Mathematical Social Sciences 90 (2017) 129–140

Contents lists available at ScienceDirect

Mathematical Social Sciences

journal homepage: www.elsevier.com/locate/mss

Immunity to credible deviations from the truth

Salvador Barberà[a], Dolors Berga[b,*], Bernardo Moreno[c]

[a] MOVE, Universitat Autònoma de Barcelona, and Barcelona GSE. Departament d'Economia i d'Història Econòmica, Edifici B, 08193 Bellaterra, Spain

[b] Departament d'Economia, C/ Universitat de Girona, 10; Universitat de Girona, 17003 Girona, Spain

[c] Departamento de Teoría e Historia Económica, Facultad de Ciencias Económicas y Empresariales, Universidad de Málaga, Campus de El Ejido, 29071 Málaga, Spain

HIGHLIGHTS

- We define the notion of immunity to credible deviations.
- We discuss alternative versions of credibility.
- We single out immune rules with multidimensional alternatives and single-peakedness.
- We identify voting by quota 1 and n as the unique GCW immune rules.

ARTICLE INFO

Article history:
Received 15 March 2016
Received in revised form
11 November 2016
Accepted 15 November 2016
Available online 30 November 2016

ABSTRACT

We study a notion of non-manipulability by groups, based on the idea that only some agreements among potential manipulators may be credible. The derived notion of immunity to credible manipulations by groups is intermediate between individual and group strategy-proofness. Our main non-recursive definition turns out to be equivalent, in our context, to the requirement that truthful preference revelation should be a strong coalition-proof equilibrium, as recursively defined by Peleg and Sudhölter (1998, 1999). We provide characterizations of strategy-proof rules separating those that satisfy it from those that do not for a large family of public good decision problems.

1. Introduction

In many contexts where the basic incentive property of strategy-proofness can be met by non-trivial social choice functions, it becomes natural to investigate whether some of them may not only be immune to manipulation by individuals, but can also resist manipulation by groups of coordinated agents. In previous work (Barberà et al., 2010, 2016) we have identified conditions under which, surprisingly, all social choice functions that are immune to manipulations by individuals will also be free from group manipulation. But this is not always the case. In particular, many interesting strategy-proof rules in separable environments[1] will indeed be group manipulable. In these cases, we shall argue that not all group manipulations represent an equally serious threat, because some strategic movements by coalitions are credible, while others are not. To make this point precise, we define several notions of immunity to credible group manipulations and characterize subclasses of social choice rules that satisfy them in specific environments. We concentrate especially in the following: we say that a deviation leading to a profitable improvement for a group is credible if no individual member of the group would gain from not following the agreed upon strategy under the assumption that all others stick to the agreement. Hence, a group manipulation is credible if the set of prescribed strategies for those individuals who plan to deviate are a Nash equilibrium in the induced game where these agents could use any other preference instead, while those of the rest of agents remain fixed.[2] And then we say that a rule is immune to credible group manipulations if no set of agents can find a profitable deviation away from the truth that is

* Corresponding author.
E-mail addresses: salvador.barbera@uab.cat (S. Barberà), dolors.berga@udg.edu (D. Berga), bernardo@uma.es (B. Moreno).

[1] We use this expression loosely here. Formal definitions of the environments we refer to are given in Section 3.

[2] Actually, the concept remains the same if the possible deviations of manipulators are limited to either following the prescription or revealing their true preferences. See Section 4 for a deeper discussion of this and related points.

http://dx.doi.org/10.1016/j.mathsocsci.2016.11.002
0165-4896/© 2016 Elsevier B.V. All rights reserved.

credible. We illustrate the strength of our new definition, which is more demanding than individual but weaker than group strategy-proofness, by characterizing some families of rules in separable environments, and distinguishing between those that can meet our new requirement and those that cannot. The issue of credibility of group deviations has been formalized in alternative ways, one of which is based on the recursive definition of coalition-proof Nash equilibrium (see Bernheim et al., 1987). In fact, Peleg and Sudhölter (1999) applied this concept to the same environment that we analyze, and concluded that all rules that are strategy-proof in that environment are also coalition-proof. In the same paper, these authors (see also Peleg, 1998) define strong coalition-proofness, again recursively based. Surprisingly, this turns out to be equivalent, in our context, to our non-recursive concept of immunity. Let us remark again that our notion of immunity, and that of strong coalition-proofness, allows for a classification and characterization of different strategy-proof rules according to their degree of group manipulability.

After this Introduction we provide notation and definitions in Section 2. Section 3 presents characterization results in two specific contexts. We start with the problem faced by voters who must select a set of entrants to a club, as described in Barberà et al. (1991). We concentrate on quota rules: voters can support all candidates they like, and then those who receive at least a fixed number of votes, q, are chosen. In the domain of separable preferences, we show that rules based on quota 1 or n (where n is the number of voters) are immune to credible deviations, while all other rules in the class are not. Hence, very extreme distributions of power among voters are needed to guarantee immunity. We then turn to a general version of choice among multi-dimensional alternatives under separable preferences, also called multidimensional single-peaked. We build on Moulin (1980), Border and Jordan (1983) and Barberà et al. (1993). The cases we consider include the previous example and many more. We restrict attention to a large class of rules that are strategy-proof in this context, and again characterize those within the class that are immune to credible deviations by groups. Again, a requirement in the form of unanimity plays a crucial role in separating these rules for all the rest, those that are credibly manipulable. Section 4 discusses alternative definitions of credibility for group manipulations, establishes the equivalence of several apparently different formulations, and the differences with other potential definitions, whose proofs are also examined. Section 5 concludes with some final remarks. Appendix contains proofs that are not essential for the continuity of our arguments.

2. The model and definitions: immunity and credibility

Let $N = \{1, \ldots, n\}$ be the set of *agents* and A be the set of *alternatives*. *Preferences* are complete, reflexive, and transitive binary relations on alternatives. Let $\mathcal{U}$ denote such set of preferences. For $i \in N$, R_i denotes agent i's *preferences* on A. As usual, P_i and I_i denote the strict and indifference preference relation induced by R_i, respectively. A *preference profile* $R_N = (R_1, \ldots, R_n) \in \mathcal{U} \times \cdots \times \mathcal{U} = \mathcal{U}^n$ is a n-tuple of preferences on A. It can also be represented by $R_N = (R_C, R_{N\setminus C}) \in \mathcal{U}^n$ when we want to stress the role of coalition C in N. We call a subprofile of agents in C as $R_C \in \times_{i\in C}\, \mathcal{U} = \mathcal{U}^c$.

A *social choice function* (or *rule*) f on $\mathcal{U}^n$ is a function $f : \mathcal{U}^n \to A$.

At this point it is worth mentioning that although we define our main concept and state our results in Sections 2 and 4 assuming that the set of preferences is the same for all agents, all definitions and results would be correct and straightforwardly obtained if we allowed agents' sets of preferences to be different. We assume equal sets of preferences since this is the case of our application in Section 3.

Let us define some incentive-related properties of social choice functions. The best known non-manipulability axiom is that of strategy-proofness. In its usual form it requires the truth to be a dominant strategy for each agent. However, we provide a more general definition which encompasses strategy-proofness and also considers the option that several agents evaluate the possibility of joint deviations.

Definition 1. Let f be a social choice function on $\mathcal{U}^n$. Let $R_N \in \mathcal{U}^n$ and $C \subseteq N$. A subprofile $R'_C \in \mathcal{U}^c$ such that $R'_i \neq R_i$ for all $i \in C$ is a **profitable deviation** of coalition C against profile R_N (for f) if $f(R'_C, R_{N\setminus C}) P_i f(R_N)$ for any agent $i \in C$.

Profitable deviations are usually called (group) manipulations in the standard definitions of group and individual strategy-proofness. Throughout the paper we shall assume that among profitable deviations for single agents there is always one that is best.[3]

Definition 2. A social choice function f on $\mathcal{U}^n$ is manipulable at $R_N \in \mathcal{U}^n$ by coalition $C \subseteq N$ if there exists a profitable deviation of coalition C against profile R_N, say $R'_C \in \mathcal{U}^c$. A social choice function is **group strategy-proof** if it is not manipulable by any coalition $C \subseteq N$.

When we consider only deviations by single agent coalitions we have strategy-proofness.

Definition 3. A social choice function f on $\mathcal{U}^n$ is manipulable at $R_N \in \mathcal{U}^n$ by agent $i \in N$ if there exists a profitable deviation of agent i against profile R_N, say $R'_i \in \mathcal{U}$. A social choice function is **strategy-proof** if it is not manipulable by any agent $i \in N$.

Remark that, formally, strategy-proofness is a much weaker condition than group strategy-proofness in any of its versions. In many environments and in spite of this definitional gap, individual strategy-proof rules end up also being group strategy-proof.[4] But, of course, in many other situations this equivalence may not hold, and even when there are attractive strategy-proof rules, they are open to manipulation by groups. In this paper, we concentrate on a form of manipulation that is intermediate between those of individual and group strategy-proofness and that is based on the notion of credible profitable deviations, those where no agent in the deviating coalition can gain by not declaring those preferences she was supposed to use as part of the group strategy. Formally,

Definition 4. Let f be a social choice function on $\mathcal{U}^n$. Let $R_N \in \mathcal{U}^n$ and $C \subseteq N$. We say that $R'_C \in \mathcal{U}^c$ a profitable deviation of C against R_N is **credible** if for all $i \in C$ and all $\overline{R}_i \in \mathcal{U}$, then $f(R'_C, R_{N\setminus C}) R_i f(\overline{R}_i, R'_{C\setminus\{i\}}, R_{N\setminus C})$.

On other terms, a profitable deviation by C from $R_N = (R_C, R_{N\setminus C})$ is credible if R'_C is a Nash equilibrium of the game among agents in C, when these agents strategies are their admissible preferences and the outcome function is $f(\cdot, R_{N\setminus C})$.

[3] The existence of a best deviation is guaranteed when the number of alternatives, and those of preferences are finite. Moreover, the condition will also hold under standard assumptions.

[4] See Le Breton and Zaporozhets (2009) and Barberà et al. (2010, 2016).

Definition 5. A social choice function f on $\mathcal{U}^n$ is **immune to credible deviations** if for any $R_N \in \mathcal{U}^n$, any $C \subseteq N$, there is no credible profitable deviation of C against R_N (that is, for any profitable deviation $R'_C \in \mathcal{U}^c$ of C against R_N there exists $i \in C$ such that $f(\overline{R}_i, R'_{C\setminus\{i\}}, R_{N\setminus C})P_if(R'_C, R_{N|C})$ for some $\overline{R}_i \in \mathcal{U}$).[5]

Immunity to credible deviations means that no profitable deviation of any coalition is credible at any profile. Observe that group strategy-proofness implies immunity to credible deviations as defined above. However, in general the converse implication fails (see Proposition 1). Moreover, as Lemma 1 shows, immunity to credible deviations implies strategy-proofness. And strategy-proofness implies immunity to credible deviations by singletons.

Lemma 1. *Any social choice function f on $\mathcal{U}^n$ that is immune to credible deviations is strategy-proof.*

Proof. By contradiction, let $R_N \in \mathcal{U}^n$, $i \in N$, and $R'_i \in \mathcal{U}$ such that $R'_i \neq R_i$ and $f(R'_i, R_{N\setminus\{i\}})P_if(R_N)$ and R'_i be such that it is a best deviation for agent i (which, as already stated, we assume to exist). By immunity to credible deviations, there exists $\overline{R}_i \in \mathcal{U}$, such that $f(\overline{R}_i, R_{N\setminus\{i\}})P_if(R'_i, R_{N\setminus\{i\}})$ which contradicts that R'_i is a best deviation for i. ■

3. Applications

We have remarked in the introduction that in some important domains one can define strategy-proof rules that are, however, not group strategy-proof. In this section we illustrate the strength of our new definition, by showing that it allows to differentiate between different rules that are all manipulable by groups, but with different degrees of credibility. In fact, we can characterize the subfamilies of strategy-proof, anonymous, and onto rules that are immune to credible group manipulations, and separate them from those that are not.

We concentrate on the analysis of strategy-proof rules in contexts where alternatives are multidimensional and preferences are multidimensional single-peaked, since it is a very well studied case admitting a large family of non-trivial, strategy-proof social choice rules which are, nonetheless, group manipulable. Hence, it is a natural testing ground for our presumption that some of them may be immune to credible manipulations while others are not.

For expositional purposes, we have chosen to first discuss a special case of the general model, one where alternatives are sets of candidates, preferences are over sets and satisfy a separability condition. After that, we turn to the general case. The reader who prefers to stick to the initial example can already appreciate the broad lines of the arguments leading to our characterizations. Likewise, the reader who prefers the general (and somewhat more involved) arguments may want to skip the special case.

