On the Values of Bayesian Cooperative Games with Sidepayments

by

Andrés Salamanca

On the Values of Bayesian Cooperative Games with Sidepayments[☆]

Andrés Salamanca

Department of Business and Economics, University of Southern Denmark

Abstract

In this paper we explore the relationship between several value-like solution concepts for cooperative games with incomplete information and utility transfers in the form of sidepayments. In our model, state-contingent contracts are required to be incentive compatible, and thus utility might not be not fully transferable (as it would be in the complete information case). When we restrict our attention to games with orthogonal coalitions (i.e., which do not involve strategic externalities), our first main result states that Myerson's [Cooperative games with incomplete information. *Int. J. Game Theory*. (1984), 13, 69-96] generalization of the Shapley NTU value and Salamanca's [A generalization of the Harsanyi NTU value to games with incomplete information. (2016), HAL 01579898] extension of the Harsanyi NTU value are interim utility equivalent. I f we allow for arbitrary informational and strategic externalities, our second main result establishes that the ex-ante evaluation of Myerson's solution equals Kalai and Kalai's [Cooperation in strategic games revisited. *Q. J. Econ.* (2013) 128, 917-966] cooperative-competitive value in two-player games with verifiable types.

Keywords: Cooperative games, incomplete information, sidepayments.

JEL Classification: C71, C78, D82.

1. Introduction

In this paper we explore the relationship between the following value-like solution concepts for cooperative games with incomplete information: Myerson's (1984b) generalization of Shapley's (1969) NTU value, Salamanca's (2016) extension of Harsanyi's (1963) NTU value, and A. Kalai and E. Kalai's (2013) cooperative-competitive solution. We consider a model in which utility transfers in the form of sidepayments are allowed. Transferable utility is a common assumption in cooperative game theory. It states that utilities are quasi-linear in money and that *unrestricted* monetary transfers can be performed (see Aumann, 1960). Our model, however, may exhibit *restricted* monetary transfers. The reason is that allowable state-contingent

contracts are required to be incentive compatible and, thus, not all state-contingent plans of sidepayments might be feasible.

We separate the analysis of our model into two parts according to the number of players and the amount of externalities involved in the game situation. In the first part, we start by considering general n-person Bayesian cooperative games in which there are no *strategic externalities* (i.e, the actions available to any particular coalition do not have an impact on the utilities of the players in the complementary coalition). In such a situation, coalitions are then said to be *orthogonal*. However, we allow a player's utility to depend on the other players' information; that is, we permit arbitrary *informational externalities*. In this specialized model, when one considers the particular case of complete information, it is widely known that both Harsanyi's (1963) NTU value and Shapley's (1969) NTU value coincide. Moreover, their common formula is given by Shapley's (1953) transferable utility (TU) value. The first part of this paper aims to provide an analogous result in a more general environment with incomplete information.

Myerson (1984a,b) developed an approach in which incentive constraints are used to define the virtual utility of players. Virtual utilities generalize the weighted-utility scales of the Harsanyi-Shapley method of fictitious transfers.[1] This approach was used in Myerson (1984b) to extend the Shapley NTU value to games with incomplete information. It has also recently been used in Salamanca (2016) to generalize the Harsanyi NTU value. Both solution concepts reflect not only the signaling costs associated with incentive compatibility, but also the fact that individuals negotiate at the interim stage (i.e., after each player has received his private information). Our first main result (Theorem 1) establishes that these two cooperative solutions are interim utility equivalent in our model with sidepayments and orthogonal coalitions. Remarkably, this result is not the consequence of the fact that utility transfers may serve as a linear activity that can be used for signaling purposes (i.e., for helping to satisfy incentive compatibility).[2] Indeed, in our model a transfer scheme will typically affect the interim utilities, which makes it impossible to transfer utility across types without affecting the incentive constraints.[3] Instead, Theorem 1 follows from the fact that coalitional agreements can be made equitable by means of an appropriate transfer scheme.

As a direct corollary of Theorem 1, we obtain a generalization of the Shapley TU value to games with incomplete information. However, its formula cannot be described by a simple closed form expression. The reason is that, due to the incentive constraints, the set of (interim incentive) efficient utility allocations is not generally described by a hyperplane as it would be in a game with complete information.

The second part of this paper is devoted to the analysis of our model when arbitrary informational and strategic externalities are allowed. However, we simplify the coalitional analysis by focusing only on two-player games. At the more general level, Myerson's (1984b) coope-

---

[1] See Myerson (1992) for a detailed explanation of the fictitious transfers method.

[2] see d'Aspremont and Gérard-Varet (1979, 1982).

[3] In particular, one cannot generally construct, corresponding to a first best interim utility allocation, a transfer scheme satisfying incentive compatibility (see, e.g., example 1 in Myerson, 2007). Similar difficulties were also encountered by Forges, Mertens and Vohra (2002) in their analysis of the incentive-compatible interim (coarse) core of an exchange economy with differential information and quasi-linear utilities.

rative solution allows for both kinds of externalities. Thus, it can be applied to this particular class of games. Indeed, his cooperative solution extends Nash's (1953) bargaining solution with variable threats. More recently, A. Kalai and E. Kalai (2013) proposed a value for two-person Bayesian games (in strategic form) with transferable utility. Their semi-cooperative solution, called the *cooperative-competitive* (or "*coco*") *value*, is based on a decomposition of the game into cooperative and competitive component games. The coco value conceptually differs from Myerson's solution in that the former ignores potential incentive compatibility issues. Also, it is defined only at the ex-ante stage. Kalai and Kalai (2013), however, conjectured the existence of a close relationship between these two solution concepts in the case where private information is verifiable ex-post.[4] Our second main result (Theorem 2) provides a positive answer to this open problem. We show that under de Clippel and Minelli's (2004) verifiable types assumption, Myerson's solution and the coco value are ex-ante utility equivalent; that is, if the players evaluate their welfare as if they were uninformed, both values prescribe the same utility allocation. Verifiable types mean that when agreements are implemented, each individual can costlessly verify the true information state. This assumption implies that incentive constraints are not required and, thus, any contract can be enforced once it is agreed upon. According to Theorem 2, Myerson's solution can thus be viewed as a formal generalization of the coco value to games in which attention should be restricted to outcomes that are incentive compatible. This result helps us to understand why extending the coco value to cooperative games in which first best outcomes cannot be made incentive compatible requires an appropriate interim framework, as in Myerson's (1984a,b) approach. In this sense, Theorem 2 might be considered as evidence in favor of the conceptual significance of Myerson's (1984b) theory.

At this point, it is reasonable to ask why it is important to study the relations that can be established between the different cooperative solutions analyzed here. On the one hand, it allows us to determine under which circumstances distinct theories of cooperation can be unified in a common framework. Indeed, this is the main direct contribution of Theorems 1 and 2. At the same time, the proofs of these two theorems serve as a device to better understand the differences and similarities between these solution concepts. On the other hand, comparing various solution concepts helps us to clarify the nature of their hypotheses, exhibiting their logic and revealing what they do and do not depend on.

The paper is organized as follows. Section 2 is devoted to formally specifying the model of a Bayesian cooperative game with sidepayments. In Section 3 we introduce the concept of incentive efficiency and its relation to the virtual utility approach. Sections 4 and 5 contain the main body of results: Section 4 analyzes n-player games with orthogonal coalitions, and finally, Section 5 focuses on two-player Bayesian games in strategic form.

2. Bayesian Cooperative Games

A *Bayesian cooperative game* (or *cooperative game with incomplete information*) is a tuple $\Gamma = \{N, (D_S)_{S \subseteq N}, (T_i, u_i)_{i \in N}, p\}$, where $N = \{1, 2, ..., n\}$ denotes the set of players and for any

---

[4]This open problem was also pointed out by Forges and Serrano (2013).

(nonempty) coalition $S \subseteq N$,[5] D_S is the set of feasible decisions for S. For any player $i \in N$, T_i denotes the (finite) set of possible types of player i, p is the prior probability distribution over $T_N = \prod_{i \in N} T_i$, and $u_i : D_N \times T_N \to \mathbb{R}$ is the utility function of player i. The sets of feasible decisions are finite and *superadditive*, namely for any two disjoint coalitions S and R, $D_R \times D_S \subseteq D_{R \cup S}$. These definitions allow for both informational and strategic externalities, since the payoffs of the members of a coalition S may depend on the types and decisions of the players in $N \setminus S$.

We assume that types are stochastically independent and $p(t_i) > 0$ for all $t_i \in T_i$ and all $i \in N$.[6] We use the notations $t_{-i} = (t_j)_{j \in N \setminus i} \in T_{-i} = \prod_{j \in N \setminus i} T_j$ and $t_S = (t_i)_{i \in S} \in T_S = \prod_{i \in S} T_i$. For simplicity, we drop the subscript N in the case of the grand coalition, so we define $D = D_N$ and $T = T_N$.

