Efficient Coalition-Proof Full Implementation

Mikhail Safronov*

University of Cambridge

February 29, 2016

Abstract

D'Apresmont Gerard-Varet (AGV) mechanism implements efficient social choice in a budget-balanced manner, however it is susceptible to a joint misreport by a coalition of agents, and it may have inefficient equilibria. This paper extends AGV mechanism by putting more structure on its monetary transfers; in the resulting direct mechanism each agent is paid the Shapley value generated from the expected externalities his report imposes on others. This makes each group of agents to be paid in total the expected externality their report imposes on others, and makes it Bayesian incentive compatible to report truthfully. Moreover, any agent can guarantee to receive his ex ante efficient payoff by reporting truthfully, making all equilibria efficient. It is generically impossible to make truthful report a dominant strategy for all coalitions.

*E-mail: mikhailsafronov2014@u.northwestern.edu. This project was started during my PhD study at Northwestern University. I am grateful to Matt Elliott, Aytek Erdil, Robert Evans, Qingmin Liu, Alessandro Pavan, Antonio Penta, Soenje Reiche, Hamid Sabourian and Asher Wolinsky for fruitful discussion. I am thankful for financial support provided by the Center for Economic Theory and the Graduate School at Northwestern University; and by the Cambridge-INET Institute. All errors are mine.

1 Introduction

In order to make an efficient social choice, a benevolent planner needs to have information about agents' preferences. However, if an agent knows that his report will affect the social choice, he may choose to strategically lie about his preferences. A natural question is when the social planner is able to design an efficient, incentive-compatible social choice, and how she can do it. Social choice design with private information is relevant in a variety of economic environments. A firm manager tests employees' abilities to better assign the tasks among them. A mayor of a city conducts a survey to learn whether the benefit of having a new park for citizens surpasses the costs of its construction. Inhabitants of a village learn about each other to exchange favors and food when the need arises. Government launches a procurement auction to find the efficient firm for a project.

The renowned Vickrey, Clark, Groves (VCG) and D'Aspresmont Gerard-Varet (AGV) mechanisms both implement the efficient social choice in an incentive-compatible way. These mechanisms assume the agent's preferences to be linear in monetary transfers, and assume *private types*: each agent's payoff from social choice depends only on own private information. In VCG (AGV) mechanism each agent is paid the (expected) externality his report imposes on the payoffs of other agents; these transfers create the required incentives to report truthfully.

Individual incentive compatibility is necessary for designing a social choice; however, what happens if several agents decide to form a coalition and jointly misreport? The coalition formation is a practical issue, for example, it has appeared many times in auctions. If the agents have common interests, together they may better affect the social choice and benefit from misreporting, even if individually each of them prefers to report own preferences truthfully. In particular, the transfer schemes in VCG and AGV mechanisms do not prevent a group deviation.

The current paper provides a solution to group deviation problem by explicitly constructing a *direct* mechanism, where each agent is only asked to report own private type, and not the coalition he belongs to (if any). In the mechanism the transfers are designed in such a way that each group of agents is paid in total the expected externality their report imposes on others. This makes truthful report Bayesian incentive compatible for any coalition of agents, assuming other agents report truthfully as well. The constructed coalition-proof mechanism is similar to AGV mechanism; it requires *independent private types* with quasilinear preferences, and achieves budget balance.

In the main part of this paper it is assumed that the process of coalition formation is exogenous and happens independently of the mechanism. In each coalition agents behave as a single player: the private types of all members is a common knowledge within a coalition, and they jointly report these private types to maximize their total payoff plus the monetary transfers. Considering the coalition as a single player is reasonable if one thinks of the agents interacting in many periods and colluding in many events, the current mechanism being one of them. This concept differs from most of the literature with endogenous process of coalition formation, and is more demanding: one does not need to incentivize agents to report truthfully their types within a coalition and therefore there is a larger set of potential joint deviations. I will show that the mechanism designed in this paper works under both exogenous and endogenous processes of coalition formation.

The goal in designing the new mechanism is to make any coalition of agents to be paid the expected externality it imposes on others, expectation is taken with respect to the types (reports) of the agents outside coalition. Since the transfer to any agent will affect the incentives of all coalitions including this agent, the incentives for coalitions are interdependent and challenging to be satisfied at the same time. However, it appears, that if one incentivizes all "big" coalitions, each missing exactly one agent, to report truthfully, then the smaller coalitions are automatically incentivized as well. This finding is intuitive: if a group of agents cannot profitably deviate, neither can the subgroup, controlling fewer reports.

The simplified problem of incentivizing only big coalitions can be solved by linear algebra techniques. In the solution each agent i a) pays the externality which others impose on him; and b) receives a Shapley value of externalities his report imposes on others. That is, for any coalition not containing agent i one can find by how much the expected total payoff changes when learning agent i's report, given the report of this coalition. When taking average of these changes (externalities) across all coalitions one obtains the required Shapley value, the transfer to agent i. This result therefore can be thought of as a connection to cooperative games.

It appears that the constructed mechanism not only solves the problem of coalition-proofness, but it also gets rid of all inefficient equilibria of the original AGV mechanism. That is, the mechanism achieves a *full implementation* of efficient social choice; and it does so even in case of some agents being in exogenously formed coalitions. This happens since any agent i by reporting truthfully can guarantee himself his ex ante efficient payoff, regardless of reports of others. The intuition behind this result is as follows. The coalition-proof mechanism assigns transfers to any coalition equal to externality it imposes on others. One can think of the mechanism as a new social choice function, where the payoffs of agents at any type profile include the monetary transfers of the mechanism. This new social choice is efficient, coalition-proof, and requires no monetary transfers. The latter means that a coalition containing everyone but agent i imposes no externality on agent i, regardless of report of this coalition. This means in turn, that if agent i reports truthfully, he always gets his ex ante efficient payoff. The same logic works for the coalition of agents: if together they report truthfully, their expected total payoff is ex ante efficient, making the mechanism to fully implement efficient social choice even in case of exogenously formed coalitions.