Now, let us comment that the distinction between those strategy-proof rules in our context that are immune to group deviations and those that do not depend on the existence of some "privileged" alternative in each dimension that will be chosen unless all agents agree otherwise.[6] In that limited sense, our immune rules satisfy a form of solidarity among agents that has been presented as a normative requirement in a different context (Thomson, 1993, 1999). However, we prefer to remain agnostic regarding the desirability of using those rules, rather than others that may be vulnerable to credible group deviations but provide a more even treatment to different alternatives. At any rate, our characterization is based on remarks that hinge upon our own definition of credibility and do not depend on any normative considerations.

Let us anticipate the reasons why we may be able to avoid credible deviations under certain rules and not in others. The explanation at this point will necessarily be sketchy, as it comes before formal statements, but we hope it may help the reader. Take an agent i who is considering to participate in a profitable deviation involving agents in coalition S, and in case of participation is asked to reveal a preference R'_i rather than her true preference R_i. Suppose that, if all other members of the deviating coalition do follow the strategies they are asked to, then i will not be able to alter the resulting outcome, regardless of whether she still declares R'_i or any other preference. Then we can say that this agent is individually redundant. And if all agents involved in a jointly beneficial deviation are individually redundant in that sense, their joint beneficial actions will constitute a Nash equilibrium and their threat will be credible. Now, how can one make voters redundant? By creating coalitions and strategies that support their objectives "in excess". Say that they need one agent to deviate in a certain manner, but actually ask two of them to do it. Then, each one becomes "redundant" for the purposes of the decision, but not of course for the credibility of the threat. Our reasoning along the proofs that follow goes in that direction. Rules that allow for threats to be reinforced by involving more agents than those who are strictly necessary to obtain a gain will be vulnerable to credible deviations.

After these motivational comments, we introduce our general framework and definitions.

Let $\mathcal{K} = \{1, \ldots, K\}$ be a finite set of $K \geq 2$ coordinates and for each $k \in \mathcal{K}$, let $B_k = [a_k, b_k]$ with $a_k < b_k$ be an integer interval. Our *alternatives* are K-dimensional vectors in $B = \prod_{k=1}^K B_k$. To stress the role of a set of coordinates k_S, we will write $x = \left(x_{k_S}, x_{\mathcal{K}\setminus k_S}\right) \in B$. We endow B with the L_1-norm. That is, for any $x \in B$,

$$\|x\| = \sum_{k=1}^{K} |x_k|.$$

Given $x, y \in B$, the *minimal box containing x and y* is defined by

$$MB(x, y) = \{z \in B : \|x - y\| = \|x - z\| + \|z - y\|\}.$$

We restrict attention to the case where individual preferences are *antisymmetric* and thus, have a unique best alternative that we denote by $\tau(R_i)$.

We now impose a restriction on preferences which is a natural extension of single-peakedness to the multidimensional setting.

Definition 6. A preference $R_i \in \mathcal{U}$ is *multidimensional* **single-peaked** if for any $z, y \in B$, if $y \in MB(z, \tau(R_i))$ then yR_iz.

Let $\mathcal{S} \subset \mathcal{U}$ be the set of multidimensional *single-peaked* preferences on B. Under this preference restriction $\tau(R_i) = (\tau_1(R_i), \ldots, \tau_K(R_i)) \in B$ where $\tau_j(R_i)$ is the best (or top) alternative of R_i in dimension j.

Multidimensional single-peakedness has two basic implications. One is that the restriction of preferences to alternatives that only differ in a single dimension, while holding the values in all other dimensions fixed, has its best element at the same value than the absolute best alternative. Informally, the projection of the top in any dimension is the top of the projection; and this is independent of the values at which we may have fixed the rest of dimensions. The other implication is that any of these one-dimensional restrictions is single-peaked.

[5] For short, we use the expression "immunity to credible deviations" instead of "immunity to credible profitable deviations" since, by Definition 4, a credible deviation is profitable.

[6] To be more precise, observe that for the setting studied in Section 3.2, such privileged alternative may not exist in one dimension. This is because in the one dimensional setting each strategy-proof rule is also group strategy-proof.

It is known in the literature (Barberà et al., 1993) that the class of multidimensional Generalized Median Voter Schemes (GMVS) are the only strategy-proof social choice functions in our setting, where multidimensional GMVS can be written as K unidimensional GMVS, one for each dimension. In this paper we restrict attention to a particular subclass of GMVS that is a K-dimensional extension of what Moulin (1980) called generalized Condorcet winner rules.

For each $k \in \mathcal{K}$, let $P_k = \{p_k^1, \dots, p_k^{n-1}\}$ be an ordered list of $n-1$ values in B_k where $p_k^1 \leq \cdots \leq p_k^{n-1}$. In what follows, we shall use K lists of such values, one for each dimension, as definitional parameters.

Definition 7. We say that $f : \mathcal{S}^n \rightarrow B, f = (f_1, \dots, f_K)$ is a generalized Condorcet winner rule if for any profile $R_N \in \mathcal{S}^n$, for any $k \in \mathcal{K}, f_k(R_N) = med\{\tau_k(R_1), \dots, \tau_k(R_n), p_k^1, \dots, p_k^{n-1}\}$, where $P_k(f) = \{p_k^1, \dots, p_k^{n-1}\}$ is a list of parameters in B_k.[7]

Remark 1. Generalized Condorcet winner rules constitute the set of anonymous, onto, strategy-proof rules in multidimensional single-peaked domains (see Barberà et al., 1993).

Remark 2. When all parameters in Definition 7 have the same value for some dimension k, we say for short that the list of parameters is degenerate in that dimension. That does not mean, however, that these rules are not interesting. See our discussion in Section 5.

3.1. *Choosing sets of candidates*

Before engaging in a full analysis of those rules that are immune in a general framework, we consider a simple case proposed in Barberà et al. (1991). These authors discuss situations where there exists a set $\mathcal{O}$ of K potential candidates or objects out of which a set of agents must choose the new members of a club, and they characterize the voting rules that may be strategy-proof when preferences are separable. For the benefit of the reader who prefers to stick to this simple case, we analyze it separately, but let us start by saying that it is simply a special case of our more general result discussed in the following subsection. This is because any set of candidates can be described by its characteristic function, assigning value 1 to those that are in the set and 0 to those outside it. Hence, in terms of the alternatives, this is a special case where B_k can take only two values, 0 and 1, in each dimension. As for the restriction of preferences, their notion of separability is equivalent to multidimensional single-peakedness when adapted to their limited context.[8] As we have observed for the general case, here again the best element in any dimension will be found at the value of the absolute best in that dimension (now 0 or 1), and the implication of single-peakedness is immediate because the variable in each dimension only takes two values.

Individual preferences are linear orders on the set $2^{\mathcal{O}}$ (including the empty set). Given any preference R on $2^{\mathcal{O}}$, we define the set of "good" objects $G(\mathcal{O}, R) = \{o_k \in \mathcal{O} : \{o_k\}P\varnothing\}$ and the set of "bad" objects $\mathcal{O} \setminus G(\mathcal{O}, R) = \{o_k \in \mathcal{O} : \varnothing P\{o_k\}\}$.

Definition 8. R is a separable preference on $2^{\mathcal{O}}$ if and only if for any set T and any object $o_l \notin T, T \cup \{o_l\}PT$ if $o_l \in G(\mathcal{O}, R)$.

In words, adding a new good object to any set makes the union better than the original set and adding a bad object makes it worse. Now $\mathcal{S}$ denote the set of all separable preferences.

In this setting there exist strategy-proof social choice functions. In particular, the set of such functions that are anonymous, neutral, and satisfy voter sovereignly coincides with the family of voting by quota rules, $f : \mathcal{S}^n \rightarrow 2^{\mathcal{O}}$ defined as follows:

Definition 9. Let $q \in \{1, \dots, n\}$. The social choice function f on $\mathcal{S}^n$ defined so that for any $R_N \in \mathcal{S}^n$,

$$f(R_N) = \{o_k \in \mathcal{O} : |\{i : o_k \in G(\mathcal{O}, R_i)\}| \geq q\}$$

is called voting by quota q.

However, none of these voting by quota rules are group strategy-proof. And yet, we will show that some of them are immune to credible deviations, while others are not.

Before providing a characterization theorem allowing to distinguish between those rules that are immune and those that are not, we present two examples with 5 voters and 2 candidates. The set of all separable preferences when $K = 2$ is detailed in Table 1.

Example 1. Voting by quota 1: each agent declares her best set of objects and any object that is declared as good by some agent is selected.

Consider the profile where $R_1 = R^3, R_2 = R^5$ and for any other agent $R_i = R^1$ the outcome would be $\{o_1, o_2\}$, whereas 1 and 2 could vote for $\varnothing$ and get a preferred outcome.

This proves that the rule is group manipulable. Notice, however, that after having agreed on voting for empty, any of the two agents could simply keep voting for their preferred candidate, and obtain an even better result, provided the other sticks to her announcement. Hence, this group manipulation will not be credible. We leave it to the reader to check that any other group manipulation under this rule will fail to be credible. Hence, in this example, voting by quota 1 is immune to credible deviations. As we shall see the result generalizes.

Example 2. Voting by quota 3: each agent declares her best set of objects and any object that is declared as good by at least three agents is selected.

Consider now the profile where $R_1 = R_2 = R^3, R_3 = R_4 = R^5$ and $R_5 = R^7$. Then, the outcome would be $\{o_1, o_2\}$. Now, if agents 1 and 2 agree to vote for $\varnothing$, and so do agents 3 and 4, the coalition of these four agents can manipulate and have the outcome to be $\varnothing$, that they all prefer to $\{o_1, o_2\}$. Hence, the rule is group manipulable. Moreover, this particular manipulation is credible, because as long as the rest of deviators complies with the agreement, no single agent can profitably deviate from it. Hence, the rule is not immune to credible deviations in this case.

Notice, however, that there would be other profitable deviations that would not be credible. For example, the one where only 1 and 3 agreed to drop their support to their preferred alternative.

In fact, we can prove the following general result.

Proposition 1. *Let $n = 2$ or $n > 3$. Then, voting by quota 1 and n are the only voting by quota rules satisfying immunity to credible deviations.*

Proof. To prove that voting by quota 1 is never subject to credible profitable deviations, notice that any profitable deviation by a group must involve agents who do not vote for some of the candidates they like (since they can always get them without anyone's help). In exchange these agents can get others not to vote for candidates that they dislike.

Let $R_N \in \mathcal{S}^n$, C be a coalition that has a profitable deviation R'_C against R_N. Note that $f(R_N) \nsubseteqq f(R'_C, R_{N\setminus C})$ (otherwise, if $f(R_N) \subsetneqq$

[7] The notation *med* denotes the median(s) of an ordered list. In the present definition this will be unique.

[8] See also Border and Jordan (1983), Le Breton and Sen (1999) and Le Breton and Weymark (1999) who have analyzed a model with separable preferences in continuous multidimensional spaces.

Table 1
The set of all separable preferences when $K = 2$.

R^1	R^2	R^3	R^4	R^5	R^6	R^7	R^8
$\varnothing$	$\varnothing$	o_1	o_1	o_2	o_2	$\{o_1, o_2\}$	$\{o_1, o_2\}$
o_1	o_2	$\varnothing$	$\{o_1, o_2\}$	$\varnothing$	$\{o_1, o_2\}$	o_1	o_2
o_2	o_1	$\{o_1, o_2\}$	$\varnothing$	$\{o_1, o_2\}$	$\varnothing$	o_2	o_1
$\{o_1, o_2\}$	$\{o_1, o_2\}$	o_2	o_2	o_1	o_1	$\varnothing$	$\varnothing$

$f(R'_C, R_{N\setminus C})$, by quota 1, for any candidate $o \in f(R'_C, R_{N\setminus C}) \setminus f(R_N)$, $o \notin G(\mathcal{O}, R_i)$ for any $i \in N$. By separability, for any $i \in N$, $f(R_N)P_i f(R'_C, R_{N\setminus C})$ and R'_C could not be a profitable deviation, which is a contradiction). Thus, there exists a candidate o such that $o \in f(R_N) \setminus f(R'_C, R_{N\setminus C})$. Observe that for each such candidate $o \in f(R_N) \setminus f(R'_C, R_{N\setminus C})$, since f is voting by quota 1, there is at least one individual $i \in C$ such that $o \in G(\mathcal{O}, R_i)$ and $o \notin G(\mathcal{O}, R'_i)$. But now if i declares a preference $\overline{R}_i$ such that $G(\mathcal{O}, \overline{R}_i) = G(\mathcal{O}, R'_i) \cup \{o\}$, the outcome $f(\overline{R}_i, R'_{C\setminus\{i\}}, R_{N\setminus C}) = f(R'_C, R_{N\setminus C}) \cup \{o\}$, which is, by separability, strictly better for i under R_i than what she would get by following the agreed upon strategy. Therefore, no profitable deviation is credible under quota 1. A similar argument applies for quota n.

This already proves the proposition for the case $n = 2$ since there only the two extreme quotas can be used. From now on we treat the case $n > 3$.

To prove that any voting by quota rule q, $q \neq \{1, n\}$ violates immunity to credible deviations we construct profiles against which there is a credible profitable deviation by some coalition. We begin by the case $K = 2$ and then argue that this can be embedded in a general profile presenting the same deviations whenever $K > 2$.

