LEMMA Working Paper

n°2024-06

Coalition certification in games of imperfect information

Emmanuel Lagrée

Université Paris-Panthéon-Assas, LEMMA

LEMMA EA 4442 - 4 rue Blaise Desgoffe 75006 Paris

Coalition certification in games of imperfect information

Emmanuel Lagrée, Lemma, Université Paris-Panthéon-Assas

4 rue Blaise Desgoffes, 75006, Paris, FRANCE

emmanuel.lagree@u-paris2.fr

October 14, 2024

Abstract

I analyse a two-stage game where a group of players have asymmetric information about the coalitional binary game they are playing, in the sense that they know for sure only the worth of the coalitions they belong to. Members of a coalition can certify its worth to the social planner if they unanimously decide to. The social planner then updates her belief and shares the worth of the grand coalition according to some solution concept. I assume that the coalitional game is superadditive and that the cooperative solution is increasing with respect to marginal contributions. Surprisingly, in equilibrium players do not necessarily certify their winning coalitions. Under conditions on the prior belief and the reasoning of the social planner–she treats her ignorance symmetrically in the players–, I characterise the set of Bayesian Nash equilibria of the two-stage game. In particular, in a strong equilibrium players certify all and only their winning coalitions.

Keywords— Coalitional games, incomplete information, certification

1 Introduction

Some projects must be carried out collectively because the resources that are necessary to the development of the project exceed the capacities of single stakeholders. As an example, consider several cities that need a water treatment plant. Cities are nearby enough so that a single plant suffices to fulfil their needs, thus building more than one plant is a waste of resources. The resources necessary to build the plant (e.g. land, human resources, raw material, money) are spread across the cities; therefore cities must cooperate.

Coalitional game theory is suitable to model this issue. Players form coalitions: the *worth* of a coalition of players is 1 (say, the coalition is "winning") if the coalition is able, as a whole and through cooperation, to carry out the project; the worth is 0 if the coalition is unable to do so (say, the coalition is "losing"). In the framework of the example, a coalition of cities is winning if, and only if, it is able to build a water treatment plant without external assistance (of the remaining cities). The coalitional game is exogenous in the sense that the worth of a coalition only transcribes objective abilities of the members and the way they fit in together[1].

To fix the ideas, take two cities represented by their mayors Adam and Bob. The coalitional game that encodes the situation is $\overline{v}$ defined by $\overline{v}(A) = 0, \overline{v}(B) = \overline{v}(A, B) = 1$. In words, Adam's city is not self-sufficient while Bob's city is.

If cities A and B were to decide on their own which coalition undertakes the construction of the treatment plant and to whom it will benefit, then it could be inefficient in terms of social welfare. Indeed, Bob could build a plant and restrict its use to his city only whereas the remaining city (Adam's) does not have the capacity to build another plant. In general, one wants to avoid two more cases: (i) more than one plant is built, (ii) no plant is built because the cities do not reach an agreement, even in the long run.

That is why the intervention of a social planner, here incarnated by the State or whichever political or administrative power above the cities, is necessary to ensure that exactly one treatment plant is built and that every city has access to it. This goes through the formation of the grand coalition: all the cities participate to the project. The social planner next distributes the costs (or benefits) of the treatment plant, that is, the dividends of cooperation, to the cities according to a cooperative solution. In the example, should a social planner use the Shapley value (which I shall define later), we have: $Sh_B(\overline{v}) = 1, Sh_B(\overline{v}) = 0$. Bob, the winner, takes it all.

Usually, cooperative game theory assumes that the coalitional game at play, call it $\overline{v}$, is common knowledge, both to the social planner and the players. In our example, it means that each city knows exactly which coalitions of cities–be it a member of the coalition or not–are winning, which are losing; and so for the social planner. This assumption seems unjustified here: the mayor of city A knows better than his peers or the State the quality of the land of his city. Likewise, mayors of cities A and B know how well the different assets they have fit in together and how fruitful their cooperation can be; something that both the mayor of city C and the social planner can ignore.

One can therefore wonder what happens when the social planner has only partial beliefs upon $\overline{v}$: she has a prior μ on a set of coalitional games $\mathcal{G}$. The social planner could dictate the formation of the grand coalition (N) and impose an allocation to the extent of her beliefs: $\phi(\mathbb{E}_\mu[v]) \in \mathbb{R}^N$ where ϕ is a predetermined function. Consider again the two-cities example and suppose that the social planner believes two coalitional games are equally likely: $\overline{v}$ (defined as before) and v with $v(A) = v(B) = 0, v(A, B) = 1$. In v, no city is self-sufficient. Then given the belief of the social planner, the payments are $Sh_B((\overline{v} + v)/2) = 3/4$ and $Sh_A((\overline{v} + v)/2 = 1/4$. Consequently, the ignorance of the social planner makes Bob worse off (and conversely it makes Adam better off).

If Bob has the possibility or the opportunity to prove the stand-alone worth of his city to the social planner, then she gains in turn full information and can allocate $Sh_B(v_2)$ to Bob. This improves Bob's payment. Naturally, such proof must be reliable to be efficient. For instance, cheap talk cannot be credible (here) because it would be in the interest of Bob to say his stand-alone worth is 1, regardless of whether it is true or not. Hence, Bob must prove his worth in a more credible manner: it must be reliable and not opposable. I call it the *certification* of a coalition. One

[1]The worth of a coalition does not encode the different bargaining powers of the players which are left aside in this paper.

can assume that Bob can gather enough reliable evidence to truthfully convince the social planner that his city is capable.

Consequently, the question I address in this paper is: Under what conditions do players have an incentive to certify their private information to the social planner?

To answer this question, I define a two-stage game of certification. n players of a coalitional game $\overline{v}$ and the social planner share a prior belief μ over a set $\mathcal{G}$ of coalitional games. In addition, each player knows the worth of the coalitions he belongs to, *i.e.* his type. In the first step of the game, the worth of each coalition can be certified if its members unanimously agree to. In the second step, the social planner updates her belief given the certifications (it yields a new belief $\hat{\mu}$), imposes the formation of the grand coalition and allocates $\phi(\mathbb{E}_{\hat{\mu}}[v])$ to the players (where ϕ is a predetermined function). Both the players and the social planner face asymmetric information. It is sound it terms of welfare as the same amount of utility is shared among the players (namely $\overline{v}(N)$) whatever the prior μ and the certifications, so there is no loss.

The two-stage game admits a Bayesian Nash equilibrium provided mild continuity requirements (Prop. 1). Then, I make three assumptions: the prior μ gives positive probability to superadditive games only, namely the worth of a disjoint union is no less than the sum of the worth; the cooperative solution is strictly strongly monotonic (following Young, 1985); coalitional games that are *possible* are given the same weight by the prior belief μ. They are not sufficient to ensure that certifying all and only one's winning coalitions is a best response (Prop 2). Finally, I investigate the role of *minimal winning coalitions*–winning coalitions whose strict subsets lose. I characterise the set of Nash equilibria in three cases: when the ignorance of the social planner is restricted to minimal winning coalitions only ("horizontal asymmetry", Prop. 3); or to one minimal winning coalition and its subsets of cardinal *minus* 1 ("downwards asymmetry", Prop. 4); or to one minimal winning coalition and its supersets of cardinal *plus* 1 ("upwards asymmetry", Prop. 5). Furthermore, players certify all and only one's winning coalitions in Bayesian strong Nash equilibrium. I discuss the tightness of the assumptions through couterexamples of relaxations.

We have seen in the two-cities example above that Bob has an incentive to reveal his type, Adam does not. It makes sense because Bob is in a much stronger position in $\overline{v}$ (he is a "dictator") than in v_1 (he is equally as strong as Adam) and that the Shapley value rewards that aspect. In Propositions 3 to 5, if a winning coalition S is certified to the social planner, then every marginal contribution of $i \in S$ gets increased (against if not certified). This is how the social planner creates an incentive for the players to reveal their own types. Obtaining this incentive is not straightforward as Proposition 2 shows.

The two-stage certification game I present involves asymmetric information in a coalitional game and hence strategic aspects. It does so in a different fashion to what has been done so far in the literature. The literature about coalitional games with incomplete information is scarce. Forges and Serrano, 2013 and Forges, 2017 are useful surveys for game with incomplete information for which binding agreements, that is cooperation, are possible. Salamanca, 2020 defines Bayesian coalitional games (with non transferable utility). In his setting the utility of a player depends on both the allocation he gets when in a given coalition and the vector of types. What a player can get from a coalition S (the set of reachable allocations is D_S) is however commonly known; the asymmetric information is encoded in the types that are independent from the coalitions. It differs from my framework in which players do not know the worth of some coalitions (and hence what allocations are reachable by that coalition): the types are "within" the coalitional game.

