# On Societies Choosing Social Outcomes, and their Memberships: Strategy-proofness*

Gustavo Bergantiños†, Jordi Massó‡ and Alejandro Neme§

January 12, 2017

Abstract: We consider a society whose members have to choose not only an outcome from a given set of outcomes but also the subset of agents that will remain members of the society. We assume that each agent is indifferent between any two alternatives (pairs of final societies and outcomes) provided that the agent does not belong to any of the two final societies, regardless of the chosen outcome. Under this preference domain restriction we characterize the class of all strategy-proof, unanimous and outsider independent rules as the family of all serial dictator rules.

*We are grateful to two referees and an Associate Editor of this journal for very useful comments and suggestions that helped us to write a better version of the paper. The work of G. Bergantiños is partially supported by research grants ECO2014-52616-R from the Spanish Ministry of Science and Competitiveness, GRC 2015/014 from "Xunta de Galicia", and 19320/PI/14 from "Fundación Séneca de la Región de Murcia". J. Massó acknowledges financial support from the Spanish Ministry of Economy and Competitiveness, through the Severo Ochoa Programme for Centres of Excellence in R&D (SEV-2015-0563) and grant ECO2014-53051, and from the Generalitat de Catalunya, through grant SGR2014-515. The paper was partly written while J. Massó was visiting the Department of Economics at Stanford University; he wishes to acknowledge its hospitality as well as financial support from the Ministerio de Educación, Cultura y Deporte through project PR2015-00408. The work of A. Neme is partially supported by the Universidad Nacional de San Luis, through grant 319502, and by the Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET), through grant PIP 112-200801-00655.

†Department of Statistics and Operations Research. Universidade de Vigo. 36310, Vigo (Pontevedra), Spain. E-mail: gbergant@uvigo.es

‡Universitat Autònoma de Barcelona and Barcelona Graduate School of Economics. Departament d'Economia i d'Història Econòmica, Campus UAB, Edifici B. 08193, Bellaterra (Barcelona), Spain. E-mail: jordi.masso@uab.es

§Instituto de Matemática Aplicada de San Luis. Universidad Nacional de San Luis and CONICET. Ejército de los Andes 950. 5700, San Luis, Argentina. E-mail: aneme@unsl.edu.ar

Journal of Economic Literature Classification Number: D71.
Keywords: Strategy-proofness; Unanimity; Outsider independence.

# 1 Introduction

A classical social choice problem is the following. A society formed by a set of agents has to choose an outcome from a given set of outcomes. Since agents may have different preferences over outcomes, and it is desirable that the chosen outcome be perceived as a compromise among their potentially different preferences, they have to be asked about them. A social choice function (a rule) collects individual preferences and selects, in a systematic and known way, an outcome taking into account the profile of revealed preferences.

This classical approach assumes that the composition of the society is fixed and, in particular, independent of the chosen outcome. There are situations for which this assumption may not be appropriate. For instance, in the case of an excludable and costly public good, agents' preferences may depend on the level of the public good and on the size of the set of agents consuming (and contributing to finance) it. Also, when membership is voluntary in a double sense: no agent can be forced to belong to the final society and any agent can be part of it, if the agent whishes to be. A prototypical example of this class of problems is a political party, whose membership may depend on the positions the party takes on issues like the death penalty, abortion or the possibility of allowing a region of a country to become independent. Or when no agent can be forced to belong to the final society but to be a member requires some kind of approval of the current members. For example, a department whose members decide upon its new members but then a professor, already a member of the department, may start looking for a position elsewhere if he considers that the recruitment of the department has not been satisfactory to his standards; and this in turn might trigger further exits. Or finally, when no stability property is required since agents can simultaneously be forced to remaining and to be excluded from the final society. For example, extremely hierarchical societies like traditional families, religious orders or criminal organizations.[1] To be able to deal with such situations the classical social choice model has to be modified. Agents' preferences have to be extended to order pairs formed by the final society and the chosen outcome.

There is a large literature that has already considered explicitly the dependence of the

[1] All these examples will be included as particular instances of our model. However, the strong incentive requirement of strategy-proofness will be incompatible with any stability notion related to voluntary membership.

final society on its choices in specific settings, in terms of two issues: the voting methods under which members choose the outcome, and the timing under which members reconsider their membership. See for instance Roberts (1999), Demange and Wooders (2005), and Berga, Bergantiños, Massó and Neme (2006) for problems related to the club model; Barberà, Sonnenschein and Zhou (1991), Barberà, Maschler and Shalev (2001), and Berga, Bergantiños, Massó and Neme (2004, 2007) for a society choosing a subset of new members; and Jackson and Nicolò (2004) for a provision of excludable public goods when agents care also about the number of other consumers. In the last section of the paper, we comment with more detail some of this literature.

In this paper we look at the general setting without being specific about the two issues, and by not requiring *a priori* any stability property on the final society. We do that by considering that the set of alternatives are all pairs formed by a subset of the original society (an element in $2^N$, the subset of the set of agents $N$ that will remain in the society) and an outcome in the set $X$. Then, we assume that agents' preferences are defined over the set of alternatives $2^N \times X$ and satisfy two requirements. First, each agent has strict preferences between any two alternatives, provided he belongs to at least one of the two corresponding societies. Second, each agent is indifferent between any two alternatives, provided he is not a member of any of the two corresponding societies; namely, agents that do not belong to the final society do not care about neither its composition nor the chosen outcome.[2]

We consider rules that operate on this restricted domain of preference profiles by selecting, for each profile, an alternative (a final society and an outcome); that is, direct revelation mechanisms. Note that the alternative chosen by a rule at a preference profile may be the consequence of the application of a potentially complex mechanism, where agents' behavior are driven by (and linked to) their preferences over the set of alternatives.[3] We abstract from this, by focusing on direct revelation mechanisms. An agent that understands the effect of his revealed preference on the chosen alternative faces the strategic problem of selecting it. Depending on the rule under consideration, the agent may realize that the solution to this problem is ambiguous because it may depend on the agent's expectations that he has about the revealed preferences of the others, and in turn he may also realize that to formulate hypothesis about those revealed preferences require hypothesis about the others' expectations,

[2] See the last section of the paper for a discussion about the consequences of requiring stronger domain restrictions.

[3] For instance, complex and sequential algorithms defined for matching problems induce rules mapping preference profiles into alternatives.

and so on. Strategy-proof rules make all these considerations unnecessary since truthtelling is a weakly dominant strategy of the direct revelation mechanism at each profile; namely, each agent's decision problem is independent of the revealed preferences by the others, and truth-telling is an optimal decision. In addition to strategy-proofness, we will also consider two weak versions of efficiency and non-bossiness. A rule is unanimous if it always selects an alternative belonging to the set of common best alternatives, whenever this set is nonempty. A rule is outsider independent if it is invariant with respect to the change of preferences of an agent who is not a member of the two final societies.

Observe that the (natural) domain restriction under consideration requires that each agent $i \in N$ is indifferent among a large subset of alternatives, all those for which $i$ does not belong to their corresponding final societies; namely, $i$ is indifferent among all alternatives in the subset $2^{S_{-i}} \times X$, where $2^{S_{-i}}$ is the family of all subsets of $N$ that do not contain $i$. Hence, the set of individual preferences over which we want the rule to operate is far from being the universal domain of preferences over the set of alternatives. Thus, the Gibbard-Satterthwaite theorem (see Gibbard (1973) and Satterthwaite (1975)) does not apply and the goal of identifying all strategy-proof rules (or a tractable subclass) remains meaningful and interesting. We want to emphasize that the reason why our model is not a particular case of the classical social choice model, where one can directly apply the Gibbard-Satterthwaite theorem, is the specific domain restriction we are interested in. It follows from the particular indifferences admitted over the set $2^N \times X$ which are natural for settings where agents, to enjoy the effects of the chosen outcome, have to remain members of the final society and, at the same time, non final members do not care about the specific chosen outcome. Of course, without this kind of indifferences, the domain of preferences would be the universal domain and the Gibbard-Satterthwaite theorem would apply, precipitating dictatorship.

