us show that $\bar{\mu}_N$ is indeed a neutral optimum for player 1 in $\Gamma^1_N(\mu^0_2)$.

Consider the virtual scales

$$\begin{gathered}\lambda^N_1(H) = \tfrac{5}{8}, \quad \lambda^N_1(L) = \tfrac{3}{8}, \quad \lambda^N_2(H) = \tfrac{3}{8}, \quad \lambda^N_2(L) = \tfrac{5}{8}, \\ \alpha^N_1(H \mid L) = \alpha^N_2(L \mid H) = \tfrac{1}{8}, \quad \alpha^N_1(L \mid H) = \alpha^N_2(H \mid L) = 0\end{gathered} \tag{4.14}$$

The utility weights λ^N_1 reflect the optimal inter-type compromise between both types of player 1. To conceal his type, player 1 must scale up the utility of type H and scale down the utility of type L as if his type H were five times more important than his type L. On the other hand, the fact that $\alpha^N_1(H \mid L) > 0$ implies that type H is *jeopardized* by type L, that is, type L has incentives to mimic type H, so that the latter faces a difficulty to credibly signal his information. Likewise, $\alpha^N_2(L \mid H) > 0$ implies that type L of player 2 is jeopardized by his type H.

Given the scales (λ^N, α^N), virtual utilities are:

(v_1, v_2)	d_0	d_1	d_2
H	$(0, 0)$	$(0, 0)$	$(-100, 100)$
L	$(0, 0)$	$(100, -100)$	$(0, 0)$

The only difference between virtual utility and actual utility is for 1's type H and 2's type L. When player 1 is type H (resp. player 2 is type L), the virtual value of the good for player 1 is \$100 (resp. \$0). Type H of player 1 exaggerates her valuation of the good, while type L of player 2 understates his valuation of the good. Incentive compatibility compels the players to incur in a costly behavior (from which they only get virtual utility) in order to reduce the misrepresentation of the jeopardizing type.

The unique solution to the warrant equations for player 1 is $\omega_1^N(H) = 10$ and $\omega_1^N(L) = 50$. The mechanism $\bar{\mu}_N$ is the unique feasible mechanism in $\Gamma_N^1(\mu_2^0)$ giving both types of player 1 their warranted claims. In fact, it yields the interim utility allocation $(\bar{U}_1^H, \bar{U}_1^L, \bar{U}_2^H, \bar{U}_2^L) = (10, 50, 0, 0)$. Then, $\bar{\mu}_N$ is a neutral optimum for player 1. Indeed, it is the unique neutral optimum for player 1 in $\Gamma_N^1(\mu_2^0)$.

Consider now the case in which player 2 is given the power of dictatorship. Exploiting the symmetry of this game, a similar analysis for player 2 in $\Gamma_N^2(\mu_1^0)$ shows that the mechanism $\hat{\mu}_N$ is the unique neutral optimum for player 2. It is supported by the same virtual scales (λ^N, α^N) in (4.14). This mechanism implements the interim allocation $(\hat{U}_1^H, \hat{U}_1^L, \hat{U}_2^H, \hat{U}_2^L) = (0, 0, 50, 10)$.

According to the conditional random dictatorship procedure, each player has equal chance to be a dictator for N. Then, the following interim allocation can be regarded as an equitable outcome for N,

$$(U_1^H, U_1^L, U_2^H, U_2^L) = \tfrac{1}{2}(10, 50, 0, 0) + \tfrac{1}{2}(0, 0, 50, 10) = (5, 25, 25, 5)$$

This allocation is implemented by the mechanism $\mu_N = \frac{1}{2}\bar{\mu}_N + \frac{1}{2}\hat{\mu}_N$. This mechanism is incentive efficient for N, thus it can be defined to be a bargaining solution of the game Γ_N. $\triangle$

Incentive efficiency of the average mechanism μ_N in the previous example is guaranteed by the fact that the *same* vectors λ^N and α^N support both neutral optima. The following lemma is due to the linearity in μ_S of the formulas (4.7) and (4.9).

Lemma 3.
Let $S \subseteq N$ be a coalition and, for each $i \in S$, let $\mu_{S\setminus i} \in \mathcal{M}_{S\setminus i}$ be a mechanism for $S \setminus i$. Let $(\lambda^S, \alpha^S) \geq 0$ (with $\lambda^S \neq 0$). For each $i \in S$, let μ_S^i be an undominated mechanism for i in $\Gamma_S^i(\mu_{S\setminus i})$ supported by (λ^S, α^S). Then, the average mechanism $\mu_S = \frac{1}{|S|}\sum_{i\in S}\mu_S^i$ is incentive efficient for S. Moreover, it satisfies (2.6)-(2.7) for (λ^S, α^S).

5. The MO-Solution

The idea in defining our bargaining solution was already anticipated in Example 1. The solutions is recursively constructed starting from singleton coalitions. For any player i, let $\mu_i \in \mathcal{M}_i$ be a mechanism satisfying (4.2). This is the best player i can do in the game Γ_N without any other player's help, thus it constitutes the obvious solution of Γ_i. Let $S \subseteq N$ be a coalition (with $|S| \geq 2$) and assume that, for each $i \in S$, the members of $S \setminus i$ have agreed on the bargaining solution $\mu_{S\setminus i}$ of the game $\Gamma_{S\setminus i}$. The bargaining solution of the game Γ_S is computed by means of the conditional random dictatorship procedure: a player $i \in S$ is picked at random, with all players having equal probability. Player i then selects a neutral optimum μ_S^i in the game $\Gamma_S^i(\mu_{S\setminus i})$. The mechanism $\mu_S = \frac{1}{|S|}\sum_{i\in S}\mu_S^i$ can thus be regarded as an "equitable" agreement for coalition S. This mechanism guarantees each type t_i of a player $i \in S$ an expected utility that is at least as large as the average expected payoff he could get from his neutral optimum μ_S^i (when he is a dictator for S) and the solutions $(\mu_{S\setminus j})_{j\in S\setminus i}$ in the subgames with $|S|-1$ players (when any other player $j \in S \setminus i$ is a dictator for S), namely,

$$U_i(\mu_S \mid t_i) \geq \frac{1}{|S|}\left[\sum_{j\in S\setminus i} U_i(\mu_{S\setminus j} \mid t_i) + U_i(\mu_S^i \mid t_i)\right] \tag{5.1}$$

If μ_S happens to be incentive efficient for S, then it is defined to be the bargaining solution of Γ_S. In view of Lemma 3, we then require all neutral optima $(\mu_S^i)_{i\in S}$ to be supported by the same virtual scales (λ^S, α^S). Some care is needed in formulating this idea if we want to proof an existence theorem. Two issues arise: First, some of the neutral optima may not exist when requiring them all to be supported by the same virtual scales. The reason is that the corresponding warranted claims may not be feasible. Feasibility of the warranted claims can only be satisfied when the interim Pareto frontier coincides with an hyperplane on the individually rational zone, as it is the case in Example 1[12]. In order to deal with this issue, we replace $U_i(\mu_S^i \mid t_i)$ in (5.1) by the warranted claim of type t_i, $\omega_i^S(t_i)$. We notice that if $\omega_i^S(t_i)$ is feasible, then $\omega_i^S(t_i) = U_i(\mu_S^i \mid t_i)$.

Second, we cannot exclude vanishing utility weights. The reason is that we cannot prevent the interim Pareto frontier to have level segments, i.e., boundary points at which the surface contains some line segment parallel to one of the coordinate axes.

Definition 4 (MO-solution).
A vector of threats $(\mu_S)_{S\subseteq N}$ is a MO-solution of the game Γ_N if, for each $S \subseteq N$, μ_S is incentive compatible and there exist vectors $(\lambda^S, \alpha^S, \omega^S)_{S\subseteq N}$ such that:

(i) $\lambda^S \in \prod_{i\in S} \mathbb{R}_+^{T_i} \setminus \{0\}$, $\alpha^S \in \prod_{i\in S} \mathbb{R}_+^{T_i\times T_i}$ *and* $\omega^S \in \prod_{i\in S} \mathbb{R}^{T_i}$.
(ii) *For every* $i \in S$, ω_i^S *is warranted by* λ^S, α^S, $\alpha^{S\setminus i}$ *and* $\mu_{S\setminus i}$.
(iii) $U_i(\mu_S \mid t_i) \geq \frac{1}{|S|}\left[\sum_{j\in S\setminus i} U_i(\mu_{S\setminus j} \mid t_i) + \omega_i^S(t_i)\right], \quad \forall t_i \in T_i, \quad \forall i \in S.$

The interim allocation $U(\mu_N) := (U_i(\mu_N \mid t_i))_{t_i\in T_i, i\in N}$ is called the MO-value of Γ_N.