Let n be odd. We distinguish three subcases.

(1) $q > \frac{n-1}{2} + 1$. Let R_N be as follows: the preferences of any agent i in a set of $\frac{n-1}{2}$ agents are such that $o_1 P_i \{o_1, o_2\} P_i \varnothing$, the preferences of any agent j in a different set of $\frac{n-1}{2}$ agents are such that $o_2 P_j \{o_1, o_2\} P_j \varnothing$, and the preference of the remaining agent l is such that $\tau(R_l) = \{o_1, o_2\}$. Observe that $f(R_N) = \varnothing$. Let C be the coalition of all agents except agent l, let R'_C be such that each agent $i \in C$, $\tau(R'_i) = \{o_1, o_2\}$. Observe that since $f(R'_C, R_{N\setminus C}) = \{o_1, o_2\}$, R'_C is a profitable deviation of C against R_N. Finally, R'_C is credible since no agent can change the outcome by a unilateral deviation since $n > 3$.

(2) $q = \frac{n-1}{2} + 1$. Let R_N be as follows: the preferences of any agent i in a set of $\frac{n-1}{2}$ agents are such that $o_1 P_i \{o_1, o_2\} P_i \varnothing$, the preferences of any agent j in a different set of $\frac{n-1}{2}$ agents are such that $o_2 P_j \{o_1, o_2\} P_j \varnothing$, and the preference of the remaining agent l is such that $\tau(R_l) = \varnothing$. Observe that $f(R_N) = \varnothing$. Let C be the coalition of all agents except agent l, let R'_C be such that each agent $i \in C$, $\tau(R'_i) = \{o_1, o_2\}$. Observe that since $f(R'_C, R_{N\setminus C}) = \{o_1, o_2\}$, R'_C is a profitable deviation of C against R_N. Finally, R'_C is credible since no agent can change the outcome by a unilateral deviation since $n > 3$.

(3) $q < \frac{n-1}{2} + 1$. Let R_N be as follows: the preferences of any agent i in a set of $\frac{n-1}{2}$ agents are such that $o_1 P_i \varnothing P_i \{o_1, o_2\}$, the preferences of any agent j in a different set of $\frac{n-1}{2}$ agents are such that $o_2 P_j \varnothing P_j \{o_1, o_2\}$, and the preference of the remaining agent l is such that $\tau(R_l) = \varnothing$. Observe that $f(R_N) = \{o_1, o_2\}$. Let C be the coalition of all agents except agent l, let R'_C be such that each agent $i \in C$, $\tau(R'_i) = \varnothing$. Observe that since $f(R'_C, R_{N\setminus C}) = \varnothing$, R'_C is a profitable deviation of C against R_N. Finally, R'_C is credible since no agent can change the outcome by a unilateral deviation since $n > 3$.

Let n be even. We distinguish two subcases.

(1) $q > \frac{n}{2}$. Let R_N be as follows: the preferences of any agent i in a set of $\frac{n}{2}$ agents are such that $o_1 P_i \{o_1, o_2\} P_i \varnothing$, the preferences of any agent j in a different set of $\frac{n}{2}$ agents are such that $o_2 P_j \{o_1, o_2\} P_j \varnothing$. Observe that $f(R_N) = \varnothing$. Let C be the coalition of all agents, let R'_C be such that each agent $i \in C$, $\tau(R'_i) = \{o_1, o_2\}$. Observe that since $f(R'_C, R_{N\setminus C}) = \{o_1, o_2\}$, R'_C is a profitable deviation of C against R_N. Finally, R'_C is credible since no agent can change the outcome by a unilateral deviation.

(2) $q \leq \frac{n}{2}$. Let R_N be as follows: the preferences of any agent i in a set of $\frac{n}{2}$ agents are such that $o_1 P_i \varnothing P_i \{o_1, o_2\}$, the preferences of any agent j in a different set of $\frac{n}{2}$ agents are such that $o_2 P_j \varnothing P_j \{o_1, o_2\}$. Observe that $f(R_N) = \{o_1, o_2\}$. Let C be the coalition of all agents, let R'_C be such that each agent $i \in C$, $\tau(R'_i) = \varnothing$. Observe that since $f(R'_C, R_{N\setminus C}) = \varnothing$, R'_C is a profitable deviation of C against R_N. Finally, R'_C is credible since no agent can change the outcome by a unilateral deviation.

This is easily extended to the case $K > 2$ by considering profiles where agents preferences are like the ones described in each case above for objects 1 and 2, while all the agents share exactly the same preferences concerning other objects for all cases analyzed (for example, $o_k \in G(\mathcal{O}, \widehat{R}_i)$ for each $o_k \in \mathcal{O} \setminus \{o_1, o_2\}$, each $i \in N$ and each individual preference $\widehat{R}_i$ used in the analyzed cases). ■

Our next proposition covers the case $n = 3$, which is not contemplated by the previous one.

Proposition 2. *When $n = 3$ and $K = 2$, any voting by quota rule is immune to credible deviations. When $n = 3$ and $K \geq 3$, voting by quotas 1 and 3 are the only voting by quota rules satisfying immunity to credible deviations.*

Proof. Let $N = \{1, 2, 3\}$ and $K = 2$. For voting by quotas 1 and 3 the same argument in Proposition 1 applies. Consider voting by quota 2. As already remarked in Barberà et al. (1991) this rule is not only strategy-proof but also efficient. Thus the only coalitions with profitable deviations consist of two agents. Let R_N, $C = \{i, j\}$, and R'_C be a profitable deviation of C against R_N.

To be a profitable deviation, observe that, by separability and voting by quota 2, either (1) both candidates are chosen under $(R'_C, R_{N|C})$ but none under R_N, or (2) no candidate is chosen under $(R'_C, R_{N|C})$ but both are chosen under R_N, or (3) only one candidate is chosen under R_N and only the other candidate is chosen under $(R'_C, R_{N|C})$.

In the first case, for each candidate, one of the agents in C considered that candidate not good under R_i but good under R'_i. In the second case, for each candidate, one of the agents in C considered that candidate good under R_i but not good under R'_i. In the third case, what was said in the second case holds for the candidate chosen under R_N and what was said in the first case holds for the candidate chosen under $(R'_C, R_{N|C})$.

In the three cases, either declaring $\overline{R}_i$ such that a good candidate under R'_i not to be under $\overline{R}_i$, or supporting a bad one will be an individual profitable deviation with respect to $(R'_C, R_{N|C})$. Thus, R'_C is not credible.

Let $N = \{1, 2, 3\}$ and $K = 3$. For voting by quotas 1 and 3 the same argument in Proposition 1 applies. To prove that voting by quota 2 violates immunity to credible deviations we provide an example of a credible profitable deviation against a profile. Let R_N be as follows: the preferences of agent 1 are such that $\tau(R_1) = o_1$ and $\{o_1, o_2, o_3\} P_1 \varnothing$, the preferences of agent 2 are such that $\tau(R_2) = o_2$ and $\{o_1, o_2, o_3\} P_2 \varnothing$, and the preferences of agent 3 are such that $\tau(R_3) = o_3$ and $\{o_1, o_2, o_3\}P_3\emptyset$. Observe that $f(R_N) = \emptyset$. Let $C = N$, and R'_N be such that each agent $i \in C$, $\tau(R'_i) = \{o_1, o_2, o_3\}$. Since $f(R'_N) = \{o_1, o_2, o_3\}$, R'_N is a profitable deviation of C against R_N. Finally, R'_N is credible since no agent can change the outcome by a unilateral deviation.

This is easily extended to the case $K > 3$ by considering profiles where agents preferences are like the ones described in each case above for objects 1, 2, and 3, while all the agents share exactly the same preferences concerning other objects for all cases analyzed (for example, $o_k \in G(\mathcal{O}, \widehat{R}_i)$ for each $o_k \in \mathcal{O} \setminus \{o_1, o_2, o_3\}$, each $i \in N$ and each individual preference $\widehat{R}_i$ used in the analyzed cases). ■

3.2. The general case: choosing from a grid

We now consider the general case where the set of possible choices in each dimension is not binary, as discussed in Barberà et al. (1993). As we mentioned above, we consider $K \geq 2$ since for $K = 1$ any strategy-proof rule is also group strategy-proof and thus immune to credible profitable deviations. The following three propositions allow us to identify the class of generalized Condorcet winner rules that are immune to credible deviations. Notice that Propositions 3 and 4 completely characterize the case with at least four agents in the society. Proposition 4 also covers the case of two agents. The case of three agents, requiring special treatment, is provided by Proposition 5.

As we have already remarked, our analysis in the preceding subsection is a special case of what comes ahead. Let us then use some of the intuition we got from the previous analysis to anticipate the results that come. In the choice of sets example (with the necessary qualifications regarding number of voters and alternatives) the rules that emerge as not being vulnerable are those where unanimity is required to either impose each alternative, or to avoid it. In a similar spirit, we will see that, for functions that are not vulnerable in the general case there must be a specific privileged value in the interval corresponding to each dimension, and that only unanimous decisions against the prevalence of this particular value will avoid its selection. Hence, immunity to credible deviations requires a privileged treatment of one alternative per dimension, and a unanimity requirement to escape it.

Here are the propositions.

Proposition 3. *Let $n > 3$. Let f be a generalized Condorcet winner rule. If f is defined by lists of parameters that are non-degenerate in at least two dimensions, then f is not immune to credible deviations.*

Proposition 4. *Let $n \geq 2$. Let f be a generalized Condorcet winner rule. If f is defined by lists of parameters that are degenerate in at least $K - 1$ dimensions, then f is immune to credible deviations.*

Proposition 5. *Let $n = 3$. Any generalized Condorcet winner rule defined by lists of parameters such that are non-degenerate in two dimensions is immune to credible deviations. Any generalized Condorcet winner rule defined by non-degenerate lists of parameters in at least three dimensions is not immune to credible deviations.*

Note that for societies with two agents, generalized Condorcet winner rules have only one parameter, thus, since the list of parameters is degenerate in each dimension, Proposition 4 covers the 2-agents case. Moreover, although we only use the result in Proposition 4 for the cases $n = 2$ and $n \geq 4$, we prove it for the 3-agents case for completeness.

Now we turn to the proof of the above propositions.

Proof of Proposition 3. Let f be a generalized Condorcet winner rule with two dimensions, say 1 and 2, for which $P_1(f)$ and $P_2(f)$ are not degenerate. Consider the median(s), $medP_1(f)$ and $medP_2(f)$ of these parameters' lists. These medians may be unique or consist of two contiguous points, say $med^-P_k(f) < med^+P_k(f)$, for each $k \in \{1, 2\}$.

In all cases below, in any profile we will define the preferences of each agent in N concerning dimensions different from 1 and 2 to be the same and with top at some point x_k in B_k, $k \in \mathcal{K} \setminus \{1, 2\}$.

Assume first that for each $k \in \{1, 2\}$, $med^-P_k(f) \neq med^+P_k(f)$. This can only happen if n is odd and thus the number of parameters is even. Consider a partition of N into three sets, $\overline{N}$, $\widetilde{N}$, and l where l is a singleton and such that $\#\overline{N} = \#\widetilde{N}$. Let the projections of R_N in dimensions 1 and 2 be as follows. For agents in $\overline{N}$, let the k-dimensional top be $med^+P_k(f)$ for $k \in \{1, 2\}$. For agents in $\widetilde{N}$, let the k-dimensional top be $med^-P_k(f)$ for $k \in \{1, 2\}$. Agent l has the 1-dimensional top at $med^+P_1(f)$ and the 2-dimensional top at $med^-P_2(f)$. Also assume for agents in $\overline{N} \cup \widetilde{N}$ that $(med^-P_1(f), med^+P_2(f), x_{\mathcal{K}\setminus\{1,2\}})P_i(med^+P_1(f), med^-P_2(f), x_{\mathcal{K}\setminus\{1,2\}})$. Observe that $f_k(R_N) = \tau_k(R_l)$ for each $k \in \{1, 2\}$ and $f_k(R_N) = x_k$ for each $k \in \mathcal{K} \setminus \{1, 2\}$. This is because, for each $k \in \{1, 2\}$, $\tau_k(R_l)$ tie-breaks when computing f_k as the median of all tops and parameters in B_k. Let $C = \overline{N} \cup \widetilde{N}$ and let R'_C be such that for each agent $i \in C$, $\tau_1(R'_i) = med^-P_1(f)$, $\tau_2(R'_i) = med^+P_2(f)$.[9] Observe that $f_k(R'_C, R_{N\setminus C}) = \tau_k(R'_i)$ for $k \in \{1, 2\}$, and $f_k(R'_C, R_{N\setminus C}) = x_k$ for each $k \in \mathcal{K} \setminus \{1, 2\}$. This is because, for each k, $f_k(R'_C, R_{N\setminus C})$ is the top for individual preferences in $(R'_C, R_{N\setminus C})$ for $n - 1$ agents and it coincides with $med^-P_1(f)$ in dimension 1 and with $med^+P_2(f)$ in dimension 2. By definition, this shows that R'_C is a profitable deviation of C against R_N.