The other direction that is close to my paper is the literature about the strategic formation of coalitions (in complete information). There, the players can make proposals to form coalitions and solve the game. The formation is an endogenous process. Chatterjee et al., 1993 develop a protocol to solve a coalitional game. It consists of a sequential order with a discount factor where players propose the formation of a coalition and an allocation thereof. They study the stationarity of equilibria and investigate when the formation of the grand coalition occurs at the first step. Ray and Vohra, 1999 add externalities, modelled by partition functions, to this framework[2]. Pérez-Castrillo and Wettstein, 2001 slightly change the protocol and make it two-stage: first a bidding stage during which players bid to get the right to make a proposal; second a proposal stage in which the winner of the auction proposes the formation of a coalition and an allocation thereof. They show that in their protocol the subgame perfect equilibrium implements the Shapley value. My paper differs from theirs because, here, a social planner exogenously imposes the formation of the grand coalition and an allocation; a player's sole objective is to modify the belief of the social planner to his advantage. While his behaviour is strategic, he does not propose endogenous coalition formations to the other players.

Finally, Konishi et al., 1999 study coalition-proof Nash equilibria for agency games, that is when several principals announce incentive schemes to a single agent. A group deviation is *weakly* profitable if it makes at least one player better off and the remaining players at least as well of; it is a *strict* deviation if it makes all the players better off. A profile is a coalition-proof equilibrium when it is immune to group deviation that are themselves self-enforcing. I allow for weak deviations and make use of strong Nash equilibria, namely profiles for which no group has an interest to deviate provided the complementary strategy is fixed. Strong Nash equilibria are more demanding.

The paper goes as follows. In Section 2, I describe formally the two-stage certification game. Section 3 presents the equilibrium analysis and a discussion of the results. Section 4 presents possible extensions of the model. The proofs are in Section 5.

2 A Two-stage Game of Certification

There are n players of a coalitional game and a social planner who all face asymmetric information about the coalitional game at play. In the first stage of the game, players have the ability to certify their private information. In the second stage, the social planner updates her belief about the coalitional game, imposes the formation of the grand coalition and an allocation to the players according to a predetermined cooperative solution.

Let $N = \{1, ..., n\}$ be the set of players. They are players of a simple coalitional game with transferable utility $(N, \overline{v})$ where $\overline{v} : 2^N \to \{0, 1\}, \overline{v}(\emptyset) = 0$ and $\overline{v}(N) = 1$. Let $\mathcal{G}$ be the set of such simple games on the set of players N.

[2] See also Konishi and Ray, 2003.

As there is no ambiguity about the set of players, one writes $v \in \mathcal{G}$ instead of (N, v).

The social planner has a prior μ on $\mathcal{G}$ such that $\mu(\overline{v}) > 0$. I define the expected coalitional game at play as

$$v_\mu := \mathbb{E}_\mu[v] = \sum_{v \in \mathcal{G}} \mu(v) \times v$$

where the linear combination $v + \lambda w$ is defined by $(v + \lambda w)(S) := v(S) + \lambda w(S)$ for all v, w coalitional games, $\lambda \in \mathbb{R}, S \in 2^N$. (One extends the definition and allows a coalition to take any real value.) Note that in general $v_\mu \notin \mathcal{G}$.

Each player $i \in N$ shares the common prior μ with the social planner and learns the worth of the coalitions he belongs to, *i.e.*, his type. The private belief of i over $\mathcal{G}$ is denoted μ_i and defined as

$$\mu_i(v) = \frac{\mu(v)}{\mu(\mathcal{G}_i)} \mathbb{1}_{\mathcal{G}_i}(v)$$

where $\mathcal{G}_i := \{v \in \mathcal{G} : \forall S \ni i, v(S) = \overline{v}(S)\}$ is the set of all possible coalitional games to player i conditional on his type[3].

Player $i \in S$ is better informed about the worth of S than the social planner if, and only if, $0 < v_\mu(S) < 1$. He can be better informed than the social planner about the worth of a coalition he does not belong to, but not necessarily: this depends on the prior μ[4].

Strategies. The social planner offers each coalition of players the opportunity to certify its worth through hard evidence. Certification occurs if, and only if, each player of the coalition agrees to certify. Players can disclose or withhold true information but cannot lie. Pre-play negotiation is not allowed.

Formally, a strategy for player i is a function $\sigma_i : \{S \subseteq N : i \in S\} \to [0, 1]$ where $\sigma_i(S)$ is the probability of i deciding to certify S. The strategies are chosen independently and simultaneously. Then, S is certified with probability $\prod_{i \in S} \sigma_i(S)$. Let $\sigma = (\sigma_i)_{i \in N}$ be a strategy profile and $a = (a_i)_{i \in N}$ be a realisation of σ. a_i is identified with a pure strategy $a_i : \{S \subseteq N : i \in S\} \to \{0, 1\}$ of player i. The social planner observes only which coalitions are effectively certified and their worth (but neither σ nor a), then updates her belief by Bayes' rule. It yields a new belief $\hat{\mu}(a, \sigma)$.

The social planner imposes the allocation $\phi(\mathbb{E}_{\hat{\mu}(a,\sigma)}[v]) \in \mathbb{R}^N$ where the cooperative solution ϕ is *efficient*, namely $\sum_{i \in N} \phi_i(v) = v(N)$ for each coalitional game v. $\phi_i(v)$ is interpreted as the share of the worth $v(N)$ of the grand coalition allocated to player i (for a game v).

Finally, the players are risk-neutral and choose their strategies *ex interim*. The expected payoff of i given σ is

$$u_i(\sigma) = \mathbb{E}\Big[\phi_i\big(\mathbb{E}_{\hat{\mu}(.,\sigma)}[v]\big)\Big]$$

[3] By construction, $\mu_i(\overline{v}) > 0$ and $\cap_{i \in N} \mathcal{G}_i = \{\overline{v}\}$.

[4] The two extreme cases are (1) the knowledge of his type does not give him additional information about the worth of the coalitions he does not belong to, namely $v_{\mu_i}(T) = v_\mu(T)$ for all $T \not\ni i$; (2) knowing his type is sufficient to get $\overline{v}$, namely knowing one's worth implies knowing the worth of all coalitions. Everything in between is possible. In the case of the leading example, it is expected that if city $i \in \{i, j, k\}$ knows how well i, j and k fit in together, how complementary the three of them are, then it gives insight to i about the joint ability of cities j and k. This is precisely encoded by the prior μ.

Timing of the game. We can summarise how the game goes:

1. Nature picks a coalitional game $\overline{v}$ according to a prior distribution μ on $\mathcal{G}$;

2. For every $S \in 2^N$, Nature reveals $\overline{v}(S)$ to every $i \in S$;

3. Every player $i \in N$ chooses a strategy σ_i;

4. The profile of actions $a = a_\sigma$ is realised (the random variables are drawn, if any);

5. The social planner allocates $\phi(\mathbb{E}_{\hat{\mu}(a,\sigma)}[v])$ to the players.

3 Equilibrium Analysis

I first prove that there always exists a Bayesian Nash equilibrium provided the functions ϕ_i are continuous. Afterwards, I assume throughout the paper that μ gives a positive probability only to simple superadditive games. I describe the structure of superadditive games and explain how (direct) certifications can entail indirect certifications. I prove that for some types of priors, certifying all and only the winning coalitions uniformly improves i's position. In words, each strategy profile induces a posterior belief to the social planner and yields an expected coalitional game. Following this strategy ensures to player i the largest possible expected marginal contributions with respect to the posterior belief (compared to the marginal contributions resulting from any other strategy of i). I provide counterexamples to this fact for extensions of these specific priors. The proofs are given in Section 5.