Now we can state our main results.

Theorem 1 (Incentive efficiency).
Let $(\mu_S)_{S\subseteq N}$ be a MO-solution of Γ_N supported by $(\lambda^S, \alpha^S, \omega^S)_{S\subseteq N}$. Then, for each $S \subseteq N$, μ_S is incentive efficient. Moreover, μ_S satisfies Proposition 2 for (λ^S, α^S).

As a byproduct of Theorem 1, we obtain that condition *(iii)* in the definition of the MO-solution can only hold as inequality for some type t_i of player $i \in S$ if $\lambda_i^S(t_i) = 0$. Equivalently, $\lambda_i^S(t_i) > 0$ implies that *(iii)* must hold as equality for type t_i.

Corollary 1.
Let $(\mu_S)_{S\subseteq N}$ be a MO-solution of Γ_N supported by $(\lambda^S, \alpha^S, \omega^S)_{S\subseteq N}$. Then, for every coalition $S \subseteq N$,

$$\sum_{i\in S}\sum_{t_i\in T_i} \lambda_i^S(t_i)\left[|S|U_i(\mu_S \mid t_i) - \sum_{j\in S\setminus i} U_i(\mu_{S\setminus j} \mid t_i) - \omega_i^S(t_i)\right] = 0. \tag{5.2}$$

In particular, condition (iii) in Definition 4 implies

$$\lambda_i^S(t_i)\left[U_i(\mu_S \mid t_i) - \frac{1}{|S|}\left(\sum_{j\in S\setminus i} U_i(\mu_{S\setminus j} \mid t_i) + \omega_i^S(t_i)\right)\right] = 0, \quad \forall t_i \in T_i,\ \forall i \in S,\ \forall S \subseteq N. \tag{5.3}$$

[12]This difficulty is not proper of our solution concept, but it also arises for the the consistent value.

Theorem 2 (Individual rationality).
Let $(\mu_S)_{S\subseteq N}$ be a MO-solution of Γ_N supported by utility weights $(\lambda^S)_{S\subseteq N}$ with $\lambda^S > 0$ for all $S \subseteq N$. Then, for each $S \subseteq N$, μ_S is individually rational., i.e., it satisfies (2.2).

Individual rationality holds for all coalitions. This is in contrast with Myerson's (1984b) generalization of the Shapley NTU value and Salamanca's (2016) generalization of the Harsanyi NTU value which are show to be individually rational only for the grand coalition.

Theorem 3 (Existence).
For any game Γ_N, there exists at least one MO-solution.

The results that are presented below follow directly from the definitions. They relate our solution concept to other cooperative solutions proposed in the literature.

Theorem 4 (Generalization of the consistent value).
Let Γ_N be a cooperative games with complete information, i.e., T_i is a singleton for every $i \in N$. If $(\mu_S)_{S\subseteq N}$ is a MO-solution supported by strictly positive utility weights, then for each coalition $S \subseteq N$, the utility allocation $(U_i(\mu_S))_{i\in S}$ is a consistent value of the game Γ_S. Conversely, if $(U_i^S)_{i\in S, S\subseteq N}$ is a (non-degenerate) consistent value payoff configuration of Γ_N, then there exists a MO-solution of Γ_N, $(\mu_S)_{S\subseteq N}$, such that $(U_i^S)_{i\in S} = (U_i(\mu_S))_{i\in S}$ for each $S \subseteq N$.

A two-person bargaining problem with incomplete information is a two-player cooperative game with incomplete information for which $D_i = \{d_i^0\}$ and $u_i(d_i^0, t_i) = 0$ for all $t_i \in T_i$ and all $i \in N$. The decision $d^0 := [d_1^0, d_1^0] \in D$ denotes the *disagreement decision* made when players fail to reach a cooperative agreement. Then, in the absence of agreement, each player i can only get zero utility[13]. Clearly, Example 1 in Section 3 is a two-person bargaining problem. For this kind of problems, Myerson (1984a) proposed a *neutral solution* generalizing the Nash bargaining solution. The following result establishes an equivalence relation between the MO-solution and the neutral solution. The result is a direct consequence of Corollary 1 and the characterization theorems 5 and 6 in Myerson (1984a).

Theorem 5 (Equivalence with Myerson's (1984a) neutral solution).
Let Γ_N be a two-person bargaining problem with incomplete information. A mechanism μ_N is a (non-degenerate) MO-solution of Γ_N if and only if it is a (non-degenerate) neutral solution of Γ_N.

In what follows, we compute the MO-solution in two intuitive examples proposed by de Clippel (2005) and Salamanca (2016). These games were constructed to show that there are instances in which Myerson's (1984b) M-solution does not reflect well enough the game situation. For his example, de Clippel proposed an alternative outcome resulting from the unique Perfect Bayesian equilibrium of an extensive form game based on the random arrival procedure described in Section 1.1. Not surprisingly, this payoff vector is the unique utility allocation supported by a MO-solution of the game. On the other hand, Salamanca (2016) showed that the H-solution generates an interesting alternative to the M-solution in the context of his example. We shall see

[13] The 0-normalization of the disagreement payoffs is a convenient normalization that can be done without loss of generality.

that the MO-solution coincides with the M-solution in Salamanca's example. The analysis of these games is intended to show that a value is just an index summarizing different qualitative features of a game. To that extent, a value that better reflects the structure of a particular game may not be the most appropriate approach for analyzing other games[14].

A Trading 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 $\{d^1_{12}, d^2_{12}\}$. When player 3 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 from trade to player 2 (resp. 3) in both states.

We notice that the game Γ_N does *not* satisfy the private values assumption: the utility of player 2 (resp. 3) from d_{23} (resp. d_{32}) depends on the state. In spite of this, all our formulas and results continue to hold in this example. The reason is that for every $S \subseteq N$ with $S \neq N$, the subgame Γ_S satisfies all our assumptions. This confirms that our assumptions cannot ever be more than sufficient.

We start the analysis of this example by considering the subgame $\Gamma_{\{1,2\}}$. Clearly, in all other subgames the players can only get zero utility. Figure 1 depicts the set of incentive efficient allocations for coalition $\{1, 2\}$ giving 1 and 2 a non-negative expected utility.
In this subgame there is a unique utility allocation that is the best for both types of player 1, namely, $(U^H_1, U^L_1, U_2) = (90, 30, 0)$. Then, demanding this allocation is the best inscrutable compromise for player 1 when she is given the power of dictatorship in $\{1, 2\}$. Similarly, the best player 2 can claim when he is chosen to be a dictator for $\{1, 2\}$ is the allocation $(U^H_1, U^L_1, U_2) = (0, 0, 72)$. Because the efficient frontier is flat, the average allocation $(U^H_1, U^L_1, U_2) = \frac{1}{2}(90, 30, 0) + \frac{1}{2}(0, 0, 72) = (45, 15, 36)$ is incentive efficient for $\{1, 2\}$, thus it is the MO-value of the subgame.

The unique MO-solution of $\Gamma_{\{1,2\}}$ is the mechanism

$$\mu_{\{1,2\}}(d^1_{12} \mid H) = \mu_{\{1,2\}}(d^2_{12} \mid H) = \tfrac{1}{2}, \quad \mu_{\{1,2\}}(d^1_{12} \mid L) = \mu_{\{1,2\}}([d_1, d_2] \mid L) = \tfrac{1}{2}$$

[14]When comparing the different cooperative solutions in these examples, we only consider *coalitionally incentive compatible* solutions. See Salamanca (2016) for further details.