A *mechanism* for coalition $S \subseteq N$ is a pair of functions (μ_S, x_S) defined by[7]

$$\mu_S : T_S \to \Delta(D_S) \qquad x_S : T_S \to \mathbb{R}^S$$
$$t_S \mapsto \mu_S(\cdot \mid t_S) \qquad t_S \mapsto (x_S^i(t_S))_{i \in S}.$$

The component μ_S is a type-contingent lottery on the set of feasible decisions for S, while x_S is a vector of type-contingent monetary transfers. Monetary transfers must satisfy the following budget feasibility condition:[8]

$$\sum_{i \in S} x_S^i(t_S) \le 0, \quad \forall t_S \in T_S. \tag{2.1}$$

When $S \ne N$, the mechanism (μ_S, x_S) stands as a *threat* to be carried out only if $N \setminus S$ refuses to cooperate with S. We let the set of *budget-feasible* mechanisms satisfying (2.1) be denoted $\mathcal{F}_S$.

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

$$U_i(\mu_N, x_N, \tau_i \mid t_i) = \sum_{t_{-i} \in T_{-i}} p(t_{-i}) \left[x_N^i(\tau_i, t_{-i}) + \sum_{d \in D} \mu_N(d \mid \tau_i, t_{-i}) u_i(d, (t_i, t_{-i})) \right].$$

Monetary transfers are added linearly to the expected utilities. As is standard, we denote $U_i(\mu_N, x_N \mid t_i) = U_i(\mu_N, x_N, t_i \mid t_i)$.

---

[5]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.

[6]The assumption of independent types is only needed to simplify our results, and it does not entail any loss of generality. Indeed, the solution concepts studied in this paper satisfy the *probability-invariance axiom* described by Myerson (1984a), and so for any game with dependent types, conditional probabilities and utilities can be jointly modified in such a way that the new game has independent types and both games impute probability and utility functions that are decision-theoretically equivalent.

[7]For any finite set A, $\Delta(A)$ denotes the set of probability distributions over A.

[8]Other forms of budget feasibility can be defined. For instance, Prescott and Townsend (1984) and Myerson (2007) consider *average* budget feasibility; i.e., $\sum_{t_S \in T_S} p(t_S) \sum_{i \in S} x_S^i(t_S) \le 0$.

A mechanism (μ_N, x_N) is *incentive compatible* if and only if

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

By the revelation principle, any Nash equilibrium of any non-cooperative game that the players could design in order to exchange information and make decisions can be equivalently represented by an incentive compatible-mechanism. Hence, there is no loss of generality in restricting our attention to such incentive compatible mechanisms. We also notice that incentive constraints only depend on the conditional expected monetary transfers. Therefore, we can restrict ourselves without loss of generality to deterministic money transfers. A budget-feasible and incentive-compatible mechanism for the grand coalition is said to be *feasible* for N. We denote by $\mathcal{F}_N^*$ the set of feasible mechanisms for N.

3. Incentive Efficiency and Virtual Utility

A mechanism is (*interim*) *incentive-efficient* for the grand coalition if and only if it is feasible and no other feasible mechanism yields strictly higher expected utilities to all types of all players. Incentive-efficient mechanisms can be characterized using the concept of virtual utility.

Given vectors $\lambda \in \prod_{i \in N} \mathbb{R}_+^{T_i}$ and $\alpha \in \prod_{i \in N} \mathbb{R}_+^{T_i \times T_i}$, the *virtual utility* of player i from decision $d \in D$ in state $t \in T$ is defined as follows:

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

The vectors λ and α are called the virtual utility scales. The virtual utility of player i is a distorted utility scale, which exaggerates the difference between his actual utility and that of the other types that would be tempted to imitate him. The following characterization results from duality theory of linear programming (a detailed reasoning is given in Myerson, 2007).

Proposition 1 (**Incentive-efficiency**). *A feasible mechanism $(\mu_N, x_N) \in \mathcal{F}_N^*$ is incentive- efficient if and only if there exist $\lambda \in \prod_{i \in N} \mathbb{R}_+^{T_i}$ and $\alpha \in \prod_{i \in N} \mathbb{R}_+^{T_i \times T_i}$, such that*

$$\alpha_i(\tau_i \mid t_i) \left[U_i(\mu_N, x_N \mid t_i) - U_i(\mu_N, x_N, \tau_i \mid t_i) \right] = 0, \quad \forall i \in N, \forall t_i \in T_i, \forall \tau_i \in T_i, \tag{3.1}$$

$$\lambda_i(t_i) + \sum_{\tau_i \in T_i} \alpha_i(\tau_i \mid t_i) - \sum_{\tau_i \in T_i} \alpha_i(t_i \mid \tau_i) = p(t_i), \quad \forall i \in N, \forall t_i \in T_i, \tag{3.2}$$

$$\sum_{d \in D} \mu_N(d \mid t) \sum_{i \in N} v_i(d, t, \lambda, \alpha) = \max_{d \in D} \sum_{i \in N} v_i(d, t, \lambda, \alpha), \quad \forall t \in T, \tag{3.3}$$

$$\sum_{i \in N} x_N^i(t) = 0, \quad \forall t \in T. \qquad (\textit{exact budget feasibility}) \tag{3.4}$$

Equation (3.1) is the usual *complementary slackness* condition. Equations in (3.2) are called *hydraulic equations* by Myerson (2007): Consider a network in which, at each node, a type t_i is located. If we interpret $p(t_i)$ as the flow into the network at t_i, $\lambda_i(t_i)$ as the flow out of the

network at t_i, and $\alpha_i(\tau_i \mid t_i)$ as the flow from τ_i to t_i, then (3.4) states that these flows are balanced at each node. Finally, (3.3)-(3.4) imply that any incentive-efficient mechanism determines an allocation that is ex-post efficient in terms of the virtual utility scales. Hence, one could say that incentive compatibility compels the players to behave according to their virtual utilities. Myerson (1984b) refers to this idea as the *virtual utility hypothesis*.[9]

4. Values for Bayesian Cooperative Games with Orthogonal Coalitions

As in most of the literature in cooperative game theory, in this section we shall assume that coalitions are *orthogonal*, namely when coalition $S \subseteq N$ chooses an action that is feasible for it, the payoffs to the members of S do not depend on the actions of the complementary coalition $N \setminus S$. Formally,

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

for every S, $i \in S$, $d_S \in D_S$, $d_{N\setminus S}, d'_{N\setminus S} \in D_{N\setminus S}$, and $t \in T$. Then we can let $u_i(d_S, t)$ denote the utility of player $i \in S$ if $d_S \in D_S$ is carried out. That is, $u_i(d_S, t) := u_i((d_S, d_{N\setminus S}), t)$ for some $d_{N\setminus S} \in D_{N\setminus S}$ (recall that $D_S \times D_{N\setminus S} \subseteq D$). This assumption excludes strategic externalities. However, the payoffs of the members of a coalition S might still depend on the types of the players in $N \setminus S$. Hence, informational externalities are allowed.

When information is complete, the orthogonal coalitions hypothesis makes it possible to describe an NTU game as a collection of feasible utility sets. This *characteristic function form* suppresses any explicit mention of the decisions generating the utilities. Although, it is implicitly assumed that an utility allocation u_S is feasible for S if the players in S together have a joint strategy that enables them to allocate u_S.[10] If utilities are linear in money and players can make unrestricted sidepayments of money, we obtain a TU game.[11] It is well known that both the Shapley NTU value and the Harsanyi NTU value of a TU game coincide and their common formula is given by the Shapley TU value.[12] Even though utility may not be fully transferable in our model because of the presence of the incentives constraints, we provide an analogous result in the class of Bayesian cooperative games.

4.1. The M-Solution

We consider the fictitious game in which the players make interpersonal utility comparisons in terms of some fixed virtual scales (λ, α). For any coalition S, we let $W_S(\mu_S, x_S, t, \lambda, \alpha)$ be the sum of virtual utilities that the members of S would expect in state t when they select the

---

[9]Myerson (1991, ch. 10) provides a detailed discussion about the meaning and conceptual significance of the virtual utility hypothesis.

[10]This may include a correlated strategy or a joint decision, discarding utility or even transferring utility.

[11]Indeed, let $V = (V(S))_{S \subseteq N}$ be an NTU game. For each $S \subseteq N$, let $P(S) := \{u \in \mathbb{R}^S \mid \sum_{i \in S} u_i \leq 0\}$ denote the set of (unrestricted) sidepayments for the members of S. Then the game $W = (V(S) + P(S))_{S \subseteq N}$ is a TU game for which the worth of coalition S is $w(S) = \max_{v \in V(S)} \sum_{i \in S} v_i$ and $W(S) = \{w \in \mathbb{R}^S \mid \sum_{i \in S} w_i \leq w(S)\}$.

[12]See Myerson (1991, pp. 470) for a detailed explanation of why the Shapley NTU value coincides with the Shapley TU value in TU games. On the other hand, Proposition 4.10 in Hart (1985) establishes the equivalence between the Harsanyi NTU value and the Shapley TU value in TU games.

mechanism (μ_S, x_S) as a threat; that is[13]

$$W_S(\mu_S, x_S, t, \lambda, \alpha) = \sum_{i \in S} \left[v_i(\mu_S, t, \lambda, \alpha) + x_S^i(t_S) \right], \tag{4.1}$$

where $v_i(\mu_S, t, \lambda, \alpha)$ is the linear extension of $v_i(\cdot, t, \lambda, \alpha)$ over μ_S. For any vector of threats $\eta = (\mu_S, x_S)_{S \subseteq N}$, we define $W(\eta, t, \lambda, \alpha) = (W_S(\mu_S, x_S, t, \lambda, \alpha))_{S \subseteq N}$ as the characteristic function game when virtual utility is fully transferable in state t. Let ϕ be the Shapley TU value operator; for $i \in N$, $\phi_i(N, W(\eta, t, \lambda, \alpha))$ will thus denote the Shapley TU value of player i in the virtual game when the vector of threats η is selected.