The coalition-proof mechanism has all its equilibria to be efficient even with endogenous process of coalition formation. That is, there is a third party, which commits to some mechanism, approaches a group of agents and proposes them to join the coalition. If all the agents agree, then third party collects their reports, and, according to the mechanism it has committed to, it submits a joint report to the social planner, and later makes a budget-balanced transfer among the members of the coalition. However, any agent can refuse to join the coalition in which case the coalition is not formed and the agents proceed to play in the original mechanism designed by social planner. Since by refusing to join the coalition and reporting truthfully afterwards an agent can guarantee to get ex ante efficient payoff, all equilibria are efficient.

The current mechanism fully implements efficient social choice in a coalition-proof manner; however it is generically impossible to have an even stronger result of truthful report being a *weakly dominant* strategy for all coalitions. The required condition is very restrictive: the total payoff in an efficient social choice as a function of overall type profile has to be a separable function across agents' types. This condition is clearly satisfied in case when the agents' reports do not affect each other and the problem of efficient social choice is independent across the agents. In all known environments such as auctions, public good provision, task allocation agents do exert externalities on each other making the coalition-proof dominant strategy implementation impossible.

The paper is organized as follows. Section 2 discusses the relevant literature. Section 3 provides an example where AGV mechanism is not resistant to group deviation and shows how to fix it with the new mechanism. Section 4 introduces the model, Section 5 constructs the mechanism and shows all equilibria to be efficient. Section 6 shows the equivalent results for endogenous coalition formation. Section 7 discusses the dominant strategy implementation for coalitions. Section 8 concludes.

2 Literature review

The problem of coalition formation in mechanism design has been studied in the papers by Laffont, Martimort (1997, 1998, 2000). In their model there are two agents whose types may be independent or correlated; and in case of no correlation the optimal outcome is coalition-proof. The paper by Che and Kim (2006) extends the model to an arbitrary number of agents and more general environment with object allocation. Che and Kim show that for any incentive compatible, individually rational mechanism there is another mechanism which in addition gives the principal the same expected payoff in case the agents form a grand coalition. The idea is that the principal can "sell" the object to the grand coalition, in a way that each agent will make the same expected payment as in the original mechanism. With additional requirement of ex-post incentive compatibility, the same result holds if a subgroup of agents can form a coalition and the principal knows at least two of the subgroup. In another paper on auctions Che and Kim (2009) show that with passive beliefs under some conditions on value distributions of bidders and impossibility of forming the grand coalition the seller can achieve the same revenue as in case of no coalition formation. Without passive beliefs, however, one needs to know the precise cartel group to adequately adjust the mechanism.

The assumption of passive beliefs, which plays a crucial role in designing the coalition-proof mechanism, was motivated in Myerson (2007). The agents report their types to the social planner, though they are not yet committed to them. The third party proposes a coalition formation, and if successful, the involved agents resubmit their reports. Otherwise, if the coalition formation fails, the reports are unchanged, and the social planner proceeds with implementing the mechanism.

The concept of allowing the coalition to behave as a single player has been considered in the paper by Chen, Micali (2012) in the auction with any arbitrary partition of the agents into coalitions. Their idea is to make each agent to report not only their type, but also the coalition they belong to (if any). Afterwards they run a modified Vickrey mechanism: if several agents report to belong to the same coalition and one of them wins the good, the price for the winner does not increase with the reports of other members in coalition. This makes the agents report truthfully both the valuation for the good and the coalition.

The problem of different aspects of the mechanism with collusion has been studied more extensively in auctions. McAfee and McMillan (1992) show that the inability of the cartel members to pay each other cuts down their payoffs. Later, Che, Condorelli and Kim (2013) show that in this case the seller is not hurt by the collusion possiblity. Erdil and Klemperer (2011) propose a new class of payment rules to make the players less willing to submit non-truthful bids if colluding. Biran and Forges (2011) consider the stability of a collusion in auctions with respect to externalities each bidder may impose on others if getting the object.

The question of coalition formation is also studied in the games with incomplete information in the papers by Demange (1987), Vohra (1999), Serrano, Vohra, Volij (2001), Forges, Mertens, Vohra (2002), Dutta and Vohra (2005), Kamishiro and Serrano (2009), Forges and Serrano (2011). The idea behind is to analyze under which conditions there exists a core in a coalitional game with agents having private information. Second branch of papers studies the coalitional games with no transfers: Barbera, Valenciano (1983), Barbera, Gerber (2003), Barbera, Berga, Moreno (2010, 2014).

An independent branch of literature is devoted to *full implementation*: it considers mechanism design in which all equilibria achieve the desired social choice. In the environment with observable types one requires Maskin monotonicity condition (described in Maskin (1998)). This condition is extended to Bayesian monotonicity in the environments with incomplete information and interdependent types, as shown in Jackson (1991). The idea behind Bayesian monotonicity is that for any outcome which is not socially desirable, there is an agent who can credibly inform the designer if this outcome is being played and get rewarded. However, one needs a non-direct mechanism for this communication to be possible. Matsushima (1993) shows that with quasilinear utilities and side payments one can replace Bayesian monotonicity with much weaker condition, which is satisfied for a generic class of social choices. This result is further developed by Chen, Kunimoto, Sun (2015) where one needs only small transfers for full implementation. A paper by Ollar, Penta (2015) shows the full implementation with boundedly interdependent types, by constructing a direct mechanism which utilizes agent's beliefs.