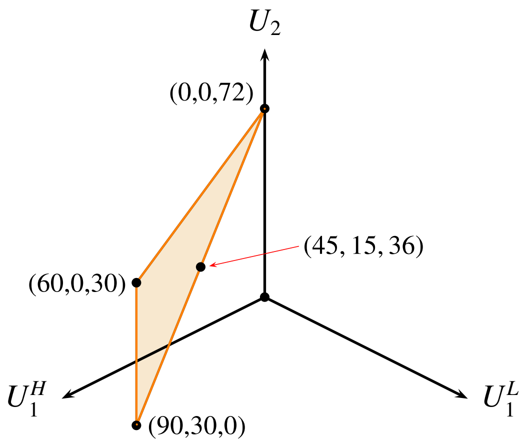

Figure 1: Incentive efficient allocations for $\{1, 2\}$

The MO-solution is supported by the virtual scales $(\lambda^{\{1,2\}}, \alpha^{\{1,2\}})$ given by

$$\lambda_1^{\{1,2\}}(H) = \tfrac{7}{10}, \quad \lambda_1^{\{1,2\}}(L) = \tfrac{3}{10}, \quad \lambda_2^{\{1,2\}} = 1,$$

$$\alpha_1^{\{1,2\}}(H \mid L) = 0, \quad \alpha_1^{\{1,2\}}(L \mid H) = \tfrac{1}{10}.$$

Indeed, any incentive efficient mechanism for $\{1, 2\}$ satisfies Proposition 2 for these virtual scales.

The only difference between virtual and real utility in coalition $\{1, 2\}$ is for type L of player 1, for which the virtual value of the good is \$0. Incentive compatibility forces type L of player 1 to behave as if her valuation of the good were lower than it really is. Incentive compatibility also limits the ability of the players to share the gains from trade. Indeed, the mechanism that gives the entire surplus to player 2 in both states is not incentive compatible: the most player 2 can get in any incentive efficient mechanism is $\$72 (< \$78 = 4/5 \times \$90 + 1/5 \times \$30)$.

When player 3 joins coalition $\{1, 2\}$, he 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 in coalition $\{1, 2\}$. To see this graphically, we compare in Figure 2 the set of incentive efficient allocation for coalition $\{1, 2\}$ (thin lines) with the projection of the incentive efficient allocations for coalition N when player 3 is rewarded \$0 (thick lines).

We observe that in coalition N, player 2 may now achieve the allocation $(U_1^H, U_1^L, U_2) = (0, 0, 78)$ corresponding to the situation in which the whole surplus of trade is given to him.

Every incentive efficient mechanism for N satisfies Proposition 2 for the virtual scales (λ^N, α^N) given by

$$\lambda_1^N(H) = \tfrac{4}{5}, \quad \lambda_1^N(L) = \tfrac{1}{5}, \quad \lambda_2^N = 1, \quad \lambda_3^N = 1,$$

$$\alpha_1^N(H \mid L) = \alpha_1^N(L \mid H) = 0.$$

The fact that $\alpha^N = 0$ implies that incentives constraints are not essential for coalition N, which corroborates that incentive compatibility is weakened inside coalition N. In addition, real and virtual utilities coincide in N.

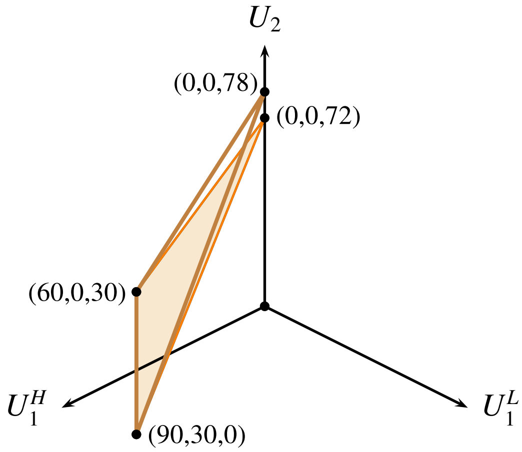

Figure 2: Incentive efficient allocations for $\{1, 2\}$ (thin line) and N (thick line)

Let us consider now the different situations in which every player is chosen to be a dictator for N. When player 1 becomes a dictator, it is clear from Figure 2 that the best allocation she may demand is $(U_1^H, U_1^L, U_2, U_3) = (90, 30, 0, 0)$. Of course, this is the unique allocation supported by neutral optimum of $\Gamma_N^1([d_2, d_3])$. Similarly, the unique neutral optimum of player 2 in $\Gamma_N^2([d_1, d_3])$ yields the allocation $(U_1^H, U_1^L, U_2, U_3) = (0, 0, 78, 0)$. Assume now that player 3 is given the power of dictatorship. The virtual utilities players 1 and 2 obtain under the scales $(\lambda^N, \alpha^{\{1,2\}})$ are

(v_1, v_2)	$[d_1, d_2]$	d_{12}^1	d_{12}^2
H	$(0, 0)$	$\left(\frac{405}{4}, 0\right)$	$(0, 90)$
L	$(0, 0)$	$(-15, 0)$	$(-60, 90)$

Therefore, the warranted claim of player 3 in $\Gamma_N^3(\mu_{\{1,2\}})$ is

$$\begin{aligned}
\omega_3^N &= \sum_{t \in \{H,L\}} p(t) \left[\max_{d \in D} \sum_{i \in N} v_i(d, t, \lambda^N, \alpha^N) - \sum_{i \in \{1,2\}} v_i(\mu_{\{1,2\}}, t, \lambda^N, \alpha^{\{1,2\}}) \right] \\
&= \tfrac{4}{5}\left(90 - \tfrac{765}{8}\right) + \tfrac{1}{5}\left(30 + \tfrac{15}{2}\right) \\
&= 3.
\end{aligned}$$

The unique utility allocation in N that is incentive compatible, gives players 1 and 2 at least what they would obtain under $\mu_{\{1,2\}}$ and that guarantees player 3 to obtain his warranted claim is $(U_1^H, U_1^L, U_2, U_3) = (45, 15, 36, 3)$. Certainly, this is the unique allocation supported by a neutral optimum of $\Gamma_N^3(\mu_{\{1,2\}})$. Applying the conditional random dictatorship procedure to coalition N we obtain the equitable allocation

$$(U_1^H, U_1^L, U_2, U_3) = \tfrac{1}{3}(90, 30, 0, 0) + \tfrac{1}{3}(0, 0, 78, 0) + \tfrac{1}{3}(45, 15, 36, 3) = (45, 15, 38, 1) \tag{5.4}$$

This allocation is incentive efficient for N, thus it is the unique MO-value of Γ_N.

Let us compare now the value allocation in (5.4) with other outcomes generated by alternative cooperative solutions. The unique allocation that can be supported by some M-solution of this game (see de Clippel (2005)) is

$$(U_1^H, U_1^L, U_2, U_3) = (45, 15, 39, 0) \tag{5.5}$$

This allocation considers player 3 as a null player. As de Clippel (2005, p. 79) points out: this happens because in the M-solution, the virtual value of coalition $\{1, 2\}$ is computed while using the vector (λ^N, α^N) specified for the grand coalition. By doing so, we act as if incentive constraints do not matter in coalition $\{1, 2\}$, although they do.

Even though player 3 does not create any additional surplus, it would be fair to give him a positive reward, as players 1 and 2 have to rely on him in order to weaken the incentive constraints they face when they cooperate in coalition $\{1, 2\}$. The MO-solution abides to this principle.

On the other hand, the unique allocation that can be supported by some H-solution of this game (see Salamanca (2016)) is

$$(U_1^H, U_1^L, U_2, U_3) = (45, 13, 38.6, 0.8) \tag{5.6}$$

Unlike the M-solution, this allocation gives player 3 a positive payoff. This is due to the fact that, in the H-solution, coalition $\{1, 2\}$ is restricted to choose an incentive compatible mechanism giving its members the same virtual gains (egalitarian criterion). This results in a weakening of the bargaining position of player 1 and 2, which increases 3's threatening power. Therefore, player 3 is rewarded 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 surplus in both states.

Thus it appears that the MO-solution reflects the structure of this game better than the other cooperative solutions.

A Collective Choice Problem.. The following cooperative game with incomplete information was proposed by Salamanca (2016). $N = \{1, 2, 3\}$, $T_3 = \{H, L\}$, $p(H) = 1 - p(L) = 9/10$. Decision options for coalitions 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 $D_N = \{D_{\{1,2\}} \times D_3\} \cup \{D_{\{1,3\}} \times D_2\} \cup \{D_{\{2,3\}} \times D_1\}$. Utility functions are as follows:

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

The game situation is interpreted a follows. Three players may invest in a work project which would cost \$10. 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 may be H with probability $9/10$ or L with probability $1/10$. If 3's type is H then the project is worth \$10 to him. If 3's type is L then the project is only worth \$5 to him. Every player $i \in N$ may decide not to cooperate (decision d_i), in which case he gets a reservation utility normalized to zero. If coalition $\{1,2\}$ forms, its members may 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. Any other financing option may be represented by a lottery on $\{d_{13}^1, d_{13}^3\}$. Players 1 and 2 are symmetric, then decision options for coalition $\{2,3\}$ are similarly interpreted. If all three players form a coalition, they may use a random device to pick a two-person coalition which must then make a decision as above.