The interim allocation $\omega \in \prod_{i \in N} \mathbb{R}^{T_i}$ is *warranted* by λ, α, and η if and only if

$$\left(\lambda_i(t_i) + \sum_{\tau_i \in T_i} \alpha_i(\tau_i \mid t_i) \right) \omega_i(t_i) - \sum_{\tau_i \in T_i} \alpha_i(t_i \mid \tau_i) \omega_i(\tau_i) =$$
$$\sum_{t_{-i} \in T_{-i}} p(t) \phi_i(N, W(\eta, t, \lambda, \alpha)), \quad \forall i \in N, \quad \forall t_i \in T_i. \tag{4.2}$$

In other words, ω corresponds to the real utility allocation that would give every type of each player his expected Shapley TU value in the virtual game.

We say that $\eta = (\mu_S, x_S)_{S \subseteq N}$ is a vector of *rational threats* (w.r.t. λ and α) if, for each $S \subseteq N$, the mechanism (μ_S, x_S) is an optimal solution to

$$\max_{(\nu_S, y_S) \in \mathcal{F}_S} \sum_{t \in T} p(t) W_S(\nu_S, y_S, t, \lambda, \alpha). \tag{4.3}$$

It follows straightforwardly that a mechanism (μ_S, x_S) is an optimal solution to (4.3) if and only if, for every $t_S \in T_S$,

$$\sum_{i \in S} v_i(\mu_S, t_S, \lambda, \alpha) = \max_{d_S \in D_S} \sum_{i \in S} v_i(d_S, t_S, \lambda, \alpha) \quad \text{and} \quad \sum_{i \in S} x_S^i(t_S) = 0,$$

where

$$v_i(d_S, t_S, \lambda, \alpha) := \sum_{t_{N \setminus S} \in T_{N \setminus S}} p(t_{N \setminus S}) v_i(d_S, t, \lambda, \alpha),$$

and $v_i(\mu_S, t_S, \lambda, \alpha)$ is the linear extension of $v_i(\cdot, t_S, \lambda, \alpha)$ over μ_S (recall that μ_S is measurable w.r.t. T_S).

Definition 1 (**M-Solution, Myerson, 1984b**). *A feasible mechanism for the grand coalition* $(\bar{\mu}_N, \bar{x}_N) \in \mathcal{F}_N^*$ *is an M-solution if there exist vectors* $\lambda > 0$, $\alpha \geq 0$, *and* $\eta = (\mu_S, x_S)_{S \subseteq N}$ *with* $(\mu_N, x_N) = (\bar{\mu}_N, \bar{x}_N)$ *such that:*

---

[13]In the virtual game, sidepayments are meant to be done in terms of the virtual scales (λ, α). Hence, an appropriate definition for W_S should be $W_S(\mu_S, x_S, t, \lambda, \alpha) = \sum_{i \in S} \left(v_i(\mu_S, t, \lambda, \alpha) + \beta_i(t_i, \lambda, \alpha) x_S^i(t_S) \right)$, where $\beta_i(t_i, \lambda, \alpha) := \left[\lambda_i(t_i) + \sum_{\tau_i \in T_i} \alpha_i(\tau_i \mid t_i) - \sum_{\tau_i \in T_i} \alpha_i(t_i \mid \tau_i) \right] / p(t_i)$. However, the scales (λ, α) are selected endogenously in such a way that the mechanism (μ_N, x_N) satisfies (3.2). Then we can set $\beta_i(t_i, \lambda, \alpha) = 1$ for all $i \in N$ and all $t_i \in T_i$.

(i) (μ_N, x_N) satisfies (3.1)-(3.4) for λ and α.

(ii) η is a vector of rational threats w.r.t. λ and α.

(iii) $U(\mu_N, x_N) = (U_i(\mu_N, x_N \mid t_i))_{t_i \in T_i, i \in N}$ is warranted by λ, α and η.

The vector $U(\mu_N, x_N)$ of warranted claims is called an *M-value*. We denote by $\mathcal{V}_M(\Gamma)$ the set of M-values of Γ.[14]

4.2. The S-Solution

A first component in the definition of the S-solution is the construction of a vector of threats called an *egalitarian solution*. An egalitarian solution requires threats to satisfy a condition of average balanced contributions (principle of equal gains) and differs from the S-solution in that the latter endogenously determines the virtual scales (λ, α) by additionally imposing a utilitarian criterion.

Given a vector of virtual scales (λ, α), a vector of threats $\eta = (\mu_S, x_S)_{S \subseteq N}$ is an *egalitarian solution* (w.r.t. λ and α) if and only if, for all $S \subseteq N$, the mechanism (μ_S, x_S) is an optimal solution to

$$\max_{(v_S, y_S) \in \mathcal{F}_S} \sum_{t \in T} p(t) W_S(v_S, y_S, t, \lambda, \alpha)$$

$$\text{s.t.} \sum_{t_{-i} \in T_{-i}} p(t_{-i}) \sum_{j \in S \setminus i} \left[v_i(v_S, t, \lambda, \alpha) + y_S^i(t_S) - v_i(\mu_{S \setminus j}, t, \lambda, \alpha) - x_{S \setminus j}^i(t_{S \setminus j}) \right]$$

$$= \sum_{t_{-i} \in T_{-i}} p(t_{-i}) \sum_{j \in S \setminus i} \left[v_j(v_S, t, \lambda, \alpha) + y_S^j(t_S) - v_j(\mu_{S \setminus i}, t, \lambda, \alpha) - x_{S \setminus i}^j(t_{S \setminus i}) \right], \quad (4.4)$$

$$\forall t_i \in T_i, \ \forall i \in S.$$

We notice that an egalitarian solution is defined recursively: Given a vector of threats $(\mu_{S \setminus j}, x_{S \setminus j})_{j \in S}$ previously computed solving (4.4), (μ_S, x_S) is determined solving (4.4). When $S = \{i\}$, for some $i \in N$, problem (4.4) reduces to (4.4). The possibility to make unrestricted sidepayments in terms of the virtual utility scales guarantees that this construction is always possible.[15]

In problem (4.4), the objective function is the same as in (4.3). In an egalitarian solution, however, optimal threats are required to be "equitable". Here, equitable means that the expected average virtual contribution of the different players in S to player i of type t_i (in coalition S) equals the expected average virtual contribution of player i to the different players in S as assessed by type t_i (see Section 4 in Salamanca (2016) for a justification of this equity criterion).

---

[14]Definition 1 involves strictly positive utility weights λ. This complicates matters for obtaining existence results of the M-solution. Myerson (1984b) solves this dilemma by slightly enlarging the solution set to include utility allocations that are reasonable as emerging from limit points.

[15]In the absence of sidepayments, the optimization problem in (4.4) may not be feasible. The difficulty is due to a lack of comprehensiveness in the set of attainable virtual utility allocations (see Section 7 in Salamanca, 2016).

Definition 2 (**S-Solution, Salamanca, 2016**). *A feasible mechanism for the grand coalition* $(\bar{\mu}_N, \bar{x}_N) \in \mathcal{F}_N^*$ *is an S-solution if there exist vectors* $\lambda > 0$, $\alpha \geq 0$, *and* $\eta = (\mu_S, x_S)_{S \subseteq N}$ *with* $(\mu_N, x_N) = (\bar{\mu}_N, \bar{x}_N)$, *such that:*

(*i*) (μ_N, x_N) *satisfies* (3.1)-(3.4) *for* λ *and* α.

(*ii*) η *is an egalitarian solution w.r.t.* λ *and* α.

The vector of interim utilities $U(\mu_N, x_N)$ *is called an S-value. We denote by* $\mathcal{V}_S(\Gamma)$ *the set of S-values of* Γ.

4.3. Reconciling the Differences

When we compare the previous solutions, both satisfy the utilitarian conditions (3.1)-(3.4). In addition, for any egalitarian solution $\eta = (\mu_S, x_S)_{S \subseteq N}$, $U(\mu_N, x_N)$ is warranted by λ, α, and η.[16] Hence, the M-solution and the S-solution differ only in the way they determine the threats of intermediate coalitions. However, we will show that this difference disappears in the present model with sidepayments. Formally, we are going to prove the following:

Theorem 1. *Let* Γ *be a Bayesian cooperative game with orthogonal coalitions and sidepayments. Then,* $\mathcal{V}_M(\Gamma) = \mathcal{V}_S(\Gamma)$.

In order to establish this result, we shall construct a particular class of threat mechanisms that will help us to establish a certain connection between the rational threats and the egalitarian solution. This relationship is stated in Lemma 1. The final conclusion of Theorem 1 follows from the double inclusion established in Propositions 2 and 3.

Fix the virtual scales (λ, α) and let $S \subseteq N$ be a coalition. Given a vector of threats $(\mu_R, x_R)_{R \subset S}$, we define

$$r_S^i(t_S, \lambda, \alpha) := \sum_{\substack{R \subset S \\ i \in R}} (-1)^{|S \setminus R|+1} \left[v_i(\mu_R, t_S, \lambda, \alpha) + x_R^i(t_R) \right], \quad \forall t_S \in T_S. \tag{4.5}$$

The quantity $r_S^i(t_S, \lambda, \alpha)$ can be thought of as the cumulated "virtual dividends" that player $i \in S$ expects in state t_S from his participation in all coalitions $R \subset S$ to which he belongs.