Proposition 1. *Let μ be a prior on $\mathcal{G}, \overline{v} \in \mathcal{G}$ a coalitional game. If ϕ_i is continuous*[5] *for each $i \in N$, then the game admits a Bayesian Nash equilibrium.*

I only consider equilibria under imperfect information. From now on, "equilibrium" stands for "Bayesian Nash equilibrium". I shall focus on superadditive games for the rest of the paper and investigate the incentive of the players to reveal their private information. Recall that a coalitional game v is *superadditive* if, for all $S, T \in 2^N$ such that $S \cap T = \emptyset, v(S \cup T) \geq v(S) + v(T)$. Call $\mathcal{G}_{SUPA}$ the set of simple superadditive games. I make the following assumption: $\overline{v}$ is superadditive, $\mu(\mathcal{G}_{SUPA}) = 1$. The belief of the social planner is consistent with the superadditive structure of $\overline{v}$: for all strategy profile σ and realisation $a_\sigma, \mathbb{E}_{\hat{\mu}(a_\sigma,\sigma)}[v]$ is also a superadditive game. Indeed,superadditive games are stable by non-negative linear combinations, $\hat{\mu}(a_\sigma, \sigma)$ is a vector of non-negative weights and taking the expectation boils down to computing a large linear combination. Naturally, the superadditive structure of the game yields information to the players. Player i can gather indirect information about the worth of the coalitions that are *unknown* without belonging to them.

Definition 1. The worth of $S \in 2^N$ is *unknown* (to the social planner) if there exists $v \in \mathcal{G}_{SUPA}$ with $\mu(v) > 0$ and $v(S) \neq \overline{v}(S)$. Equivalently, $0 < v_\mu(S) < 1$. I denote by $\mathcal{X} = \mathcal{X}_\mu$ the set of unknown coalitions.

[5] A characteristic function v is identified with a vector of $\mathbb{R}^{2^N}$. $\phi_i : \mathbb{R}^{2^N} \to \mathbb{R}$ is continuous with respect to the Euclidean distance.

Here is how indirect information works. Take $S, T \in \mathcal{X}$ such that $i \in S, i \notin T$. If $S \subseteq N \setminus T$ and S is winning, then player i knows that T is losing. Indeed, out of superadditivity, $1 = \overline{v}(S \cup T) \geq \overline{v}(S) + \overline{v}(T) = 1 + \overline{v}(T)$. Superadditivity therefore encodes what a player can deduce from his type. If coalition (ijk) is winning, then the superadditive structure gives information to player i about how likely (jk) itself is winning (such sharp information the social planner does not have)[6].

Likewise, the certification of a coalition has indirect effects. Certifying a winning coalition S certifies indirectly (as winning) every superset of S and (as losing) every subset of $N \setminus S$. Consequently, when choosing his strategy, i must take into account the indirect certifications it entails, not only the direct ones. Important is the fact that two disjoint coalitions can both lose at the same time. Thus, considering superadditive games gives rise to interesting second order phenomena and incentives for the players to reveal their private information[7].

A winning coalition for which every strict subset loses cannot be certified indirectly. I call such a coalition *minimal winning*: $S \subseteq N$ is minimal winning for $v \in \mathcal{G}_{SUPA}$ if, and only if, $v(S) = 1$ and $T \subsetneq S \implies v(T) = 0$. Call $\mathcal{M}(v)$ the set of minimal winning coalitions for v. Knowing $\mathcal{M}(v)$ is sufficient to have a complete description of v.

Example 1. $N = (1234), \mathcal{M}(\overline{v}) = \{(12), (23)\}, \mathcal{X} = 2^N \setminus \{N, \emptyset\}$. Suppose player 1 must respond to the strategy profile σ_{-1}: $\forall i \in \{2, 3, 4\}, \forall S \in 2^N$ such that $i \in S, \sigma_i(S) = 1$

If player 1 wants (12) to be certified (surely), he must set $\sigma_1(12) = 1$ while he is indifferent about $\sigma_1(123)$ if he wants (123) to be certified: (123) is already certified indirectly through the direct certification of (12). The strategies $(\sigma_1(12), \sigma_1(123)) \in \{(1,0), (1,1)\}$ (all other things remaining equal) yield the same superadditive game $\mathbb{E}_{\hat{\mu}(a,\sigma)}$, hence the same payoff. They are not distinguishable from the perspective of the social planner[8]. However, no indirect certification of (12) is possible. The reason is, all the subsets of (12) may lose. (A precision is needed. $\overline{v}$ is such that (1) and (2) lose. Yet, player 1 is situated *ex interim* and (2) can win with a positive probability.)

Finally, we need to define what the properties of the solution ϕ are. First, ϕ need be defined only on the convex hull of $\mathcal{G}$, that is, $Co(\mathcal{G}) = \{v : 2^N \to [0, 1] : v(N) = 1, v(\emptyset) = 0\}$, because the players cannot induce expected games (given the belief of the social planner) outside this set. Take then a coalitional game v. The *marginal contribution* of agent i to coalition $S \subseteq N \setminus i$ for the game v is $v(S \cup i) - v(S)$.

Definition 2.

- For all $i \in N, v, w$ coalitional games, say i prefers v to w and write $v \succeq_i w$ if for all coalition $S \subseteq N \setminus i, v(S \cup i) - v(S) \geq w(S \cup i) - w(S)$. In words, i prefers v to w if, and only if, every marginal contribution of i for v is weakly greater than the corresponding marginal contribution for w. The preference is strict ($\succ_i$) if at least one of the inequalities is strict.
- ϕ is *strictly strongly monotonic* (SSM) if for all player $i \in N, v \succ_i w \implies \phi_i(v) > \phi_i(w)$.

The Shapley value[9], defined for all coalitional game v by

[6] $(i_1...i_t)$ a shortcut for $\{i_1, ..., i_t\}$.

[7] Besides, $\overline{v}$ being superadditive, the worth of N is the maximum total worth a partition of the players can produce. In that respect, imposing the formation of the grand coalition creates no loss of welfare.

[8] And they are payoff equivalent for the players.

[9] Without ambiguity, s stands for the cardinality of coalition S.

$$\forall i \in N, Sh_i(v) = \sum_{S \subseteq N \setminus i} \frac{s!(n-s-1)!}{n!} \big(v(S \cup i) - v(S)\big)$$

satisfies efficiency and strictly strong monotonicity[10,11].

Without loss of generality, we can restrict the strategies to the *unkwown* coalitions. Certifying a coalition which worth is known by the social planner does not influence her belief (neither directly, nor indirectly). Denote $\mathcal{G}(\mathcal{X}) = \{v \in \mathcal{G}_{SUPA} : \forall S \notin \mathcal{X}, v(S) = \overline{v}(S)\}$ the set of possible games given that $\mathcal{X}$ is the set of unknown coalitions.

Recall that an equilibrium is *strong* if no group of players has an incentive to deviate as a coalition.

In light of the second order phenomena described above when only superadditive games are given a positive probability, one can think that it is always a best response to certify one's winning coalitions, not one's losing coalitions. Therefore, it seems that the corresponding profile is an equilibrium. Indeed, the reasoning goes, consider $i \in S, S$ winning. If S is certified as winning, then so are all the supersets, to which i belongs; and all the subsets of $N \setminus S$, to which i does not belong, are certified as losing. It seems to improve i's every marginal contributions (compared with i not certifying S). This, however, is false in general: revelation of information does not always happen in equilibrium:

Proposition 2. *(Impossibility to reveal information in general.) There exists a triple $(N, \mu, \overline{v}), i \in N$ and a strategy profile σ such that, σ is a Nash equilibrium; $i \in S$ does not certify the winning coalition S; it would make i strictly worse off to do so.*

Proof. It is sufficient to provide an example. Take $N = (1234), \mathcal{M}(v_1) = \{(12)\}, \mathcal{M}(v_2) = \{(13), (14)\}, \mu(v_1) = \mu(v_2) = 1/2$. Then if $\overline{v} = v_1$, it makes player 1 worse off to certify (12) when $\phi = Sh$.

Indeed, the set of unknown coalitions is $\mathcal{X} = \{(12), (13), (14), (134)\}$. First, coalition (13) (resp. (14)) is not certified in equilibrium because it would inform the social planner that player 3 (resp. 4) is a null player. The same holds for (134). Second, player 2 wants (12) to be certified because it improves his position. Yet, it is detrimental for player 1 when the social planner knows the worth of (12) because 1 is "stronger" overall in v_2 than in $v_1 : Sh_1(v_1) = \frac{1}{2} < Sh_1(v_2) = \frac{2}{3}$. To sum up, only player 2 wants some information to be transmitted whereas players 1, 3 and 4 want the social planner to remain uninformed. Therefore, the following strategy profile (restricted to the unknown coalitions) is a Nash equilibrium: $\sigma_2(12) = 1, \sigma_i(S) = 0$ otherwise. □

[10] One can wonder how large the set of compatible cooperative solutions ϕ is. Young, 1985 shows that the Shapley value is the only cooperative solution defined on the set of all cooperative games that satisfies efficiency, strong monotonicity (replace for all $i, \succ_i$ by $\succeq_i$ and $>$ by $\geq$ in my definition) and symmetry (the allocation of a player in the permuted game is the allocation of the permuted player in the original game). Here, I do not require the solutions ϕ to be symmetric. Besides, I do not require the solution to be defined on the set of all coalitional games. In particular, the size of the cake is always the same in my setup: there is always $v(N) = 1$. Therefore, I allow for many more cooperative solutions than affine transformations of the Shapley value. For instance, $v \to Sh(v^\alpha)$ fulfills the criteria for all $\alpha > 0$.