We begin the analysis of this example by studying the different subgames. The subgame $\Gamma_{\{1,2\}}$ is a two-person bargaining problem with *complete* information. Clearly, the unique value (Nash bargaining solution) is the allocation $(U_1, U_2) = (5,5)$ achieved by the decision d_{12}. The virtual scales supporting this bargaining solution can be taken to be $\lambda_1^{\{1,2\}} = \lambda_2^{\{1,2\}} = 1$.[15] Let $i \in \{1,2\}$ be a fixed player and consider the subgame $\Gamma_{\{i,3\}}$. Figure 3 depicts the set of incentive efficient allocations for $\{i,3\}$ giving i and 3 a non-negative expected utility. Assume that player 3 is chosen to be a dictator for $\{i,3\}$. Then, visibly the best inscrutable inter-type compromise for both types of player 3 is to demand the allocation $(U_i, U_3^H, U_3^L) = (0,10,5)$. Likewise, the allocation $(U_i, U_3^H, U_3^L) = (9,0,0)$ is the best player i can achieve when he is given the power of dictatorship in $\{i,3\}$.

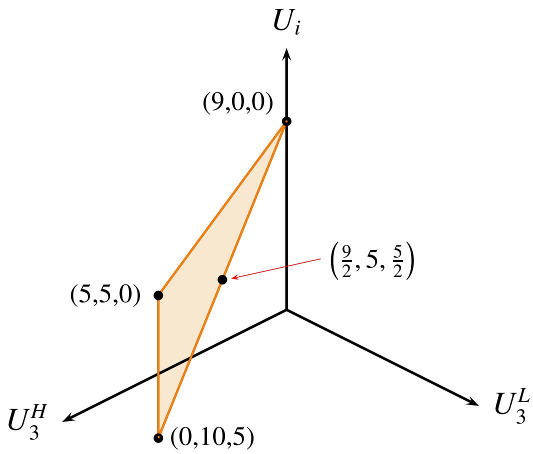

Figure 3: Incentive efficient allocations for $\{i,3\}$

Applying the conditional random dictatorship procedure, we obtain the allocation $(U_i, U_3^H, U_3^L) = \frac{1}{2}(9,0,0) + \frac{1}{2}(0,10,5) = (9/2, 5, 5/2)$. This payoff vector is incentive efficient for $\{i,3\}$, thus it is the unique MO-value of the subgame $\Gamma_{\{i,3\}}$. It is supported by the virtual scales $(\lambda^{\{i,3\}}, \alpha^{\{i,3\}})$ given by

$$\lambda_3^{\{i,3\}}(H) = \tfrac{4}{5}, \quad \lambda_3^{\{i,3\}}(L) = \tfrac{1}{5}, \quad \lambda_i^{\{i,3\}} = 1,$$
$$\alpha_3^{\{i,3\}}(H \mid L) = 0, \quad \alpha_3^{\{i,3\}}(L \mid H) = \tfrac{1}{10}.$$

The unique MO-solution of $\Gamma_{\{i,3\}}$ is the mechanism

$$\mu_{\{i,3\}}(d_{i3}^i \mid H) = \mu_{\{i,3\}}(d_{i2}^3 \mid H) = \tfrac{1}{2}, \quad \mu_{\{i,3\}}(d_{i3}^i \mid L) = \mu_{\{1,2\}}([d_i, d_3] \mid L) = \tfrac{1}{2}.$$

[15]Players 1 and 2 do not face any incentive constraint, so there are no signaling costs inside $\{1,2\}$.

We observe that by forming a coalition with player 3, 1 and 2 cannot expect to get more than \$9/2 in an equitable allocation. Therefore, players 1 and 2 are better off in coalition $\{1,2\}$, in which case they both get \$5 each. Salamanca (2016) then argues that coalition $\{1,2\}$ should be more likely to form, thus leaving the informed player with a low expected payoff.

Let us consider now the whole game Γ_N. We start by noticing that every incentive efficient mechanism for N satisfies Proposition 2 for the virtual scales (λ^N, α^N) given by

$$\lambda_1^N = 1, \quad \lambda_2^N = 1, \quad \lambda_3^N(H) = \tfrac{9}{10}, \quad \lambda_3^N(L) = \tfrac{1}{5},$$
$$\alpha_3^N(H \mid L) = \alpha_3^N(L \mid H) = 0.$$

The unique difference between real and virtual utility in coalition N is for type L of player 3, for which the virtual value of the good is \$0.

Assume that player 3 is chosen to be a dictator for N. The warranted claims of player 3 in $\Gamma_N^3(d_{12})$ are given by

$$\lambda_3^N(t)\omega_3^N(t) = \max_{d\in D} \sum_{i\in N} v_i(d,t,\lambda^N,\alpha^N) - \sum_{i\in\{1,2\}} u_i(d_{12}) = 0, \quad \forall t \in \{H, L\}.$$

It therefore follows that $\omega_3^N(H) = \omega_3^N(L) = 0$. The unique feasible allocation in $\Gamma_N^3(d_{12})$ that guarantees player 3's warranted claims is $(U_1, U_2, U_3^H, U_3^L) = (5,5,0,0)$. We deduce that this is the unique allocation supported by a neutral optimum of $\Gamma_N^3(d_{12})$.

Suppose now that player 1 is given the power of dictatorship in N. The virtual utilities that players 2 and 3 obtain under the scales $(\lambda^N, \alpha^{\{2,3\}})$ are

(v_2, v_3)	$[d_2, d_3]$	d_{23}^2	d_{23}^3
H	$(0,0)$	$\left(0, \frac{100}{9}\right)$	$(10,0)$
L	$(0,0)$	$(0,0)$	$(10,-10)$

Therefore, the warranted claim of player 1 in $\Gamma_N^1(\mu_{\{2,3\}})$ is

$$\begin{aligned}\omega_1^N &= \sum_{t\in\{H,L\}} p(t) \left[\max_{d\in D} \sum_{i\in N} v_i(d,t,\lambda^N,\alpha^N) - \sum_{i\in\{2,3\}} v_i(\mu_{\{2,3\}}, t, \lambda^N, \alpha^{\{2,3\}})\right] \\ &= \tfrac{9}{10}\left(10 - \tfrac{95}{9}\right) + \tfrac{1}{10}(10-0) \\ &= \tfrac{1}{2}.\end{aligned}$$

The unique feasible allocation in $\Gamma_N^1(\mu_{\{2,3\}})$ that guarantees player 1 to obtain his warranted claim is $(U_1, U_2, U_3^H, U_3^L) = (1/2, 9/2, 5, 5/2)$. This allocation is the unique payoff vector supported by a neutral optimum of $\Gamma_N^1(\mu_{\{2,3\}})$. By the symmetry of players 1 and 2, a similar analysis yields the allocation $(U_1, U_2, U_3^H, U_3^L) = (9/2, 1/2, 5, 5/2)$ when player 2 is chosen to be a dictator for N.

Applying the conditional random dictatorship procedure we obtain the allocation

$$(U_1, U_2, U_3^H, U_3^L) = \tfrac{1}{3}\left(\tfrac{9}{2}, \tfrac{1}{2}, 5, \tfrac{5}{2}\right) + \tfrac{1}{3}\left(\tfrac{1}{2}, \tfrac{9}{2}, 5, \tfrac{5}{2}\right) + \tfrac{1}{3}(5,5,0,0) = \left(\tfrac{10}{3}, \tfrac{10}{3}, \tfrac{10}{3}, \tfrac{5}{3}\right). \quad (5.7)$$

This allocation is incentive efficient for N, thus it is the unique MO-value of Γ_N.