Given the vector $r_S(\lambda, \alpha) = (r_S^i(t_S, \lambda, \alpha))_{i \in S, \, t_S \in T_S}$, consider a threat mechanism $(\bar{\mu}_S, \bar{x}_S)$ for coalition S defined by[17]

$$\sum_{i \in S} v_i(\bar{\mu}_S, t_S, \lambda, \alpha) = \sum_{i \in S} v_i^*(t_S, \lambda, \alpha), \quad \forall t_S \in T_S, \tag{4.6a}$$

$$\bar{x}_S^i(t_S) = v_i^*(t_S, \lambda, \alpha) - v_i(\bar{\mu}_S, t_S, \lambda, \alpha), \quad \forall i \in S, \ \forall t_S \in T_S, \tag{4.6b}$$

where $v^*(\lambda, \alpha) = (v_i^*(t_S, \lambda, \alpha))_{i \in S, t_S \in T_S}$ is the solution to

$$v_i^*(t_S, \lambda, \alpha) - r_S^i(t_S, \lambda, \alpha) = v_j^*(t_S, \lambda, \alpha) - r_S^j(t_S, \lambda, \alpha), \quad \forall i, j \in S, \ \forall t_S \in T_S, \tag{4.7a}$$

---

[16]See Remark 3 in Salamanca (2016).

[17]It is worth noting that $(\bar{\mu}_S, \bar{x}_S)$ is not *uniquely* determined by (4.6a)-(4.6b). Indeed, there may be several random joint decisions μ_S satisfying (4.6a). Yet, once μ_S is determined, there exists a unique $\bar{x}_S$ satisfying (4.6b) and (4.7a)-(4.7b).

$$\sum_{i\in S} v_i^*(t_S, \lambda, \alpha) = \max_{d_S \in D_S} \sum_{i\in S} v_i(d_S, t_S, \lambda, \alpha), \quad \forall t_S \in T_S. \tag{4.7b}$$

The system of linear equations in (4.7a)-(4.7b) is always solvable, and its solution is unique; hence $v_i^*(\lambda, \alpha)$ is well defined.[18] Notice also that for each $t_S \in T_S$, the transfers $\bar{x}_S$ are exactly balanced

$$\sum_{i\in S} \bar{x}_S^i(t_S) = \sum_{i\in S} v_i^*(t_S, \lambda, \alpha) - \sum_{i\in S} v_i(\bar{\mu}_S, t_S, \lambda, \alpha) = 0.$$

Lemma 1. *Let (λ, α) and $(\mu_R, x_R)_{R\subset S}$ be fixed. For a given coalition $S \subseteq N$, let $(\bar{\mu}_S, \bar{x}_S)$ be defined by (4.6a)-(4.6b) and (4.7a)-(4.7b). Then $(\bar{\mu}_S, \bar{x}_S)$ is an optimal solution to (4.3). If, in addition, for every $R \subset S$, (μ_R, x_R) is feasible in (4.4) w.r.t $(\mu_{R\setminus j}, x_{R\setminus j})_{j\in S}$, then $(\bar{\mu}_S, \bar{x}_S)$ is also an optimal solution to (4.4) w.r.t. $(\mu_{S\setminus j}, x_{S\setminus j})_{j\in S}$.*

Proof. The fact that $(\bar{\mu}_S, \bar{x}_S)$ is an optimal solution to (4.3) is straightforward. Let (μ_S, x_S) be an optimal solutions of (4.4) w.r.t. $(\mu_{S\setminus j}, x_{S\setminus j})_{j\in S}$. We notice that

$$\sum_{t_S\in T_S} p(t_S)\left[\sum_{i\in S}\left(v_i(\bar{\mu}_S, t_S, \lambda, \alpha) + \bar{x}_S^i(t_S)\right)\right] = \sum_{t_S\in T_S} p(t_S) \max_{d_S\in D_S} \sum_{i\in S} v_i(d_S, t_S, \lambda, \alpha)$$

$$\geq \sum_{t_S\in T_S} p(t_S)\left[\sum_{i\in S} v_i(\mu_S, t_S, \lambda, \alpha) + x_S^i(t_S)\right].$$

Then it suffices to show that $(\bar{\mu}_S, \bar{x}_S)$ is feasible in (4.4) (w.r.t. $(\mu_{S\setminus j}, x_{S\setminus j})_{j\in S}$). By construction, for any $t_S \in T_S$, $(\bar{\mu}_S, \bar{x}_S)$ satisfies

$$v_i(\bar{\mu}_S, t_S, \lambda, \alpha) + \bar{x}_S^i(t_S) - r_S^i(t_S, \lambda, \alpha) = v_j(\bar{\mu}_S, t_S, \lambda, \alpha) + \bar{x}_S^j(t_S) - r_S^j(t_S, \lambda, \alpha), \quad \forall i, j \in S. \tag{4.8}$$

Let $t_i \in T_i$ be a fixed type of a player $i \in S$. Multiplying both sides of (4.8) by $p(t_{S\setminus i})$, summing over all $t_{S\setminus i} \in T_{S\setminus i}$ and all $j \in S \setminus i$, and rearranging terms yields

$$\sum_{t_{-i}\in T_{-i}} p(t_{-i}) \sum_{j\in S\setminus i} [v_i(\bar{\mu}_S, t, \lambda, \alpha) + \bar{x}_S^i(t_S) - v_i(\mu_{S\setminus j}, t, \lambda, \alpha) - x_{S\setminus j}^i(t_{S\setminus j})]$$

$$- \sum_{t_{-i}\in T_{-i}} p(t_{-i}) \sum_{j\in S\setminus i} [v_j(\bar{\mu}_S, t, \lambda, \alpha) + \bar{x}_S^j(t_S) - v_j(\mu_{S\setminus i}, t, \lambda, \alpha) - x_{S\setminus i}^j(t_{S\setminus i})]$$

$$= \sum_{\substack{R\subset S \\ i\in R \\ |R|\geq 2}} (-1)^{|S\setminus R|}\left[\sum_{t_{-i}\in T_{-i}} p(t_{-i}) \sum_{j\in R\setminus i} [v_j(\mu_R, t, \lambda, \alpha) + x_R^j(t_R) - v_j(\mu_{R\setminus i}, t, \lambda, \alpha) - x_{R\setminus i}^j(t_{R\setminus i})]\right.$$

$$\left. - \sum_{t_{-i}\in T_{-i}} p(t_{-i}) \sum_{j\in R\setminus i} [v_i(\mu_R, t, \lambda, \alpha) + x_R^i(t_R) - v_i(\mu_{R\setminus j}, t, \lambda, \alpha) - x_{R\setminus j}^i(t_{R\setminus j})]\right]. \tag{4.9}$$

If, for every $R \subset S$, (μ_R, x_R) is feasible in (4.4) w.r.t. $(\mu_{R\setminus j}, x_{R\setminus j})_{j\in S}$, then the right-hand side (and therefore also the left-hand side) of (4.9) is zero. This concludes the proof. $\square$

---

[18]Consider the homogeneous system. For each $t_S \in T_S$, the system has exactly $|S|$ linearly independent equations, namely $v_1(t_S) - v_j(t_S) = 0$ for each $j \in S \setminus 1$ and $\sum_{i\in S} v_i(t_S) = 0$. The unique solution is $v_i(t_S) = 0$ for all $i \in S$. Hence, the non-homogeneous system is solvable and its solutions is unique.

The following inclusion readily follows from Lemma 1.

Proposition 2. $\mathcal{V}_M(\Gamma) \supseteq \mathcal{V}_S(\Gamma)$.

Proof. Let (μ_N, x_N) be an S-solution supported by $\eta = (\mu_S, x_S)_{S \subseteq N}$, λ and α. For each $S \subseteq N$, Lemma 1 implies that (μ_S, x_S) attains the optimal value of (4.3). Hence, η is a vector of rational threats.[19] $\square$

We shall now prove the reverse inclusion. The basic idea will be to show that, given an M-solution with corresponding rational threats $\eta = (\mu_S, x_S)_{S \subseteq N}$, one can construct a transfer scheme $(\tilde{x}_S)_{S \subseteq N}$ as in (4.6a)-(4.6b), such that the vector of threats $(\mu_S, \tilde{x}_S)_{S \subseteq N}$ is an egalitarian solution. This is possible thanks to Lemma 1.

Proposition 3. $\mathcal{V}_M(\Gamma) \subseteq \mathcal{V}_S(\Gamma)$.