[11] One can also wonder why I do not consider power indices (for a survey, see Andjiga et al., 2003; in their article, "simple" stands for "binary and monotonic for the inclusion", "proper" for "superadditive"). Index powers indeed measure the power of each player in a voting game–a framework close to my simple superadditive games. The reason is twofold. First, while the real game $\overline{v}$ is in $\{0, 1\}$, the social planner applies ϕ to an expected game given her belief. In general, this game is no longer binary; power indexes are not defined over such games in general. Besides, many famous indexes consider only minimal winning coalitions (e.g. the Banhzaf power index that measures the probability to change the outcome of a vote). However, it is impracticable in my setting because the social planner knows *not* what the minimal winning coalitions are in general.

What happens in the example of the proof is that the heuristics–certify all and only one's winning coalitions–is true only all other things being equal. Yet, certifying a winning coalition can be detrimental to other orthogonal marginal contributions. Here, certifying (12) indeed improves player 1's marginal contribution to (2) but worsens his marginal contributions to (3) and (4).

I give conditions under which players indeed want to certify all and only their winning conditions. I always assume that $(H1): \phi$ is (SSM); $(H2): \mu$ is such that $v \in \mathcal{G}(\mathcal{X}) \implies \mu(v) = 1/\#\mathcal{G}(\mathcal{X})$. Property $(H2)$ means that the prior distribution assigns the same probability to every possible game: the social planner follows a principle of insufficient reason. In Proposition 3, a subset of what are the minimal winning coalitions of $\overline{v}$ is unknown.

Proposition 3. *(Horizontal asymmetry.) Suppose that $\mathcal{X} \subseteq \mathcal{M}(\overline{v})$, namely coalition S is unknown only if it is minimal winning*[12]*. Suppose $(H1)$ and $(H2)$ hold. Then a profile σ is a Nash equilibrium if, and only if, for all $S \in \mathcal{X}$, one of the two conditions is satisfied:*

- $(C1): \forall i \in S, \sigma_i(S) = 1$;
- $(C2): \exists i \neq j \in S$ *s.t.* $\sigma_i(S) = \sigma_j(S) = 0$.

Furthermore, σ is a strong equilibrium if, and only if, for all $S \in \mathcal{X}$ condition $(C1)$ is satisfied.

Under condition $(C1)$, every player of a winning coalition S decide to certify it (with full probability) and the certification takes place. It turns out to be optimal for all the players of S. However, in equilibrium, it may happen that S is not certified (with zero probability) even though every player of S would like to. Indeed, because of the unanimity requirement, if two players decide to certify S with probability zero, the deviation of one single player is not sufficient to have it certified with positive probability. I call this situation a "lack of coordination": this is precisely described by condition $(C2)$.

In Proposition 4, one minimal coalition S and its subsets of cardinal $s-1$ are unknown: the asymmetry of information is downwards.

Proposition 4. *(Downwards asymmetry.) Take one S minimal winning for $\overline{v}$. Suppose that $\mathcal{X} = \{S, S \setminus i : i \in S\}$, $(H1)$ and $(H2)$ hold. Then a profile σ is a Nash equilibrium if, and only if, (1) for every losing coalition $S \setminus i$, there exists $j \in S \setminus i$ s.t. $\sigma_j(S \setminus i) = 0$; and (2) coalition S satisfies one of conditions $(C1)$ or $(C2)$ as defined in Proposition 3. The profile is a strong equilibrium when condition $(C1)$ is satisfied.*

Finally, the asymmetry of information of Proposition 5 is upwards: one minimal winning coalition S and its supersets of cardinal $s+1$ are unknown. It is a converse statement to Proposition 4.

Proposition 5. *(Upwards asymmetry.) Take one S minimal winning for $\overline{v}$. Suppose that $\mathcal{X} = \{S, S \cup k : k \in N \setminus S\}, (H1)$ and $(H2)$ hold. Then a profile σ is a Nash equilibrium if, and only if, either (α) coalition S satisfies condition $(C1)$ as defined in Proposition 3; and for all $k \in N \setminus S$, one of the two following conditions holds:*

- $\sigma_k(S \cup k) = 1$;

[12] The result holds for one up to every minimal coalitions that are unknown.

- $\exists i \in S$ *s.t.* $\sigma_i(S \cup k) = 0$.

or (β) *coalition* S *satisfies* ($C2$) *while for all* $k \in N \setminus S$*, coalition* $S \cup k$ *satisfies one of conditions* ($C1$) *or* ($C2$)*.*

Furthermore, the profile σ *is a strong Nash equilibrium when* S *is certified (situation* (α)*).*

Remark 1. The requirement that μ is uniform on $\mathcal{G}(\mathcal{X})$ in the propositions is a sufficient condition. It is possible to prove that a small perturbation of the uniform prior works when the ϕ_i are assumed continuous (see Proposition 6). I do not know whether there exist priors "far" from the uniform prior for which the result holds.

Discussion of the results. Propositions 3, 4 and 5 study different situations but are very close in spirit. They hold for every ϕ that is (SSM). Indeed, the proofs establish that when a winning S is certified, it increases *every* marginal contribution of every player $i \in S$[13]. The same holds for i succeeding *not* to certify a losing coalition. In that sense, a player uniformly improves his position in the sense of Definition 2.

In the proof, I show that in each situation, a player of a winning coalition S is better off when S gets certified to the social planner. However, if the strategy profile is such that at least two players decide not to certify S and due to the requirement of unanimity, the deviation of a single player cannot make the certification of the coalition possible (condition ($C2$). This is a "lack of coordination".

At the equilibrium, it is possible for the players to use mixed strategies, namely to certify some coalitions with a probability in $]0,1[$. However, this possibility to mix is somewhat illusory. Indeed, as far as losing coalitions are concerned, one can mix only if another player certifies with probability zero. Likewise, if a player knows that his strategy certifies indirectly a coalition, he is indifferent about his strategy restricted to this very coalition and can mix. All in all, players can mix at the equilibrium if, and only if, their mixing has no impact on which coalitions will be certified (directly and indirectly). They can mix if, and only if, they are indifferent in their response to a strategy profile. For instance, in Proposition 5, $i \in S$ is better off when $S \cup k$ gets known by the social planner. However, provided S is certified directly, he is indifferent about his $\sigma_i(S \cup k)$ and can mix : $S \cup k$ shall be certified indirectly[14].

In Proposition 4, a player is worse off when a losing coalition he belongs to is certified. The player can unilaterally prevent the certification of the coalition thanks to unanimity. Therefore, no losing coalition is certified at the equilibrium.

Finally, while there are infinitely many equilibrium profiles (because of the many possible indifference), one can list all the equilibrium payoffs. For instance, there are exactly $2^{|\mathcal{X}|}$ equilibrium payoffs in Proposition 3: for each unknown coalition S, there are two possibilities, ($C1$) and ($C2$). Likewise, there are exactly two equilibrium payoffs in Proposition 4: no losing coalition is certified, there are two possibilities for S, ($C1$) and ($C2$). Last, it is slightly more cumbersome to list the equilibrium payoffs in Proposition 5 but one can show that there are exactly $1 + 2^{n-s}$ of

[13] Conversely, it makes the $k \notin S$ worse off. One can view the certification (or not) of S as a zero-sum game between the members of S and those of $N \setminus S$. It is zero-sum because the amount to be shared, namely $\overline{v}(N)$ is constant. Of course, one can view as many zero-sum games as there are unknown coalitions.

[14] Suppose that ϕ is linear (e.g. $\phi = Sh$), then the impossibility to mix at the equilibrium unless when indifferent can be given a general proof. Indeed, the payoff of a player is then ϕ taken over a linear combination of several coalitional games. The linear coefficients are given by the strategy profile. Therefore, the several possible games are the extreme points of a polytope and, because ϕ_i is linear, player i maximises in his favourite coalitional game/direction. Thus, he can mix only when he is "stuck" on a facet regardless of his strategy.