Our result, Theorem 1, characterizes the class of all strategy-proof, unanimous and outsider independent rules as the family of all serial dictator rules. A serial dictator rule, relative to an ordering of the agents, gives to the first agent the power to select his best alternative, and only if this agent has many indifferent alternatives at the top of his preference, the second agent in the order has the power to select his best alternative among those declared as being at the top and indifferent by the first agent, and proceeds similarly following the ordering of the agents. A serial dictator rule moves away from just dictatorship by using the loophole left by the potential indifferences, present in the domain, and it does so by allocating the power among agents to break the indifferences sequentially. Often, this can be done in a strategy-proof way and satisfying at the same time other desirable properties like weak notions of efficiency (unanimity), non arbitrariness (non-bossiness or order independence), or neutrality, consistency, and so on. Serial dictator rules have been characterized as the family of strategy-proof rules (satisfying in addition some other properties) in many different settings. See, for instance, Satterthwaite and Sonnenschein (1981), Svensson (1999), Papai (2001) and Bade (2015).

In a companion note (Bergantiños, Massó and Neme (2016)) we consider the same setting but assume that the preference profile is common knowledge (and hence, the strategic revelation of agents' preferences is not an issue) and focus on the properties of internal stability and consistency, which guarantee that the chosen alternative is indeed the final one in a double sense. Internal stability says that nobody can force an agent to remain in the society if the agent does not want to do so. Consistency says that if the rule would be applied again to the final society it would choose the same alternative, so there is no need to do so. We exhibit the difficulties of finding rules satisfying the two properties; however, we show that approval voting, adapted to our setting, not only satisfies internal stability and consistency but it also satisfies efficiency and neutrality.

The paper is organized as follows. In Section 2 we describe the model. Section 3 contains the definitions of the basic properties of rules that we will be interested in. In Section 4 we state, as Theorem 1, the characterization of the class of all strategy-proof, unanimous and outsider independent rules as the family of all serial dictator rules. Section 5 contains the proof of Theorem 1. Section 6 concludes with several final remarks.

# 2 Preliminaries

Let $N = \{1, \ldots, n\}$ be the set of *agents*, with $n \geq 2$, and let $X$ be the finite set of possible *outcomes*. We are interested in situations where some agents may not be part of the final society, perhaps as the consequence of the chosen outcome. To model such situations, let $A = 2^N \times X$ be the set of *alternatives* and assume that each $i \in N$ has preferences over $A$.[4] We will often use the notation $a$ for a generic $(S, x) \in A$; *i.e.*, $a \equiv (S, x)$, $a' \equiv (S', x')$, and so on. Let $R_i$ denote agent $i$'s (weak) *preference* over $A$, where for any pair $a, a' \in A$, $aR_ia'$ means that $i$ considers $a$ to be at least as good as $a'$. Let $P_i$ and $I_i$ denote the strict and indifference relations over $A$ induced by $R_i$, respectively; namely, for any pair $a, a' \in A$, $aP_ia'$ if and only if $aR_ia'$ and $\neg a'R_ia$, and $aI_ia'$ if and only if $aR_ia'$ and $a'R_ia$. We assume that

[4] Note that we are admitting the possibility that the society selects all outcomes with no agent in the final society; *i.e.*, for all $x \in X$, $(\varnothing, x) \in A$.

each $i$ does not care about all alternatives at which he does not belong to their corresponding final societies and $i$ is not indifferent between any pair of alternatives at which he belongs to at least one of the two corresponding final societies. Namely, we assume that $R_i$ satisfies the following two properties: for all $S, T \in 2^N$ and $x, y \in X$,

(P.1) if $i \notin S \cup T$, then $(S, x)\, I_i\, (T, y)$ ; and

(P.2) if $i \in S \cup T$ and $(S, x) \neq (T, y)$, then either $(S, x)\, P_i\, (T, y)$ or $(T, y)\, P_i\, (S, x)$ .

The fact that agents' preferences satisfy (P.1) is the reason why our model cannot mechanically be embedded into the classical model and a specific analysis is required. We see property (P.1) as being a natural assumption for our setting, and it is a critical requirement for our results to hold. Let $\mathcal{R}_i$ be the set of $i$'s preferences satisfying (P.1) and (P.2), and let $\mathcal{R} = \times_{i \in N} \mathcal{R}_i$ be the set of (preference) *profiles*. Given $S \subset N$, we denote a profile $R \in \mathcal{R}$ by $(R_S, R_{-S})$ where $R_S \in \times_{j \in S} \mathcal{R}_j \equiv \mathcal{R}_S$ and $R_{-S} \in \times_{j \in N \setminus S} \mathcal{R}_j \equiv \mathcal{R}_{-S}$. If $S = \{i\}$ or $S = \{i, j\}$, we write $(R_i, R_{-i})$ and $(R_i, R_j, R_{-i,j})$, respectively.

We denote the subset of alternatives with the property that $i$ is not a member of the corresponding final society by $[\varnothing]_i = \{(S, x) \in A \mid i \notin S\}$. By (P.1), $i$ is indifferent among them; *i.e.*,

$$[\varnothing]_i = \{a \in A \mid a I_i (\varnothing, x) \text{ for some } x \in X\} .$$

By (P.1), $(\varnothing, x) I_i (\varnothing, y)$ for all $x, y \in X$ and $[\varnothing]_i$ can be seen as the indifference class generated by the empty society. Observe that $[\varnothing]_i$ may be at the top of $i$'s preferences. With an abuse of notation we often treat, when listing a preference ordering, the indifference class $[\varnothing]_i$ as if it were an alternative; for instance, given $R_i$ and $a \in A$ we write $a R_i [\varnothing]_i$ to represent that $a R_i a'$ for all $a' \in [\varnothing]_i$ .

To clarify the model, we relate it with two of the examples used in the introduction. The set of initial members of the political party corresponds to the set of agents, the set of outcomes to the set of choices ($X$ could be written as $\{0, 1\}^3$ where for instance, $x = (0, 1, 0)$ would correspond to the choices of not supporting the death penalty, admitting abortion, and standing against the independence of the region) and the set $S$, if the chosen alternative is $(S, x)$, corresponds to the set of final members of the party that want to stay after it supports outcome $x$. Similarly, all professors in the department correspond to the agents, the set of outcomes $X$ to all subsets of hired candidates (again, an outcome $x \in X$ could be identified with $x = (x_1, \ldots, x_K) \in \{0, 1\}^K$, where $K$ is the number of candidates and $x_k = 1$ if and only if candidate $k$ is hired) and the set $S$, if the chosen alternative is $(S, x)$, corresponds to the set of professors who remain in the department after $x$ has been selected.

Given $A' \subseteq A$ and $R_i$, the *choice* of $i$ in $A'$ at $R_i$ is the set of best alternatives in $A'$ according to $R_i$; namely,

$$C(A', R_i) = \{a \in A' \mid aR_i a' \text{ for all } a' \in A'\}.$$

Since the set $2^N \times X$ is finite, the choice set is well-defined and non-empty.

We define three different sets that we will use later on, all related to $R_i$. The *top* of $R_i$, denoted by $\tau(R_i)$, is the set of all best alternatives according to $R_i$; namely,

$$\tau(R_i) = \{a \in A \mid aR_i a' \text{ for all } a' \in A\}.$$

Of course, $C(A, R_i) = \tau(R_i)$. The *lower contour set of* $R_i$ *at* $a$, denoted by $L(a, R_i)$, is the set of alternatives that are at least as bad as $a$ according to $R_i$; namely,

$$L(a, R_i) = \{a' \in A \mid aR_i a'\}.$$

The *upper contour set of* $R_i$ *at* $a$, denoted by $U(a, R_i)$, is the set of alternatives that are at least as good as $a$ according to $R_i$; namely,

$$U(a, R_i) = \{a' \in A \mid a'R_i a\}.$$

A *rule* is a social choice function $f : \mathcal{R} \to A$ selecting, for each profile $R \in \mathcal{R}$, an alternative $f(R) \in A$. To be explicit about the two components of the alternative chosen by $f$ at $R$, we will often write $f(R)$ as $(f_N(R), f_X(R))$, where $f_N(R) \in 2^N$ and $f_X(R) \in X$.