The current paper unites both coalition-proofness and full implementation. Like the original AGV mechanism, I consider independent private types and do not impose a participation constraint. These assumptions allow to create a mechanism, attractive from many prospects. The mechanism is efficient and budget-balanced. The mechanism is explicit and direct, which would simplify its actual implementation. The agents neither have to report their coalition, nor signal that bad equilibrium is being played; nevertheless the mechanism gives the required incentives to all possible coalitions, regardless of whether the process of coalition formation is exogenous or endogenous. Finally, all equilibria in the mechanism are efficient.

3 An example when AGV mechanism is not coalition-proof

In this section I show an example with three agents where AGV mechanism is not coalition-proof. [1] That is, if two agents form a coalition, they can jointly misreport their types and benefit. Then, I construct a new mechanism which enforces an efficient

[1] This example is taken from the joint work with Wojciech Olszewski

allocation and for each coalition it is Bayesian incentive compatible to report true types. The construction will put more structure on the transfers of AGV mechanism, though from since agent's point of view, the expected transfer is the same as in original AGV mechanism.

Let there be three airlines, numbered by 1,2,3. Each airline operates on one of two routes: A or B. Each airline is known to be overbooked on *exactly* one route, with probability 1/2-1/2, independently from other airlines. The route which the airline is overbooked in, is referred to as the *type* of the airline.

The airlines may decide to accommodate passengers from other airlines to reduce the costs of overbooking. Each airline may choose to accommodate the passengers from other airlines on one of its routes. The costs and benefits of such an action depend on the types of the airlines, and are described below. If airline 1, being type A, decides to accommodate passengers from other airlines on route B, it carries the costs of 3 (thousand dollars). However, each other airline of type B, will get a benefit of 2 (and will not benefit at all if being of type A). If airline 1 decides to accommodate passengers on route A, which it is itself overbooked, it will fail to do so and will carry the reputation costs of 100, though other airlines will still benefit. The costs and benefits are symmetric across the airlines and the routes.

With these parameters one has the following efficient social choice. If all three airlines are of the same type, then no airline does the accommodation. If one of three airlines (say, airline 1) is of unique type, then airline 1 does the accommodation. The total payoff is positive $2 * 2 - 3 = 1$, and dominates the choice of no accommodation.

Knowing the efficient social choice, one can find the required monetary transfers of AGV mechanism, which are shown below to be always zero. The payoff of airline 1 in efficient allocation is as follows: with probability 1/4 all three airlines are of the same type, and every airline gets 0. With probability 1/4 airline 1 has the unique type, has to do the accommodation, and gets a payoff of -3. With probability 1/2 some other airline is of unique type and airline 1 gets a payoff of 2. The ex ante payoff of airline 1 is therefore 1/4, and it coincides with other airlines due to symmetry with respect to the airlines.

Let's assume that all airlines report truthfully. In order to find transfers for AGV mechanism, one takes the report of airline 1 and calculates the expected payoff of other airlines given this report. This payoff is paid to airline 1 and is taken with equal shares from airlines 2 and 3. Notice, however, that due to the symmetry with respect to the types, regardless of airline 1's report, the expected payoff of each other airline will be the same of 1/4. This means that the AGV transfer to airline 1 is always 1/2. Since the same method applies to other airlines, no matter of the report profile, each airline will get a transfer of 1/2 from other airlines, and in total each airline receives *zero transfer*.

Due to the symmetry of the problem, no transfers are required to implement an efficient social choice as a Bayesian Nash equilibrium. Indeed, if one of the airlines misreports, it risks accommodating passengers on the route it itself is overbooked and carrying huge costs. However, the nature of the social choice - airline with unique type helping two others, creates incentives for two airlines to unite together and misreport.

Let's suppose that airlines 2 and 3 behave as a single player (the deviation described below will also work if two airlines engage in an endogenous coalition formation process). With no transfers, they only care about the payoff from the possible accommodation. If airlines 2 and 3 have different types, and they both report truthfully, one of them will have the unique report and therefore has to do the accommodation. The total payoff of airlines 2 and 3 will be $-3 + 2 = -1$. On the other hand, if airlines 2 and 3 report the same type, there will be either no accommodation (if airline 1 makes the same report), or airline 1 has to do the accommodation. In any case the total payoff of two airlines will be non-negative and therefore they are incentivized to report the same type. AGV transfers are therefore not coalition-proof. [2]

Let's now find the required transfers to make efficient social choice a Bayesian Nash equilibrium for coalitions as well as for the individual airlines. If one airline makes the unique report, let's make two other airlines to pay each 7/4 to the airline with unique report. This transfers scheme is budget-balanced, and it will be shown to incentivize coalitions to report truthfully. The sum of payoff (benefits minus costs) and monetary transfers is called profit.

Let two airlines 2 and 3 behave as a single player, and let's find their optimal behavior, given that airline 1 always reports truthfully. If airlines 2 and 3 have the same type A and report it truthfully, then with probability 1/2 airline 1 reports the same type and the total payoff is 0. With probability 1/2 airline 1 reports type B and does the accommodation. In this case each of airlines 2, 3 gets a payoff of 2 and has to transfer 7/4 to airline 1. In total, both airlines get a profit of 1/2. If both airlines 2, 3 deviate and report each type B, then the expected transfer they pay to airline 1 is the same. However, they get no benefit from accommodation making the deviation not profitable. If one of the airlines - say, 2, misreports its type to be B, and airline 3 reports truthfully, then with probability 1/2 airline 2 has to accommodate passengers on route A and carry the costs of 100, making this deviation not profitable. Thus, with equal types, airlines 2 and 3 are incentivized to report truthfully.