	(12)	(23)	μ
v_1	1	0	1/4
v_2	0	1	1/8
v_3	1	1	1/4
v_4	0	0	3/8

Table 1: Counterexample to Proposition 3 for a general prior

them.

The assumption that the prior μ gives the same probability to every possible game is very restrictive. It does not mean that the prior is symmetric (anonymous) in the players. Indeed, the social planner may have more knowledge about some players than others. For instance, in Proposition 4, the social planner has full knowledge about the worth of k's coalitions, $k \notin S$; not about i's coalitions, $i \in S$. Rather, it means that the *ignorance* of the social planner must be symmetric in the players that are concerned. If the social planner does not know the worth of S and $S \setminus i$, then she does not know the worth of every $S \setminus j, j \in S$. In addition, what she does not know is given the same weight. It is in this sense only that μ is symmetric/uniform. It follows a sort of principle of insufficient reason.

What can we conclude from this result? In strong equilibrium σ^* every winning coalition is certified (be it directly or indirectly). *A contrario*, the social planner understands that a coalition that is *not* certified is losing. Thus, the information is fully revealed: the social planner knows that the superadditive game at play is $\overline{v}$, *i.e.* $\hat{\mu}(.,\sigma^*)(\overline{v}) = 1$, and can allocate $\phi(\overline{v})$.

Tightness of the result. Proposition 2 shows that, in general, it is not a best response to certify one's winning coalitions, not one's losing coalitions. Propositions 3 to 5 depend crucially on hypothesis $(H2) : \mu(v) = 1/\#\mathcal{G}(\mathcal{X}), v \in \mathcal{G}(\mathcal{X})$. The following examples show that there exist priors μ over $\mathcal{X}$ for which the propositions do not hold. Let a_i^* denote the strategy "i certifies all and only his winning coalitions (with probability 1)".

First, a counterexample for the horizontal asymmetry of information:

Example 2. Consider $N = (123), \mathcal{M}(\overline{v}) = \{(12),(23)\}$ and suppose that players 2 and 3 certify their winning coalitions (as player 1 is situated *ex interim*, he does not know that (23) is winning). There are four possible games given with their respective prior probabilities as depicted in Table 1.

If player 1 faces a_{-1}^* and plays the strategy "1 does not certify (12)", then the expected worth of (23) for the social planner is $3/8 < 1/2$ which would be the expected value of (23) had player 1 played a_1^* (the expectations are *ex interim*: 1 does not know whether (23) is a winning coalition). Therefore, if ϕ is sufficiently biased towards 1's marginal contribution to coalition (23), then player 1 is better off when he does *not* certify his winning coalition (12)[15].

[15]One can check that, if $\phi = Sh$, then whatever the positive prior μ, 1 is better off when (12) is certified. What 1 gains at not certifying (12)–his marginal contribution to (23)–is offset by his loss–his marginal contribution to (2).

	(12)	(13)	(23)		(12)	(13)	(23)
v_1	1	1	1	v_5	0	1	1
v_2	1	1	0	v_6	0	1	0
v_3	1	0	1	v_7	0	0	1
v_4	1	0	0	v_8	0	0	0

Table 2: Counterexample to Proposition 4 for a general prior

The weakness of this counterexample lies in the fact that it does not work for $\phi = Sh$: one needs a biased cooperative solution. The existence of a counterexample for the horizontal asymmetry, a general prior and the Shapley value as a cooperative solution is yet to be found or denied. Second, Proposition 4 does not hold when $(H2)$ is not verified. The corresponding counterexample works with $\phi = Sh$.

Example 3. Take $N = (123), \mathcal{M}(\overline{v}) = \{(123)\}, \mathcal{X} = \{(12),(13),(23)\}$. Suppose $\phi = Sh$. Consider player 1 and suppose player 2 (resp. 3) certifies (12) but not (23) (resp. (13), not (23)) (strategy σ_{-1}). There are eight possible games as depicted in Table 2.

Call τ_1 the strategy "1 certifies (13) only (with full probability)". Consider

$$\mu = \left(\varepsilon^2, \varepsilon^2, \frac{1-\varepsilon}{2}, \frac{1-\varepsilon}{2}, \varepsilon^2, \varepsilon - 5\varepsilon^2, \varepsilon^2, \varepsilon^2\right)$$

This prior is such that the social planner believes that v_3 and v_4 are equally likely and that one of the two is $\overline{v}$ with probability close to one (up to an ε). It is such that, if the social planner knows that (12) loses (namely, $\overline{v} \in \{v_5, ..., v_8\}$), then she believes that with a small error $\overline{v} = v_6$ (as ε is small, we have $\varepsilon - 5\varepsilon^2 >> 3\varepsilon^2$).

On the one hand, if player 1 does not certify his coalitions, then[16] $\mathbb{E}_{\hat{\mu}(a_1^*,\sigma_{-1})}[v] = v_\mu$ is close to the average $(v_3 + v_4)/2$. It entails that $Sh_1(\mathbb{E}_{\hat{\mu}(a_1^*,\sigma_{-1})}[v]) = 1/3 + O(\varepsilon)$ with $O(\varepsilon)$ a rest of magnitude ε. On the other hand, $\mathbb{E}_{\hat{\mu}(\tau_1,\sigma_{-1})}[v] \approx v_6$, where, 1 and 3 are (almost) symmetrical and $Sh_i(E_{\hat{\mu}(\tau_1,\sigma_{-1})}[v]) = 1/2 + O(\varepsilon)$.

Consequently, with this μ and when $\phi = Sh$, player 1 prefers to certify (12) (against σ_{-1}). (A similar counterexample can be given for Proposition 5 and a prior μ that is not uniform.)

Now, the result holds for both downwards and upwards asymmetries of information, namely for $\mathcal{X} = \{S, S \setminus i\}$ and for $\mathcal{X} = \{S, S \cup k\}$ provided μ is uniform on $\mathcal{G}(\mathcal{X})$. It cannot be generalised to an asymmetry both ways, namely when $\mathcal{X} \subseteq \{S, S \cup k, S \setminus i : k \in N \setminus S, i \in S\}$[17]. The reason is straightforward. Take $k \in N \setminus S$. If k can and does certify $S \cup k$, then it improves (all other things being equal) the estimated worth of $S \setminus i$ (for $i \in S$) but leaves the belief about $S \setminus i \cup k$ unchanged (this coalition is known by assumption). Consequently, the certification of $S \cup k$ decreases k's marginal contribution to $S \setminus i$. Finally, it suffices to consider a solution ϕ that is (SSM) but strongly biased towards this very marginal contribution $v(S \setminus i \cup k) - v(S \setminus i)$.

Example 4. Take $N = (1234)$ and $\overline{v}$ defined by $\mathcal{M}(\overline{v}) = \{(12),(13),(14)\}$. Take $\mathcal{X} = \{(1),(12),(123),(124)\}$. In particular, it is commonly known that (13) is winning. Suppose μ is uniform on $\mathcal{G}(\mathcal{X})$. One can show that if 3 certifies

[16] The strategies a_1^*, σ_{-1} and τ_1 are pure. I therefore identify the profile with its realisation and omit the latter in the notations.

[17] One cannot have $\mathcal{X} = \{S, S \cup k, S \setminus i : k \in N \setminus S, i \in S\}$: this definition is ill-posed. Indeed, take $i \in S, k \in N \setminus S$. By assumption, $S \cup k \setminus i \notin \mathcal{X}$: this coalition is known. By superadditivity, this coalition reveals the worth of either $S \cup k$ or $S \setminus i$.

(123), then the estimated worth of (1) is 1/4; if 3 does not certify (123), then the estimated worth of (1) is 1/6. Therefore, certifying (123) decreases 3's marginal contribution to (1).

Interestingly, player 3 prefers to certify his winning coalitions and not his losing coalitions if $\phi = Sh$. It means that, even though he deteriorates his marginal contribution to (1), he improves his overall position. This should not come as a surprise because μ is uniform (on $\mathcal{G}(\mathcal{X})$) and the Shapley value is particularly balanced and rewards the players irrespective of their labels and gives the same weight to coalitions of the same cardinal.

This leads to the following open question. Say μ treats the players equally if $S, T \in \mathcal{X}, \#S = \#T \implies v_\mu(S) = v_\mu(T)$: two unknown coalitions of the same cardinal have the same expected worth. In this setting, is it always optimal for a player to certify his winning coalitions and not his losing coalitions when μ treats all the players equally and the cooperative solution is the Shapley value?