Notice that we allow rules that choose outcomes $(f_X(R))$ and agents $(f_N(R))$ simultaneously but also rules that first choose outcomes (or agents) and later agents (or outcomes).

# 3 Basic properties of rules

We present three properties that a rule $f : \mathcal{R} \to A$ may satisfy, and that we will use in our characterization result. The first one imposes a minimal efficiency requirement at each profile.

A rule is unanimous if it selects an alternative in the intersection of all tops, whenever this intersection is nonempty.

UNANIMITY For all $R \in \mathcal{R}$ such that $\bigcap_{i \in N} \tau(R_i) \neq \varnothing$, $f(R) \in \bigcap_{i \in N} \tau(R_i)$.

The next two properties impose conditions by comparing the alternatives chosen by the rule at two different profiles. A rule is strategy-proof if it is always in the best interest of the agents to reveal their preferences truthfully; namely, truth-telling is a weakly dominant strategy in the direct revelation game induced by the rule.

STRATEGY-PROOFNESS For all $R \in \mathcal{R}$, all $i \in N$ and all $R'_i \in \mathcal{R}_i$,

$$f(R_i, R_{-i})\, R_i f(R'_i, R_{-i}).$$

If otherwise, *i.e.,* $f(R'_i, R_{-i}) P_i f(R_i, R_{-i})$, we will say that $i$ *manipulates* $f$ at $(R_i, R_{-i})$ via $R'_i$.

Outsider independence requires that any modification of the preferences of an agent who was not a member of the society and remain nonmember once she changes her preferences, does not modify the alternative chosen.

OUTSIDER INDEPENDENCE For all $R \in \mathcal{R}$, all $i \in N$ and all $R'_i \in \mathcal{R}_i$ such that $i \notin f_N(R) \cup f_N(R'_i, R_{-i})$, $f(R'_i, R_{-i}) = f(R)$.

The notion of outsider independence was used in the context of public goods with exclusion by Deb and Razzolini (1999), under the name of non-bossiness of excluded individuals, and by Jackson and Nicolò (2004). It constitutes a weakening of non-bossiness, a property introduced by Satterthwaite and Sonnenschein (1981). Different variants of non-bossiness have been intensively used in the literature. Often, to eliminate arbitrary (and hence, difficult to describe) rules in axiomatic characterizations where strategy-proofness plays a salient role. Thomson (2016) contains a systematic analysis of non-bossiness by giving alternative definitions and interpretations of it, and by relating them to a large family of allocation problems. Outsider independence requires that the rule does not change *only* after a change of preferences of an agent that is not a member of the two final societies, for which he is indifferent to. We think that this is a natural requirement in our setting; otherwise, by changing his preferences the agent could lead the rule to select different alternatives, inducing welfare changes to the other agents but not to himself.[5]

[5] See the final section of the paper for an indication of the class of strategy-proof and unanimous rules that are not outsider independent.

# 4 Characterization result

In this section we state the result of the paper characterizing the class of all strategy-proof, unanimous and outsider independent rules.[6] This class coincides with the family of all serial dictator rules. To define a serial dictator rule we need some preliminaries. Let $\pi : N \to \{1, \ldots, n\}$ be an ordering (one-to-one mapping) of the set of agents. Given $i \in N$, $\pi(i)$ is the position in the order assigned to $i$ after applying $\pi$ to $N$. The set of all orderings $\pi : N \to \{1, \ldots, n\}$ will be denoted by $\Pi$. For $\pi \in \Pi$ and $1 \leq k \leq n$, we write $\pi_k$ to denote the agent $\pi^{-1}(k)$; *i.e.*, the $k$-th agent in the ordering $\pi$.

A serial dictator rule induced by $\pi \in \Pi$ and $x \in X$, denoted by $f^{\pi,x}$, proceeds (in up to $n$ steps) as follows. Fix a profile $R \in \mathcal{R}$ and look for any alternative $(S_1, x_1)$ in the best indifference class of agent $\pi_1$, the first agent in the ordering induced by $\pi$. If $\pi_1 \in S_1$, set $f^{\pi,x}(R) = (S_1, x_1)$. Otherwise, look for any alternative $(S_2, x_2)$ in the best indifference class of agent $\pi_2$, the second agent in the ordering induced by $\pi$, only among those classes satisfying the property that $\pi_1 \notin S_2$, If $\pi_2 \in S_2$, set $f^{\pi,x}(R) = (S_2, x_2)$. Otherwise, proceed similarly until the $n-$th step, if reached, by looking for any alternative $(S_n, x_n)$ in the best indifference class of agent $\pi_n$, the last in the ordering induced by $\pi$, only among those classes satisfying the property that for each $k \in \{1, \ldots, n-1\}$, $\pi_k \notin S_n$. If $\pi_n \in S_n$, set $f^{\pi,x}(R) = (S_n, x_n)$. Otherwise, and since no agent wants to stay in the society whatever element of $X$ is selected, set $f^{\pi,x}(R) = (\varnothing, x)$. So, $x$ plays the role of the residual outcome only when no agent wants to stay in the society under any circumstance.

We now define a *serial dictator rule* formally. Fix $\pi \in \Pi$ and $x \in X$, and let $R \in \mathcal{R}$ be a profile. Define $f^{\pi,x}(R)$ recursively, as follows.
<u>Step 1</u>. Let $A_1 = A$. Consider two cases:

1. $|C(A_1, R_{\pi_1})| = 1$. Then, $C(A_1, R_{\pi_1}) = \tau(R_{\pi_1})$. Set $(S_1, x_1) = C(A_1, R_{\pi_1})$ and observe that $\pi_1 \in S_1$. Define
$$f^{\pi,x}(R) = (S_1, x_1).$$
2. $|C(A_1, R_{\pi_1})| > 1$. Then, $C(A_1, R_{\pi_1}) = \{(S, x') \in A \mid \pi_1 \notin S \text{ and } x' \in X\}$. Go to Step 2.

We now define Step $k$ $(1 < k \leq n)$, assuming that at Step $k-1$, $\left|C\left(A_{k-1}, R_{\pi_{k-1}}\right)\right| > 1$.
<u>Step $k$</u>. Let $A_k = C(A_{k-1}, R_{\pi_{k-1}})$. Consider two cases.

[6] Observe again that the preferences we are considering satisfy (P.1) and hence, rules do not operate on the universal domain of preferences over $A$. Thus, the Gibbard-Satterthwaite theorem can not be applied.

1. $|C(A_k, R_{\pi_k})| = 1$. Then, $C(A_k, R_{\pi_k}) = \tau(R_{\pi_k})$. Set $(S_k, x_k) = C(A_k, R_{\pi_k})$ and observe that $\pi_k \in S_k$. Define
$$f^{\pi,x}(R) = (S_k, x_k).$$

2. $|C(A_k, R_{\pi_k})| > 1$.

   (a) If $k < n$, then $C(A_k, R_{\pi_k}) = \{(S, x') \in A \mid \pi_i \notin S \text{ for all } i \leq k \text{ and } x' \in X\}$. Go to Step $k+1$.

   (b) If $k = n$, then $C(A_n, R_{\pi_n}) = \{(\varnothing, x') \in A \mid x' \in X\}$. Define
$$f^{\pi,x}(R) = (\varnothing, x).$$

Example 1 below illustrates this procedure.