Moreover, for each dimension $k \in \{1, 2\}$, since $f_k(R'_C, R_{N\setminus C})$ is either $med^-P_k(f)$ or $med^+P_k(f)$ and, given that $n > 3$, there are at least two parameters smaller than or equal to $f_k(R'_C, R_{N\setminus C}) = med^-P_k(f)$ or greater than or equal to $f_k(R'_C, R_{N\setminus C}) = med^+P_k(f)$.

Therefore, $f_k(R'_C, R_{N\setminus C})$ receives at least $n+1$ total votes for each $k \in \{1, 2\}$. Hence, the profitable deviation R'_C is credible and f is not immune to credible deviations.

Assume now that for at least some $k \in \{1, 2\}$, $med^-P_k(f) = med^+P_k(f) = medP_k(f)$.

Remember that $med^-P_k(f) \neq med^+P_k(f)$ can only hold if n is odd and therefore the number of parameters is even. Because of that in the case where the above equality holds for only one of the two dimensions but not for the other can only happen when n is odd. This distinction is used along the rest of the proof because in one case a partition will only use two sets of agents $\overline{N}$, $\widetilde{N}$ while in other cases we will have to add a singleton l to it.

For n odd, let $\overline{N}$, $\widetilde{N}$, and agent l be the elements of a partition of N such that $\#\overline{N} = \#\widetilde{N} = \frac{n-1}{2}$. For n even, let, $\overline{N}$, $\widetilde{N}$ a partition of N such that $\#\overline{N} = \#\widetilde{N} = \frac{n}{2}$. Let R_N be as follows. The preferences of agents in $\overline{N}$ are such that in the dimension 1 the top is either $med^+P_1(f)$ when $med^-P_1(f) \neq med^+P_1(f)$, or $medP_1(f)$, otherwise. In dimension 2 the top is either $med^+P_2(f)$ when $med^-P_2(f) \neq med^+P_2(f)$, or the highest parameter strictly smaller than $medP_2(f)$ if it exists and $med^-P_2(f) = med^+P_2(f)$, or the lowest parameter strictly greater than $medP_2(f)$, otherwise. The preferences of agents in $\widetilde{N}$ are such that in dimension 1 the top is either $med^-P_1(f)$ when $med^-P_1(f) \neq med^+P_1(f)$, or the highest parameter strictly smaller than $medP_1(f)$ if it exists and $med^-P_1(f) = med^+P_1(f)$, or the lowest parameter strictly greater than $medP_1(f)$,

[9] In words, to define R'_C notice that by changing their vote the agents in $\overline{N}$ vote for the tops of those in $\widetilde{N}$ in dimension 1, while agents in $\widetilde{N}$ vote for the top of those in $\overline{N}$ in dimension 2.

otherwise. In dimension 2 the top is either $med^{-}P_2(f)$ when $med^{-}P_2(f) \neq med^{+}P_2(f)$, or $medP_2(f)$, otherwise.

Preferences of agent l (only required if n is odd) are defined as follows: R_l is such that in dimension 1 agent l's top is either $\tau_1(R_j)$, $j \in \overline{N}$ when $med^{-}P_1(f) \neq med^{+}P_1(f)$, or the highest parameter strictly smaller than $medP_1(f)$ if such parameter exists or the lowest parameter strictly greater than $medP_1(f)$, otherwise. In dimension 2 the top of agent l is either $\tau_2(R_i)$, $i \in \widetilde{N}$ when $med^{-}P_2(f) \neq med^{+}P_2(f)$, or the highest parameter strictly smaller than $medP_2(f)$ if such parameter exists or the lowest parameter strictly greater than $medP_2(f)$, otherwise.

From now on let $i \in \widetilde{N}$ and $j \in \overline{N}$. We also assume that for any agent $m \in \overline{N} \cup \widetilde{N}$, $(\tau_1(R_i), \tau_2(R_j), x_{\mathcal{K}\setminus\{1,2\}})\ P_m(\tau_1(R_j), \tau_2(R_i), x_{\mathcal{K}\setminus\{1,2\}})$. Observe that $f_k(R_N) = \tau_k(R_l)$ if $med^{-}P_k(f) \neq med^{+}P_k(f)$, and $f_k(R_N) = medP_k(f)$ otherwise for each $k \in \{1, 2\}$, and that $f_k(R_N) = x_k$ for each $k \in \mathcal{K} \setminus \{1, 2\}$. This is because, for each $k \in \{1, 2\}$, $\tau_k(R_l)$ tie-breaks when computing f_k as the median of all tops and parameters in B_k in the case where only for one $k \in \{1, 2\}$, $med^{-}P_k(f) = med^{+}P_k(f) = medP_k(f)$ and thus n is odd. And for each $k \in \{1, 2\}$, $medP_k(f)$ tie-breaks when computing f_k as the median of all tops and parameters in B_k in the case where for both $k \in \{1, 2\}$, $med^{-}P_k(f) = med^{+}P_k(f) = medP_k(f)$. Let $C = \overline{N} \cup \widetilde{N}$ and let R'_C be such that for each agent $j \in \overline{N}$, $\tau_1(R'_j) = \tau_1(R_i)$ and $\tau_k(R'_j) = \tau_k(R_j)$ for each $k \in \mathcal{K} \setminus \{1\}$, and for each $i \in \widetilde{N}$, $\tau_2(R'_i) = \tau_2(R_j)$ and $\tau_k(R'_i) = \tau_k(R_i)$ for each $k \in \mathcal{K} \setminus \{2\}$. Observe that $f_k(R'_C, R_{N\setminus C}) = \tau_k(R'_i)$ for $k \in \{1, 2\}$, and $f_k(R'_C, R_{N\setminus C}) = x_k$ for each $k \in \mathcal{K} \setminus \{1, 2\}$. This is because, for each k where $med^{-}P_k(f) = med^{+}P_k(f)$, $f_k(R'_C, R_{N\setminus C})$ is the top for individual preferences in $(R'_C, R_{N\setminus C})$ for n agents. For each k where $med^{-}P_k(f) \neq med^{+}P_k(f)$, $f_k(R'_C, R_{N\setminus C})$ is the top for the preferences for $n - 1$ agents in $(R'_C, R_{N\setminus C})$ and coincides either with $med^{-}P_k(f)$ or $med^{+}P_k(f)$. By definition, this shows that R'_C is a profitable deviation of C against R_N.

Moreover, for the dimensions where $med^{-}P_k(f) = med^{+}P_k(f)$ there is a parameter at $f_k(R'_C, R_{N\setminus C})$. For the dimensions where $med^{-}P_k(f) \neq med^{+}P_k(f)$, $f_k(R'_C, R_{N\setminus C})$ is either $med^{-}P_k(f)$ or $med^{+}P_k(f)$ and, given that $n > 3$, there are at least two parameters smaller than or equal to $f_k(R'_C, R_{N\setminus C}) = med^{-}P_k(f)$ or greater than or equal to $f_k(R'_C, R_{N\setminus C}) = med^{+}P_k(f)$.

Therefore, $f_k(R'_C, R_{N\setminus C})$ receives at least $n+1$ total votes for each $k \in \{1, 2\}$. Hence, the profitable deviation R'_C is credible and f is not immune to credible deviations. ■

Before we prove Propositions 4 and 5, we need some definitions, claims and lemmas.

Let $\mathcal{S}_{B_k}$ be the set of all unidimensional (strict) single-peaked preferences on B_k (for a formal definition see, for example, Definition 6 when B is an integer interval).

Definition 10. Let f be a generalized Condorcet winner rule. For any $k \in \mathcal{K}$, define $F_k : \left(\mathcal{S}_{B_k}\right)^n \to B_k$ such that for any $\widetilde{R}_N \in \left(\mathcal{S}_{B_k}\right)^n$, $F_k(\widetilde{R}_N) = f_k(R_N)$ for any $R_N \in \mathcal{S}^n$ such that $\tau_k(R_i) = \tau(\widetilde{R}_i)$ for any $i \in N$.

Note that F_k is well-defined since f_k is tops-only and any R_N as defined will work. Moreover, F_k is a unidimensional generalized Condorcet winner rule as in Definition 7.

Definition 11. Let f be a generalized Condorcet winner rule, R'_C be a profitable deviation of C against R_N for f and let $k \in \mathcal{K}$ such that $f_k(R'_C, R_{N\setminus C}) \neq f_k(R_N)$. We say that agent $i \in C$ is losing according to R_i at $(R'_C, R_{N\setminus C})$ in dimension $k \in \mathcal{K}$ if

$$\left(f_k(R_N), f_{\mathcal{K}\setminus\{k\}}(R'_C, R_{N\setminus C})\right) P_i f(R'_C, R_{N\setminus C}).$$

Define $L_k(f, R_{C'}, R_N) = \big\{i \in C : i \text{ is losing according to } R_i \text{ at } (R'_C, R_{N\setminus C}) \text{ in dimension } k \in \mathcal{K}\big\}$.

Similarly, we say that agent $i \in C$ is winning according to R_i at $(R'_C, R_{N\setminus C})$ in dimension $k \in \mathcal{K}$ if

$$f(R'_C, R_{N\setminus C}) P_i \left(f_k(R_N), f_{\mathcal{K}\setminus\{k\}}(R'_C, R_{N\setminus C})\right).$$

Define $W_k(f, R_{C'}, R_N) = \big\{i \in C : i \text{ is winning according to } R_i \text{ at } (R'_C, R_{N\setminus C}) \text{ in dimension } k \in \mathcal{K}\big\}$.

In short, when no confusion may arise, we will say that i is losing or winning in dimension k and we denote the sets of losers and winners as L_k and W_k, respectively. Note that L_k and W_k are a partition of C by definition since preferences are strict. That is, $L_k \cup W_k = C$ and $L_k \cap W_k = \varnothing$.

Claim 1. *Let $i \in C \subseteq N$, $R_i \in \mathcal{S}$, $x \in B$. Define $\widehat{R}_i \in \mathcal{S}$ such that (1) $\tau(\widehat{R}_i) = \tau(R_i)$, (2) for any $k \in \mathcal{K}$, for any $z_k, w_k \in B_k$, for $x_{\mathcal{K}\setminus\{k\}}$,*

$$\left(z_k, x_{\mathcal{K}\setminus\{k\}}\right) \widehat{P}_i \left(w_k, x_{\mathcal{K}\setminus\{k\}}\right) \Leftrightarrow \left(z_k, x_{\mathcal{K}\setminus\{k\}}\right) P_i \left(w_k, x_{\mathcal{K}\setminus\{k\}}\right), \quad \textit{and}$$

(3) for any $k \in \mathcal{K}$, for any $z_{\mathcal{K}\setminus\{k\}}, w_{\mathcal{K}\setminus\{k\}} \in B_{\mathcal{K}\setminus\{k\}}$, for any $y_k, v_k \in B_k$,

$$\left(y_k, z_{\mathcal{K}\setminus\{k\}}\right) \widehat{P}_i \left(v_k, z_{\mathcal{K}\setminus\{k\}}\right) \Leftrightarrow \left(y_k, w_{\mathcal{K}\setminus\{k\}}\right) \widehat{P}_i \left(v_k, w_{\mathcal{K}\setminus\{k\}}\right).$$

We leave the proof to the reader. Note however, that $\widehat{R}_i$ exists and additively representable separable preferences work. A separable preference (as defined in Le Breton and Sen, 1999) induces the same ordering over dimension j for any alternative. Claim 1 says that starting from any preference relation R and any alternative x, we can construct a separable preference where the ordering in each dimension is the one induced by R over that dimension relative to alternative x.

Claim 2. *Let f be a generalized Condorcet winner rule and let R'_C be a profitable deviation of C against R_N for f. Then:*

(i) *For any agent $i \in C$, there exists $k \in \mathcal{K}$ for which $f_k(R'_C, R_{N|C}) \neq f_k(R_N)$ such that $f(R'_C, R_{N\setminus C})P_i \left(f_k(R_N), f_{\mathcal{K}\setminus\{k\}}(R'_C, R_{N\setminus C})\right)$.*

(ii) *If agent $i \in C$ is such that $f(R'_C, R_{N\setminus C})P_i\left(f_k(R_N), f_{\mathcal{K}\setminus\{k\}}(R'_C, R_{N\setminus C})\right)$ for some k, then there exists another agent $j \in C$ such that $\left(f_k(R_N), f_{\mathcal{K}\setminus\{k\}}(R'_C, R_{N\setminus C})\right) P_j f(R'_C, R_{N\setminus C})$.*

(iii) *For any $k \in \mathcal{K}$, there is an agent $i \in C$ such that $\left(f_k(R_N), f_{\mathcal{K}\setminus\{k\}}(R'_C, R_{N\setminus C})\right)P_i f(R'_C, R_{N\setminus C})$.*

In words: (i) any agent in the profitable deviating coalition is winning in some dimension, (ii) if an agent in the profitable deviating coalition is winning in some dimension, there is another agent in the deviating coalition losing in the same dimension. Part (iii) ensures that there is at least a losing agent in each dimension.