Let's suppose now that airline 2 has type A and airline 3 has type B. If they report truthfully, then one of them has to do the accommodation. For example, if airline 1 reports type A, then airline 3 does the accommodation on route A. The total payoff from accommodation for two airlines is -1, plus there is a transfer of 7/4 from airline 1. The total profit is therefore 3/4. If both airlines misreport, one of them will have to do the accommodation on self-overbooked route and incur large reputation costs.

[2] Notice that even if two airlines explicitly form a coalition, without knowing the types of other, it is still a optimal strategy for each to coordinate on the same report.

Let's see that the other deviation of reporting the same type (say, type A) is not profitable either. If airline 1 reports type A as well, two airlines get a zero payoff. If airline 1 reports type B, it will do the accommodation. Only airline 2 will get a benefit 2 from accommodation, and together airlines $2, 3$ have to pay $7/2$. The total profit is therefore $-3/2$, making the deviation not profitable.

The new transfer scheme incentivizes each coalition of two airlines to report truthfully. From the point of view of single airline, the expected transfer it has to make is zero, it coincides with the original AGV transfer. Therefore, each individual airline is incentivized to report truthfully as well. Finally, were all three airlines a single player, they would also choose to report truthfully: the mechanism is efficient, and the total transfer is zero.

One can check that for the coalition of two airlines the expected transfer it has to make equals to the externality it imposes on the outside airline. That is, one calculates the difference between the outside airline's expected payoff, given the coalition's report, and ex ante payoff. The difference is the expected transfer to the coalition, and this fact incentivizes the coalition to report truthfully. In the next sections it will be shown how to construct a mechanism for a general environment, so that the same idea will hold.

4 Model

Let there be several agents denoted as $1, 2, ..., n$ with the total set of agents denoted as I. Each agent i has a private type θ_i, which is independently distributed from types of other agents. Vector of types is denoted as θ. There is a set of public choices S with typical element $s \in S$. Each agent i gets a payoff from a public choice, equal to $u_i(\theta_i, s)$. That is, the payoff for each agent depends on own private type and public choice. Monetary transfers are allowed in the model. The agents are assumed to have quasilinear utility: if agent i receives amount x of money, her total utility equals $u_i(\theta_i, s) + x$. I will refer to "payoff" as the payoff from social choice, and "utility" as the payoff plus transfers.

It is assumed that for any type profile θ there exists a public choice $s^*(\theta)$, which maximizes the sum of agent's payoffs $\sum_i u_i(\theta_i, s)$, given θ. The choice $s^*(\theta)$ is called *efficient* public choice. The payoff of agent i at choice $s^*(\theta)$ is denoted as $u_i(\theta)$. There is a mechanism designer whose goal is to implement the efficient public choice given θ. In the mechanism the agents simultaneously report their private types. Then the mechanism designer implements the efficient public choice, assuming all reports are true. In order to incentivize the agents to report truthfully, mechanism designer introduces monetary transfers: each agent receives transfer $x_i(\theta)$, which depends on the *total* profile of reports θ.

It is assumed that some agents might be in a coalition C. The event of agents forming a coalition is assumed to be exogenous and independent of the mechanism. In case of coalition C being formed its members behave as one player: their private types are a common knowledge within coalition C, and the members of C collectively report their types to maximize the sum of their payoffs together with monetary transfers. The transfers can be freely redistributed among the members of coalition. With this concept of coalition formation, the coalition-proof mechanism is defined as follows:

DEFINITION 1 *A mechanism is coalition-proof, if for any coalition C it is optimal to report truthfully, given that the agents outside C report truthfully as well.*

The goal is to find transfers $x_i(\theta)$ such that the resulting efficient mechanism is coalition-proof and budget-balanced. Two conditions are shown below which happen to be sufficient for such a mechanism. The first condition is budget balance:

$$\sum_i x_i(\theta) = 0, \forall\theta \tag{1}$$

The second condition is that for any proper coalition of agents $C \subset I$ and any report θ_C of types of the agents in coalition C one has:

$$E_{\theta_{I/C}}(\sum_{i\in C} x_i(\theta_C, \theta_{I/C})) = E_{\theta_{I/C}}(\sum_{j\in I/C} u_j(\theta_C, \theta_{I/C})) + D_C \tag{2}$$

where I/C is the set of all agents outside coalition C, D_C is a constant, which may depend on set C but is independent of type profile θ, and operator $E_{\theta_{I/C}}$ means that the expectation is taken over types in I/C (assuming the agents in I/C report truthfully).

PROPOSITION 1 *Expressions* (1)- (2) *are sufficient for the mechanism to be efficient, coalition-proof and budget-balanced.*

Proof.

The equation (1) means that the mechanism is budget-balanced. The equation (2) is written for the total transfer to coalition C and it resembles the idea of AGV mechanism. The first term in right-hand side of (2) means that any coalition C when reporting θ_C will receive (in expectation) the expression $E_{\theta_{I/C}}(\sum_{j\in I/C} u_j(\theta_C, \theta_{I/C}))$ as transfers. This incentivizes the coalition C to report truthfully: coalition C will maximize the expected payoff of its own members, plus the expected payoff of the agents outside C which it receives as transfers. Thus, coalition C will maximize the total expected payoff, making its incentives to coincide with social planner, and making it report truthfully. The second term in right-hand side of (2) is a constant D_C which does not depend on reports and therefore does not affect incentives. *Q.E.D.*

5 Coalition-proof mechanism.

This section is devoted to construct transfers satisfying (1)- (2). First let's introduce some additional notations, which will be used later.