**Example 1** Let $N = \{1, 2\}$ and $X = \{a, b, c\}$ be respectively the set of agents and the set of outcomes, and consider the ordering $\pi$, where $\pi_1 = 1$ and $\pi_2 = 2$, and $x = a$. We apply the serial dictator rule $f^{\pi,a}$ to the following four preference profiles, where we give the list of the alternatives in decreasing order from the top and we only order the alternatives needed to compute $f^{\pi,a}$ at the four profiles.

| $R_1$ | $R_1'$ | $R_2$ | $R_2'$ |
|---|---|---|---|
| $(N, b)$ | $\{(S, y) \in A \mid 1 \notin S, y \in X\}$ | $(N, a)$ | $(N, a)$ |
| | | $(N, b)$ | $(N, b)$ |
| | | $(\{2\}, c)$ | $\{(S, y) \in A \mid 2 \notin S, y \in X\}$ |

Then,

$$\begin{aligned}
f^{\pi,a}(R_1, R_2) &= (N, b), \\
f^{\pi,a}(R_1, R_2') &= (N, b), \\
f^{\pi,a}(R_1', R_2) &= (\{2\}, c), \text{ and} \\
f^{\pi,a}(R_1', R_2') &= (\varnothing, a).
\end{aligned}$$
□

We are now ready to state Theorem 1, the characterization of the class of all strategy-proof, unanimous and outsider independent rules as the family of all serial dictator rules. Section 5 contains the proof of Theorem 1 and three examples of rules indicating the independence of the three properties used in the characterization.

**Theorem 1** *Assume $|X| \geq 3$. A rule $f : \mathcal{R} \to A$ is strategy-proof, unanimous and outsider independent if and only if $f$ is a serial dictator rule for some ordering $\pi \in \Pi$ and alternative $x \in X$.*

# 5 Proof of Theorem 1

We start by presenting an additional notion, two properties that a rule may satisfy and a sketch of the proof that follows. Given $f : \mathcal{R} \to A$, the *option set* of $i \in N$ at $R_{-i} \in \mathcal{R}_{-i}$, denoted by $o_i(R_{-i})$, is the set of alternatives that may be chosen by $f$ when the other agents declare the subprofile $R_{-i}$; namely,

$$o_i(R_{-i}) = \{a \in A \mid a = f(R_i, R_{-i}) \text{ for some } R_i \in \mathcal{R}_i\}.$$

Notice that the option set of $i$ at $R_{-i}$ does not depend on $R_i$.

A rule is efficient if it always selects a Pareto optimal allocation.

Efficiency For each $R \in \mathcal{R}$ there is no $a \in A$ with the property that $aR_i f(R)$ for all $i \in N$ and $aP_j f(R)$ for some $j \in N$.

Monotonicity requires that the chosen alternative at a profile is still selected at a new profile if the alternative improves in the ordering of an agent.

Monotonicity For all $R \in \mathcal{R}$, all $i \in N$ and all $R'_i \in \mathcal{R}_i$ such that $L(f(R), R_i) \subset L(f(R), R'_i)$, $f(R) = f(R'_i, R_{-i})$.

Since the set of indifferent alternatives for an agent coincides in all of his preferences, monotonicity could be reformulated in an equivalent way by stating that for all $R \in \mathcal{R}$, all $i \in N$ and all $R'_i \in \mathcal{R}_i$ such that $U(f(R), R_i) \supset U(f(R), R'_i)$, $f(R) = f(R'_i, R_{-i})$.

The proof that, for any $\pi$ and $x$, the serial dictator rule $f^{\pi,x}$ is strategy-proof, unanimous and outsider independent is easy. The main idea of the proof of the other implication is as follows. We first show that any strategy-proof, unanimous and outsider independent rule is efficient and monotonic; moreover, at every profile $R$, the rule selects the alternative that is simultaneously the best alternative on the option set of each agent $i$ at $R_{-i}$. These three facts will be useful later on. The main step of the proof is to construct from $f$, and for every subset of agents $N^* \subseteq N$, a rule $g$ on the set of strict preferences over the set of outcomes $X$ *only*, which depends on $N^*$. Since $|X| \geq 3$ (here is when this assumption plays a crucial role) and $g$ is onto (because $f$ is unanimous), by the Gibbard-Satterthwaite theorem, $g$ is dictatorial; denote by $d(N^*)$ its dictator. The remainder of proof consists of two last steps (the structure of the options sets plays an important role here). First, a preliminary extension in which we show that $f$ has to be also dictatorial on a subdomain of profiles (over $A$) related with the universal domain of preferences over $X$ (which depends on $N^*$) under which $d(N^*)$ is the dictator of $g$. Second, we obtain the series of dictators by applying the above result sequentially to $N^* = N$, and setting $\pi_1 = d(N)$, to $N^* = N \backslash \{\pi_1\}$, and setting $\pi_2 = d(N \backslash \{\pi_1\})$, and so on. Finally, the default outcome $x$, needed to define a serial dictator rule, is obtained by looking at the outcome chosen by $f$ (together with the empty society) at any profile $R$ for which $\tau(R_i) = [\varnothing]_i$ for all $i \in N$.

We proceed formally by presenting some lemmata that will be used in the proof.

**Lemma 1** *Let $f : \mathcal{R} \to A$ be a strategy-proof, unanimous and outsider independent rule. Then, the following hold.*

(1) *$f$ satisfies monotonicity.*

(2) *$f$ satisfies efficiency.*

(3) *For all $R \in \mathcal{R}$ and $i \in N$, $f(R) = C(o_i(R_{-i}), R_i)$.*

**Proof** Assume $f : \mathcal{R} \to A$ is strategy-proof, unanimous and outsider independent. We prove the three statements.

(1) Suppose $R \in \mathcal{R}$, $i \in N$, and $R'_i \in \mathcal{R}_i$ are such that $L(f(R), R_i) \subset L(f(R), R'_i)$ and $f(R) \neq f(R'_i, R_{-i})$. Three cases are possible:

1. $f(R) P_i f(R'_i, R_{-i})$. Since $L(f(R), R_i) \subset L(f(R), R'_i)$, $f(R'_i, R_{-i}) \in L(f(R), R'_i)$ and hence $f(R) P'_i f(R'_i, R_{-i})$. Thus, $i$ manipulates $f$ at $(R'_i, R_{-i})$ via $R_i$, which contradicts strategy-proofness.

2. $f(R'_i, R_{-i}) P_i f(R)$. Similarly, this contradicts strategy-proofness of $f$ since $i$ manipulates $f$ at $R$ via $R'_i$.

3. $f(R'_i, R_{-i}) I_i f(R)$. Then, by (P.2), $i \notin f_N(R'_i, R_{-i}) \cup f_N(R)$. By outsider independence, $f(R'_i, R_{-i}) = f(R)$ which is a contradiction.

(2) Suppose $f$ is not efficient. Namely, there exist $R \in \mathcal{R}$ and $a \in A$ such that $aR_i f(R)$ for all $i \in N$ and $aP_j f(R)$ for some $j \in N$. Let $R' \in \mathcal{R}$ be such that for each $k \in N$, $\tau(R'_k) = \{a' \in A \mid a' I_k a\}$ and orders the rest of alternatives as $R_k$ does. Consider the profile $(R'_1, R_{-1}) \in \mathcal{R}$ and suppose that $f(R'_1, R_{-1}) \neq f(R)$. If $f(R'_1, R_{-1}) I_1 f(R)$ then $1 \notin f_N(R'_1, R_{-1}) \cup f_N(R)$, but this contradicts outsider independence. If $f(R'_1, R_{-1}) P_1 f(R)$ then $f$ is not strategy-proof. If $f(R) P_1 f(R'_1, R_{-1})$ then $f(R) P'_1 f(R'_1, R_{-1})$, which means that 1 manipulates $f$ at $(R'_1, R_{-1})$ via $R_1$. Repeating this argument sequentially for agents $k = 2, \ldots, n$ we obtain that $f(R') = f(R)$. However, by unanimity, $f(R') \in \bigcap\limits_{k \in N} \tau(R'_k)$. Since $f(R') = f(R)$, $f(R)$ can not be dominated by $a$, implying that $f$ is efficient.