We now turn to stating Lemmas 2 and 3.

Lemma 2. *Let f be a generalized Condorcet winner rule and let R'_C be a profitable deviation of C against R_N for f. There exist $k, k' \in \mathcal{K}$ and $i, j \in C$ such that the following hold: $\left(f_k(R_N), f_{\mathcal{K}\setminus\{k\}}(R'_C, R_{N\setminus C})\right)\ P_i f(R'_C, R_{N\setminus C})$, $\left(f_{k'}(R_N), f_{K\setminus\{k'\}}(R'_C, R_{N\setminus C})\right) P_j f(R'_C, R_{N\setminus C})$, $\left(f_{k'}(R'_C, R_{N\setminus C}), f_{K\setminus\{k'\}}(R'_C, R_{N\setminus C})\right)P_i \left(f_{k'}(R_N), f_{K\setminus\{k'\}}(R'_C, R_{N\setminus C})\right)$, and $\left(f_k(R'_C, R_{N\setminus C}), f_{\mathcal{K}\setminus\{k\}}(R'_C, R_{N\setminus C})\right)P_j \left(f_k(R_N), f_{\mathcal{K}\setminus\{k\}}(R'_C, R_{N\setminus C})\right)$.*

Lemma 2 formalizes the idea that in any profitable deviation there exist two agents and two dimensions where agents are exchanging roles and helping each other.

Lemma 3. *Let f be a generalized Condorcet winner rule, R'_C be a profitable deviation of C against R_N for f and k be such that $P_k(f)$ is degenerate. If there exists $i \in C$ winning at $(R'_C, R_{N\setminus C})$ in dimension k, then, the profitable deviation R'_C is not credible.*

Lemma 3 says that a credible profitable deviation relative to a generalized Condorcet winner rule cannot involve an agent winning in a degenerate dimension.

The proofs of Claim 2, and those of Lemmas 2 and 3 are in Appendix. Finally, we can provide a proof for Propositions 4 and 5.

Proof of Proposition 4. Consider f as in the statement. By contradiction, let $R_N \in \mathcal{S}^n$, $C \subseteq N$, and $R'_C \in \mathcal{S}^c$ be a profitable deviation of C against R_N. By Lemma 2, there exist $k, k' \in \mathcal{K}$ and $i, j \in C$ such that

$$\left(f_k(R_N), f_{\mathcal{K}\setminus\{k\}}(R'_C, R_{N\setminus C})\right) P_i \left(f_k(R'_C, R_{N\setminus C}), f_{\mathcal{K}\setminus\{k\}}(R'_C, R_{N\setminus C})\right),$$
$$\left(f_k(R'_C, R_{N\setminus C}), f_{\mathcal{K}\setminus\{k\}}(R'_C, R_{N\setminus C})\right) P_j \left(f_k(R_N), f_{\mathcal{K}\setminus\{k\}}(R'_C, R_{N\setminus C})\right),$$
$$\left(f_{k'}(R'_C, R_{N\setminus C}), f_{K\setminus\{k'\}}(R'_C, R_{N\setminus C})\right) P_i \left(f_{k'}(R_N), f_{K\setminus\{k'\}}(R'_C, R_{N\setminus C})\right), \quad \text{and}$$
$$\left(f_{k'}(R_N), f_{K\setminus\{k'\}}(R'_C, R_{N\setminus C})\right) P_j \left(f_{k'}(R'_C, R_{N\setminus C}), f_{K\setminus\{k'\}}(R'_C, R_{N\setminus C})\right).$$

By hypothesis, either $P_k(f)$ or $P_{k'}(f)$ is degenerate (or both).[10] By Lemma 3, R'_C is not a credible profitable deviation. This ends the proof. ■

Proof of Proposition 5. To prove the first statement, let f be a generalized Condorcet winner rule with lists of parameters, denoted in each dimension k as $p_k^- \leq p_k^+$, and such that they are non-degenerate in exactly two dimensions. To prove that f is immune to credible deviations, let $R_N \in \mathcal{S}^3$, $C \subseteq N = \{1, 2, 3\}$, and $R'_C \in \mathcal{S}^c$ be a profitable deviation of C against R_N. If the profitable deviation is such that there is an agent that is winning in a dimension k for which $P_k(f)$ is degenerate. Thus, by Lemma 3, R'_C would not be a credible profitable deviation. Now assume that agents are winning at $(R'_C, R_{N\setminus C})$ in dimensions k for which $P_k(f)$ is not degenerate. By Lemma 2, there must be agents in C winning at $(R'_C, R_{N\setminus C})$ in the two dimensions with non-degenerate list of parameters $P_k(f)$. Without loss of generality, assume that these dimensions are 1 and 2. By strategy-proofness, C has at least two agents. Without loss of generality, by anonymity and Lemma 2, suppose that agents 1 and 2 belong to C and that agent 1 is winning while agent 2 is losing at $(R'_C, R_{N\setminus C})$ in dimension 1 and the opposite holds in dimension 2. Consider the following two cases depending on the size of the deviating coalition.

<u>Case 1</u>: $C = \{1, 2\}$.
Since agent 3's preferences are fixed, we can assume now that we have 3 fixed parameters in each dimension, two of them different: p_k^-, p_k^+, and $\tau(R_3)$. Consider dimension 1. First, observe that

$$f_1(R_N), f_1(R'_{\{1,2\}}, R_3) \in \left[\min\left\{\tau_1(R_3), p_1^-\right\}, \max\left\{\tau_1(R_3), p_1^+\right\}\right].$$

Since agent 1 is winning and agent 2 is losing at $(R'_{\{1,2\}}, R_3)$ in dimension 1, then $\tau_1(R_1)$ must be strictly placed on one side of $f_1(R_N)$, while $\tau_1(R_2)$ must be weakly placed on the other side of $f_1(R_N)$. Moreover, for both $i \in \{1, 2\}$, $\tau_1(R'_i)$ must be strictly on the same side of $f_1(R_N)$ and in fact on the same side as $\tau_1(R_1)$ is. Then, note that agent 2 announcing $\overline{R}_2$ such that $\tau_1(\overline{R}_2) = \tau_1(R_2)$ and for each $k \in \mathcal{K} \setminus \{1\}$, $\tau_k(\overline{R}_2) = \tau_k(R'_2)$ would be winning at $(\overline{R}_2, R'_1, R_3)$ in dimension 1 and by separability she would be better off $f(\overline{R}_2, R'_1, R_3) P_2 f(R'_{\{1,2\}}, R_3)$ which means that $R'_{\{1,2\}}$ is not a credible profitable deviation.

<u>Case 2</u>: $C = \{1, 2, 3\}$.
Since $R'_N \in \mathcal{S}^n$ is a profitable deviation of C against R_N, agent 3 is winning at R'_N in some dimension. Without loss of generality, suppose that agent 3 is winning at R'_N in dimension 1 (otherwise, by anonymity a similar argument would apply). We already assumed that agent 1 is winning and agent 2 is losing at R'_N in dimension 1. If $f_1(R_N) \in \left(p_1^-, p_1^+\right)$, then $\tau_1(R_2) = f_1(R_N)$ since agent 2 is losing at R'_N, and also both $\tau_1(R_1)$ and $\tau_1(R_3)$ must be strictly on a different side of $f_1(R_N)$. But then, R'_C would not be a profitable deviation where agents 1 and 3 are winning at R'_N in dimension 1. Thus, $f_1(R_N) \leq p_1^-$ or $f_1(R_N) \geq p_1^+$. Suppose $f_1(R_N) \leq p_1^-$ (a symmetric argument would apply for the other case). Observe first that for each $i \in \{1, 2, 3\}$, $\tau_1(R_i) \leq f_1(R_N)$. Since agent 2 is the only agent losing at R'_N in dimension 1, $\tau_1(R_2) = f_1(R_N)$ and $\tau_1(R_1), \tau_1(R_3) < f_1(R_N)$, and thus $f_1(R'_C, R_{N\setminus C}) < f_1(R_N)$. Then, note that agent 2 announcing $\overline{R}_2$ such that $\tau_1(\overline{R}_2) = \tau_1(R_2)$ and for each $k \in \mathcal{K} \setminus \{1\}$, $\tau_k(\overline{R}_2) = \tau_k(R'_2)$ would be winning at $(\overline{R}_2, R'_{\{1,3\}})$ in dimension 1 and by separability she would be better off $f(\overline{R}_2, R'_{\{1,3\}}) P_2 f(R'_{\{1,2,3\}})$ which means that $R'_{\{1,2,3\}}$ is not a credible profitable deviation.

For all cases we obtain that the profitable deviation is not credible which shows the first statement in the proposition.

To prove the second statement, let f be a generalized Condorcet winner rule with lists of parameters, denoted in each dimension k as $p_k^- \leq p_k^+$, and such that they are non-degenerate in at least three dimensions. Assume they are dimensions 1, 2, and 3. To prove that f is not immune to credible deviations, we provide an example of a credible profitable deviation against a profile. In any profile we will define the preferences of each agent in N concerning dimensions different from 1, 2, and 3 to be the same and with top at some point x_k in B_k, $k \in \mathcal{K} \setminus \{1, 2, 3\}$.

Let $R_N \in \mathcal{S}^3$ be as follows in dimensions 1, 2, and 3: define the preferences of agent 1 such that $\tau(R_1) = \left(p_1^+, p_2^-, p_3^-\right)$ and $\left(p_1^+, p_2^+, p_3^+, x_{\mathcal{K}\setminus\{1,2,3\}}\right) P_1 \left(p_1^-, p_2^-, p_3^-, x_{\mathcal{K}\setminus\{1,2,3\}}\right)$, the preferences of agent 2 such that $\tau(R_2) = \left(p_1^-, p_2^+, p_3^-\right)$ and $\left(p_1^+, p_2^+, p_3^+, x_{\mathcal{K}\setminus\{1,2,3\}}\right) P_2 \left(p_1^-, p_2^-, p_3^-, x_{\mathcal{K}\setminus\{1,2,3\}}\right)$, and the preferences of agent 3 such that $\tau(R_3) = \left(p_1^-, p_2^-, p_3^+\right)$ and $\left(p_1^+, p_2^+, p_3^+, x_{\mathcal{K}\setminus\{1,2,3\}}\right) P_3 \left(p_1^-, p_2^-, p_3^-, x_{\mathcal{K}\setminus\{1,2,3\}}\right)$. Observe that $f(R_N) = \left(p_1^-, p_2^-, p_3^-, x_{\mathcal{K}\setminus\{1,2,3\}}\right)$. Let $C = N$, and R'_N such that each agent $i \in C$, $\tau(R'_i) = \left(p_1^+, p_2^+, p_3^+\right)$. Since $f(R'_N) = \left(p_1^+, p_2^+, p_3^+\right)$, R'_N is a profitable deviation of C against R_N. Finally, R'_N is credible since no agent can change the outcome by a unilateral deviation.

This ends the proof. ■

4. Some alternative formulations of credibility, and their consequences

We believe that our definition of a credible deviation is quite attractive. But others could be conceivable, and in this section we shall discuss other possible proposals, and relate them to ours.

We could take several directions to obtain alternative definitions of immunity. First, we concentrate on varying the notion of credibility to which we devote more attention.

To favor the comparison, let us go back to the interpretation of credibility that we already proposed after Definition 4. A profitable deviation by C from $R_N = (R_C, R_{N\setminus C})$ is credible if R'_C is a Nash equilibrium of the game among agents in C, when these agents strategies are their admissible preferences and the outcome function is $f(\cdot, R_{N\setminus C})$. Starting from this, we shall discuss, then, three possible variants of the credibility concept.

The first variant will be one where, instead of letting agents in C to have any choice of preferences as a strategy, we restrict them to either use strategy R'_i or to revert to strategy R_i. The resulting notion of a credible deviation will be weaker than ours. However, we will show that the set of rules that are immune to credible deviations will be the same (after a minimal qualification) under either definition. This is expressed in Proposition 6.

A second variant will require that in order to be (extensively) credible, the deviation R'_C should be a Nash equilibrium for the

[10] For $n = 2$ any list of parameters is degenerate in all dimensions since all parameters take the same value.

Table 2
The set $\mathcal{S}$ of all separable preferences when $K = 2$.

R^1	R^2	R^3	R^4	$R^{1'}$	$R^{2'}$	$R^{3'}$	$R^{4'}$
$\varnothing$	o_1	o_1	$\{o_1, o_2\}$	$\varnothing$	o_2	o_2	$\{o_1, o_2\}$
o_1	$\varnothing$	$\{o_1, o_2\}$	o_1	o_2	$\varnothing$	$\{o_1, o_2\}$	o_2
o_2	$\{o_1, o_2\}$	$\varnothing$	o_2	o_1	$\{o_1, o_2\}$	$\varnothing$	o_1
$\{o_1, o_2\}$	o_2	o_2	$\varnothing$	$\{o_1, o_2\}$	o_1	o_1	$\varnothing$

game where all agents (whether or not they are part of C) can play any preference, and f is the outcome function. If the initial function f is assumed to be strategy-proof (an assumption that we do not need under our original definition), then again the set of rules immune to credible deviations will still be the same under either definition (see Proposition 7). However, the equivalence is not true if our f function is not a priory restricted to be strategy-proof, as shown in Example 3.