Incidentally, in this example, (5.7) is also the unique payoff vector supported by some M-solution of Γ_N. According to this allocation, both types of player 3 obtain a significantly high expected payoff. In particular, type H of player 3 is rewarded the same as players 1 and 2. We notice, however, that conditional on state H, the players are not symmetric. Although, they all have the same valuation for the good in state H, because information is not verifiable, 1 and 2 are adversely affected by the likely presence of 3's type L. Salamanca (2016) provides a rationale for (5.7): when player 2 drops out of the game, player 3 becomes "surprisingly strong", since he has the informational advantage and, at the same time, player 1 cannot go to close a deal with 2. Such a surprisingly strong position is reflected in the value of the subgame $\Gamma_{\{1,3\}}$, which favors the informed player. Myerson (1991, p. 523) calls this property of the bargaining solution *arrogance of strength*. A similar reasoning applies when player 1 leaves the game. As a consequence, under the conditional random dictatorship procedure, player 3 has a two-thirds chance of finding herself in such a surprisingly strong position.

Salamanca (2016) offers an alternative outcome for this game. The unique utility allocation supported by some H-solution of Γ_N is

$$(U_1, U_2, U_3^H, U_3^L) = \left(\tfrac{41}{12}, \tfrac{41}{12}, \tfrac{40}{12}, \tfrac{10}{12}\right). \tag{5.8}$$

The allocation (5.8) gives less to both types of player 3 than (5.7). This is because, in the H-solution, the members of $\{i, 3\}$ ($i = 1, 2$) have to settle for a threat giving payoffs $(U_i, U_3^H, U_3^L) = (19/4, 5, 5/4)$. This payoff vector is more "equitable" than the MO-value of $\Gamma_{\{i,3\}}$ in the sense that type L of player 3 bears the efficiency losses originated on the adverse selection problem. Consequently, the strong position of player 3 in coalition $\{i, 3\}$ is weakened. It turns out that the H-solution prescribes a more appealing outcome in this game.

6. Summary and Final Comments

In this paper we have have provided a bargaining solution for cooperative game with incomplete information. Our solution concept – the MO-solution – is inspired in the conditional random dictatorship procedure studied by de Clippel, Peters and Zank (2004) and in Myerson's (1983) theory on the informed principal problem. It has the essential feature of generalizing Maschler and Owen's (1989, 1992) consistent value. We have studied its properties, as well as its behavior in some eloquent examples. We have also seen that it coincides with Myerson's (1984a) generalization of the Nash bargaining solution in two-person bargaining problems with incomplete information.

The main properties of the MO-solution are individual rationality and existence. These results are obtained under two important hypothesis: stochastically independent types and private values. As it was already commented in the introduction, independent types is a simplifying assumption that we can make without loss of generality since the MO-solution satisfies the *invariance probability axiom* described by Myerson (1984a). So for any game with dependent types, prior probabilities and utilities can be jointly modified in a way that the new game has independent types and both games impute probability and utility functions that are decision-theoretically equivalent. In contrast, private values is a restrictive assumption that rules out many interesting applications, yet it may be used as a first step toward the construction of a more general and complete theory.

For the MO-solution to be individually rational, a sufficient condition is that for every player $i \in N$ and every coalition $S \subseteq N$, each type of player i extracts an expected surplus in the virtual mechanism selection problem associated to $\Gamma_S^i(\mu_{S\setminus i})$ that is at least what he would get without any other player's help, namely,

$$\sum_{t_{S\setminus i}\in T_{S\setminus i}} p(t_{S\setminus i})\left(\max_{d_S\in D_S}\sum_{j\in S} v_j(d_S,t_j,\lambda^S,\alpha^S) - \sum_{j\in S\setminus i} v_j(\mu_{S\setminus i},t_j,\lambda^S,\alpha^{S\setminus i})\right) \geq v_i(\hat{\mu}_i,t_i,\lambda^S,\alpha^S), \quad \forall t_i \in T_i, \tag{6.1}$$

where $\hat{\mu}_i$ is defined as in (4.2). Unfortunately, superadditivity might not be sufficient to guarantee this inequality. The reason is that incentive compatibility in coalition $S \setminus i$ has an indirect impact on the participation constraints that player i has to meet when he is a dictator for coalition S. Private values exclude the informational externalities that the members of $N \setminus S$ exert on the incentive compatibility of coalitions S and $S \setminus i$, which help us to estimate the magnitude of the effect that the signaling costs associated to incentive compatibility in coalition $S \setminus i$ have on the residual virtual surplus that i can extract in coalition S. The inequality (6.1) is at the basis of Lemma 2. The requirement to exclude informational externalities is not surprising, if we recall that the consistent value treats every intermediate coalition independently of the grand coalition.

Our existence theorem is based on different arguments used to prove the existence of the Shapley NTU value (see Shapley (1969), Kern (1985) and Myerson (1991a, sec. 9.9)). One important element in the proof is the capacity to confine the warranted claims to a compact set. This is basically achieved by Lemma 2.

The difficulties identified above are not present either in Myerson's (1984b) M-solution or Salamanca's (2016) H-solution. This is so because these two cooperative solutions define the principles for equitable agreements only regarding the incentives and inter-type compromises inside the grand coalition. Indeed, both solution concepts define the virtual utility uniquely with respect to the utility weights λ^N and the signaling costs α^N of the grand coalition. In contrast, the MO-solution adjusts the virtual scales (λ^S, α^S) according to the incentives and inter-type compromises inside every coalition $S \subseteq N$. The adjustment of the virtual scales is done in such a way that "optimal" threats for intermediate coalitions are determined in exactly the same way the solution is determined for the whole game.

Equitable agreements in Myerson's (1984b) and Salamanca's (2016) theories are constructed upon normative criteria of *distributive justice*. The M-solutions reward players in proportions to their marginal contributions to all coalitions which they can join. The H-solutions are *egalitarian-based*, i.e., they require that if a player leaves a coalition, then the surplus variation for another player in the same coalition must be equal to his own surplus variation if this other player leaves the coalition. In contrast, the MO-solution adopts the view of fairness described by the *procedural justice*: a fair procedure is one that affords every individual an opportunity to participate in making the decision. In our model, "random dictatorship" gives every player an equal chance to act as a dictator. For the procedural justice, the perception of fairness is not given by the final outcome, as in the distributive justice, but by the process leading to it. From the viewpoint of the distributive justice, random dictatorship may be considered not equitable, since the outcome once a dictator has been selected is manifestly unfair. Nevertheless, if we think the different "dictatorial" mechanisms $(\mu_S^i)_{i\in S}$ in coalition S as a measure of its members individual bargaining ability, the average mechanism $\mu_S := \frac{1}{|S|}\sum_{i\in S}\mu_S^i$ may be considered as a *power-based* equitable agreement for S. To close, it is worth saying that random dictatorship is also an essential element of Myerson's (1984a) theory of two-person bargaining. Indeed, our extension of the conditional random dictatorship procedure can be considered as a generalization of Myerson's random dictatorship approach to games with more than two players.

7. Proofs

7.1. Proof of Lemma 2

Since $\alpha^S \geq 0$ and $\mu_{S\setminus i}$ is incentive compatible for $S \setminus i$, then for all $j \in S \setminus i$ we have that

$$\alpha_j^S(\tau_j \mid t_j)\Big[U_j(\mu_{S\setminus i} \mid t_j) - U_j(\mu_{S\setminus i}, \tau_j \mid t_j)\Big] \geq 0, \quad \forall t_j, \tau_j \in T_j. \tag{7.1a}$$

Therefore, the following chain of inequalities hold:

$$\begin{aligned}
&\sum_{t_{S\setminus i}\in T_{S\setminus i}} p(t_{S\setminus i}) \sum_{j\in S\setminus i} v_j(\mu_{S\setminus i}, t_j, \lambda^S, \alpha^{S\setminus i}) \\
&\quad= \sum_{j\in S\setminus i}\sum_{t_j\in T_j} \lambda_j^S(t_j) U_j(\mu_{S\setminus i} \mid t_j) \\
&\quad\leq \sum_{j\in S\setminus i}\sum_{t_j\in T_j} \lambda_j^S(t_j) U_j(\mu_{S\setminus i} \mid t_j) + \sum_{j\in S\setminus i}\sum_{t_j\in T_j}\sum_{\tau_j\in T_j} \alpha_j^S(\tau_j \mid t_j)\Big[U_j(\mu_{S\setminus i} \mid t_j) - U_j(\mu_{S\setminus i}, \tau_j \mid t_j)\Big] \\
&\quad= \sum_{t_{S\setminus i}\in T_{S\setminus i}} p(t_{S\setminus i}) \sum_{j\in S\setminus i} v_j(\mu_{S\setminus i}, t_j, \lambda^S, \alpha^S)
\end{aligned} \tag{7.1b}$$

where, the first equality is due to (4.5), the inequality in the second line follows from (7.1a) and finally the last equality is obtained from the definition of virtual utility in (2.4).