Proof. Let (μ_N, x_N) be an M-solution supported by $\eta = (\mu_S, x_S)_{S \subseteq N}$, λ, and α. Recursively define for each $S \subset N$

$$\tilde{x}_S^i(t_S) := \tilde{v}_i^*(t_S, \lambda, \alpha) - v_i(\mu_S, t_S, \lambda, \alpha), \quad \forall i \in S, \ \forall t_S \in T_S$$

where $\tilde{v}^*(\lambda, \alpha) = (\tilde{v}_i^*(t_S, \lambda, \alpha))_{i \in S, t_S \in T_S}$ is the solution to (4.7a)-(4.7b) with $\tilde{r}_S(\lambda, \alpha)$ computed using $(\mu_R, \tilde{x}_R)_{R \subset S}$ (already defined in the recursion). Define $\tilde{x}_N = x_N$ and $\tilde{\eta} = (\mu_S, \tilde{x}_S)_{S \subseteq N}$. Obviously, $(\mu_N, \tilde{x}_N)$ is a feasible mechanism satisfying (3.1)-(3.4) w.r.t. λ and α, and $U(\mu_N, \tilde{x}_N) = U(\mu_N, x_N)$. Hence, we only need to show that $\tilde{\eta}$ is an egalitarian solution w.r.t. λ and α. Notice that, for each $S \subset N$, $(\mu_S, \tilde{x}_S)$ satisfies (4.6a)-(4.6b) and (4.7a)-(4.7b) with $(\mu_R, \tilde{x}_R)_{R \subset S}$. Then Lemma 1 (applied inductively) implies that, for each $S \subset N$, $(\mu_S, \tilde{x}_S)$ is an optimal solution of (4.4) w.r.t. $(\mu_{S \setminus j}, \tilde{x}_{S \setminus j})_{j \in S}$. It remains to show that $(\mu_N, \tilde{x}_N)$ is an optimal solution of (4.4) w.r.t. $(\mu_{N \setminus j}, \tilde{x}_{N \setminus j})_{j \in N}$. To do this, we first notice that, since $\tilde{x}_S$ is exactly balanced for all subcoalitions, then $W_S(\mu_S, x_S, t, \lambda, \alpha) = W_S(\mu_S, \tilde{x}_S, t, \lambda, \alpha)$ for every $S \subseteq N$. Therefore, $U(\mu_N, \tilde{x}_N)$ is warranted by λ, α, and $\tilde{\eta}$.[20] Remark 3 in Salamanca (2016) then implies that $(\mu_N, \tilde{x}_N)$ is feasible in (4.4) w.r.t. $(\mu_{S \setminus j}, \tilde{x}_{S \setminus j})_{j \in S}$. But,

$$\sum_{i \in N} \left(v_i(\mu_N, t, \lambda, \alpha) + \tilde{x}_N^i(t) \right) = \max_{d \in D} \sum_{i \in N} v_i(d, t, \lambda, \alpha), \quad \forall t \in T.$$

Therefore, as required, $(\mu_N, \tilde{x}_N)$ solves (4.4) w.r.t. $(\mu_{N \setminus j}, \tilde{x}_{N \setminus j})_{j \in N}$. We conclude that $\tilde{\eta}$ is an egalitarian solution w.r.t. λ and α. $\square$

Theorem 1 states that, in our model with sidepayments, the M-solution and the S-solution are interim utility equivalent. Moreover, as can be deduced from the proof of Propositions 2 and 3, any M-solution is an S-solution and vice-versa.[21] Henceforth, a cooperative solution will simply be called an *MS-solution*.

Notice that if Γ is a game with complete information (i.e., T_i is a singleton for every $i \in N$), there are no incentive constraints (or equivalently $\alpha = 0$), and consequently (3.2) implies that $\lambda_i = 1$ for every $i \in N$. Hence, all efficient mechanisms are supported by the same utility weights $\lambda_i = 1$, which means that the Pareto efficient frontier is thus characterized by an hyperplane.

---

[19]Indeed, we must have that $\sum_{i \in S} x_S^i(t_S) = 0$ for every $t_S \in T_S$, since $\sum_{t_S \in T_S} p(t_S) \sum_{i \in S} x_S^i(t_S) = 0$ and $\sum_{i \in S} x_S^i(t_S) \leq 0$ for all $t_S \in T_S$.

[20]In particular, $(\mu_N, \tilde{x}_N)$ is also an M-solution supported by $\tilde{\eta}$, λ and α.

[21]Notice, however, that Definitions 1 and 2 are not equivalent: An optimal solution to (4.3) is not necessarily also an optimal solution to (4.4), unless an appropriate transfers scheme is used (see proof of Proposition 3).

From Definition 1, it then follows that an MS-value coincides with the Shapley TU value. An MS-value is thus a valid generalization of the Shapley TU value to Bayesian cooperative games with sidepayments.

Corollary 1. *Let Γ be a Bayesian cooperative game with orthogonal coalitions, sidepayments and complete information. Denote as (N, w) the characteristic function game associated with Γ, i.e., $w(S) = \max \left\{ \sum_{i \in S} U_i(\mu_S) \mid \mu_S \in \Delta(D_S) \right\}$ for every $S \subseteq N$. If (μ_N, x_N) is an MS-solution of Γ, then $U(\mu_N, x_N)$ equals the Shapley TU value of (N, w). Conversely, let $\phi_i(N, w)$ denote the Shapley TU value of player i in (N, w). Then there exists an MS-solution of Γ, (μ_N, x_N), such that $U_i(\mu_N, x_N) = \phi_i(N, w)$ for every $i \in N$.*

Proof. Because $\alpha = 0$ and $\lambda_i = 1$ for every $i \in N$, the rational threats criterion (see *(ii)* in Definition 1) implies that the transferable virtual utility game coincides with the characteristic function game (N, w). Hence, the warranted claims (see *(iii)* in Definition 1) specializes to $U_i(\mu_N, x_N) = \phi_i(N, w)$. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\square$

Under incomplete information, however, the hydraulic equations in (3.2) only imply that

$$\sum_{t_i \in T_i} \lambda_i(t_i) = \sum_{t_i \in T_i} p(t_i) = 1, \quad \forall i \in N,$$

and therefore $\lambda_i(t_i) > 0$ for some t_i for each player $i \in N$, but we still have $\sum_{i \in N} (|T_i| - 1)$ degrees of freedom for choosing λ. With incomplete information, the restricted nature of the utility transfers implies that the interim incentive-efficient frontier is not generally a hyperplane. This indeterminacy makes utility transfers less useful for games with incomplete information than they are for games with complete information. In particular, it makes impossible to get a general and simple closed form expression for the MS-value as in the case of the Shapley value for TU games with complete information.

Example 1. Consider the following Bayesian cooperative game introduced by Salamanca (2017). Let r be a parameter with $0 < r < 1/2$. The set of players is $N = \{1, 2, 3\}$. Only player 1 has private information about one of two possible states in $T = \{H, L\}$, with prior probabilities $p(H) = 1 - p(L) = 4/5$. Feasible decisions for each coalition are: $D_{\{i\}} = \{d_i\}$ $(i = 1, 2, 3)$, $D_{\{i,j\}} = \{[d_i, d_j], d_{ij}\}$ $(i \neq j)$ and $D_N = [D_{\{1,2\}} \times D_{\{3\}}] \cup [D_{\{1,3\}} \times D_{\{2\}}] \cup [D_{\{2,3\}} \times D_{\{1\}}]$. For each value of r, utility functions are given by

(u_1, u_2, u_3)	$[d_1, d_2, d_3]$	$[d_{12}, d_3]$	$[d_{13}, d_2]$	$[d_{23}, d_1]$
H	$(0, 0, 0)$	$(50, 50, 0)$	$(100r, 0, 100(1-r))$	$(0, 100r, 100(1-r))$
L	$(0, 0, 0)$	$(40, 40, 0)$	$(40r, 0, 40(2-r))$	$(0, 40r, 40(2-r))$

Decision $[d_1, d_2, d_3]$ denotes the noncooperative outcome which leaves every player with his reservation utility (normalized to zero), whereas $[d_{ij}, d_k]$ denotes the cooperative outcome in which players i and j form a coalition (leaving player k alone) and they share the proceeds from cooperation as specified above.

Let us first consider the game in which sidepayments are *not* allowed; that is, we set $x_S^i(t_S) = 0$ for all $i \in S$, $t_S \in T_S$ and $S \subseteq N$. A mechanism for a coalition S is thus only composed by a

state-dependent lottery $\mu_S : T_S \to \Delta(D_S)$.[22] In such a situation, irrespective of the state, player 3 can only offer players 1 and 2 a payoff that is strictly lower than what they both can get by acting together in coalition $\{1, 2\}$ (lack of transferability). Hence, a reasonable outcome for this game should reward player 3 strictly less than players 1 and 2. On the other hand, because the incentives of player 1 are fully aligned in both states, incentive constraints are not essential (i.e., we can set $\alpha = 0$). This implies that player 1 has no information rents, and therefore, players 1 and 2 must be treated symmetrically. In addition, as r goes to zero, player 3's ability to transfer payoffs decreases; thus his reward should also reduce.

For any given r, the unique M-value of this game is the interim allocation[23]

$$\left(U_1^H, U_1^L, U_2, U_3 \right) = \left(\tfrac{100}{3}, \tfrac{80}{3}, 32, 32 \right). \tag{4.10a}$$

We notice that, regardless of the value of r, the M-value treats all players symmetrically, as if utility were fully transferable.

On the other hand, as shown by Salamanca (2017) the unique S-value of this game is[24]

$$\left(U_1^H, U_1^L, U_2, U_3 \right) = \left(50 - \tfrac{100}{3} r \left(\tfrac{88-88r}{96-88r} \right), 40 - \tfrac{80}{3} r \left(\tfrac{88-44r}{96-88r} \right), 48 - \tfrac{88}{3} r, \tfrac{176}{3} r \right). \tag{4.10b}$$

The arguments presented above are better reflected by this allocation. In particular, as r decreases to 0, the S-value converges to $(50, 40, 48, 0)$.