A third variant of our definition of credibility would result from simply changing our original one, but ask the deviation to be a strong Nash, rather than a Nash equilibrium. The rationale for such proposal would be to allow for several agents to coordinate when defecting from the agreed upon joint manipulation. We will show that under this definition, all of the rules we consider will be immune to credible deviations (see Proposition 8).[11]

We now present formal arguments to make the preceding discussion more precise. We also state some results whose proofs are included in Appendix.

Definition 12. Let f be a social choice function on $\mathcal{U}^n$. Let $R_N \in \mathcal{U}^n$ and $C \subseteq N$. We say that $R'_C \in \mathcal{U}^c$ a profitable deviation of C against R_N is (type 1) **credible** if $f(R'_C, R_{N|C}) R_i f(R_i, R'_{C\setminus\{i\}}, R_{N\setminus C})$ for all $i \in C$. A social choice function f on $\mathcal{U}^n$ is **immune to** (type 1) **credible deviations** if for any $R_N \in \mathcal{U}^n$, any $C \subseteq N$, there is no (type 1) credible profitable deviation of C against R_N.

Proposition 6. *Any social choice function f on $\mathcal{U}^n$ is immune to credible deviations if and only if f is immune to (type 1) credible deviations.*

Definition 13. Let f be a social choice function on $\mathcal{U}^n$. Let $R_N \in \mathcal{U}^n$ and $C \subseteq N$. We say that $R'_C \in \mathcal{U}^c$ a profitable deviation of C against R_N is (type 2) **credible** if $f(R'_C, R_{N|C}) R_i f(\bar{R}_i, R'_{C\setminus\{i\}}, R_{N\setminus(C\cup\{i\})})$ for all $i \in N$ and all $\bar{R}_i \in \mathcal{U}$. A social choice function f on $\mathcal{U}^n$ is **immune to** (type 2) **credible deviations** if for any $R_N \in \mathcal{U}^n$, any $C \subseteq N$, there is no (type 2) credible profitable deviation of C against R_N.

Proposition 7. *Any strategy-proof social choice function f on $\mathcal{U}^n$ is immune to credible deviations if and only if f is immune to (type 2) credible deviations.*

The following example shows that the latter immunity concept does not imply strategy-proofness. Therefore, the concept may be useful to apply in contexts where strategy-proofness is not to be expected, but one may still be interested in discussing the diversity of manipulative actions by groups of voters.

Example 3. Immunity to (type 2) credible deviations does not imply strategy-proofness under appropriately restricted domains.[12]

Let $K = 2$, $N = \{1, 2\}$, and for $i \in N$, let the set of admissible preferences for both agents be $\mathcal{S}$, that is, preferences defined as in Table 2. Consider the social choice function f defined as in Table 3.

Table 3
The social choice function f defined on $\mathcal{S}^2$.

f	$R_2^{1'}, R_2^{2'}$	$R_2^{3'}, R_2^{4'}$	R_2^3, R_2^4	R_2^1, R_2^2
R_1^3, R_1^4	o_1	o_2	o_1	o_1
R_1^1, R_1^2	o_2	o_1	o_1	o_1
$R_1^{1'}, R_1^{2'}$	o_2	o_2	o_1	o_2
$R_1^{3'}, R_1^{4'}$	o_2	o_2	o_2	o_1

Note that in the direct revelation game induced by this social choice function, no agent has a dominant strategy. Hence, the rule is not strategy-proof (thus, violating immunity to both credible and (type 1) credible deviations). Also notice that the grand coalition has no profitable deviation. Hence, all profitable deviations involve a single agent, and for each one of them, the remaining agent can respond with a new profitable deviation. Hence, the social choice function is immune to (type 2) credible deviations, even if not strategy-proof.

Definition 14. Let f be a social choice function on $\mathcal{U}^n$. Let $R_N \in \mathcal{U}^n$ and $C \subseteq N$. We say that $R'_C \in \mathcal{U}^c$ a profitable deviation of C against R_N is **strongly credible** if $f(R'_C, R_{N|C}) R_i f(\bar{R}_S, R'_{C\setminus S}, R_{N\setminus C})$ for all $S \subseteq C$, for all $\bar{R}_S \in \mathcal{U}^s$ and for some $i \in S$. A social choice function f on $\mathcal{U}^n$ is **immune to strongly credible deviations** if for any $R_N \in \mathcal{U}^n$, any $C \subseteq N$, there is no strongly credible profitable deviation of C against R_N.

Proposition 8. *All generalized Condorcet winner rules are immune to strongly credible deviations.*

Let us say that we are aware that the idea of credibility may have other expressions. As already noted in the Introduction, Bernheim et al. (1987) introduced the concept of coalition-proof Nash equilibrium, and Peleg and Sudhölter (1999) studied its application to the set of strategy-proof voting rules in multidimensional single-peaked preference domains. This equilibrium notion turns out not to be discriminating, since all Generalized Median Voter Schemes satisfy it. That conclusion is the same as the one we obtain under our notion of strong credibility (Proposition 8) although they obtain it under weaker assumptions: we assume anonymity and ontoness. Peleg (1998) and Peleg and Sudhölter (1999) proposed the notion of strong coalition-proofness. The latter paper comments, by means of an example, that not all strategy-proof rules defined on the multidimensional single-peaked domain satisfy this condition. One can show that, in spite of its complicated formulation, their notion of strong coalition-proofness essentially boils down to requiring our basic notion of immunity to credible deviations in our context.[13] Hence, we can read their example as an announcement that there is room for the analysis we have just provided, identifying and characterizing those functions among generalized Condorcet winner rules that satisfy these conditions and those that do not. Another related paper is due to Serizawa (2006), who defines an immunity notion in the line in our paper that only considers profitable deviations by pairs of agents.

[11] The same will hold if instead of allowing agents to use any preferences, they are only assumed to use their true and the manipulative one.

[12] This example can be straightforwardly generalized when agents have different sets of preferences: Let $\mathcal{R}_1 = \{R^1, R^2, R^3, R^4\}$, $\mathcal{R}_2 = \{R^{1'}, R^{2'}, R^{3'}, R^{4'}\}$ in Table 2, and f defined by the first two rows and columns in Table 3.

[13] For a formal proof of this, see the proof of Proposition 9 in Appendix.

Alternative definitions of credibility could also be obtained by weakening the notion of a profitable deviation allowing some, though not all, deviators to report their original preference. In this case, the new set of profitable deviations would be larger than ours. However, the resulting immunity notion would be equivalent to the one we use (see Definition 5 or Definition 14). Also notice that if indifferences are allowed one could also consider additional types of deviations and allowing some of the agents in the deviating coalition to weakly gain by deviating jointly. We do not go further in this direction because in our application all preferences are strict (see Serizawa, 2006).

5. Final remarks

We have studied the incentives of groups of agents to cooperate in manipulating social choice functions, by formalizing different notions of credibility, and characterized subclasses of strategy-proof rules that may be immune to credible manipulations by groups in multidimensional single-peaked preference domains.

The voting rules we have identified are interesting in several respects.

One interesting aspect is efficiency. It is clear that strategy-proof rules cannot be fully efficient unless they satisfy a strong notion of group strategy-proofness. Yet, those that satisfy our intermediate property have the interesting feature that any departure from their prescribed outcomes leading to an efficient one would not be credible. Thus, they are, in that sense, efficient up to credibility constraints.

Another interesting conclusion of our analysis is that those rules that imply extreme distributions of voting power are immune to credible deviations from truth-telling. One could think that this distribution is uneven or unfair. However, the class of Generalized Condorcet winner rules that are obtained when the definitional parameters are concentrated in a single point do coincide, in each dimension, with those characterized by Thomson (1993, 1999) as being the only methods that satisfy an attractive normative property. His property, that Thomson calls "welfare domination under preference replacement", requires that when one agent changes preferences and modifies the social outcome, all other agents' welfare must change in the same direction. Hence, we not only found exactly what are the conditions that allow immunity, but also discovered that they may be partially justified in terms of pre-existing normative concepts.

Finally, let us acknowledge that the treatment of strategic considerations by the different agents is somewhat asymmetric. Indeed, groups are allowed to form in order to manipulate, but our main concept of credibility only considers single-agent non-cooperative departures from cooperative agreements, à la Nash. This invites for further reflection regarding these and other issues of coalition formation, that we hope to keep developing in further work.

Acknowledgments

We thank the two referees, Matthew Jackson, and Albin Erlanson for their helpful suggestions. S. Barberà acknowledges financial support through grants ECO2014-53052-P and SGR2014-515, and Severo Ochoa Programme for Centres of Excellence in R&D (SEV-2015-0563). D. Berga acknowledges the support from grants ECO2013-45395-R and 2014-SGR-1360. B. Moreno acknowledges financial support from grants SEJ4941, SEJ-5980, and ECO2014-53767. D. Berga and B. Moreno thank the MOMA network under project ECO2014-57673-REDT.

Appendix

We first present the proofs of Claim 2, and Lemmas 2 and 3 used in the proofs of Propositions 4 and 5 in Section 3.2. For that, we need Definition 15.

Definition 15. Let $i \in N$, $R_i \in \mathcal{S}$, $k \in \mathcal{K}$, let $\mathcal{S}_k(R_i) = \left\{R_i^k \in \mathcal{S}_{B_k} : \tau(R_i^k) = \tau_k(R_i)\right\}$ where $\mathcal{S}_{B_k}$ is the set of all unidimensional (strict) single-peaked preferences on B_k and $\tau(R_i^k)$ is the best alternative of R_i^k in B_k.

We call a $R_i^k \in \mathcal{S}_k(R_i) \subseteq \mathcal{S}_{B_k}$ a unidimensional single-peaked preference on B_k induced from R_i.

Proof of Claim 2. Let f be a generalized Condorcet winner rule and let R'_C be a profitable deviation of C against R_N for f. To prove part (i), define $\overline{\mathcal{K}} = \left\{k \in \mathcal{K} : f_k(R'_C, R_{N|C}) \neq f_k(R_N)\right\}$. Without loss of generality, let $\overline{\mathcal{K}} = \{1, \ldots, \overline{k}\}$, $\overline{k}$ denoting its cardinality. By contradiction, suppose that there exists $i \in C$ such that for any $k \in \overline{\mathcal{K}}$, agent i is not winning according to R_i at $(R'_C, R_{N\setminus C})$ in dimension k. That is, agent i is losing at $(R'_C, R_{N\setminus C})$ in dimension k: $\left(f_k(R_N), f_{\mathcal{K}\setminus\{k\}}(R'_C, R_{N\setminus C})\right) P_i f(R'_C, R_{N\setminus C})$. Let $\widehat{R}_i$ be defined by Claim 1 given $x = f(R'_C, R_{N\setminus C})$. By part (2) of Claim 1, we have that for any $k \in \overline{\mathcal{K}}$, $\left(f_k(R_N), f_{\mathcal{K}\setminus\{k\}}(R'_C, R_{N\setminus C})\right) \widehat{P}_i f(R'_C, R_{N\setminus C})$. In particular, for $k = 1$, we obtain $\left(f_1(R_N), f_{\mathcal{K}\setminus\{1\}}(R'_C, R_{N\setminus C})\right) \widehat{P}_i f(R'_C, R_{N\setminus C})$.

Now, proceed as follows: define $\overline{k} - 1$ steps and in each step $t \in \{1, \ldots, \overline{k} - 1\}$ replace $f_{t+1}(R'_C, R_{N\setminus C})$ by $f_{t+1}(R_N)$ in $\big(f_t(R_N), f_{\{1,\ldots,t-1\}}(R_N), f_{\overline{\mathcal{K}}\setminus\{t+1,\ldots,\overline{k}\}}(R'_C, R_{N\setminus C}), f_{\mathcal{K}\setminus\overline{\mathcal{K}}}(R'_C, R_{N\setminus C})\big)$.

By transitivity of preferences we will obtain that $f(R_N)\widehat{P}_i f(R'_C, R_{N\setminus C})$ which will be the desired contradiction. To check the contradiction, note that by part (1) in Claim 1 and by tops-onliness of the generalized Condorcet winner rule f, if R'_C is a profitable deviation of C against R_N, R'_C is also a profitable deviation of C against $(\widehat{R}_C, R_{N\setminus C})$.