On the other hand, let $\hat{\mu}_i \in \mathcal{M}_i$ be defined as in (4.2). Then, for any $t_i \in T_i$ we have

$$\begin{aligned}
&\left(\lambda_i^S(t_i) + \sum_{\tau_i\in T_i} \alpha_i^S(\tau_i \mid t_i)\right)\omega_i^S(t_i) - \sum_{\tau_i\in T_i} \alpha_i^S(t_i \mid \tau_i)\omega_i^S(\tau_i) \\
&\quad= p(t_i) \sum_{t_{S\setminus i}\in T_{S\setminus i}} p(t_{S\setminus i}) \left(\max_{d_S\in D_S} \sum_{j\in S} v_j(d_S, t_S, \lambda^S, \alpha^S) - \sum_{j\in S\setminus i} v_j(\mu_{S\setminus i}, t_S, \lambda^S, \alpha^{S\setminus i})\right) \\
&\quad\geq p(t_i) \sum_{t_{S\setminus i}\in T_{S\setminus i}} p(t_{S\setminus i}) \left(\sum_{j\in S} v_j((\hat{\mu}_i, \mu_{S\setminus i}), t_S, \lambda^S, \alpha^S) - \sum_{j\in S\setminus i} v_j(\mu_{S\setminus i}, t_S, \lambda^S, \alpha^{S\setminus i})\right) \\
&\quad= p(t_i) v_i(\hat{\mu}_i, t_i, \lambda^S, \alpha^S) \\
&\qquad\quad+ p(t_i) \sum_{t_{S\setminus i}\in T_{S\setminus i}} p(t_{S\setminus i}) \left(\sum_{j\in S\setminus i} v_j(\mu_{S\setminus i}, t, \lambda^S, \alpha^S) - \sum_{j\in S\setminus i} v_j(\mu_{S\setminus i}, t, \lambda^S, \alpha^{S\setminus i})\right)
\end{aligned}$$

$$\geq p(t_i) v_i(\hat{\mu}_i, t_i, \lambda^S, \alpha^S)$$

$$= \left(\lambda_i^S(t_i) + \sum_{\tau_i \in T_i} \alpha_i^S(\tau_i \mid t_i)\right) \max_{d_i \in D_i} u_i(d_i, t_i) - \sum_{\tau_i \in T_i} \alpha_i^S(t_i \mid \tau_i) \max_{d_i \in D_i} u_i(d_i, \tau_i)$$

In this chain, the first equality follows from the fact that ω_i^S is warranted by λ^S, α^S, $\alpha^{S\setminus i}$ and $\mu_{S\setminus i}$; the inequality in the second line is due to the max operator in the first line and the fact that $(\hat{\mu}_i, \mu_{S\setminus i}) \in \mathcal{M}_S$; the equality in the third line uses the definition of $(\hat{\mu}_i, \mu_{S\setminus i})$ together with the orthogonal coalitions assumption; the inequality in the forth line is due to (7.1b); and finally, the equality in the fifth line uses the definition of the virtual utility. The result thus follows from Lemma 1. □

7.2. *Proof of Proposition 4*

Since $U_i(\mu_S)$ is warranted by λ^S, α^S, $\alpha^{S\setminus i}$ and $\mu_{S\setminus i}$, summing the warrant equations we obtain

$$\begin{aligned}\sum_{t_i \in T_i} \lambda_i^S(t_i) U_i(\mu_S \mid t_i) &= \sum_{t_S \in T_S} p(t_S) \left[\max_{d_S \in D_S} \sum_{j \in S} v_j(d_S, t_j, \lambda^S, \alpha^S) - \sum_{j \in S\setminus i} v_j(\mu_{S\setminus i}, t_j, \lambda^S, \alpha^{S\setminus i})\right] \\ &= \sum_{t_S \in T_S} p(t_S) \max_{d_S \in D_S} \sum_{j \in S} v_j(d_S, t_j, \lambda^S, \alpha^S) - \sum_{j \in S\setminus i} \sum_{t_j \in T_j} \lambda_j^S(t_j) U_j(\mu_{S\setminus i} \mid t_j)\end{aligned} \tag{7.2a}$$

The right-hand side of (7.2a) is the objective function of the dual problem associated to (4.1) evaluated at (λ^S, α^S). Because μ_S is feasible in the primal problem for i in $\Gamma_S^i(\mu_{S\setminus i})$ w.r.t. λ_i^S and $(\lambda^S, \alpha^S) \geq 0$ is feasible in the corresponding dual problem, duality theory implies that μ_S and α^S are optimal solutions of the primal and dual respectively. Hence, conditions (4.7)-(4.9) are satisfied for λ^S, α^S, $\alpha^{S\setminus i}$ and $\mu_{S\setminus i}$. □

7.3. *Proof of Theorem 1*

We proceed by induction. We consider first a singleton coalition $\{i\}$, with $i \in N$. The warrant equations for this coalition are:

$$\frac{1}{p(t_i)} \left[\left(\lambda_i^{\{i\}}(t_i) + \sum_{\tau_i \in T_i} \alpha_i^{\{i\}}(\tau_i \mid t_i)\right) \omega_i^{\{i\}}(t_i) - \sum_{\tau_i \in T_i} \alpha_i^{\{i\}}(t_i \mid \tau_i) \omega_i^{\{i\}}(\tau_i)\right] = \max_{d_i \in D_i} v_i(d_i, t_i, \lambda^{\{i\}}, \alpha^{\{i\}}), \quad \forall t_i \in T_i.$$

Summing the warrant equations over all $t_i \in T_i$ and using condition (*iii*) in Definition 4 yields

$$\sum_{t_i \in T_i} \lambda_i^{\{i\}}(t_i) U_i(\mu_i \mid t_i) \geq \sum_{t_i \in T_i} \lambda_i^{\{i\}}(t_i) \omega_i^{\{i\}}(t_i) = \sum_{t_i \in T_i} p(t_i) \max_{d_i \in D_i} v_i(d_i, t_i, \lambda^{\{i\}}, \alpha^{\{i\}}) \tag{7.3a}$$

Notice that μ_i is feasible in the primal for $\{i\}$ w.r.t. $\lambda^{\{i\}}$ and $\alpha^{\{i\}} \geq 0$ is feasible in the corresponding dual problem. By duality theory,

$$\sum_{t_i \in T_i} \lambda_i^{\{i\}}(t_i) U_i(\mu_i \mid t_i) \leq \sum_{t_i \in T_i} p(t_i) \max_{d_i \in D_i} v_i(d_i, t_i, \lambda^{\{i\}}, \alpha^{\{i\}})$$

Therefore, (7.3a) holds as equality. Weak duality then implies that μ_i and $\alpha^{\{i\}}$ are optimal solutions of the primal and dual problems for $\{i\}$ w.r.t. $\lambda^{\{i\}}$, respectively. Thus, μ_i is incentive efficient for $\{i\}$, and (2.6) and (2.7) are satisfied for μ_i with $(\lambda^{\{i\}}, \alpha^{\{i\}})$.

Let $S \subseteq N$ be a coalition with $|S| \geq 2$. Assume that, for each $i \in S$, $\mu_{S\setminus i}$ satisfies Proposition 2 for $(\lambda^{S\setminus i}, \alpha^{S\setminus i})$. Summing the warrant equations for coalition S yields

$$\begin{aligned}
&\sum_{i\in S}\sum_{t_i\in T_i}\lambda_i^S(t_i)\omega_i^S(t_i) \\
&\quad= \sum_{i\in S}\sum_{t_S\in T_S} p(t_S)\left(\max_{d_S\in D_S}\sum_{j\in S} v_j(d_S, t_j, \lambda^S, \alpha^S) - \sum_{j\in S\setminus i} v_j(\mu_{S\setminus i}, t_j, \lambda^S, \alpha^{S\setminus i})\right) \\
&\quad= |S|\sum_{t_S\in T_S} p(t_S)\max_{d_S\in D_S}\sum_{j\in S} v_j(d_S, t_j, \lambda^S, \alpha^S) - \sum_{i\in S}\left(\sum_{t_S\in T_S} p(t_S)\sum_{j\in S\setminus i} v_j(\mu_{S\setminus i}, t_j, \lambda^S, \alpha^{S\setminus i})\right) \\
&\quad= |S|\sum_{t_S\in T_S} p(t_S)\max_{d_S\in D_S}\sum_{j\in S} v_j(d_S, t_j, \lambda^S, \alpha^S) - \sum_{i\in S}\left(\sum_{j\in S\setminus i}\sum_{t_j\in T_j}\lambda_j^S(t_j)U_j(\mu_{S\setminus i} \mid t_j)\right)
\end{aligned} \tag{7.3b}$$

where the last equality follows from (4.5) and the fact that (4.4) is satisfied for $\mu_{S\setminus i}$ and $\alpha^{S\setminus i}$ by the induction hypothesis. On the other hand, condition (*iii*) in Definition 4 gives

$$\sum_{i\in S}\sum_{t_i\in T_i}\lambda_i^S(t_i)\omega_i^S(t_i) \leq |S|\sum_{i\in S}\sum_{t_i\in T_i}\lambda_i^S(t_i)U_i(\mu_S \mid t_i) - \sum_{i\in S}\sum_{j\in S\setminus i}\sum_{t_i\in T_i}\lambda_i^S(t_i)U_i(\mu_{S\setminus j} \mid t_i) \tag{7.3c}$$