1) The operator $E_{\theta_{i_1},\theta_{i_2},...,\theta_{i_k}} X(\theta)$ is the operator, which takes the expectation of variable X at the type profile θ, with respect to types of agents $i_1, i_2, ..., i_k$. In other words, at type profile θ one fixes the reports of all agents other than $i_1, ..., i_k$ and takes expectation with respect to types of $i_1, ..., i_k$.

2) The value $C_n^k = \frac{n!}{k!(n-k)!}$ is amount of different subsets of size k out of set of size n.

3) The variable $a(\theta)$ is defined as:

$$a(\theta) = \sum_i E_{\theta_i} u_i(\theta)$$

The required transfers will be constructed in two steps. First step is to show that if the mechanism is budget-balanced then a relatively simple condition is enough to show that the mechanism is coalition-proof. This new condition means that with budget balance incentives for big coalitions to report truthfully imply incentives for small coalitions to report truthfully. Second step is the explicit construction of monetary transfers which satisfy both the budget-balance and the new condition.

PROPOSITION 2 *Let two conditions hold:*

a) budget balance condition (1)*;*

b) for any θ and any agent i one has:

$$E_{\theta_i} x_i(\theta_i, \theta_{-i}) = -E_{\theta_i} u_i(\theta_i, \theta_{-i}) + D_i \tag{3}$$

where D_i is a constant. (D_i depends on agent index i but is independent of type profile θ).

Then the mechanism is coalition-proof: equation (2) *is satisfied for any coalition.*

Proof.

Let's notice that, considering any i, equation (3) means that the coalition $C = I/i$, which contains anyone but agent i, has incentives to report truthfully. Indeed: budget balance means that the transfer agent i *receives* in equation (3) is the total amount the coalition I/i *pays*. Thus, coalition I/i gets the negative of expression (3) and it coincides with the expression (2) for coalition $C = I/i$, and $D_C = -D_i$.

Now let's consider any smaller coalition which misses more than one agent. That is, let coalition $C = I/(\{i_1, i_2, ..., i_k\})$ be the coalition for which incentives (equation (2)) is checked. Let's also define the agents outside coalition C as $S = \{i_1, i_2, ..., i_k\}$. One has the following:

$$E_{\theta_S} \sum_{i \in S} x_i(\theta) = -E_{\theta_S} \sum_{i \in S} u_i(\theta) + D_S \tag{4}$$

Expression (4) is derived from expression (3) as follows: one writes expression (3) for any agent $i \in S$, takes expectations over θ_S and sums all of them. The related constant is denoted as D_S.

However, due to budget-balance, the amount coalition S receives is the amount the coalition $C = I/S$ pays. One can see that the expression (4) is equivalent to expression (2) for a coalition C, with relabeling of constants: $D_C = -D_S$. *Q.E.D.*

Remark 1 *As one can see, the idea of proposition 2 is that with budget balance the incentives for large coalitions lead to the incentives for small coalitions. The intuition is that if the bigger coalition cannot profitably deviate to receive a better utility, neither* *can the smaller coalition, controlling fewer reports.*

Now let us consider the transfers below and show that they are budget-balanced and satisfy (3) for any i:

$$
\begin{aligned}
x_i(\theta) = E_\theta u_i(\theta) - E_{\theta_i} u_i(\theta) + \\
+\frac{1}{n}[a(\theta) - E_{\theta_i} a(\theta)] + \\
+\frac{1}{n(n-1)} \sum_{j \neq i} [E_{\theta_j} a(\theta) - E_{\theta_i,\theta_j} a(\theta)] + \\
+\frac{1}{nC_{n-1}^2} \sum_{j,k \neq i} [E_{\theta_j,\theta_k} a(\theta) - E_{\theta_i,\theta_j,\theta_k} a(\theta)] + ... + \\
+\frac{1}{nC_{n-1}^m} \sum_{j_1,j_2,...,j_m \neq i} [E_{\theta_{j_1},...,\theta_{j_m}} a(\theta) - E_{\theta_i,\theta_{j_1},...,\theta_{j_m}} a(\theta)] + ... + \\
+\frac{1}{n}[E_{\theta_{-i}} a(\theta) - E_{\theta_i,\theta_{-i}} a(\theta)]
\end{aligned}
\tag{5}
$$

The equation (5) can be written in more compact way:

$$
\begin{aligned}
x_i(\theta) = E_\theta u_i(\theta) - E_{\theta_i} u_i(\theta) + \\
+\sum_{m=0}^{m=n-1} [\frac{1}{nC_{n-1}^m} \sum_{j_1,j_2,...,j_m \neq i} (E_{\theta_{j_1},...,\theta_{j_m}} a(\theta) - E_{\theta_i,\theta_{j_1},...,\theta_{j_m}} a(\theta))]
\end{aligned}
\tag{6}
$$

THEOREM 1 *Transfers defined by* (6) *constitute efficient, coalition-proof, budget-balanced mechanism.*

Proof. First let's show that the equation (3) holds. Indeed, if one takes expectation over θ_i, the second line from transfers given by (6) disappears. The first line of (6) consists of a constant $E_\theta u_i(\theta)$ and the required incentive term $-E_{\theta_i} u_i(\theta)$ from (3).