Consider the first step, $t = 1$, and change $f_2(R'_C, R_{N\setminus C})$ by $f_2(R_N)$. By part (3) of Claim 1 applied for $k = 2$, $z_{\mathcal{K}\setminus\{2\}} = (f_1(R_N), f_{\mathcal{K}\setminus\{1,2\}}(R'_C, R_{N\setminus C}))$ and $w_{\mathcal{K}\setminus\{2\}} = f_{\mathcal{K}\setminus\{2\}}(R'_C, R_{N\setminus C})$ we obtain the following:

$$
\begin{aligned}
&(f_2(R_N), z_{\mathcal{K}\setminus\{2\}})\widehat{P}_i(f_2(R'_C, R_{N\setminus C}), z_{\mathcal{K}\setminus\{2\}}) \\
&\quad\Leftrightarrow (f_2(R_N), w_{\mathcal{K}\setminus\{2\}})\widehat{P}_i(f_2(R'_C, R_{N\setminus C}), w_{\mathcal{K}\setminus\{2\}}).
\end{aligned}
$$

Note that $(f_2(R_N), w_{\mathcal{K}\setminus\{2\}})\widehat{P}_i(f_2(R'_C, R_{N\setminus C}), w_{\mathcal{K}\setminus\{2\}})$ (equivalently, $(f_2(R_N), f_{\mathcal{K}\setminus\{2\}}(R'_C, R_{N\setminus C}))\ \widehat{P}_i f(R'_C, R_{N\setminus C}))$ holds since by part (2) of Claim 1 applied for $k = 2$, $z_2 = f_2(R_N)$, $w_2 = f_2(R'_C, R_{N\setminus C})$ we obtain the following:

$$
\begin{aligned}
&(f_2(R_N), f_{\mathcal{K}\setminus\{2\}}(R'_C, R_{N\setminus C}))\widehat{P}_i f(R'_C, R_{N\setminus C}) \\
&\quad\Leftrightarrow (f_2(R_N), f_{\mathcal{K}\setminus\{2\}}(R'_C, R_{N\setminus C}))P_i f(R'_C, R_{N\setminus C}).
\end{aligned}
$$

And moreover, the latter strictness preference relationship holds by hypothesis at the beginning of this proof. Thus, we get that

$$
\begin{aligned}
&(f_2(R_N), f_1(R_N), f_{\mathcal{K}\setminus\{1,2\}}(R'_C, R_{N\setminus C})) \\
&\quad\times \widehat{P}_i(f_2(R'_C, R_{N\setminus C}), f_1(R_N), f_{\mathcal{K}\setminus\{1,2\}}(R'_C, R_{N\setminus C})).
\end{aligned}
$$

Repeating exactly the same argument for any $t \in \{2, \ldots, \overline{k}-1\}$, we will obtain our desired contradiction: $f(R_N)\widehat{P}_i f(R'_C, R_{N\setminus C})$ which ends the proof of part (i).

To prove part (ii), by contradiction let $i \in C$ be winning at $(R'_C, R_{N\setminus C})$ in some dimension k and suppose that for any other agent $j \in C \setminus \{i\}$, j is also winning at $(R'_C, R_{N\setminus C})$ in k. That is, suppose that for any $j \in C\setminus\{i\}$, $f(R'_C, R_{N\setminus C})P_j \left(f_k(R_N), f_{\mathcal{K}\setminus\{k\}}(R'_C, R_{N\setminus C})\right)$ holds.

We now proceed to define unidimensional single-peaked preferences for dimension k using Definition 15 as follows: for any

$j \in N \setminus C$, let $R_j^k \in \mathcal{S}_k(R_j)$; for any $i \in C$, let $R_i'^k \in \mathcal{S}_k(R_i')$, and for any $l \in C$ let $R_l^k \in \mathcal{S}_k(R_l)$ such that for any $x_k, y_k \in B_k$,

$$x_k P_l^k y_k \Leftrightarrow (x_k, f_{\mathcal{K}\setminus\{k\}}(R_C', R_{N\setminus C}))P_l(y_k, f_{\mathcal{K}\setminus\{k\}}(R_C', R_{N\setminus C})).$$

Note that R_l^k is well-defined. Let $x_k = f_k(R_C', R_{N\setminus C})$ and $y_k = f_k(R_N)$. Then, we obtain by definition of R_l^k that for any $j \in C \setminus \{i\}$, $f_k(R_C', R_{N\setminus C})P_l^k f_k(R_N)$. By Definition 10, by any $l \in C$, $F_k(R_C'^k, R_{N\setminus C}^k)P_l^k$ $F_k(R_N^k)$. Observe that this last expression is a contradiction to group strategy-proofness of generalized Condorcet winner rules. This ends the proof of part (ii).

To prove part (iii), by contradiction suppose that there exists $k \in \mathcal{K}$ such that for any agent $i \in C, f(R_C', R_{N\setminus C})P_i\big(f_k(R_N), f_{\mathcal{K}\setminus\{k\}}(R_C', R_{N\setminus C})\big)$ (*). Now, by Definition 15, let $R_i^k, R_i'^k$ be the induced one dimensional preferences on B_k as follows:

For $i \in N \setminus C$, take any $R_i^k \in \mathcal{S}_k(R_i)$. For $i \in C$, take $R_i'^k \in \mathcal{S}_k(R_i')$. For $i \in C$, take $R_i^k \in \mathcal{S}_k(R_i)$ such that $x = f(R_C', R_{N\setminus C})$. That is, for any $x_k, y_k \in B_k$,

$$y_k P_i^k x_k \Leftrightarrow \left(y_k, f_{\mathcal{K}\setminus\{k\}}(R_C', R_{N\setminus C})\right) P_i^k \left(x_k, f_{\mathcal{K}\setminus\{k\}}(R_C', R_{N\setminus C})\right).$$

Observe that by Definition 10, $F_k(R_N^k) = f_k(R_N)$ and $F_k(R_C'^k, R_{N\setminus C}^k) = f_k(R_C', R_{N\setminus C})$. Thus, combining the latter equality with expression (*), we obtain that for any k and any agent $i \in C$, $F_k(R_C', R_{N\setminus C})$ $P_i^k F_k(R_N)$ holds. And this contradicts that F_k be a group strategy-proof. This end the proof of part (iii). ■

Proof of Lemma 2. Let f be a generalized Condorcet winner rule and let R_C' be a profitable deviation of C against R_N for f. Consider, for each dimension k, the sets $L_k \equiv L_k(f, R_{C'}, R_N)$ and $W_k \equiv W_k(f, R_{C'}, R_N)$, that is, the partition of agents in C defined by the ones winning and the ones losing at $(R_C', R_{N\setminus C})$ in dimension k (see Definition 11). By part (i) of Claim 2 for any agent in C there is $k \in \mathcal{K}$ such that $i \in W_k$. By part (ii) of Claim 2 there exists a different agent j in C such that $j \in L_k$. Again, by part (i) of Claim 2, $j \in W_{k'}$ for some $k' \in \mathcal{K} \setminus \{k\}$. If some agent $i \in W_k$ belongs to $L_{k'}$, the result holds.

Otherwise, suppose that for any $i \in W_k, i \notin L_{k'}$. Since $L_k \cup W_k = L_{k'} \cup W_{k'} = C$, then $W_k \subsetneqq W_{k'}$ and $L_{k'} \subsetneqq L_k$ (note that $L_{k'} \neq L_k$ since $j \in L_k \cap W_{k'}$). Take now an agent $l \in L_{k'}$ which exists by part (ii) of Claim 2. By part (i) of Claim 2, $l \in W_{k''}$ for some $k'' \in \mathcal{K}$. If some agents $i \in W_{k'}, i \in L_{k''}$, then the result holds.

Otherwise, suppose that for any $i \in W_{k'}, i \notin L_{k''}$. Since $L_{k'} \cup W_{k'} = L_{k''} \cup W_{k''} = C$, then $W_{k'} \subsetneqq W_{k''}$ and $L_{k''} \subsetneqq L_{k'} \subsetneqq L_k$ (note that $L_{k''} \neq L_{k'}$ since $j \in L_{k'} \cap W_{k''}$). Since there are a finite number of agents in L_k and a finite number of k, we will obtain the result for some k. Otherwise, there would be a k^* such that $L_{k*} = \varnothing$ which is a contradiction to part (iii) of Claim 2. ■

Proof of Lemma 3. Let f be a generalized Condorcet winner rule, R_C' be a profitable deviation of C against R_N for f and k be such that $P_k(f)$ is degenerate. Since $i \in W_k$, then $f_k(R_C', R_{N\setminus C}) \neq f_k(R_N)$. Consider two cases.

In the first case, in R_N each agent's k-dimensional top is placed on the same side of the parameters' unique position. Without loss of generality, assume that the k-dimensional tops are to the left of the parameter. Note that $f_k(R_N)$ is the top closest to the single parameter. Since $i \in W_k$, then $f_k(R_C', R_{N\setminus C}) < f_k(R_N)$. Otherwise, if $f_k(R_C', R_{N\setminus C}) > f_k(R_N)$, by single-peakedness,[14] we obtain $\big(f_k(R_N), f_{\mathcal{K}\setminus\{k\}}(R_C', R_{N\setminus C})\big)$ P_i $\big(f_k(R_C', R_{N\setminus C}), f_{\mathcal{K}\setminus\{k\}}(R_C', R_{N\setminus C})\big)$ contradicting that $i \in W_k$. Then, by definition of f, for any $l \in N$ such that $\tau_k(R_l) = f_k(R_N)$ then $l \in C$ and $\tau_k(R_l') \leq f_k(R_C', R_{N\setminus C}) < f_k(R_N)$. And, by single-peakedness, for any such $l \in C$, $\big(f_k(R_N), f_{\mathcal{K}\setminus\{k\}}(R_C', R_{N\setminus C})\big)$ P_l $\big(f_k(R_C', R_{N\setminus C}), f_{\mathcal{K}\setminus\{k\}}(R_C', R_{N\setminus C})\big)$. Take any of such agents, say $j \in C$. Take profile $(R_C', R_{N\setminus C})$ and let $j \in C$ announce $\overline{R}_j$ such that $\tau_k(\overline{R}_j)$ $= \tau_k(R_j)$ and for each $k' \in \mathcal{K} \setminus \{k\}$, $\tau_{k'}(\overline{R}_j) = \tau_{k'}(R_j')$. By definition of f, $f_k(\overline{R}_j, R'_{C\setminus\{j\}}, R_{N\setminus C}) = f_k(R_N)$ and $f_{\mathcal{K}\setminus\{k\}}(\overline{R}_j, R'_{C\setminus\{j\}}, R_{N\setminus C}) = f_{\mathcal{K}\setminus\{k\}}$ $(R_C', R_{N\setminus C})$. Therefore, $f(\overline{R}_j, R'_{C\setminus\{j\}}, R_{N\setminus C})P_j f(R_C', R_{N\setminus C})$ by single-peakedness, which means that R_C' is not a credible profitable deviation.

In the second case, on both sides of the single parameter there is at least one agent's top given R_N. Thus, $f_k(R_N)$ is the single parameter. Suppose, without loss of generality, that $f_k(R_C', R_{N\setminus C}) < f_k(R_N)$. Observe that by definition of f, there exists $j \in C$ such that $\tau_k(R_j') < f_k(R_C', R_{N\setminus C}) < \tau_k(R_j)$. Then, by single-peakedness, for agent $j \in C$, $\big(f_k(R_N), f_{\mathcal{K}\setminus\{k\}}(R_C', R_{N\setminus C})\big)$ $P_j\big(f_k(R_C', R_{N\setminus C}), f_{\mathcal{K}\setminus\{k\}}(R_C', R_{N\setminus C})\big)$. As in the above case, take profile $(R_C', R_{N\setminus C})$ and let agent $j \in C$ announce $\overline{R}_j$ such that $\tau_k(\overline{R}_j) = \tau_k(R_j)$ and for each $k' \in \mathcal{K} \setminus \{k\}$, $\tau_{k'}(\overline{R}_j) = \tau_{k'}(R_j')$. By definition of f, $f_k(\overline{R}_j, R'_{C\setminus\{j\}}, R_{N\setminus C}) = f_k(R_N)$ and $f_{\mathcal{K}\setminus\{k\}}(\overline{R}_j, R'_{C\setminus\{j\}}, R_{N\setminus C}) = f_{\mathcal{K}\setminus\{k\}}(R_C', R_{N\setminus C})$. Thus, $f(\overline{R}_j, R'_{C\setminus\{j\}}, R_{N\setminus C})P_j f(R_C', R_{N\setminus C})$ by single-peakedness meaning that R_C' is not a credible profitable deviation. This ends the proof. ■

We now prove Propositions 6, 7, and 8 stated in Section 4.