Let us now turn to the situation in which players are allowed to transfer payoffs using state-contingent sidepayments. In this case, because incentive constraints are not essential, utility is fully transferable.[25] Therefore, in any two-person coalition with player 3, the members can agree on an arbitrary distribution of the proceeds. In particular, if coalition $\{1, 3\}$ or $\{2, 3\}$ forms, its members can get half of the gains each, something that is not possible in the absence of sidepayments. There is no puzzle in the present case; the three players are symmetric, and thus the allocation (4.10a) is a reasonable outcome. In fact, (4.10a) is the unique MS-value of the game with sidepayments.

5. Values for Two-person Bayesian Games

In this section, we shall study in detail both the M-solution and the S-solution in two-player games ($n = 2$). In order to explore all the particularities of this case, we first consider games

---

[22]When sidepayments are not allowed, the hydraulic equations in (3.2) are removed. Moreover, condition (3.4) is automatically satisfied.

[23]This value is supported by the vectors $\lambda = (4/5, 1/5, 1, 1)$ and $\alpha = (0, 0)$.

[24]This value allocation is supported by the vectors $\lambda = (4/5, 1/5, 1, 1)$ and $\alpha = (0, 0)$. An important remark about the computation of the S-value in this example is in place. As already mentioned in footnote 15, in the absence of sidepayments an egalitarian solution may fail to exist. To deal with this issue in the present example, Salamanca (2017) considers a class of mechanisms that enlarge the set of attainable (virtual) utility allocations by allowing players to discard utility (*free disposal assumption*). These mechanisms, however, do not have an impact on the incentives structure of the game or on the M-value allocation.

[25]Since $\alpha = 0$, condition (3.2) then implies that $\lambda_i(t_i) = p(t_i)$ for all $t_i \in T_i$ of every player $i \in N$. Real and virtual utilities coincide; hence, the interim Pareto frontier is a hyperplane.

with orthogonal coalitions in which sidepayments are not allowed (as in Example 1). We then extend the analysis to allow for both strategic and informational externalities. In this latter situation, we investigate an additional semi-cooperative solution introduced by Kalai and Kalai (2013) under the further assumptions of sidepayments and ex-post verifiable information.

In the previous section we showed that in our model with orthogonal coalitions and sidepayments, the M-solution and the S-solution are interim utility equivalent. The same conclusion also holds in the absence of sidepayments, provided that $n = 2$. Indeed, when there are only two players, all intermediate coalitions correspond to singletons $\{i\}$ ($i \in \{1, 2\}$). For any of those coalitions, the optimization problem in (4.3) is the same as in (4.4), irrespective of whether sidepayments are allowed or not. Therefore, it follows immediately that Definitions 1 and 2 are equivalent.[26] This reasoning is summarized in the following proposition (which already appeared in Salamanca, 2016).

Proposition 4. *Let Γ be a two-player Bayesian cooperative game with orthogonal coalitions (with or without sidepayments). Then, the M-solution and the S-solution for Γ coincide.*

Thus, we may continue to call a M(or S)-solution simply a MS-solution.

Let Γ be a two-person cooperative game with orthogonal coalitions. Suppose that, for every $i \in \{1, 2\}$, $D_i = \{d_i\}$, and $u_i(d^*, t) = 0$ for all $t \in T$, where $d^* := [d_1, d_2](= D_1 \times D_2 \subseteq D_N)$ is called the *disagreement decision*. To make this game interesting, we further assume that there exists at least one joint decision in D_N that is beneficial for both individuals. Such a game is called a *two-person bargaining problem with incomplete information*. Clearly, the issue of coalitional threats does not arise in this model. For this kind of problems, Myerson (1984a) defined a *neutral bargaining solution*, which generalizes Nash's (1950) bargaining solution with fixed threats. The next result follows from the characterization theorem 5 in Myerson (1984a).[27]

Proposition 5. *Let Γ be a two-person bargaining problem with incomplete information. Then the MS-solution and the neutral solution for Γ coincide.*

In the subsequent analysis we allow for strategic externalities; that is, we let the utility of every player depend on the choices of the other individual. We will represent the underlying game situation by a *non-cooperative* Bayesian game in strategic form, $\Gamma = \{N, (T_i, D_i, u_i)_{i \in N}, p\}$, where $N = \{1, 2\}$. All components of Γ have the same interpretation as in the model of Section 4, but the decision set D is now defined to be $D = D_1 \times D_2$. For this specific model, we extend the MS-solution. We follow the method developed by Nash (1953) in the proposal to modify his bargaining solution (Nash, 1950) by making the threat point endogenous. The generalization presented here is not new. Indeed, Myerson (1984b) offers a more general approach to the case $n \geq 2$ without sidepayments.

Let (λ, α) be some fixed virtual scales and consider the virtual game in which players make interpersonal utility comparison in the scales (λ, α). As in the previous section, before entering into negotiations within the grand coalition, each player i commits to a threat strategy (mecha-

---

[26]Recall from Section 4 that the M-solution and the S-solution differ only in the way they determine the threats for intermediate coalitions.

[27]See also Myerson (1991, sec. 10).

nism) $\mu_i : T_i \to \Delta(D_i)$ to be used in case the players fail to reach a cooperative agreement.[28] Then $v_i((\mu_i, \mu_j), t, \lambda, \alpha)$ is the disagreement virtual payoff to player $i \in \{1, 2\}$ in state $t \in T$. The total transferable virtual utility in state t available to both players if they cooperate is

$$w_{12}(t, \lambda, \alpha) := \max_{d \in D} (v_1(d, t, \lambda, \alpha) + v_2(d, t, \lambda, \alpha)).$$

Therefore, the Shapley TU value (or equivalently, the Nash bargaining solution) of player i in the virtual game conditional on state t is

$$\phi_i(W((\mu_1, \mu_2), t, \lambda, \alpha)) = \frac{1}{2} w_{12}(t, \lambda, \alpha) + \frac{1}{2} \left[v_i((\mu_i, \mu_j), t, \lambda, \alpha) - v_j((\mu_i, \mu_j), t, \lambda, \alpha) \right]. \quad (5.1)$$

After rearranging terms on the right-hand side, this equation simply states that the total virtual surplus "$w_{12} - v_i - v_j$" is equally divided among the two players.

Suppose now that players expect to reach a cooperative agreement (μ_N, x_N) giving every type of each player his conditionally expected Shapley TU value from the virtual game. Then the final interim utility allocation, $U(\mu_N, x_N) = (U_i(\mu_N, x_N \mid t_i))_{t_i \in T_i, i \in N}$, must be warranted by λ, α, and (μ_i, μ_j); that is,[29]

$$\left(\lambda_i(t_i) + \sum_{\tau_i \in T_i} \alpha_i(\tau_i \mid t_i) \right) U_i(\mu_N, x_N \mid t_i) - \sum_{\tau_i \in T_i} \alpha_i(t_i \mid \tau_i) U_i(\mu_N, x_N \mid \tau_i) =$$

$$\sum_{t_j \in T_j} p(t) \phi_i(W((\mu_1, \mu_2), t, \lambda, \alpha)), \quad \forall t_i \in T_i, \quad i = 1, 2. \quad (5.2)$$

It must be the case that players should evaluate their threats only in terms of their impact on the payoffs granted by the final agreement (μ_N, x_N). By Lemma 1 in Myerson (1983), the final warranted payoffs (weakly) increase as the right-hand side of (5.2) increases. Then each player i should want to choose his threat strategy μ_i so as to maximize his expected Shapley TU value in the virtual game. From equation (5.1) we have that player i's virtual Shapley TU value increases as the disagreement virtual payoff to player j decreases. Therefore, for a given threat strategy μ_j, an optimal threat strategy for player i should solve

$$\max_{\mu_i \in \mathcal{F}_i} \sum_{t \in T} p(t) \left(v_i((\mu_i, \mu_j), t, \lambda, \alpha) - v_j((\mu_i, \mu_j), t, \lambda, \alpha) \right).$$

That is, optimal threats form an equilibrium of a two-person Bayesian zero-sum game in which the players' utility functions are given by

$$w_i((\mu_i, \mu_j), t, \lambda, \alpha) := v_i((\mu_i, \mu_j), t, \lambda, \alpha) - v_j((\mu_i, \mu_j), t, \lambda, \alpha). \quad (5.3)$$

We let $\Gamma^{\text{ad}}(\lambda, \alpha)$ denote this zero-sum game. The superscript "ad" stands for "advantage", since the utility scale $w_i(\cdot, \lambda, \alpha)$ is a measure of the payoff advantage of player i in the virtual game.[30]

---

[28] The use of transfers inside a single person coalition is unnecessary for the coalition member.

[29] Myerson (1984a, sec. 9) provides a rationale for equation (5.2).

[30] Myerson (1991, sec. 8) refers to $\Gamma^{\text{ad}}(\lambda, \alpha)$ as the *difference game*. We keep the designation "advantage" in order to be consistent with Kalai and Kalai's (2013) terminology.