Now let's show budget-balance:

A) When one sums over i the first line in equation (6), one gets:

$$\sum_i E_\theta u_i(\theta) - \sum_i E_{\theta_i} u_i(\theta) \tag{7}$$

B) Let's now look at the second line of equation (6) at the term $m = 0$:

$$\frac{1}{n}(a(\theta) - E_{\theta_i} a(\theta))$$

The first part when being summed over i, gives:

$$a(\theta) = \sum_i E_{\theta_i} u_i(\theta) \tag{8}$$

The second part when being summed over i, gives:

$$-\frac{1}{n} \sum_i E_{\theta_i} a(\theta) \tag{9}$$

C) Let's now look at second line of expression (6), term $m = 1$:

$$\frac{1}{n(n-1)} \sum_{j \neq i} (E_{\theta_j} a(\theta) - E_{\theta_i, \theta_j} a(\theta))$$

The first part, $\frac{1}{n(n-1)} \sum_{j \neq i} (E_{\theta_j} a(\theta))$, when being summed up over i, gives:

$$\frac{1}{n} \sum_i E_{\theta_i} a(\theta)$$

which is exactly the negative of expression (9). Thus one gets the following: in the second line of expression (6), when being summed over i, the second part of the term

$m = 0$ cancels the first part of the term $m = 1$. One repeats this procedure for future terms ($m \geq 1$), and as a result when being summed over i, second part of any term m cancels the first part of term $m + 1$. Thus, when the second line of expression (6) is being summed over i, only first part of term $m = 0$ and second part of term $m = n - 1$ are left. The latter expression (before summing over i) is

$$\frac{1}{n}(-E_{\theta_i,\theta_{-i}}a(\theta)) = -\frac{1}{n}E_\theta a(\theta)$$

and when being summed over i, it gives:

$$-E_\theta a(\theta) = -\sum_i E_\theta u_i(\theta) \tag{10}$$

When one sums up equations (7), (8) and (10), one gets 0. Thus, transfers $x_i(\theta)$ are budget-balanced. *Q.E.D.*

The intuition of construction of transfers (6) is that one needs to have expression (3) for incentives. However, if one just makes the transfers equal the first line of expression (6), the budget-balance condition will not be satisfied. Therefore one starts to add terms of the form

$$\frac{1}{nC_{n-1}^m} \sum_{j_1,j_2,\ldots,j_m \neq i} (E_{\theta_{j_1},\ldots,\theta_{j_m}}a(\theta) - E_{\theta_i,\theta_{j_1},\ldots,\theta_{j_m}}a(\theta)) \tag{11}$$

which constitute the second line of (6). These terms do not affect incentive part (equation (3)), and at the same time help to establish budget-balance.

Another way to think of those transfers is as follows. For any agent i the first line in expression (6) is how much she should pay to other agents as their reports affect her payoff. The second line in expression (6) indicates how much agent i should be paid as her report affects the payoffs of others. The term $a(\theta)$ represents the sum of payoffs of all agents. Any term (11) in the second line of expression (6) is the difference, and in that difference the payoff of agent i as a part of $a(\theta)$ gets cancelled, making it the externality agent i's report imposes on others. When one sums terms (11) to get the second line of (6), as result agent i gets a weighted sum of externalities of her report. The transfer to agent i is constructed as the Shapley value from coalition games: one takes any subset S of agents other than i and calculates the externality of agent i's report given the reports of agents from S; and then takes the average across all possible S.

REMARK 2 *One can also see that ex-ante any agent receives zero transfer in constructed mechanism. The transfers for each agent can be changed though at a type-independent constant, as this will not affect the incentives.*

5.1 Uniqueness of equilibrium

It appears that the constructed coalition-proof mechanism fully implements efficient social choice, and it does so with exogenously formed coalitions as well. That is, all equilibria in the mechanism are efficient. The result follows from the observation that any agent i (any coalition C) can guarantee herself the ex ante efficient payoff $E_\theta u_i(\theta)$, regardless of the report of other agents (of agents outside C):

PROPOSITION 3 *1. If all agents except for i always report the same profile θ'_{-i}, and agent i always reports truthfully, the ex ante utility of agent i is $E_\theta u_i(\theta)$;*

2. If all agents outside C always report the same profile $\theta'_{I/C}$, and agents in C always report truthfully, the total ex ante utility of agents in C is $E_\theta \sum_{i\in C} u_i(\theta)$.

Proof.

The proof is presented for a single agent i, and is the same for coalition C. For any report θ'_{-i}, the expected transfer x_i to agent i constructed in the mechanism satisfies the condition (3) with constant D_i being equal to $E_\theta u_i(\theta)$, making ex ante value of x_i zero:

$$E_{\theta_i} x_i(\theta_i, \theta'_{-i}) = -E_{\theta_i} u_i(\theta_i, \theta'_{-i}) + E_\theta u_i(\theta)$$

If agent i always reports truthfully, then she will get in expectation (before learning her own type) the value $E_{\theta_i} u_i(\theta_i, \theta'_{-i})$ as a payoff from the social choice. Adding this to her expected transfer yields the result.