4 Extensions

I summarise the results of the article and sketch three possible extensions of my model. The first extension investigates the possibility to extend the model to non simple games. The second extension enlarges the set of coalitionnal games by giving a positive probability to monotonic games instead of superadditive games only. The third extension changes the role of the social planner and allows her to have personal preferences; the proximity to the principal-agent paradigm is highlighted.

In the article, I drop the assumption of common knowledge about the coalitional game $\overline{v}$ at play; there only is a prior μ over a set of simple coalitional games. Players have the possibility to certify the worth of their coalitions to the social planner provided a unanimity condition is fulfilled. Proposition 1 shows that there always exists a (Bayesian) Nash equilibrium under mild continuity requirements. I then focus on superadditive games. Contrary to what the intuition suggests, it is not always a best response to certify winning coalitions and not losing coalitions, even though the cooperative solution that the social planner uses to "reward" the players is strictly strongly monotonic (called (SSM) in the body of the paper).

In three different situations, called horizontal, downwards and upwards asymmetry of information, I characterise the set of Nash equilibria and, among them, the subset of strong equilibria (Propositions 3, 4, 5). It relies heavily on the assumption that the prior gives the same probability to every possible game; it is false in general when one does not make this assumption. In that case, players certify their winning coalitions and not their losing ones in a strong equilibrium. It means that the social planner gains full knowledge through the certifications.

Finally, I would like to sketch three possible extensions of the model. The first one is obvious: I consider only simple games, which model very well the undertaking of large-scale projects. However, there are countless situations in which it is preferable to consider coalitional games not restricted to $\{0, 1\}$. A first step would be to allow for three worth, say $\{0, 1, 2\}$. I believe in many similar cases as in the article, players would have interest to certify their best coalitions (of worth 2) and not their worst (of worth 0). Yet, there could be issues for the social planner to interpret a

Profile / Coalition	(1)	(2)	(3)	(12)	(13)	(23)	(123)
a^*	1/2	**2/5**	2/5	1	1	4/5	1
(σ_1, a^*_{-1})	5/13	**5/13**	4/13	1	10/13	10/13	1

Table 3: Undesirable beliefs in case of monotonic games

coalition that is not certified because she might not know whether its worth is 0 or 1. This would mean that it would be even harder to reason counterfactually and to deduce full information about $\overline{v}$.

The second extension would be to drop the assumption of superadditivity and suppose that μ gives positive probabilities to monotonic simple games only[18]. In the motivation, it would mean that the required resources are not so scarcely distributed across the cities: this does not seem unrealistic. However, the issue of considering monotonic games is that it can yield highly undesirable belief updates. In general, if S gets certified, then this can increase the expected worth of $T \subseteq N \setminus S$ (according to the new belief). Consider the following three-players example. Here, the certification of (13) increases the expected worth of (2) from 5/13 to 2/5:

Example 5. $N = (123)$, coalitions $(1), (12), (13)$ and (123) are winning, the others are losing. Suppose that μ gives the same (positive) probability to each monotonic simple game[19]. Suppose that player 1 responds to a^*_{-1}; call σ_1 the strategy "player 1 certifies only (12)". The results are given in Table 3.

One can see that under strategy a^*_1, the expected worth of coalition (2) is higher (namely 2/5) than under strategy σ_1 (namely 5/13). Therefore, certifying (13) is detrimental to player 1's marginal contribution to (2) even though (13) and (2) are independent with respect to monotonicity/superadditivity. The underlying mathematical reason is that there are proportionally more coalitional games v with $v(2) = 0$ when $v(13) = 0$. It is also worth checking that the coalitional game $\overline{v}$ is superadditive: this undesirable belief only occurs because the social planner considers that all the monotonic games are possible.

This example shows that monotonicity is a property hard to work with in this model. It appears difficult for me to ignore this belief updating behaviour. Better understanding to what extent it is possible to consider monotonic priors is therefore of great interest to generalise this model.

The third extension is slightly different to the problem I investigated. I have supposed throughout the paper that the social planner is benevolent. There she has no personal interest and she only wants to allocate a vector that is as close as possible to the value of the real game $\overline{v}$. In particular, the social planner has neither available actions, nor well-defined preferences: it is not a player but a mere algorithm. This assumption is highly realistic in many economic cases–as is the State in her relationship with the different cities–but lacks credibility in other cases.

I rapidly sketch a tentative construction of a genuine player in *lieu* of a social planner. The idea is that, if the social planner has mixed interests between the revelation of information and other orthogonal interests, she can choose the cooperative solution ϕ that best realises her orthogonal interests.

Suppose for example that the social planner wants to give the same allocation to all the players. Then pick

[18] A coalitional game v is monotonic if $S \subseteq T \implies v(S) \leq v(T)$.

[19] Here, the social planner knows that $\overline{v}$ is monotonic (and that $\overline{v}(N) = 1, \overline{v}(\emptyset) = 0$), but has no further information.

any (SSM) solution ϕ that is a function of the marginal contributions only (*e.g.*, the Shapley value). Then $\psi_\delta := \delta\phi + (1-\delta)\mathbf{1}/n$ is (SSM), no matter how small δ is (and remains positive). There, the social planner has the best of both world for a very small cost (encoded in the δ): she gets the information about the game $\overline{v}$ and satisfies her orthogonal interests (under the hypotheses $(H1)$ and $(H2)$). Note that this example is credible in our leading example. Indeed, the wish to reward the cities exactly in the same way regardless of their strength/ability to build the treatment plant amounts to favour the weakest cities: this is a distributive policy oriented towards the less well-off.

Another example sees the social planner favouring a particular player. Suppose indeed that, for orthogonal reasons, the social planner wants player $i \in N$ to have a larger share than his peers. Then the solution φ^i_δ defined by $\varphi^i_{\delta,i}(v) = (1-\delta) + \delta\phi_i(v), \varphi^i_{\delta,j}(v) = \delta\phi_j(v), j \neq i$, fulfills both interests (up to a small δ). Obviously, this example no longer fits with our leading example. Indeed, the planner is here no longer benevolent and favors player i due to his label only, not due to his objective characteristics. There, the proximity with the principal-agent model arises. One can indeed view the $\delta\phi_j(v)$ payment as satisfying an incentive constraint of the players (for them to reveal their winning coalitions–there is no participation constraint here, the players are compelled to participate to the two-stage game). Then, among all the mechanisms/cooperative solutions that permit the principal/planner to obtain the information she wants, she selects the best mechanism to satisfy orthogonal preferences[20].

This idea of giving preferences to the planner, while not perfectly suitable to the leading example I have in mind and to some extent not completely aligned to the question I address, seems to be an interesting extension of this paper to be pursued.

5 Proofs

In this Section, I shall prove the results stated in Section 3. To that end, I shall need some intermediary results here and there.

Proof. (Of Prop 1.)

The game is quite standard. Certifying with probability 0 or 1 every coalition is a pure strategy. It is extended with mixed strategies. In its mixed extension, the sets of strategies of the players are compact, convex and non empty. The function u_i is a composed function. For all σ,

$$u_i(\sigma) = \mathbb{E} \circ \phi_i \circ \mathbb{E}_{\hat{\mu}(.,\sigma)}[v]$$

where the first expectancy is taken over σ. It is continuous in σ. ϕ_i is a continuous function of the cooperative game that is played by assumption. Hence, u_i is a continuous function of the strategy profile σ.

We have the necessary hypotheses to apply the standard Nash equilibrium theorem for games in incomplete information. □

[20] In my setting, the players must be strictly better off to certify a coalition. It would not make sense, here, to assume that a player wants to certify his coalition by default–that is, unless it strictly harms him. It means that the inequalities must be strict. Therefore, there is not best mechanism for the principal (the δ must be positive); it is at best ε-optimal. Mathematically, we lack closedness.

Proofs of Propositions 3, 4 and 5. In order to prove the three propositions, I first need to define what independent coalitions are. Let $S, T \in 2^N$. Say that S and T are *independent* (with respect to superadditivity) when: $S \cap T \neq \emptyset, S \not\subseteq T$ and $T \not\subseteq S$. Two independent coalitions can both win at the same time, both lose at the same time, can have different worth. Two minimal winning coalitions are independent. So are $S \setminus i, S \setminus j$ (provided $\#S \geq 3$) and $S \cup k, S \cup k'$.

In order to prove Propositions 3, 4 and 5, we need an intermediary lemma to understand what is at work.