Definition 3 (**MS-Solution**). *Let Γ be a two-player non-cooperative Bayesian game with side-payments. A feasible mechanism for the grand coalition (μ_N, x_N) is an MS-solution of Γ if there exist vectors $\lambda > 0$, $\alpha \geq 0$, and threat strategies (μ_1, μ_2) such that:*

(*i*) *(μ_N, x_N) satisfies (3.1)-(3.4) for λ and α.*

(*ii*) *(μ_1, μ_2) is a profile of minmax strategies of $\Gamma^{ad}(\lambda, \alpha)$.*

(*iii*) *The vector $U(\mu_N, x_N)$ of interim utilities satisfies (5.2) w.r.t. λ, α, and (μ_1, μ_2).*

The vector $U(\mu_N, x_N)$ is called an MS-value. In the case that sidepayments are not allowed, so that $x_N^i(t) = 0$ for all $i \in \{1, 2\}$ and $t \in T$, conditions (3.2) and (3.4) are removed from (i).

It is worth noticing that Definition 3 generalizes Nash's (1953) bargaining solution with variable threats (cf. Myerson, 1991, sec. 8).

Recently, A. Kalai and E. Kalai (2013) proposed a semi-cooperative solution for the class of two-player non-cooperative Bayesian games with sidepayments. Their solution, called the *cooperative-competitive* (or "*coco*") *value*, is based on a decomposition of the game Γ into cooperative and competitive components. Let Γ^{ad} be the Bayesian zero-sum game defined as Γ, except that player i's utility function is now $u_i - u_j$. In a similar fashion, we define the game Γ^{eq}, which differs from Γ in that both players share the same utility function given by $u_i + u_j$. The superscript "eq" stands for "equal" payoffs. Clearly, $u_i = \frac{1}{2}(u_i + u_j) + \frac{1}{2}(u_i - u_j)$. Thus, the game Γ^{eq} describes the *cooperative* component of Γ in which the interests of both players are aligned. On the other hand, the game Γ^{ad} reflects each player's *competitive* advantage in Γ. Incentives to reveal private information are opposed in both component games. While in Γ^{eq} the obvious incentive is to truthfully disclose any private information, in Γ^{ad} each player would like to prevent any information disclosure that increases the opponent's payoff advantage.

The *team optimum* of Γ^{eq} is defined by

$$Opt(\Gamma^{eq}) = \sum_{t \in T} p(t) \max_{d \in D} \sum_{i \in N} u_i(d, t).$$

In words, the team optimum is the maximum expected utility that each player can get in Γ^{eq} when they all share their information truthfully.

The game Γ^{ad} is a zero-sum game in which each player is trying to maximize the difference of his payoff and that of his opponent. This game has a unique minmax (ex-ante) expected value denoted $Val_i(\Gamma^{ad})$.

Definition 4 (**Coco Value, Kalai and Kalai, 2013**). *Let Γ be a two-player non-cooperative Bayesian game with sidepayments. The coco value of player i in Γ, denoted $\kappa_i(\Gamma)$, is defined as*

$$\kappa_i(\Gamma) = \frac{1}{2}\left(Opt(\Gamma^{eq}) + Val_i(\Gamma^{ad})\right).$$

The MS-solution differs significantly from the coco value. On the one hand, the coco value assumes that the incentives when playing each of the component games are independent of each other, and thus it ignores possible incentive compatibility issues. As a consequence, the

coco value is ex-post (first best) efficient. In contrast, the MS-solution takes into account the tension that may exist between ex-post efficiency and incentive compatibility (see Myerson and Satterthwaite, 1983). On the other hand, the coco value is defined only at the ex-ante stage (i.e., before any player learns his private information). Consequently, the interim evaluation of the coco value cannot be well determined without specifying an implementation protocol at the interim stage. Finally, in the MS-solution players bargain over mechanisms and not directly over payoffs as in the coco value. Notwithstanding the above, there is a close relationship between these two solution concepts provided that incentive constraints are not essential for the implementation of the final agreement.

The simplest bargaining situation in which incentive constraints are not essential is when all private information becomes publicly verifiable at the implementation stage. In principle, therefore, any budget-feasible mechanism in $\mathcal{F}_N$ can be enforced once it is agreed upon. We summarize this situation in the following assumption introduced by de Clippel and Minelli (2004):

Definition 5 (**Verifiable Types**). *A Bayesian cooperative game* Γ *satisfies the verifiable types assumption if, when agreements are implemented, each individual can costlessly verify the true information state.*[31]

The verifiable types assumption is satisfied in games where private information relates to signals about states that eventually become public; for instance, differential forecasts about weather conditions, as in Kalai and Kalai's (2013) vendors example. This assumption is particularly strong in the sense that it rules out some game situations, such as auctions or trading problems where payoffs depend on valuations that are not observable ex-post. Kalai and Kalai's (2013) noncooperative interim implementation of the coco value heavily relies on this observability assumption, which severely limits its practical applicability.

Let (μ_N, x_N) be an MS-solution supported by the virtual scales (λ, α) and the minmax strategies (μ_1, μ_2). Mathematically, the verifiable types assumption is equivalent to setting $\alpha = 0$, since incentive constraints can be dispensed. Under such circumstances, equation (3.1) is straightforwardly satisfied. Condition (3.2) becomes $\lambda_i(t_i) = p(t_i)$ for all $t_i \in T_i$ and $i \in N$. Hence, virtual utilities and real utilities coincide. Conditions (3.3) and (3.4) become:

$$\sum_{i\in N} v_i(\mu_N, t, \lambda, \alpha) = w_{12}(t, \lambda, \alpha) = \max_{d\in D} \sum_{i\in N} u_i(d, t) \quad \text{and} \quad \sum_{i\in N} x_N^i(t) = 0, \quad \forall t \in T.$$

On the other hand, the zero-sum game $\Gamma^{\mathrm{ad}}(\lambda, \alpha)$ coincides with the competitive component game Γ^{ad}. It then follows that

$$\sum_{t_j\in T_j} p(t_j) \left(u_i((\mu_i, \mu_j), t) - u_j((\mu_i, \mu_j), t) \right) = Val_i(\Gamma^{\mathrm{ad}}), \quad \forall i \in N.$$

---

[31]In our model, the entire game structure, Γ, is commonly known to all players, so that each player knows how much each player would have received in every information state given any combination of strategies. Therefore, the states of nature are completely determined by the information states. As a consequence, in the present context, the verifiable types assumption is equivalent to Kalai and Kalai's (2013) *(unrestricted) revealed-payoff assumption.*

Therefore, the warrant equations (5.2) simplify to

$$U_i(\mu_N, x_N \mid t_i) = \frac{1}{2} \sum_{t_j \in T_j} p(t_j) \max_{d \in D} \sum_{k \in N} u_k(d, t) + \frac{1}{2} Val_i(\Gamma^{ad}), \quad \forall t_i \in T_i, \forall i \in N. \qquad (5.4)$$

Taking expectations on both sides of (5.4) w.r.t. the different types of player i we obtain:

$$\sum_{t_i \in T_i} p(t_i) U_i(\mu_N, x_N \mid t_i) = \kappa_i(\Gamma), \quad i \in N.$$

The following result is deduced:

Theorem 2. *Let Γ be a two-player Bayesian non-cooperative game with sidepayments satisfying the verifiable types assumption. Then the MS-solution and the coco value of Γ are ex-ante utility equivalent.*

Notice that each player might have multiple minmax strategies in the game $\Gamma^{ad}(= \Gamma^{ad}(\lambda, \alpha))$. The multiplicity of equilibria in the competitive zero-sum component game may lead to different interim (i.e., type-conditional) vector payoffs.[32] Therefore, there might exist various MS-values, of course, all of them guaranteeing the same ex-ante expected payoffs. This indeterminacy of the interim payoffs makes difficult to get an appropriate definition of the "interim" coco value.

The following bilateral trading problem, adapted from Myerson (1985, 2007), analyzes how the MS-solution relates to Myerson and Satterthwaite's (1983) result about the impossibility to achieve an ex-post efficient trade. In particular, it illustrates a situation in which the applicability of the coco value is limited by its assumptions.

Example 2. There are two players $N = \{1, 2\}$. Player 1 is the seller of a single good whose quality may be high (H) with probability $p = 1/2$ or low (L) with probability $1 - p = 1/2$. Player 2 is the only potential buyer. The quality can only be observed by the seller; that is, $T_1 = \{H, L\}$ is the set of types of player 1. If the quality of the good is L, then the value of the good is \$1 to player 1 and \$2 to player 2. If the quality is H then the value of the good is \$5 to player 1 and is \$6 to player 2. Players must decide whether to trade (T) or not (NT) and at which price. Utilities from no-trade are normalized to be 0. The utility functions are:

	L				H	
u_1, u_2	T	NT		u_1, u_2	T	NT
T	$-1, 2$	$0, 0$		T	$-5, 6$	$0, 0$
NT	$0, 0$	$0, 0$		NT	$0, 0$	$0, 0$

For the game under consideration, the verifiable types assumption states that the quality of the good becomes publicly observable at the execution of the contract. In particular, it implies that if trade does not occur, the buyer would still observe the quality of the good. This might be considered not reasonable in this particular game situation. For the sake of completeness, we

---

[32]See the payoff matrices proposed in Section IV.C. of Kalai and Kalai (2013).

shall compute the MS-solution under the verifiable types assumption (and a fortiori also the coco value).