(3) Let $R \in \mathcal{R}$ and $i \in N$ be arbitrary and consider $a = (S, x) \in C\left(o_i\left(R_{-i}\right), R_i\right)$. Since by definition $f(R) \in o_i(R_{-i})$, we must have that $aR_i f\left(R\right)$. If $f(R) = a$, then the equality in the statement follows (either as singleton sets or as the indifference class $[\varnothing]_i$). Assume $f\left(R\right) \neq a$. Two cases are possible:

1. $i \in S$. Then, $aP_i f\left(R\right)$. Since $a \in o_i\left(R_{-i}\right)$, $a = f\left(R_i', R_{-i}\right)$ for some $R_i' \in \mathcal{R}_i$, which means that $i$ manipulates $f$ at $R$ via $R_i'$. A contradiction.

2. $i \notin S$. Assume $i \notin f_N(R)$. Then, $f(R) = [\varnothing]_i = a$, a contradiction. Hence, $i \in f_N\left(R\right)$ and $aP_i f\left(R\right)$. Now, we obtain a contradiction with strategy-proofness of $f$ by proceeding in a similar way as we did in the previous case. ■

For the next steps in the proof, it will be useful to consider the set $\mathcal{F}$ of all complete, transitive and antisymmetric binary relations over $X$. Namely, $\mathcal{F}$ can be seen as the set of all strict preferences over $X$. Now, for each $N^* \subset N$, each $i \in N$ and each strict preference $\succ_i$ over $X$ we associate a preference over $2^N \times X$ (namely, an element of $\mathcal{R}_i$), denoted by $R_{N^*,\succ_i}$, by selecting one among those satisfying the following features.

- If $i \in N^*$, consider several cases:
  - If $i \in S \cap T \subset N^*$, then $(S, x)\, P_{N^*,\succ_i}\, (T, y)$ if and only if $x \succ_i y$.
  - If $i \in T \subsetneq S \subset N^*$, then $(S, x)\, P_{N^*,\succ_i}\, (T, x)$ for all $x \in X$.
  - If $i \in S \subset N^*$, then $(S, x)\, P_{N^*,\succ_i}\, (\varnothing, y)$ for all $x, y \in X$.
  - If $i \in S$ and $S \cap (N \backslash N^*) \neq \varnothing$, then $(\varnothing, x)\, P_{N^*,\succ_i}\, (S, y)$ for all $x, y \in X$.
- If $i \notin N^*$, then $(\varnothing, x)\, P_{N^*,\succ_i}\, (S, y)$ for all $S \subset N$ such that $i \in S$ and for all $x, y \in X$.[7]

Note that for each $N^*$, each $i \in N$ and each $\succ_i$ over $X$ there are many preferences in $\mathcal{R}_i$ satisfying the above conditions. We just select one of them, and denote it by $R_{N^*,\succ_i}$.

Fix $N^* \subseteq N$ and define a rule $g : \mathcal{F}^{N^*} \to X$ as follows. For each subprofile $\left(\succ_i\right)_{i \in N^*} \in \mathcal{F}^{N^*}$ of preferences over $X$ set

$$g((\succ_i)_{i \in N^*}) = f_X((R_{N^*,\succ_i})_{i \in N}).$$

Lemma 2 below says that if $f$ is strategy-proof, unanimous and outsider independent, then $g$ is dictatorial; namely, there exists $j \in N^*$ such that for all $(\succ_i)_{i \in N^*} \in \mathcal{F}^{N^*}$, $g((\succ_i)_{i \in N^*}) = \tau(\succ_j)$ where $\tau(\succ_j) \succ_j y$ for all $y \in X \backslash \{\tau\left(\succ_j\right)\}$.

[7]The preference $R_{N^*,\succ_i}$ may not depend on $\succ_i$, but for simplicity we maintain the notation $R_{N^*,\succ_i}$.

**Lemma 2** *Let $f : \mathcal{R} \rightarrow A$ be strategy-proof, unanimous and outsider independent. Then, for all $N^* \subseteq N$, the rule $g$ is dictatorial.*

**Proof** Fix $N^* \subseteq N$. Since $f$ is unanimous, $|X| \geq 3$ and $g$ is defined on the universal domain of strict preference profiles over $X$, the Gibbard-Satterthwaite theorem says that if $g$ is onto (for each $x \in X$ there exists $(\succ_i)_{i \in N^*}$ such that $g((\succ_i)_{i \in N^*}) = x$) and strategy-proof, then $g$ is dictatorial.

We first prove that $g$ is onto. Let $x$ and $(\succ_i)_{i \in N^*} \in \mathcal{F}^{N^*}$ be such that for each $i \in N^*$, $\tau(\succ_i) = x$. By definition of $R_{N^*,\succ_i}$, $\tau(R_{N^*,\succ_i}) = (N^*, x)$ if $i \in N^*$ and $(N^*, x) \in \tau(R_{N^*,\succ_i})$ if $i \notin N^*$. Since $f$ is unanimous and $\bigcap_{i \in N} \tau(R_{N^*,\succ_i}) = (N^*, x)$, $f((R_{N^*,\succ_i})_{i \in N}) = (N^*, x)$. Thus, $g((\succ_i)_{i \in N^*}) = f_X((R_{N^*,\succ_i})_{i \in N}) = x$.

We now prove that $g$ is strategy-proof. Suppose otherwise. Then, there exist $(\succ_i)_{i \in N^*}$, $j \in N^*$ and $\succ'_j$ such that

$$g(\succ'_j, \succ_{-j}) \succ_j g(\succ_j, \succ_{-j}). \tag{1}$$

By definition of $g$, $f_X((R_{N^*,\succ_i})_{i \in N}) = g(\succ_j, \succ_{-j})$ and $f_X(R_{N^*,\succ'_j}, (R_{N^*,\succ_i})_{i \neq j}) = g(\succ'_j, \succ_{-j})$. By definition of $R_{N^*,\succ_i}$ we know that for each $i \in N^*$, each $\succ_i \in \mathcal{F}$, each $x \in X$, and each $S \subset N$ with $S \neq N^*$, we have that $(N^*, x) \, P_{N^*,\succ_i} \, (S, x)$. Besides, for each $i \in N \backslash N^*$, each $\succ_i \in \mathcal{F}$, each $x \in X$, and each $S \subset N$ with $i \in S$, we have that $(N^*, x) \, P_{N^*,\succ_i} \, (S, x)$. Since $f$ is efficient,

$$\begin{aligned} f((R_{N^*,\succ_i})_{i \in N}) &= (N^*, g(\succ_j, \succ_{-j})) \text{ and} \\ f(R_{N^*,\succ'_j}, (R_{N^*,\succ_i})_{i \neq j}) &= (N^*, g(\succ'_j, \succ_{-j})). \end{aligned}$$

By definition of $R_{N^*,\succ_j}$ and (1)

$$(N^*, g(\succ'_j, \succ_{-j})) P_{N^*,\succ_j} (N^*, g(\succ_j, \succ_{-j})),$$

which contradicts that $f$ is strategy -proof. ■

Fix $R_i \in \mathcal{R}_i$ and $a \in A$. Denote by $R_{a,i}$ the preference over $A$ obtained from $R_i$ by just placing $a$, and all its indifferent alternatives (if any), at the bottom of the ordering. Formally, $R_{a,i}$ is defined so that $a' R_{a,i} a$, for all $a' \in A$ and, for all $a', a'' \in A \backslash \{a\}$, $a' R_{a,i} a''$ if and only if $a' R_i a''$. Similarly, $R_i^a$ denotes the preference over $A$ obtained from $R_i$ by just placing $a$, and all its indifferent alternatives (if any), at the top of the ordering. Formally, $R_i^a$ is defined so that $a R_i^a a'$, for all $a' \in A$ and, for all $a', a'' \in A \backslash \{a\}$, $a' R_i^a a''$ if and only if $a' R_i a''$.