Lemma 1. *Suppose that the coalitions of $\mathcal{X}$ are pairwise independent. Then $|\mathcal{G}(\mathcal{X})| = 2^{|\mathcal{X}|}$ and for each $S \in \mathcal{X}$, there are exactly $2^{|\mathcal{X}|-1}$ possible games for which S is a winning coalition.*

Proof. Let us prove the lemma by induction over the cardinal of $\mathcal{X}$. If $\mathcal{X} = \{S\}$, it is easy: all the coalitions but S are known and S can take two values. Then, $\mathcal{G}(\mathcal{X}) = \{v_1, v_2\}$ with $v_1 = v_2 = \overline{v}$ on $2^N \setminus \{S\}, v_1(S) = 0, v_2(S) = 1$.

Let p be an integer. Suppose that the result holds for $|\mathcal{X}| = p$. Let us prove that it holds for $|\mathcal{X}| = p+1$. To that end, write $\mathcal{X} = \{S_1, ..., S_{p+1}\}$. By definition, $\mathcal{G}(\mathcal{X}) = \left\{v \in \mathcal{G}_{SUPA} : \forall T \notin \mathcal{X}, v(T) = \overline{v}(T)\right\}$. We can partition $\mathcal{G}(\mathcal{X})$ in two sets $\mathcal{G}(\mathcal{X})^+$ and $\mathcal{G}(\mathcal{X})^-$ where

$$v \in \mathcal{G}(\mathcal{X})^+ \text{ (resp. } v \in \mathcal{G}(\mathcal{X})^-) \iff v(S_{p+1}) = 1 \text{ (resp. } v(S_{p+1}) = 0)$$

Pose $\psi : \mathcal{G}(\mathcal{X})^- \to \mathcal{G}(\mathcal{X})^+, v \to \psi(v)$ with $v = \psi(v)$ on $2^N \setminus \{S_{p+1}\}$. As S_{p+1} is independent of each $S_t, t \in [p]$, S_{p+1} and S_t can be winning at the same time: $\psi(v)$ is indeed a superadditive game. ψ is thus a well-defined function. It is easy to see that ψ is injective. By noting with a slight abuse of notation $\psi^- : \mathcal{G}(\mathcal{X})^+ \to \mathcal{G}(\mathcal{X})^-$ the function that makes the reverse operation, we observe that $\psi \circ \psi^{-1} = Id$.

Hence $\mathcal{G}(\mathcal{X})^-$ and $\mathcal{G}(\mathcal{X})^+$ are in bijection and have the same cardinal.

Now, call $\mathcal{Y} = \{S_1, ..., S_p\}$. If, without loss of generality, $\overline{v}(S_{p+1}) = 1$, then $\mathcal{G}(\mathcal{Y}) = \mathcal{G}(\mathcal{X})^+$. By assumption, the $S_t, t \in [p]$ are pairwise independent and we can apply the induction hypothesis, $|\mathcal{G}(\mathcal{X})^+| = 2^p$. Finally, $\mathcal{G}(\mathcal{X})^-$ and $\mathcal{G}(\mathcal{X})^+$ partition $\mathcal{G}(\mathcal{X})$ so $|\mathcal{G}(\mathcal{X})| = 2^p + 2^p = 2^{p+1}$. There are half of the coalitional games in $\mathcal{G}(\mathcal{X})^+$ for which $S_t, t \in [p]$, wins by the induction hypothesis. The same holds in $\mathcal{G}(\mathcal{X})^-$ in virtue of ψ. Therefore, we have the same result in $\mathcal{G}(\mathcal{X})$. This obviously holds for S_{p+1} in virtue of ψ. □

Proof. (Of Proposition 3.)

It is now fairly easy. The coalitions $S_1, S_2, ... \in \mathcal{X}$ are pairwise independent. Hence certifying or not S_t does not have any influence on the estimated value of $S_{t'}, t \neq t'$, by the social planner. We can isolate the strategy regarding each coalition from the others.

By Lemma 1, without certification of S_t, its expected value by the social planner is 1/2. Hence players of S_t strictly improve its expected value by certifying it – and their marginal contribution accordingly. This means that a player is uniformly better off (in the sense of Definition 2) when a winning coalition he belongs to gets known by the social planner. This said and recalling that ϕ is (SSM), it is easy to characterise the set of Nash equilibria, and the strict subset of strong Nash equilibria. □

	$v(S)$	$v(i)$	$v(j)$
v_1	1	0	0
v_2	1	1	0
v_3	1	0	1
v_4	0	0	0

Table 4: Case $|S| = 2$

Remark 2. Take $S_1, ..., S_p \in \mathcal{M}(\overline{v})$. It is not possible to extend this result to $\mathcal{X} = \{S_t, N \setminus S_t : 1 \leq t \leq p\}$. Indeed, the fact that two coalitions S and T are independent does not entail that the property of independence holds for the complements. In detail,

$$S \text{ and } T \text{ independent } \implies S \cap N(\setminus T) \neq \emptyset, S \not\subseteq (N \setminus T) \text{ but } (N \setminus T) \subseteq S \text{ is possible}$$

The argument of the proof of Proposition 3 falls down and a counterexample exists.

We have all the necessary tools to prove Propositions 4 and 5. Recall that a_i^* is the strategy "i certifies all and only his winning coalitions"; call σ_{-i} the following adverse strategy: $\sigma_j(S) = 1$ if S wins *or* $i \in S, \sigma_j(S) = 0$ else (for all $j \neq i$). σ_{-i} is so defined that, on every $S \ni i, i$'s response "decides" whether S is certified or not. I shall prove that i's *strict* best response precisely is a_i^*.

Proof. (Of Proposition 4.)

Notice that only the $i \in S$ are strategic players. Indeed, $k \notin S$ does not belong to any unknown coalition. He thus plays no role and can be skipped in the proof

I shall consider three cases: S is a singleton, a pair, none of the two first cases. If $S = \{i\}$, then i is a dictator and it is trivial to prove that it is a strict best response to certify he is a dictator to the social planner. If $S = (ij)$ and j certifies (ij) but not (j), $\mathcal{G}(\mathcal{X}) = \{v_1, v_2, v_3, v_4\}$ where $v_k = \overline{v}$ on $2^N \setminus \mathcal{X}$. The v_k's are detailed in Table 4. For player i, $v_2 \succ_i v_k, k \in \{1, 3, 4\}$. Any strategy yields a convex combination of these four games. Playing a_i^* ensures the best vector of coefficients for the v_k's. It is therefore a strict best response. The spirit is the same for player j[21].

Let us finally suppose that $|S| \geq 3$. Take $i \in N$ and suppose he faces σ_{-i}. Notice that $i \in S$ knows that the $S \setminus j, j \neq i$ lose, but he does not know whether $S \setminus i$ loses. Call τ_i the following strategy: i does certify S but does also certify the $S \setminus i_l, (i_l) \subseteq S \setminus i$ (l such coalitions are not certified). Let us prove that[22]

$$V^* := \mathbb{E}_{\hat{\mu}(a_i^*, \sigma_{-i})}[v] \succ_i \mathbb{E}_{\hat{\mu}(\tau_i, \sigma_{-i})}[v] =: V \tag{1}$$

By assumption, $|S| \geq 3$. It entails that the $S \setminus i_l$ are pairwise independent. Besides, the $S \setminus i_l$ are unknown. It means that the $N \setminus (S \setminus i_l)$ lose. In the same spirit as Lemma 1, there are therefore 2^l possible coalitional games. *Ex interim*, $V^*(S \setminus i) = V(S \setminus i) = 1/2$ out of pairwise independence and uniformity of μ: whatever i does, does not have an influence on the belief of the social planner regarding the worth of $S \setminus i$.

[21] μ uniform is not necessary for $|S| = 1, 2$.

[22] The strategies a_i^*, τ_i and σ_{-i} are pure strategies. I therefore identify the profile with its realisation and omit the latter in the notations.

If $S \setminus i_j$ is certified (losing), then $V(S \setminus i_j) = 0 < V^*(S \setminus i_j)$. Consequently, V and V^* are equal on $2^N \setminus \mathcal{X}$ and on $\{S \setminus i\}$ (to 1/2, this is where we use the uniformity of μ), $1/2 = V^*(S \setminus i_j) > V(S \setminus i_j)$. This shows Equation 1 for this class of strategies τ_i.

Finally, there is a single possible game v_0 where S loses. It is easy to prove that any $v \in \mathcal{G}(\mathcal{X}), v \neq v_0$ is uniformly strictly preferred to v_0. Consequently, not certifying S is a strictly dominated response. This proves Equation 1 in general.