Hence, noting that $\sum_{i\in S}\sum_{j\in S\setminus i}\sum_{t_i\in T_i}\lambda_i^{S,k}(t_i)U_i(\mu^k_{S\setminus j} \mid t_i) = \sum_{i\in S}\sum_{j\in S\setminus i}\sum_{t_j\in T_j}\lambda_j^{S,k}(t_j)U_j(\mu^k_{S\setminus i} \mid t_j)$, (7.3b) and (7.3c) imply that

$$\sum_{t_S\in T_S} p(t_S)\max_{d_S\in D_S}\sum_{j\in S} v_j(d_S, t_j, \lambda^S, \alpha^S) \leq \sum_{i\in S}\sum_{t_i\in T_i}\lambda_i^S(t_i)U_i(\mu_S \mid t_i) \tag{7.3d}$$

Since μ_S is feasible in the primal for S w.r.t. λ^S and $\alpha^S \geq 0$ is feasible in the corresponding dual problem. By duality theory,

$$\sum_{t_S\in T_S} p(t_S)\max_{d_S\in D_S}\sum_{j\in S} v_j(d_S, t_j, \lambda^S, \alpha^S) \geq \sum_{i\in S}\sum_{t_i\in T_i}\lambda_i^S(t_i)U_i(\mu_S \mid t_i)$$

Therefore, (7.3d) holds as equality. Weak duality then implies that μ_S and α^S are optimal solutions of the primal and dual problems for S w.r.t. λ^S, respectively. Thus, μ_S is incentive efficient for S, and (2.6) and (2.7) are satisfied for μ_S with (λ^S, α^S). □

7.4. Proof of Theorem 2

By Theorem 1, for any coalition $S \subseteq N$, $\mu_{S\setminus i}$ satisfies (4.4) for $\alpha^{S\setminus i}$. Then, Lemma 2 implies that

$$\omega_i^S(t_i) \geq \max_{d_i\in D_i} u_i(d_i, t_i), \quad \forall t_i \in T_i,\ \forall i \in S,\ \forall S \subseteq N.$$

Therefore, condition (*iii*) in the definition of a bargaining solution implies that for any $i \in N$,

$$U_i(\mu_i \mid t_i) \geq \omega_i^{\{i\}}(t_i) \geq \max_{d_i\in D_i} u_i(d_i, t_i), \quad \forall t_i \in T_i.$$

Hence, (2.2) holds for all singleton coalitions. Let $S \subseteq N$ be a coalition with $|S| \geq 2$. Assume that, for each $i \in S$, (2.2) is satisfied for coalition $S \setminus i$. Then, condition (*iii*) yields the desired result. $\square$

7.5. Proof of Theorem 3

The following proof is in the spirit of de Clippel's (2002) existence theorem of the virtual utility solutions.

Let $k \geq \sum_{i \in N} |T_i|$. For each $S \subseteq N$ we define

$$\Lambda_S^k = \left\{ \lambda \in \prod_{i \in S} \mathbb{R}^{T_i} \middle| \sum_{i \in S} \sum_{t_i \in T_i} \lambda_i(t_i) = 1,\ \lambda_i(t_i) \geq \tfrac{1}{k},\ \forall t_i \in T_i,\ \forall i \in S \right\}.$$

For each $S \subseteq N$, there exists a compact and convex set $A_S \subseteq \prod_{i \in S} \mathbb{R}_+^{T_i \times T_i}$ such that, for each $\lambda^S \in \prod_{i \in S} \mathbb{R}_+^{T_i} \setminus \{0\}$, A_S contains at least one optimal solution of the dual problem (2.8) (see proof of Theorem 3 in Myerson (1983)).

For each k larger than $\sum_{i \in N} |T_i|$, we define a correspondence

$$\Phi^k : \prod_{S \subseteq N} \Lambda_S^k \times \prod_{S \subseteq N} A_S \times \prod_{S \subseteq N} \mathcal{M}_S \rightrightarrows \prod_{S \subseteq N} \Lambda_S^k \times \prod_{S \subseteq N} A_S \times \prod_{S \subseteq N} \mathcal{M}_S$$

so that $((\lambda^S)_{S \subseteq N}, (\alpha^S)_{S \subseteq N}, (\mu_S)_{S \subseteq N}) \in \Phi^k((\hat{\lambda}^S)_{S \subseteq N}, (\hat{\alpha}^S)_{S \subseteq N}, (\hat{\mu}_S)_{S \subseteq N})$ iff for each $S \subseteq N$:

$$\lambda^S \in \arg\min_{\lambda \in \Lambda_S^k} \sum_{i \in S} \sum_{t_i \in T_i} \lambda_i(t_i) \left[|S| U_i(\hat{\mu}_S \mid t_i) - \sum_{j \in S \setminus j} U_i(\hat{\mu}_{S \setminus j} \mid t_i) - \hat{\omega}_i^S(t_i) \right] \tag{7.5a}$$

$$\alpha^S \in \arg\min_{\alpha \in A_S} \sum_{t_S \in T_S} p(t_S) \max_{d_S \in D_S} \sum_{i \in S} v_i(d_S, t_i, \hat{\lambda}^S, \alpha) \tag{7.5b}$$

$$\mu_S \in \arg\max_{\mu \in \mathcal{M}_S^*} \sum_{i \in S} \sum_{t_i \in T_i} \hat{\lambda}_i^S(t_i) U_i(\mu \mid t_i), \tag{7.5c}$$

where $\hat{\omega}_i^S = (\hat{\omega}_i^S(t_i))_{t_i \in T_i}$ is the unique allocation warranted by $\hat{\lambda}^S$, $\hat{\alpha}^S$, $\hat{\alpha}^{S \setminus i}$ and $\hat{\mu}_{S \setminus i}$.

For any value of k, the correspondence Φ^k is non-empty convex valued and upper-hemicontinuous. Then by the Kakutani fixed point theorem, for each k there exits some $((\lambda^{S,k})_{S \subseteq N}, (\alpha^{S,k})_{S \subseteq N}, (\mu_S^k)_{S \subseteq N}, (\delta^{S,k})_{S \subseteq N})$ such that

$$((\lambda^{S,k})_{S \subseteq N}, (\alpha^{S,k})_{S \subseteq N}, (\mu_S^k)_{S \subseteq N}) \in \Phi^k((\lambda^{S,k})_{S \subseteq N}, (\alpha^{S,k})_{S \subseteq N}, (\mu_S^k)_{S \subseteq N})$$

Since this sequence of fixed points lies on a compact domain, we may assume w.l.g. that it converges to some $((\bar{\lambda}^S)_{S \subseteq N}, (\bar{\alpha}^S)_{S \subseteq N}, (\bar{\mu}_S)_{S \subseteq N})$. We shall see that $(\bar{\mu}_S)_{S \subseteq N}$ is a bargaining solution.