For coalition C one can notice that the transfer for the group I/C, satisfies the expression (2) with $D_{I/C} = -E_\theta \sum_{i \in C} u_i(\theta)$. Since the transfer to the group I/C is minus transfer to coalition C, the expected transfer to coalition C, given the report $\theta'_{I/C}$ and assuming agents in C report truthfully, is:

$$E_{\theta_C} \sum_{i \in C} x(\theta_C, \theta'_{I/C}) = -E_{\theta_C} \sum_{i \in C} u(\theta_C, \theta'_{I/C}) + E_\theta \sum_{i \in C} u_i(\theta)$$

If agents in coalition C always report truthfully, their total expected payoff is $E_{\theta_C} \sum_{i \in C} u(\theta_C, \theta'_{I/C})$, and their total expected utility is $E_\theta \sum_{i \in C} u_i(\theta)$.

Q.E.D.

The intuition behind this proposition is as follows. The mechanism prescribes a social choice and transfers depending on the report θ of type profile. Therefore, one can think of this mechanism as another social choice, where the payoff of the agent includes both the non-monetary payoff, and transfer x. This new social choice is efficient and coalition-proof, though now there are no monetary transfers (as they are included in the new payoffs). This means that a coalition which contains all but agent i, imposes no externality on agent i, regardless of its report, θ'_{-i}. This in turn means that agent i can guarantee herself the efficient ex ante payoff if she reports truthfully. The same logic applies to any coalition C. One thus has:

THEOREM 2 *In the coalition-proof mechanism all equilibria are efficient, regardless of whether some of the agents have exogenously formed a coalition.*

6 Concept of coalition formation

So far coalition formation has been assumed to form exogenously, and once being formed, the coalition behaves as a single player. The mechanism constructed in the paper incentivizes any coalition to report truthfully their types, if agents outside of coalition report truthfully as well. This concept is demanding: there is no problem of reporting private type within a coalition and therefore the set of possible misreports for any coalition is larger. The transfers can be made within a coalition, making a misreport profitable if it increases the total utility, rather than the utility of each coalition member individually.

Since the notion of coalition formation is demanding, one should expect the designed mechanism to work if coalition is formed endogenously, after the announcement of the mechanism. During the endogenous coalition formation one considers any group C of agents and third party, which proposes a coalition formation to all members of C. Third party makes a commitment that it will collect all reports from the agents in C, then submit a joint report θ'_C to the mechanism designer as a function of joint report θ_C it received from the agents, and make a budget-balanced transfer within C. If all the agents in C agree to form a coalition, third party fulfills its commitment, otherwise the agents play in the original mechanism.

From proposition 3 any agent can guarantee to get her ex ante expected payoff by reporting truthfully as long as she has an action which will make the other agents to behave independently of her type. With the endogenous process of coalition formation if an agent refuses to join the coalition, other agents will behave independently of her type. This yields:

Theorem 3 *All equilibria in the constructed coalition-proof mechanism are efficient with endogenous process of coalition formation.*

7 General impossibility of coalition-proof dominant implementation.

The mechanism constructed in the paper allows to extend AGV mechanism to be coalition-proof and fully implement the efficient social choice. A natural question to arise is whether it is possible to make truthful report a weakly dominant strategy for all (exogenously formed) coalitions. However, the following proposition shows that it is generically impossible:

Proposition 4 *In the efficient mechanism a truthful report is a weakly dominant strategy for all exogenously formed coalitions, if and only if there exist a set of functions $f(\theta_i)$ such that*

$$\sum_i u_i(\theta) = \sum_i f(\theta_i) \tag{12}$$

That is, the proposition says that truthful report is a weakly dominant strategy for all coalitions, if the efficient total payoff as a function of the overall type profile θ can be separated into n functions, each depending only on type of one agent. This is a very strong condition, which does not hold for many environments: auctions, good allocation. Basically it requires the social choice to be separable across agents and to impose no externalities.

Proof.

The transfer x_i to agent i depends on the total report profile: $x_i(\theta)$. If one wants the coalition C_{-i}, which contains everyone but agent i, to have a weakly dominant strategy to report truthfully, it has to be that[3]:

$$\sum_{j \neq i} x_j(\theta) = u_i(\theta) - f_i(\theta_i) \tag{13}$$

That is, the coalition C_{-i} gets as total transfer the payoff of agent i minus some term $f_i(\theta_i)$. which does not depend on the report θ_{-i} of coalition C_{-i}. The equation (13) has to hold for all i.

If one considers the grand coalition, then in order to make truthful report a weakly dominant strategy, the total transfer $\sum_i x_i(\theta)$ has to be constant, independent of θ. This means, that if one sums up the equations (13) for all i, then on left hand side there will be a constant. On right hand side one will have the total efficient payoff $\sum_i u_i(\theta)$ and sum of terms $-f_i(\theta_i)$, each depending only on the type of agent i, yielding the result. *Q.E.D.*

REMARK 3 *In the proof the incentives for a grand coalition were checked, and one may find the assumption for grand coalition (and any big coalition) to be formed as non-realistic. However, even if one wants truthful report to be a weakly dominant strategy for two agents i, j and a coalition of these two agents, it has to be that the total efficient payoff has to be separable across types θ_i, θ_j, which is a restrictive condition.*

[3]These conditions are similar to VCG mechanism, they are necessary in case of continuously distributed types, though they may not be required in case of discrete types.

8 Discussion

This paper provides an efficient, coalition-proof, budget-balanced mechanism in the environment of making a public choice with agents having independent private types and quasi-linear preferences. The idea behind this mechanism is similar to AGV mechanism, and it is extended to coalitions. The extension is made through a Shapley value: if in AGV mechanism each agent receives an expected externality her report imposes on others, in the coalition-proof mechanism the agent receives a Shapley value created from such externalities. Moreover, from each agent's individual point of view the expected transfer is the same as in AGV mechanism. The coalition-proof extension also leads to full implementation of an efficient outcome.