**Lemma 3** *Let $f : \mathcal{R} \rightarrow A$ be strategy-proof, unanimous and outsider independent, and let $R \in \mathcal{R}$ and $i, j \in S \subseteq N$ be such that $i \neq j$, $f(R) = (S, x)$ and $|o_i(R_{-i})| \geq 3$. Then, $|o_j(R_{-j})| = 1$.*

**Proof** Set $a = (S, x)$. Since $f(R) = a$, $a \in o_i(R_{-i}) \cup o_j(R_{-j})$. Suppose $|o_j(R_{-j})| \geq 2$ holds. Since $|o_i(R_{-i})| \geq 3$, we can find $a' \in o_i(R_{-i}) \setminus \{a\}$ and $a'' \in o_j(R_{-j}) \setminus \{a\}$ such that $a' \neq a''$. Consider any $R'_i \in \mathcal{R}_i$, where

$$R'_i = \begin{cases} R_{a'',i} & \text{if } aP_i a'' \\ R_i^{a''} & \text{if } a'' P_i a. \end{cases}$$

Notice that $a'' I_i a$ does not hold since $i \in S$ and $a = (S, x)$. Symmetrically, consider any $R'_j \in \mathcal{R}_j$, where

$$R'_j = \begin{cases} R_{a',j} & \text{if } aP_j a' \\ R_j^{a'} & \text{if } a' P_j a. \end{cases}$$

Again, $a' I_j a$ does not hold since $j \in S$ and $a = (S, x)$. By monotonicity, $f(R) = f\left(R'_j, R_{-j}\right) = f\left(R'_i, R_{-i}\right) = f\left(R'_i, R'_j, R_{-i,j}\right) = a$, where remember that $R_{-i,j}$ means $R_{N \setminus \{i,j\}}$.

Claim 1: (i) $o_i(R_{-i}) = o_i(R'_j, R_{-i,j})$ *and* (ii) $o_j(R_{-j}) = o_j(R'_i, R_{-i,j})$.

Proof: We only prove (i) (the proof of (ii) is similar and we omit it). Suppose otherwise and assume $o_i(R_{-i}) \setminus o_i(R'_j, R_{-i,j}) \neq \varnothing$ (the proof of the other case $o_i(R'_j, R_{-i,j}) \setminus o_i(R_{-i}) \neq \varnothing$ is similar and we omit it). Take any $\widetilde{a} \in o_i(R_{-i}) \setminus o_i(R'_j, R_{-i,j})$. Since $\widetilde{a} \in \tau(R_i^{\widetilde{a}})$, by (3) of Lemma 1, $f(R_i^{\widetilde{a}}, R_{-i}) = C(o_i(R_{-i}), R_i^{\widetilde{a}})$. Hence, $f(R_i^{\widetilde{a}}, R_{-i}) = \widetilde{a}$. Since $\widetilde{a} \notin o_i(R'_j, R_{-i,j})$, $f(R_i^{\widetilde{a}}, R'_j, R_{-i,j}) \neq \widetilde{a}$. Hence, $L(f(R_i^{\widetilde{a}}, R'_j, R_{-i,j}), R_i^{\widetilde{a}}) \subseteq L(f(R_i^{\widetilde{a}}, R'_j, R_{-i,j}), R_i)$. Since $f$ is monotone, $f(R_i^{\widetilde{a}}, R'_j, R_{-i,j}) = f(R_i, R'_j, R_{-i,j}) = a$. We now distinguish between two cases (observe that $\widetilde{a} I'_j a$ does not hold because $j \in S$ and $a = (S, x)$).

Case 1: $\widetilde{a} P'_j a$. Then,

$$f(R_i^{\widetilde{a}}, R_j, R_{-i,j}) = \widetilde{a} P'_j a = f(R_i^{\widetilde{a}}, R'_j, R_{-i,j}).$$

Thus, $j$ manipulates $f$ at $(R_i^{\widetilde{a}}, R'_j, R_{-i,j})$ via $R_j$, which is a contradiction.

Case 2: $a P'_j \widetilde{a}$. By definition of $R'_j$, $a P_j \widetilde{a}$. Then,

$$f(R_i^{\widetilde{a}}, R'_j, R_{-i,j}) = a P_j \widetilde{a} = f(R_i^{\widetilde{a}}, R_j, R_{-i,j}).$$

Thus, $j$ manipulates $f$ at $(R_i^{\widetilde{a}}, R_j, R_{-i,j})$ via $R'_j$, which is also a contradiction. □

We now define two new preferences, $\tilde{R}_i$ and $\tilde{R}_j$, where

$$\tilde{R}_i = \begin{cases} (R_{a'',i})^{a'} & \text{if } aP_i a'' \\ (R_i^{a'})^{a''} & \text{if } a'' P_i a \end{cases}$$

and

$$\tilde{R}_j = \begin{cases} (R_{a',j})^{a''} & \text{if } aP_j a' \\ (R_j^{a''})^{a'} & \text{if } a' P a. \end{cases}$$

Claim 2: (i) $f(\tilde{R}_i, R'_j, R_{-i,j}) = a'$ *and* (ii) $f(R'_i, \tilde{R}_j, R_{-i,j}) = a''$.

Proof: We only prove (i) (the proof of (ii) is similar and we omit it). Since $a' \in o_i(R_{-i})$, by (i) in Claim 1, $a' \in o_i(R'_j, R_{-i,j})$. If $aP_ia''$, $C(o_i(R'_j, R_{-i,j}), \tilde{R}_i) = a'$. Since $f$ is strategy-proof, $f(\tilde{R}_i, R'_j, R_{-i,j}) = a'$. If $a''P_ia$, $a'' \in \tau(\tilde{R}_i)$. Assume first that $a'' \in o_i(R'_j, R_{-i,j})$. Since $f$ is strategy-proof, $f(\tilde{R}_i, R'_j, R_{-i,j}) = a''$. Since $f(R'_i, R'_j, R_{-i,j}) = a$, $i$ manipulates $f$ at $(R'_i, R'_j, R_{-i,j})$ via $\tilde{R}_i$, a contradiction. Hence, $a'' \notin o_i(R'_j, R_{-i,j})$. Since $a' \in o_i(R'_j, R_{-i,j})$ and $f$ is strategy-proof, $f(\tilde{R}_i, R'_j, R_{-i,j}) = a'$ because $a' = C(A \backslash \{a''\}, \tilde{R}_i)$. □

We now proceed with the proof of Lemma 3 by considering four different cases:

(1) Assume $aP_ia''$. Since $f(R'_i, \tilde{R}_j, R_{-i,j}) = a''$ by (ii) in Claim 2, $U(f(R'_i, \tilde{R}_j, R_{-i,j}), R'_i) = A$. Hence, $U(f(R'_i, \tilde{R}_j, R_{-i,j}), \tilde{R}_i) \subseteq U(f(R'_i, \tilde{R}_j, R_{-i,j}), R'_i)$, and by monotonicity,

$$f(\tilde{R}_i, \tilde{R}_j, R_{-i,j}) = f(R'_i, \tilde{R}_j, R_{-i,j}) = a''.$$

(2) Assume $a''P_ia$. Since $f(R'_i, \tilde{R}_j, R_{-i,j}) = a''$ by (ii) in Claim 2, $L(f(R'_i, \tilde{R}_j, R_{-i,j}), \tilde{R}_i) = A$. Hence, $L(f(R'_i, \tilde{R}_j, R_{-i,j}), R'_i) \subset L(f(R'_i, \tilde{R}_j, R_{-i,j}), \tilde{R}_i)$, and by monotonicity,

$$f(\tilde{R}_i, \tilde{R}_j, R_{-i,j}) = f(R'_i, \tilde{R}_j, R_{-i,j}) = a''.$$