Proof of Proposition 6. By definition immunity to (type 1) credible deviations implies immunity to credible deviations. To prove the converse, let $R_N \in \mathcal{U}^n$, $C \subseteq N$, and $R_C' \in \mathcal{U}^c$ be a profitable deviation of C against R_N. Suppose that for all $i \in C, f(R_C', R_{N|C})$ $R_i f(R_i, R'_{C\setminus\{i\}}, R_{N\setminus C})$. By Lemma 1 f is strategy-proof, thus $f(R_C', R_{N|C})$ $I_i\, f(R_i, R'_{C\setminus\{i\}}, R_{N\setminus C})$ for all $i \in C$. By immunity to credible deviations, there exists $i \in C$ such that $f(\overline{R}_i, R'_{C\setminus\{i\}}, R_{N\setminus C})P_i f(R_C', R_{N|C})$ for some $\overline{R}_i \in \mathcal{U}$. By these two facts, for some $i \in C, f(\overline{R}_i, R'_{C\setminus\{i\}}, R_{N\setminus C})P_i f(R_i, R'_{C\setminus\{i\}}, R_{N\setminus C})$ which contradicts strategy-proofness. ■

Proof of Proposition 7. By definition immunity to credible deviations implies immunity to (type 2) credible deviations. To prove the converse, let $R_N \in \mathcal{U}^n$, $C \subseteq N$, and $R_C' \in \mathcal{U}^c$ be a profitable deviation of C against R_N. Suppose that for all $i \in C$, $f(R_C', R_{N\setminus C})R_i f(\overline{R}_i, R'_{C\setminus\{i\}}, R_{N\setminus C})$ for all $\overline{R}_i \in \mathcal{U}$. Thus, $f(R_C', R_{N\setminus C})$ $R_i f(R_i, R'_{C\setminus\{i\}}, R_{N\setminus C})$ for all $i \in C$. By immunity to (type 2) credible deviations, there exists $i \in N$ such that $f(\overline{R}_i, R'_{C\setminus\{i\}}, R_{N\setminus(C\cup\{i\})})P_i f(R_C', R_{N\setminus C})$ for some $\overline{R}_i \in \mathcal{U}$. Suppose that agent $i \in C$. Since f is strategy-proof, $f(R_C', R_{N\setminus C})I_i f(R_i, R'_{C\setminus\{i\}}, R_{N\setminus C})$ for all $i \in C$. Then, we have that for some $i \in C, f(\overline{R}_i, R'_{C\setminus\{i\}}, R_{N\setminus C})P_i f(R_i, R'_{C\setminus\{i\}}, R_{N\setminus C})$ which contradicts strategy-proofness. Therefore, it must be that agent $i \in N \setminus C$. Thus, $f(R_i, R_C', R_{N\setminus(C\cup\{i\})})R_i f(\overline{R}_i, R_C', R_{N\setminus(C\cup\{i\})})$ by strategy-proofness, and therefore we obtain $f(R_i, R_C', R_{N\setminus(C\cup\{i\})})$ $P_i f(R_C', R_{N\setminus C})$ for some $i \in N \setminus C$ which contradicts that f is single-valued and $R_i' = R_i$. ■

Proof of Proposition 8. Let f be a generalized Condorcet winner rule with lists of parameters, denoted in each dimension k as $p_k^- \leq p_k^+$. To prove that f is immune to strongly credible deviations, let $R_N \in \mathcal{S}^n$, $C \subseteq N$, and $R_C' \in \mathcal{S}^c$ be a profitable deviation of C against R_N. Since R_C' is a profitable deviation, by part (i) of Claim 2, there must exist at least one dimension k in which $f_k(R_N) \neq f_k(R_C', R_{N\setminus C})$ and some agent $i \in C$ is winning. By part (ii) of Claim 2, there is an agent $j \in C \setminus \{i\}$ who is losing in that dimension k. Let $\overline{C}$ and $\widetilde{C}$ be a partition of C such that $\overline{C} = \{i \in C$: is winning in dimension $k\}$ and $\widetilde{C} = \{j \in C$: is losing in dimension $k\}$. Suppose, without loss of generality, that $f_k(R_C', R_{N\setminus C}) < f_k(R_N)$. By definition of $\overline{C}$, for any $i \in \overline{C}$, $\tau_k(R_i) < f_k(R_N)$. By definition of $\widetilde{C}$, $\tau_k(R_j) > f_k(R_C', R_{N\setminus C})$. We distinguish two cases:

[14] In this proof and that of Proposition 8, when we say "by single-peakedness" we mean both, unidimensional single-peaked on B_k and multidimensional single-peaked on B.

Case 1. For any $j \in \widetilde{C}$, $\tau_k(R_j) \leq f_k(R_N)$. Since $f_k(R'_C, R_{N\setminus C}) < f_k(R_N)$, then for any $l \in C$, $\tau_k(R'_l) \leq f_k(R_N)$. Also, it must happen that for some $j \in \widetilde{C}$, $\tau_k(R_j) = f_k(R_N)$. Let $S = \{j \in \widetilde{C} : \tau_k(R_j) = f_k(R_N)\}$. Then for any $l \in S$, let $\overline{R}_l$ be such that $\tau_k(\overline{R}_l) = \tau_k(R_l)$ and $\tau_{k'}(\overline{R}_l) = \tau_{k'}(R'_l)$ for any $k' \in \mathcal{K} \setminus \{k\}$. Thus $f_k(\overline{R}_S, R'_{C\setminus S}, R_{N\setminus C}) = f_k(R_N)$ and by single-peakedness $f(\overline{R}_S, R'_{C\setminus S}, R_{N\setminus C}) P_l f(R'_C, R_{N\setminus C})$ for any $l \in S$, which means that R'_C is not strongly credible.

Case 2. For some $j \in \widetilde{C}$, $\tau_k(R_j) > f_k(R_N)$. Let $S = \{j \in \widetilde{C} : \tau(R_j) \geq f_k(R_N)$ and $\tau_k(R'_l) \neq \tau_k(R_l)\}$. Since $f_k(R'_C, R_{N\setminus C}) < f_k(R_N)$, S is not empty. Then for any $l \in S$, let $\overline{R}_l$ be such that $\tau_k(\overline{R}_l) = \tau_k(R_l)$ and $\tau_{k'}(\overline{R}_l) = \tau_{k'}(R'_l)$ for any $k' \in \mathcal{K} \setminus \{k\}$. By definition of f, $f_k(\overline{R}_S, R'_{C\setminus S}, R_{N\setminus C}) = f_k(R_N)$ and for any $k' \in \mathcal{K} \setminus \{k\}$, $f_{k'}(\overline{R}_S, R'_{C\setminus S}, R_{N\setminus C}) = f_{k'}(R'_C, R_{N\setminus C})$. Thus, by single-peakedness, $f(\overline{R}_S, R'_{C\setminus S}, R_{N\setminus C}) P_l f(R'_C, R_{N\setminus C})$ for any $l \in S$, meaning that R'_C is not strongly credible. ■

As promised in Section 4, we now prove the equivalence in our context between strong coalition-proofness as defined by Peleg and Sudhölter (1999) and our immunity to credible deviations of f. To do so, we first introduce the notion of strong coalition-proofness for revelation games.

A *revelation game in strategic form* is a system $G(f, \widehat{R}_N) = (N, A, (\mathcal{U})_{i\in N}, f, \widehat{R}_N)$ where N is the set of players, A is the set of outcomes, $\mathcal{U}$ the (non-empty) set of strategies of each agent (the set of all possible preferences as defined in Section 2), f is a function from preference profiles to A, and $\widehat{R}_N$ is a specific profile of preferences. Then, given a game $G(f, \widehat{R}_N)$, a coalition $C \subseteq N$, $C \neq \varnothing$, and a profile $\widetilde{R}_N \in \mathcal{U}^n$, *the reduced game of* $G(f, \widehat{R}_N)$ *with respect to* C *and* $\widetilde{R}_N$ is the game in strategic form $G^{C,\widetilde{R}_N}(f^{\widetilde{R}_{N\setminus C}}, \widehat{R}_C) = (C, A, (\mathcal{U})_{i\in C}, f^{\widetilde{R}_{N\setminus C}}, \widehat{R}_C)$ where C is the set of players, A is the set of outcomes, $\mathcal{U}$ the (non-empty) set of strategies of each agent, $f^{\widetilde{R}_{N\setminus C}} : \mathcal{U}^c \to A$ is a function such that $f^{\widetilde{R}_{N\setminus C}}(R_C) = f(R_C, \widetilde{R}_{N\setminus C})$ for all $R_C \in \mathcal{U}^c$, and $\widehat{R}_C$ is the profile of preferences.

Definition 16. Let $G(f, \widehat{R}_N) = (N, (\mathcal{U})_{i\in N}, f, \widehat{R}_N)$ be a revelation strategic game. We say that $\widetilde{R}_N \in \mathcal{U}^n$ is a strong coalition-proof Nash equilibrium if (1) $\widetilde{R}_N$ is a Nash equilibrium of $G(f, \widehat{R}_N)$; and (2) for every $C \subseteq N$, $C \neq \varnothing$, and every Nash equilibrium R'_C of $G^{C,\widetilde{R}_N}(f^{\widetilde{R}_{N\setminus C}}, \widehat{R}_C)$, there exists $i \in C$ such that $f^{\widetilde{R}_{N\setminus C}}(\widetilde{R}_C) = f(\widetilde{R}_N)\widehat{R}_i f(R'_C, \widetilde{R}_{N\setminus C}) = f^{\widetilde{R}_{N\setminus C}}(R'_C)$.

Definition 17. A social choice function f on $\mathcal{U}^n$ is **strong coalition-proof** if for any $R_N \in \mathcal{U}^n$ truth telling is a strong coalition-proof Nash equilibrium of $G(f, R_N)$.

The strategy space for each agent is $\mathcal{U}$, and the outcome function is f.

Proposition 9. *Any strong coalition-proof social choice function f is immune to credible deviations. The converse holds when best deviations exist.*

Proof of Proposition 9. By contradiction, suppose that f is not immune to credible deviations. That is, there exist R_N, and a deviation R'_C of C against R_N that is profitable, $f(R'_C, R_{N\setminus C}) P_i f(R_N)$, and credible, $f(R'_C, R_{N\setminus C}) R_i f(\overline{R}_i, R'_{C\setminus\{i\}}, R_{N\setminus C})$ for all $\overline{R}_i \in \mathcal{U}$ and for all $i \in C$. The latter implies that R'_C is a Nash equilibrium of the reduced game G^{C,R_N}. Since f is strong coalition-proof, then R_N is a strong coalition-proof Nash equilibrium of $G(f, R_N)$. By condition (2) in Definition 16 there must exist an agent $i \in C$ such that $f(R_N) R_i f(R'_C, R_{N\setminus C})$ contradicting that R'_C is a profitable deviation of C against R_N.

To show the converse, assume that there exist best deviations. Let f be immune to credible deviations and take any R_N. We have to show that truth telling is a strong coalition-proof Nash equilibrium of $G(f, R_N)$. By Lemma 1, f is strategy-proof. Therefore, R_N is a Nash equilibrium of $G(f, R_N)$ (part (1) in Definition 16). Take $C \subseteq N$, $C \neq \varnothing$, and any Nash equilibrium R'_C of G^{C,R_N}. Since f is immune to credible deviations, no profitable deviation can be credible, that is, there exists an agent $i \in C$ such that $f(R_N) R_i f(R'_C, R_{N\setminus C})$.

References

Barberà, S., Berga, D., Moreno, B., 2010. Individual versus group strategy-proofness: when do they coincide? J. Econom. Theory 145, 1648–1674.

Barberà, S., Berga, D., Moreno, B., 2016. Group strategy-proofness in private good economies. Amer. Econ. Rev. 106, 1073–1099.

Barberà, S., Gul, F., Stacchetti, E., 1993. Generalized median voter schemes and committees. J. Econom. Theory 61, 262–289.

Barberà, S., Sonnenschein, H., Zhou, L., 1991. Voting by committees. Econometrica 59, 595–609.

Bernheim, B.D., Peleg, B., Whinston, M.D., 1987. Coalition-proof Nash equilibria. I. Concepts. J. Econom. Theory 42, 1–12.

Border, K., Jordan, J., 1983. Straightforward elections, unanimity and phantom voters. Rev. Econom. Stud. 50, 153–170.

Le Breton, M., Sen, A., 1999. Separable preferences, decomposability and strategyproofness. Econometrica 67, 605–628.

Le Breton, M., Weymark, J., 1999. Strategy-proof social choice with continuous separable preferences. J. Math. Econom. 32, 47–85.

Le Breton, M., Zaporozhets, V., 2009. On the equivalence of coalitional and individual strategy-proofness properties. Soc. Choice Welf. 33, 287–309.

Moulin, H., 1980. On strategy-proofness and single peakedness. Pub Choice 35, 437–455.

Peleg, B., 1998. Almost all equilibiria in dominant strategies are coalition-proof. Econom. Lett. 60, 157–162.

Peleg, B., Sudhölter, P., 1999. Single-peakedness and coalition-proofness. Rev. Econ. Des. 4, 381–387.

Serizawa, S., 2006. Pairwise strategy-proofness and self-enforcing manipulation. Soc. Choice Welf. 26, 305–331.

Thomson, W., 1993. The replacement principle in public good economies with single-peaked preferences. Econom. Lett. 42, 31–36.

Thomson, W., 1999. Welfare-domination under preference-replacement: A survey and open questions. Soc. Choice Welf. 16, 373–394.