Let $S \subseteq N$ be a coalition. For any k and for each $i \in S$, let $\omega_i^{S,k}$ be the unique allocation warranted by $\lambda^{S,k}$, $\alpha^{S,k}$, $\alpha^{S\setminus i,k}$ and $\mu_{S\setminus i}^k$. Then, for any $\lambda \in \Lambda_S^k$ we have that

$$\begin{aligned}
&\sum_{i\in S}\sum_{t_i\in T_i}\lambda_i(t_i)\left[|S|U_i(\mu_S^k \mid t_i) - \sum_{j\in S\setminus i} U_i(\mu_{S\setminus j}^k \mid t_i) - \omega_i^{S,k}(t_i)\right] \\
&\quad \geq \sum_{i\in S}\sum_{t_i\in T_i}\lambda_i^{S,k}(t_i)\left[|S|U_i(\mu_S^k \mid t_i) - \sum_{j\in S\setminus i} U_i(\mu_{S\setminus j}^k \mid t_i) - \omega_i^{S,k}(t_i)\right] \\
&\quad = |S|\sum_{i\in S}\sum_{t_i\in T_i}\lambda_i^{S,k}(t_i)U_i(\mu_S^k \mid t_i) - \sum_{i\in S}\sum_{j\in S\setminus i}\sum_{t_i\in T_i}\lambda_i^{S,k}(t_i)U_i(\mu_{S\setminus j}^k \mid t_i) - \sum_{i\in S}\sum_{t_i\in T_i}\lambda_i^{S,k}(t_i)\omega_i^{S,k}(t_i) \\
&\quad = |S|\sum_{i\in S}\sum_{t_i\in T_i}\lambda_i^{S,k}(t_i)U_i(\mu_S^k \mid t_i) - \sum_{i\in S}\sum_{j\in S\setminus i}\sum_{t_i\in T_i}\lambda_i^{S,k}(t_i)U_i(\mu_{S\setminus j}^k \mid t_i) \\
&\qquad - \left[|S|\sum_{t_S\in T_S} p(t_S)\max_{d_S\in D_S}\sum_{i\in S} v_i(d_S, t_i, \lambda^{S,k}, \alpha^{S,k}) - \sum_{i\in S}\sum_{j\in S\setminus i}\sum_{t_j\in T_j}\lambda_j^{S,k}(t_j)U_j(\mu_{S\setminus i}^k \mid t_j)\right] \\
&= \quad |S|\left[\sum_{i\in S}\sum_{t_i\in T_i}\lambda_i^{S,k}(t_i)U_i(\mu_S^k \mid t_i) - \sum_{t_S\in T_S} p(t_S)\max_{d_S\in D_S}\sum_{i\in S} v_i(d_S, t_i, \lambda^{S,k}, \alpha^{S,k})\right] \\
&= 0,
\end{aligned} \tag{7.5d}$$

where the inequality in the first line is due to the fixed point condition together with (7.5a); the equality in the third line follows from summing the warrant equations as in (7.3b); and finally, the equality in the fifth line follows from duality theory since by the fixed point condition, (7.5c) implies that μ_S^k is an optimal solution of the primal for S w.r.t. $\lambda^{S,k}$ and (7.5b) implies that $\alpha^{S,k}$ is an optimal solution of the corresponding dual problem.

We shall now use (7.5d) to show that the sequence $\{(\omega^{S,k})_{S\subseteq N}\}_k$ is contained in a compact set. We define $M := \max_{i\in N}\max_{t_i\in T_i}\max_{d\in D}|u_i(d, t_i)|$. Then, for any $\lambda \in \Lambda_S^k$ we have that

$$\sum_{i\in S}\sum_{t_i\in T_i}\lambda_i(t_i)|S|U_i(\mu_S^k \mid t_i) \leq |S|M$$

and

$$\sum_{i\in S}\sum_{t_i\in T_i}\lambda_i(t_i)\sum_{j\in S\setminus i} U_i(\mu_{S\setminus j}^k \mid t_i) \geq -M(|S|-1).$$

Then, (7.5d) implies that for each k

$$\sum_{i\in S}\sum_{t_i\in T_i}\lambda_i(t_i)\omega_i^{S,k}(t_i) \leq M(2|S|-1), \quad \forall\lambda \in \Lambda_S^k.$$

Thus, for each $S \subseteq N$, the sequence $\{\omega^{S,k}\}_k$ is bounded above. By Lemma 2,

$$\omega_i^{S,k}(t_i) \geq \max_{d_i\in D_i} u_i(d_i, t_i), \qquad \forall t_i \in T_i,\ \forall i \in S,\ \forall S \subseteq N.$$

Hence, for any $S \subseteq N$, the sequence $\{\omega^{S,k}\}_k$ is also bounded below. Then, we may assume w.l.g. that $\{(\omega^{S,k})_{S\subseteq N}\}_k$ converges to some $(\bar{\omega}^S)_{S\subseteq N}$. By the continuity of the warranted claims, for each $S \subseteq N$ and every $i \in S$, $\bar{\omega}_i^S$ is warranted by $\bar{\lambda}^S$, $\bar{\alpha}^S$, $\bar{\alpha}^{S\setminus i}$ and $\bar{\mu}_{S\setminus i}$.

Finally, by (7.5d), we have that for each $S \subseteq N$

$$\sum_{i\in S}\sum_{t_i\in T_i}\lambda_i(t_i)\left[|S|U_i(\bar{\mu}_S \mid t_i) - \sum_{j\in S\setminus i} U_i(\bar{\mu}_{S\setminus j} \mid t_i)\right] \geq \sum_{i\in S}\sum_{t_i\in T_i}\lambda_i(t_i)\bar{\omega}_i^S(t_i), \quad \forall \lambda \in \prod_{i\in S}\mathbb{R}_+^{T_i}\setminus\{0\}.$$

In particular,

$$|S|U_i(\bar{\mu}_S \mid t_i) - \sum_{j\in S\setminus i} U_i(\bar{\mu}_{S\setminus j} \mid t_i) \geq \bar{\omega}_i^S(t_i), \quad \forall t_i \in T_i,\ \forall i \in S,\ \forall S \subseteq N.$$

□

8. References

1. de Clippel, G. (2002). Cooperative games with non-transferable utility and under asymmetric information. Ph.D. Thesis, Université Catholique de Louvain, Louvain-la-Neuve, Belgium.
2. de Clippel, G. (2005). “Values for cooperative games with incomplete information: An eloquent example”. *Games and Economic Behavior*, 53 pp. 73-82.
3. de Clippel, G., Hans, P. and Zank, H. (2004). “Axiomatizing the Harsanyi solution, the symmetric egalitarian solution and the consistent solution for NTU-games”. *International Journal of Game Theory*, 33, pp. 145-158.
4. Harsanyi, J. (1963). “A simplified bargaining model for the *n*-person cooperative game”. *International Economic Review*, 4, pp. 194-220.
5. Hart, S. (1994). On prize games. In: *Essays in Game Theory*. N. Megiddo (Ed.). Springer New York, pp. 111-121.
6. Hart, S. (2005). “An axiomatization of the consistent non-transferable utility value”. *International Journal of Game Theory*, 33, pp 355-366.
7. Hart, S. and Mas-Colell, A. (1996). “Bargaining and the value”. *Econometrica*, 64, pp. 357-380.
8. Kern, R. (1985). “The Shapley transfer value without zero-weights”. *International Journal of Game Theory*, 14, pp 73-92.
9. Maschler, M. and Owen, G. (1989). “The consistent Shapley value for hyperplane games”. *International Journal of Game Theory*, 18, pp 389-407.
10. 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.
11. Myerson, R. (1983). “Mechanism design by an informed principal”. *Econometrica*, 51, pp. 1767-1797.
12. Myerson, R. (1984a). “Two-person bargaining problems with incomplete information”. *Econometrica*, 52, pp. 461-488.
13. Myerson, R. (1984b). “Cooperative games with incomplete information”. *International Journal of Game Theory*, 13, pp. 69-96.
14. Myerson, R. (1991a). *Game Theory: Analysis of Conflict*, Harvard University Press.

15. Myerson, R. (1991b). Analysis of incentives in bargaining and mediation. In: *Negotiation Analysis*. P. Young (Ed.). The University of Michigan Press, pp. 67-85.
16. Myerson, R. (1992). Fictitious-transfers solutions in cooperative game theory. In: *Rational Interaction*. R. Selten (Ed.). New York: Springer-Verlag, pp. 13-33.
17. Nash, J.-F. (1950). "The bargaining problem". *Econometrica*, 18, pp. 155-162.
18. Salamanca, A. (2016). An egalitarian value for cooperative games with incomplete information. TSE Working paper n. 16-620. Toulouse School of Economics, Université de Toulouse.
19. 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.
20. Shapley, L. (1969). Utility comparisons and the theory of games. In: *La Decision*. Paris: Editions du CNRS, pp. 251-263.