This shows that every player $i \in S$ is uniformly better off when S is certified, when the $S \setminus j, j \neq i$ are not. It is then easy to characterise the set of Nash equilibria and to identify which equilibria are strong. □

Proof. (Of Proposition 5.)

If $|S| = n - 1$, it is trivial. Suppose that $|S| \leq n - 2$. I shall prove a slightly stronger result: S and some $S \cup k, k \in N \setminus S$ are unknown (but not necessarily for all k). There are two steps: (1) prove that playing a_i^* is a strict better response to a_{-i}^* for $i \in S$ than any strategy τ_i that does not certify S; (2) prove that playing a_j^* is a strict best response to a_{-j}^* for $j \in N \setminus S$, always in the sense of Definition 2.

(1). Let $i \in S$. Let τ_i be the following (pure) strategy: i does not certify S, does not certify the $S \cup k_l, (k_l) \subseteq N \setminus S$ (the inclusion can be strict, l such coalitions are *not* certified). Call $V^* := \mathbb{E}_{\hat{\mu}(a_i^*, a_{-i}^*)}[v], V = \mathbb{E}_{\hat{\mu}(\tau_i, a_{-i}^*)}[v]$. Let us show that $V^* \succ_i V$. First, $V^* = \overline{v}$.

There are $1 + 2^l$ games compatible with the strategy profile (τ_i, a_{-i}^*) , $1 + 2^{l-1}$ for which $S \cup k_l$ wins. Indeed, there is $\overline{v}$, the only one for which S wins. Then, the $S \cup k_l, k_l \in (k_l)$ are pairwise independent. We can apply Lemma 1. Therefore, we have $V(S) < 1 = \overline{v}(S), V(S \cup k_l) < 1 = \overline{v}(S \cup k_l)$. V and $\overline{v}$ are equal elsewhere. Thus $\overline{v} = V^* \succ_i V$. (We do not need the uniformity of μ for Point (1)).

(2). Take $j \in N \setminus S$, now. j does not know whether S and the $S \cup k_l, k_l \neq j$ win. For j, if S loses, then $S \cup j$ cannot be certified indirectly. Certifying this coalition does not change the belief of the social planner regarding the worth of $S \cup k_l, k_l \neq j$ *because* the prior μ is uniform[23]. It is therefore sufficient to prove that the certification of $S \cup j$ strictly improves the marginal contribution of player j to coalition S.

Recall j is situated *ex interim*. He believes this: for the social planner, $1 + 2^{n-s}$ coalitional games are possible. These games are partitionable in three sets: $\{\overline{v}\}, A^+, A^-$. A^+ (resp. A^-) is the set of coalitional games for which $S \cup j$ wins (resp. loses). A^- is non empty. Call τ_j the strategy "j does not certify $S \cup j$".

Then one can verify that *ex interim* for j

$$V^*(S \cup j) - V^*(S) = \frac{\#A^+}{1 + \#A^+} > \frac{\#A^+}{1 + \#A^+ + \#A^-} = V(S \cup j) - V(S)$$

hence the desired result.

So far, we known that certifying a winning coalitions increases or leaves unchanged every marginal contribution, certifying a losing coalition does the opposite. It is therefore easy to characterise the set of Nash equilibria and to identify which Nash equilibria are strong. □

[23]It would be false in general without the assumption of uniformity.

Remark 3. The proofs show that it is possible for players to play mixed strategies at the equilibrium, but *only* when they are indifferent and they cannot change the belief of the social planner. That is, when at least one other player certifies a coalition with probability zero (whatever the others do, the coalition is not certified directly), or when a player knows that he need not certify a winning coalition $S \cup k$ because it is certified indirectly by the strategy profile (this holds for $i \in S$).

Therefore, the possibility to mix at the equilibrium is illusory. In particular, if it is a best response for player i to certify a coalition with probability $p \in]0,1[$, then it is a best response to certify this coalition with probability $p' \in [0,1], p' \neq p$. In other words, it cannot be a *strict* best response to certify one coalition with probability p.

Finally, the condition "μ is uniform on $\mathcal{G}(\mathcal{X})$" is a sufficient condition, provided the ϕ_i are continuous in addition.

Proposition 6. *Suppose the ϕ_i are continuous. Proposition 4 holds in an open neighbourhood of the uniform prior $\overline{\mu}$.*

Proof. Let us formalise this. Consider a prior μ such that $\mu(\mathcal{G}(\mathcal{X})) = 1$. One can see μ as a vector of $\mathbb{R}^{|\mathcal{G}(\mathcal{X})|}$ with non-negative entries that sum to 1.

Take a pure strategy profile b. Denote

$$\mathcal{G}(b) = \big\{v \in \mathcal{G}_{SUPA} : \forall S \text{ certified by } b, v(S) = \overline{v}(S)\big\}$$

and with a slight abuse of notations

$$\mathbb{E}_\mu[\mathcal{G}(b)] := \sum_{v \in \mathcal{G}(b)} \frac{\mu(v)}{\mu(\mathcal{G}(b))} \times v$$

Now, for all $i \in S$, for all pure strategy b_i,

$$F_{b_i} : \begin{cases} \mathbb{R}^{|\mathcal{G}(\mathcal{X})|} & \to \mathbb{R} \\ \mu & \to \phi_i\big(\mathbb{E}_\mu[\mathcal{G}(a^*)]\big) - \phi_i\big(\mathbb{E}_\mu[\mathcal{G}(b_i, a^*_{-i})]\big) \end{cases}$$

is a continuous function. Indeed, for every $\mathcal{F} \subseteq \mathcal{G}, \mu \to \mathbb{E}_\mu[\mathcal{F}]$ is continuous and so is ϕ_i by assumption. Therefore, the set $\{F_{b_i} > 0\}$ is an open set (as the reciprocal image of $\mathbb{R}^*_+$). A *finite* intersection of open sets is open:

$$O := \bigcap_{i \in S} \bigcap_{\substack{b_i \text{ pure} \\ \text{strategy}}} \{F_{b_i} > 0\}$$

is open[24]. By construction, a prior $\mu \in \mathbb{R}^{|\mathcal{G}(\mathcal{X})|}$ makes Proposition 4 true if $\mu \in O$. This proposition proves that $\overline{\mu} \in O$. Therefore an open neighbourhood of $\overline{\mu}$ lies in O. Such elements of the neighbourhood are *not* uniform over $\mathcal{G}(\mathcal{X})$.

(The same proof works for $\mathcal{X} = \{S, S \cup k\}$.) □

[24] We can restrict ourselves to pure strategies as only pure strategies are played at the equilibrium provided a player is not indifferent. Therefore, there are finitely many pure strategies and the intersection of open sets is indeed finite.

6 Acknowledgements

I am grateful to Lucie Ménager for her guidance. I also received much help from Sylvain Béal and Philippe Solal. I benefited from discussions with seminar participants of Panthéon-Assas University, IHP Junior Game Theory Seminar, SING 19 and NDGSC 2023. Financial support from the project Labex MME-DII (ANR11-LBX-0023-01) is gratefully acknowledged.

Bibliography

Andjiga, N.-G., Chantreuil, F., & Lepelley, D. (2003). La mesure du pouvoir de vote. *Mathématiques et sciences humaines*, (163).

Chatterjee, K., Dutta, B., Ray, D., & Sengupta, K. (1993). A noncooperative theory of coalitional bargaining. *The Review of Economic Studies*, *60*(2), 463.

Forges, F. (2017). Coopération en information incomplète : Quelques modèles stratégiques: *Revue d'économie politique*, *Vol. 127*(4), 467–493.

Forges, F., & Serrano, R. (2013). Cooperative games with incomplete information : Some open problems. *International Game Theory Review*, *15*(2), 1340009.

Konishi, H., Le Breton, M., & Weber, S. (1999). On coalition-proof nash equilibria in common agency games. *Journal of Economic Theory*, *85*(1), 122–139.

Konishi, H., & Ray, D. (2003). Coalition formation as a dynamic process. *Journal of Economic Theory*, *110*(1), 1–41.

Pérez-Castrillo, D., & Wettstein, D. (2001). Bidding for the surplus : A non-cooperative approach to the shapley value. *Journal of Economic Theory*, *100*(2), 274–294.

Ray, D., & Vohra, R. (1999). A theory of endogenous coalition structures. *Games and Economic Behavior*, *26*(2), 286–336.

Salamanca, A. (2020). On the values of bayesian cooperative games with sidepayments. *Mathematical Social Sciences*, *108*, 38–49.

Young, H. P. (1985). Monotonic solutions of cooperative games. *International Journal of Game Theory*, *14*(2), 65–72.