There are, however, issues which may arise in the actual implementation of the mechanism. In order to determine the transfers, one needs to know the type distribution for each agent, and the types have to be independent. Moreover, with many agents participating in the mechanism it may be cumbersome to calculate the required transfers since one needs to consider all possible coalitions. On the other hand, large number of people may reduce the problem of coalition formation itself since it will be harder for a large group of people to coordinate.

An interesting question for further development is whether one can extend the constructed mechanism in this paper to more general type-interdependent environments and/or with individually rational constraint. In particular, it would be nice to apply the idea of Shapley value since it "connects" individual and coalitional incentives. Another direction is to consider the implementation in weakly dominant strategies, despite the general impossibility result: one can restrict the model to specific coalitions, and assume the specific forms of endogenous coalition formation.

References

Arrow, K. (1979) "The Property Rights Doctrine and Demand Revelation Under Incomplete Information," in , *Economics and Human Welfare*, ed. by M. Boskin. New York: Academic Press.

Ausubel, L., and O. Baranov. (2013) "Core-Selecting Auctions with Incomplete Information," *working paper*.

Barbera, S., and F. Valenciano (2003) "Collective Probabilistic Judgements," *Econometrica*, Vol. 51, pp. 1033-1046.

Barbera, S., and A. Gerber. (2003) "On Coalition Formation: Durable Coalition Structures," *Mathematical Social Sciences*, Vol. 45, pp. 185-203.

Barbera, S., Berga, D., and B. Moreno. (2010) "Individual Versus Group Strategy-Proofness: When Do They Coincide?," *Journal of Economic Theory*, Vol. 145, pp. 1648-1674.

——— (2014) "Group Strategy-Proofness in Private Good Economies," *working paper*, Barselona GSE.

Biran, O., and F. Forges (2011) "Core-Stable Rings in Auctions with Independent Private Values," *Games and Economic Behavior*, Vol. 73, pp. 52-64.

Che, Y., and J. Kim (2006) "Robustly Collusion-Proof Implementation," *Econometrica*, Vol. 74, pp. 1063-1107.

——— (2009) "Optimal Collusion-Proof Auctions," *Journal of Economic Theory*, Vol. 144, pp. 565-603.

Che, Y., Condorelli, D., and J. Kim (2013) "Weak Cartels and Collusion-Proof Auctions," *working paper*.

Chen, J., and S. Micali (2012) "Collusive Dominant-Strategy Truthfullness," *Journal of Economic Theory*, Vol. 147, pp. 1300-1312.

Chen, Y.-C., Kunimoto, T., and Y. Sun (2015) "Implementation with Transfers," *Discussion Paper No, 2015-04*.

D'Aspremont, C., and L. Gerard-Varet. (1979) "Incentives and Incomplete Information," *Journal of Public Economics*, Vol. 11, pp. 25-45.

Demange, G. (1987) "Nonmanipulable Cores," *Econometrica*, Vol. 55, pp. 1057-1074.

Dutta, B., and R. Vohra (2005) "Incomplete Information, Credibility and the Core," *working paper*, Brown University.

Erdil, A., and P. Klemperer (2011) "A New Payment Rule For Core-Selecting Package Auctions," *Journal of the European Economic Association*, Vol. 8, pp. 537-547.

Forges, F., Mertens, J., and R. Vohra (2002) "The Ex Ante Incentive Compatible Core in the Absence of Wealth Effects," *Econometrica*, Vol. 70, pp. 1865-1892.

Forges, F., and R. Serrano (2011) "Cooperative Games with Incomplete Information: Some Open Problems," *SSRN working paper*.

Jackson, M. (1991) "Bayesian Implementation," *Econometrica*, Vol. 59, pp. 461-477.

Kamishiro, Y., and R. Serrano (2009) "Equilibrium Blocking in Large Quasilinear Economies," *CEMFI working paper*, No. 0911.

Laffont, J., and D. Martimort (1997) "Collusion Under Asymmetric Information," *Econometrica*, Vol. 65, pp. 875-911.

——— (1998) "Collusion and Delegation," *The Rand Journal of Economics*, Vol. 29, pp. 280-305.

——— (2000) "Mechanism Design With Collusion and Correlation," *Econometrica*, Vol. 68, pp. 309-342.

Maskin, E. (1998) "Nash Equilibrium and Welfare Optimality," *Review of Economic Studies*, Vol. 66, pp. 23-38.

Matsushima, H. (1993) "Bayesian Monotonicity with Side Payments," *Journal of Economic Theory*, Vol. 59, pp. 107-121.

McAfee, R. and J. McMillan (1992) "Bidding Rings," *American Economic Review*, Vol. 82, pp. 579-599.

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

Ollar, M. and A. Penta (2015) "Full Implementation and Belief Restrictions," under revision in *American Economic Review*.

Palfrey, T. and S. Srivastava (1989) "Implementation with Incomplete Information in Exchange Economies," *Econometrica*, Vol. 57, pp. 115-134.

Palfrey, T. and S. Srivastava (1989) "Mechanism Design with Incomplete Information: A Solution to the Implementation Problem," *Journal of Political Economy*, Vol. 97, pp. 668-691.

Serrano, R., Vohra, R., and O. Volij (2001) "On the Failure of Core Convergence in Economies with Asymmetric Information," *Econometrica*, Vol. 69, pp. 1685-1696.

Vohra, R. (1999) "Incomplete Information, Incentive Compatibility, and the Core," *Journal of Economic Theory*, Vol. 86, pp. 123-147.