(3) Assume $aP_ja'$. Since $f(\tilde{R}_i, R'_j, R_{-i,j}) = a'$ by (i) in Claim 2, $U(f(\tilde{R}_i, R'_j, R_{-i,j}), R'_j) = A$. Hence, $U(f(\tilde{R}_i, R'_j, R_{-i,j}), \tilde{R}_j) \subset U(f(\tilde{R}_i, R'_j, R_{-i,j}), R'_j)$, and by monotonicity,

$$f(\tilde{R}_i, \tilde{R}_j, R_{-i,j}) = f(\tilde{R}_i, R'_j, R_{-i,j}) = a'.$$

(4) Assume $a'P_ja$. Since $f(\tilde{R}_i, R'_j, R_{-i,j}) = a'$ by (i) in Claim 2, $L(f(\tilde{R}_i, R'_j, R_{-i,j}), \tilde{R}_j) = A$. Hence, $L(f(\tilde{R}_i, R'_j, R_{-i,j}), R'_j) \subset L(f(\tilde{R}_i, R'_j, R_{-i,j}), \tilde{R}_j)$, and by monotonicity,

$$f(\tilde{R}_i, \tilde{R}_j, R_{-i,j}) = f(\tilde{R}_i, R'_j, R_{-i,j}) = a'.$$

Thus, in each of the four possible cases $aP_ia''$ and $aP_ja'$, $aP_ia''$ and $a'P_ja$, $a''P_ia$ and $aP_ja'$, and $a''P_ia$ and $a'P_ja$, we have that $f(\tilde{R}_i, \tilde{R}_j, R_{-i,j}) = a''$ and $f(\tilde{R}_i, \tilde{R}_j, R_{-i,j}) = a'$, which is a contradiction with $a' \neq a''$. ■

Given $N^* \subseteq N$, by Lemma 2, the rule $g$ (induced by $f$) is dictatorial on its domain $\mathcal{F}^{N^*}$. Let $d(N^*) \in N^*$ be the dictator. Using the identification described just before Lemma 2, for every $i \in N$ and any $(N^*, \succ_i)$, choose a particular $R_{N^*,\succ_i} \in \mathcal{R}_i$. Consider the subdomain

$$\mathcal{R}^{N^*} = \{R \in \mathcal{R} \mid R = (R_{N^*,\succ_i})_{i \in N} \text{ for some } (\succ_i)_{i \in N^*} \in \mathcal{F}^{N^*}\}.$$

**Lemma 4** *Let $N^* \subseteq N$ and $R \in \mathcal{R}$ be such that* (i) $\tau(R_{d(N^*)}) = (N^*, x)$ *for some $x \in X$ and* (ii) *for all $j \in N \backslash N^*$, $R_j \in \mathcal{R}_j^{N^*}$. Then, $f(R) = (N^*, x)$.*

**Proof** Assume the hypothesis of Lemma 4 holds. Suppose first that $R \in \mathcal{R}^{N^*}$, and let $(\succ_i)_{i\in N^*} \in \mathcal{F}^{N^*}$ be such that, for each $i \in N$, $R_i = R_{N^*,\succ_i}$. By Lemma 2, $f_X(R) = g((\succ_i)_{i\in N^*}) = x$. By the definition of $(R_{N^*,\succ_i})_{i\in N}$, $(N^*,x) \in \bigcap_{j\in N\setminus N^*} \tau(R_j)$. By efficiency, $f_N(R) = N^*$. Hence,

$$f(R) = (N^*, x) \tag{2}$$

for all $R \in \mathcal{R}^{N^*}$ such that $\tau(R_{d(N^*)}) = (N^*, x)$.

Now, let $R \in \mathcal{R}$ be such that (i) $\tau(R_{d(N^*)}) = (N^*, x)$ and (ii) there exists $i \in N^*\setminus\{d(N^*)\}$ such that for all $j \in N\setminus\{d(N^*), i\}$, $R_j \in \mathcal{R}_j^{N^*}$.

Claim A: $f(R) = (N^*, x)$.

Proof: Consider any $R_i' \in \mathcal{R}_i^{N^*}$. By (2), for all $y \in X$, $(N^*, y) \in o_{d(N^*)}\left(R_i', R_{-i,d(N^*)}\right)$. Since $|X| \geq 3$, $\left|o_{d(N^*)}\left(R_i', R_{-i,d(N^*)}\right)\right| \geq 3$. By Lemma 3, $|o_i(R_{-i})| = 1$. Since $(N^*, x) \in o_i(R_{-i})$ and $|o_i(R_{-i})| = 1$, $o_i(R_{-i}) = (N^*, x)$. Thus, $f(R) = (N^*, x)$. □

Applying successively Claim A above we obtain that for all $R \in \mathcal{R}$ satisfying (i) $\tau(R_{d(N^*)}) = (N^*, x)$ and (ii) for all $j \in N\setminus N^*$, $R_j \in \mathcal{R}_j^{N^*}$, we have that $f(R) = (N^*, x)$ and the statement of Lemma 4 follows. ■

**Lemma 5** *Let $N' \subsetneq N'' \subset N$ be such that $d(N'') \in N'$. Then, $d(N') = d(N'')$.*

**Proof** Suppose not. Let $x \in X$ and consider $R \in \mathcal{R}$ where (i) $\tau(R_{d(N'')}) = (N'', x)$, (ii) $\tau(R_{d(N')}) = (N', x)$ and (iii) for each $i \in N\setminus\{d(N'), d(N'')\}$, $R_i$ is any preference in the subdomain $\mathcal{R}_i^{\{d(N'),d(N'')\}}$. By Lemma 4, with $N^* = N''$, $f(R) = (N'', x)$. Since $d(N'') \in N'$, we can apply Lemma 4 with $N^* = N'$, and obtain that $f(R) = (N', x)$, a contradiction with $N' \neq N''$. ■

**Proof of Theorem 1** Let $\pi \in \Pi$ and $x \in X$ be given. It is easy to show that the serial dictator rule $f^{\pi,x}$ is strategy-proof, unanimous and outsider independent. To prove the other implication, assume $f : \mathcal{R} \to A$ is strategy-proof, unanimous and outsider independent. We will identify from $f$ an ordering $\pi \in \Pi$ and $x \in X$ such that $f = f^{\pi,x}$. The ordering $\pi$ will be recursively defined by setting $\pi_1 = d(N)$ and, for all $i = 2, \ldots, n$, $\pi_i = d(N\setminus\{\pi_1, \ldots, \pi_{i-1}\})$. To identify $x \in X$, let $R \in \mathcal{R}$ be such that, for all $i \in N$, $\tau(R_i) = [\varnothing]_i$. Thus,

$$\bigcap_{i\in N} \tau(R_i) = \{(\varnothing, x') \in A \mid x' \in X\}.$$

By unanimity, $f(R) \in \bigcap_{i\in N} \tau(R_i)$. Set $x = f_X(R)$. We now prove that $f = f^{\pi,x}$. Let $R \in \mathcal{R}$ be arbitrary and set $\pi_1 = d(N)$. Two cases are possible.

Case 1. $|\tau(R_{\pi_1})| = 1$ (*i.e.*, $\tau(R_{\pi_1}) \notin [\varnothing]_{\pi_1}$). Thus, $\tau(R_{\pi_1}) = (S_1, x_1)$ and $\pi_1 \in S_1$. By definition, $f^{\pi,x}(R) = (S_1, x_1)$. If $S_1 = N$, by Lemma 4, $f(R) = (S_1, x_1)$. Hence, $f(R) = f^{\pi,x}(R)$. Assume $S_1 \subsetneq N$. For each $j \in N \backslash S_1$, let $R'_j$ be any preference in the subdomain $\mathcal{R}_j^{S_1}$. Since $\pi_1 \in S_1$, by Lemma 5, $d(S_1) = \pi_1$. By Lemma 4, $f(R_{S_1}, R'_{-S_1}) = (S_1, x_1)$. Let $i \in N \backslash S_1$. By Lemma 4 again, $(S_1 \cup \{i\}, y) \in o_{\pi_1}(R_{(S_1 \cup \{i\}) \backslash \{\pi_1\}}, R'_{-(S_1 \cup \{i\})})$ for all $y \in X$. By Lemma 3, $\left| o_i(R_{S_1}, R'_{-(S_1 \cup \{i\})}) \right| = 1$. Since $f(R_{S_1}, R'_{-S_1}) = (S_1, x_1)$, $o_i(R_{S_1}, R'_{-(S_1 \cup \{i\})}) = (S_1, x_1)$. Hence, $f(R_{S_1 \cup \{i\}}, R'_{-(S_1 \cup \{i\})}) = (S_1, x_1)$. Similarly, $f(R_{S_1 \cup \{i,j\}}, R'_{-(S_1 \cup \{i,j\})}) = (S_1, x_1)$ holds when $j \in N \backslash (S_1 \cup \{i\})$. Repeating this process for the rest of the agents in $N \backslash S_1$, we obtain that $f(R) = (S_1, x_1)$. Hence, $f(R) = f^{\pi,x}(R)$.