The cooperative component game, Γ^{eq}, has a team optimum equal to 1, whereas the competitive component has a minmax of value 0. Therefore, the coco value of this game is $\kappa_1(\Gamma) = \kappa_2(\Gamma) = \frac{1}{2}$. It can be easily checked that the conditional minmax value is also 0 in both states. Thus, the unique MS-value is the (ex-post efficient) interim allocation

$$(U_1^H, U_1^L, U_2) = \left(\tfrac{1}{2}, \tfrac{1}{2}, \tfrac{1}{2}\right)$$

This allocation is implemented by the mechanism

$$\mu_N((T,T) \mid H) = \mu_N((T,T) \mid L) = 1, \quad x_N^1(H) = \tfrac{11}{2} = -x_N^2(H), \quad x_N^1(L) = \tfrac{3}{2} = -x_N^2(L).$$

That is, trade occurs in both states at a price that gives each player an equal gain from the trade. This is an efficient and fair solution, yet it cannot be implemented whenever information is not verifiable. Indeed, this mechanism is not incentive compatible: Type L will always gain by reporting that the good is of high quality.

Let us now consider the more reasonable situation in which the quality of the good is not verifiable before consumption takes place. It can be shown that all incentive-efficient mechanisms in this example satisfy Proposition 1 for the virtual scales (see Myerson, 1985):

$$\lambda_1^H = \tfrac{5}{8}, \quad \lambda_1^L = \tfrac{3}{8}, \quad \lambda_2 = 1,$$

$$\alpha_1(L \mid H) = 0, \quad \alpha_1(H \mid L) = \tfrac{1}{8}.$$

The virtual utility game is thus described by the following payoff matrices:

v_1, v_2	T	NT
T	$-1,2$	$0,0$
NT	$0,0$	$0,0$

L

v_1, v_2	T	NT
T	$-6,6$	$0,0$
NT	$0,0$	$0,0$

H

The only difference between virtual utilities and actual utilities is for player 1's type H, for which the virtual value of the good is 6. The advantage game, $\Gamma^{ad}(\lambda, \alpha)$, has a conditional minmax value that equals 0 in both states. Hence, the unique solution to the warrant equations in (5.2) is

$$(U_1^H, U_1^L, U_2) = \left(\tfrac{1}{10}, \tfrac{1}{2}, \tfrac{1}{4}\right). \tag{5.5}$$

The virtually equitable and incentive-efficient allocation in (5.5) is the unique MS-value of this game. It is implemented by the mechanism

$$\begin{aligned} \mu_N((T,T) \mid H) &= \tfrac{1}{10}, \quad x_N^1(H) = 6 = -x_N^2(H), \\ \mu_N((T,T) \mid L) &= 1, \quad x_N^1(L) = \tfrac{3}{2} = -x_N^2(L). \end{aligned} \tag{5.6}$$

According to this mechanism, if player 1 announces that the quality is L, then with probability 1 they trade at a price \$3/2. Yet if player 1 reports a high quality then, with probability 1/10

19

they trade at a price \$6. In state H no-trade has probability 9/10, but the seller never trades at a price lower than his virtual valuation of the good, which guarantees that he extracts the whole surplus from trade.

At this point, it should be noted that the unresolvable tension between incentives and ex-post efficiency in Myerson and Satterthwaite's (1983) impossibility theorem holds only if we insist on (interim) individually rational outcomes. In the current model, individual rationality does not appear as an explicit constraint in the set of feasible mechanisms. However, it does appear as an indirect restriction imposed by the definition of the MS-solution. Indeed, the MS-solutions are individually rational (see theorem 3 in Myerson (1984b)).[33] To illustrate this issue, consider again the incentive compatible mechanism in (5.6). Notice that if the actual quality of the good is high, and of course type H is telling the truth, there is a chance of no-trade. This efficiency loss reduces type H's expected utility from cooperation, which tightens individual rationality. Therefore, the whole surplus from cooperation must be given to type H whenever trade occurs.

Some final comments. The coco value has been constructed with the aim of being a semi-cooperative solution with a tractable definition. Its conceptual significance, however, depends on restrictive conditions, such as ex-post verifiability of private information. Also, the fact that it is conceived only for two-person games is an important requirement for obtaining its description by a simple closed form expression. Theorem 2 exhibits why the relaxation of these restrictions cannot be easily done while preserving the tractable nature of the coco value.[34] Firstly, keeping track of the incentive constraints requires referring to an appropriate interim framework. In particular, bargaining directly over payoffs, as in the coco value, does not allow us to determine what the utility of a player would be when he reports a different type. In contrast, incentive-compatible mechanisms include all equilibria that can be achieved by any communication system, and so this set can be viewed as the feasible set. Moreover, interim preferences matter when defining a suitable notion of an equitable cooperative agreement under incomplete information. As Myerson (1983, 1984a,b) argues, identifying a cooperative solution among the set of incentive-efficient mechanisms requires the definition of some principles for equitable compromises, not only among the different players (as in the case of complete information) but also among the different possible types that a player pretends to be. Secondly, even under the verifiable types assumption, there is no a unique way of extending the coco value to games with $n > 2$. Myerson's (1984b) approach provides a reference in this direction. However, other cooperative solutions can be constructed to achieve this goal. For instance, if we proceed as in Harsanyi's (1963) solution for n-person strategic games with complete information, the S-solution can be straightforwardly generalized to allow for both strategic and informational externalities.

---

[33]Even though we could restrict our attention to individually rational mechanisms, we cannot apply Myerson and Satterthwaite's (1983) result to our example, simply because they require a continuum of types.

[34]We see Theorem 2 as discouraging evidence against the following argument advanced by Kalai and Kalai (2013, p. 962) regarding a possible extension of the coco value to general environment with no verifiable information: "It seems also plausible that such a generalized coco value would have a natural and easy to interpret closed-form expression, defined through an appropriate modification of the coco decomposition".

6. References

1. Aumann, R. (1960). "Linearity of unrestrictedly transferable utilities". *Naval Research Logistics Quarterly*, 7, pp. 281-284.

2. d'Aspremont, C. and Gérard-Varet, L.-A. (1979). "Incentives and incomplete information". *Journal of Public Economics*, 11, pp. 25-54.

3. d'Aspremont, C. and Gérard-Varet, L.-A. (1982). "Bayesian incentive compatible beliefs". *Journal of Mathematical Economics*, 10, pp. 83-103.

4. de Clippel, G. and Minelli, E. (2004). "Two-person bargaining with verifiable information". *Journal of Mathematical Economics*, 40, pp. 799-813.

5. Forges, F., Mertens, J.-F. and Vohra, R. (2002). "The ex ante incentive compatible core in the absence of wealth effects". *Econometrica*, 70, pp. 1865-1892.

6. Forges, F. and Serrano, R. (2013). "Cooperative games with incomplete information: Some open problems". *International Game Theory Review*, 15, pp. 1340009-1-17.

7. Harsanyi, J. (1963). "A simplified bargaining model for the *n*-person cooperative game". *International Economic Review*, 4, pp. 194-220.

8. Hart, S. (1985). "An axiomatization of Harsanyi's nontransferable utility value". *Econometrica*, 53, pp. 1295-1313.

9. Kalai, A. and Kalai, E. (2013). "Cooperation in strategic games revisited". *The Quarterly Journal of Economics*, 128, pp. 917-966.

10. Myerson, R. (1983). "Mechanism design by an informed principal". *Econometrica*, 51, pp. 1767-1797.

11. Myerson, R. (1984a). "Two-person bargaining problems with incomplete information". *Econometrica*, 52, pp. 461-488.

12. Myerson, R. (1984b). "Cooperative games with incomplete information". *International Journal of Game Theory*, 13, pp. 69-96.

13. Myerson, R. (1985). Analysis of two bargaining problems with incomplete information. In: *Game Theoretic Models of Bargaining*. A.E. Roth (Ed.). Cambridge University Press, Cambridge, pp. 115-147.

14. Myerson, R. (1991). *Game Theory: Analysis of Conflict*, Harvard University Press.

15. Myerson, R. (1992). Fictitious-transfer solutions in cooperative game theory. In: *Rational Interaction*. R. Selten (Ed.). Springer, Berlin, Heidelberg, pp. 13-33.

16. Myerson, R. (2007). "Virtual utility and the core for games with incomplete information". *Journal of Economic Theory*, 136, pp. 260-285.

17. Myerson, R. and Satterthwaite, M. (1983). "Efficient mechanisms for bilateral trading". *Journal of Economic Theory*, 29, pp. 265-281.

18. Nash, J. (1950). "The bargaining problem". *Econometrica*, 18, pp. 155-162.

19. Nash, J. (1953). "Two-person cooperative games". *Econometrica*, 21, pp. 128-140.

20. Prescott, E. and Townsend, R. (1984). "General competitive analysis in an economy with private information". *International Economic Review*, 25, pp. 1-20.

21. Salamanca, A. (2016). A generalization of the Harsanyi NTU value to games with incomplete information. HAL Working paper 01579898.

22. SALAMANCA, A. (2017). A comparison of NTU values in a cooperative game with incomplete information. HAL Working paper 01500966.

23. SHAPLEY, L. (1953). A Value for *n*-person Games. In: *Contributions to the Theory of Games, Vol. II.* H. W. Kuhn and A. W. Tucker (Eds.). Princeton University Press, Princeton, pp. 307-318.

24. SHAPLEY, L. (1969). Utility comparisons and the theory of games. In: *La Décision: Agrégration et Dynamique des Ordres de Préférence.* Editions du CNRS, Paris, pp. 251-263.