Case 2. $|\tau(R_{\pi_1})| > 1$. Then, $\tau(R_{\pi_1}) = [\varnothing]_{\pi_1}$ and set $\pi_2 = d(N \backslash \{\pi_1\})$. We consider two subcases separately.

Case 2.1. $|C(\tau(R_{\pi_1}), R_{\pi_2})| = 1$ (*i.e.*, $\tau(R_{\pi_2}) \notin [\varnothing]_{\pi_2}$). Set $C(\tau(R_{\pi_1}), R_{\pi_2}) = (S_2, x_2)$ and observe that $\pi_2 \in S_2 \subseteq N \backslash \{\pi_1\}$. By Lemma 5, $d(S_2) = \pi_2$. It is immediate to see that $f^{\pi,x}(R) = (S_2, x_2)$. We now argue that $f(R) = (S_2, x_2)$. For each $j \in N \backslash S_2$, let $R'_j$ be any preference in the subdomain $\mathcal{R}_j^{S_2}$. Note that $R_{\pi_1}$ belongs to the subdomain $\mathcal{R}_{\pi_1}^{S_2}$. Using arguments similar to those used in Case 1 above, we can show that $f(R) = (S_2, x_2)$.

Case 2.2. $|C(\tau(R_{\pi_1}), R_{\pi_2})| > 1$. Thus,

$$C(\tau(R_{\pi_1}), R_{\pi_2}) = \{(S, y) \in A \mid \pi_1 \notin S, \ \pi_2 \notin S \text{ and } y \in X\}.$$

We would consider again two subcases separately depending on whether $|C(C(\tau(R_{\pi_1}), R_{\pi_2}), R_{\pi_3})|$ is equal to one or strictly larger, where $\pi_3 = d(N \backslash \{\pi_1, \pi_2\})$.

Continuing with this procedure, at the end we would reach agent $\pi_n = d(N \backslash \{\pi_1, ..., \pi_{n-1}\})$ and we would need to consider two subcases separately depending on whether $|C(A_{\pi_n}, R_{\pi_n})|$ is equal to one or strictly larger, where

$$A_{\pi_n} = \{(\{\pi_n\}, y) \in A \mid y \in X\} \cup \{(\varnothing, y) \in A \mid y \in X\}.$$

If $|C(A_{\pi_n}, R_{\pi_n})| = 1$ then $C(A_{\pi_n}, R_{\pi_n}) = (\{\pi_n\}, x_n)$. Thus, $f^{\pi,x}(R) = (\{\pi_n\}, x_n)$. Using arguments similar to those used above we can show that $f(R) = (\{\pi_n\}, x_n)$.

If $|C(A_{\pi_n}, R_{\pi_n})| > 1$ then $C(A_{\pi_n}, R_{\pi_n}) = \{(\varnothing, y) \in A \mid y \in X\}$. Then, $f^{\pi,x}(R) = (\varnothing, x)$. By definition of $x$, $f(R) = (\varnothing, x)$. ■

The three properties used in the characterization of Theorem 1 are independent.

Consider the Approval Voting rule $f^{AV,\rho}$ defined as follows. Each $i \in N$ votes for the subset $A_i = \{a \in A \mid aR_i [\varnothing]_i\}$. For each $a \in A$, compute the number of votes received by $a$; namely,

$|\{i \in N : a \in A_i\}|$ . The outcome with more votes is selected. The tie-breaking rule $\rho$ is applied whenever several alternatives obtain the largest number of votes, where $\rho : 2^A \backslash \{\emptyset\} \to A$ is such that for all $A' \in 2^A \backslash \{\emptyset\}$, $\rho(A') \in A'$. It is easy to see that, for any tie-breaking rule $\rho$, $f^{AV,\rho}$ is unanimous and outsider independent but it is not strategy-proof.

Any constant rule satisfies strategy-proofness and outsider independence but it is not unanimous.

Let $x, y \in X$ with $x \neq y$. Define

$$f(R) = \begin{cases} f^{\pi,x} & \text{if } \tau(R_{\pi_1}) = [\varnothing]_{\pi_1} \text{ and } (\{\pi_1\}, x) \, P_1 \, (\{\pi_1\}, y) \\ f^{\pi,y} & \text{otherwise.} \end{cases}$$

It is easy to see that $f$ is strategy-proof and unanimous but it does not satisfy outsider independence.

# 6 Final remarks

Before finishing the paper several remarks are in order.

First, the equivalence between the class of strategy-proof, unanimous and outsider independent rules and the set of serial dictator rules is a negative result. Serial dictator rules are not appealing for many reasons. In particular, in our setting, they do no satisfy any stability property: some agents may be forced to belong to the society and others may be excluded, all against their own will. Although our general model was able to encompass problems where the composition of the final society could be endogenous (by being a consequence of the chosen outcome) and stable, our result says that if we insist on requiring strategy-proofness only serial dictators remain.

Second, this negative result holds because our domain of preferences under which we want the non-trivial and strategy-proof rule to operate is, although restrictive, still very large. And hence, one may ask whether there are interesting and meaningful subdomains admitting non-trivial and strategy-proof rules. That is, what kind of possibilities can arise from imposing further preference restrictions than just condition (P.2). Some papers have followed this line of inquiry by considering stronger domain restrictions. For instance, for the case of an excludable public good (in a linearly ordered set $X$ of outcomes) when agents also care about the size of the set of its users, Jackson and Nicolò (2004) identifies classes of strategy-proof and efficient rules on the domain of single-peaked preferences on the level of the public good and alternative specifications of how agents order different sets of users. Roughly, those classes (which depend on the particular preference specification and may require that the rule be in addition outsider independent) consists of selecting the size of the set of users independently of the preference profile and then, choose the level of the public good according to a generalized median voter scheme. Although some of the results in Jackson and Nicolò (2004) are also negative, they identify restricted domains that admit more appealing rules; for instance, when preferences on the size of the set of users come from cost-sharing considerations. Berga, Bergantiños, Massó and Neme (2004) studies a particular instance of our model. It consists of a society choosing a subset of new members, from a finite set of candidates. They consider explicitly the possibility that initial members of the society (founders) may want to exit, if they do not like the resulting new society. They show that, if founders have separable (or additive) preferences, the unique strategy-proof, stable and onto rule is the one where candidates are chosen unanimously and no founder exits. In this case, the restricted domain of preferences admits (non-dictatorial) strategy-proof rules that are in addition anonymous and stable. But there are many other restricted domain possibilities, among which two appear as promising. They are based on betweenness and separability requirements, as follows. First, assume agent $i$ has $(x, S)$ as top-ranked and $i \in S$. If $i \in T$, then $(x, T)$ should be ranked higher than $(x, T\backslash\{i\})$; that is, $(x, T)$ lies somehow between $(x, S)$ and $(x, T\backslash\{i\})$, and since $(x, S)$ is the top-ranked alternative, agent $i$ should prefer the alternative that is closer to the top. Second, agent $i$ has a strict order over $X$ and a strict order over $N\backslash\{i\}$ and then, the ranking of alternatives of the form $(x, S)$, where $i \in S$, must be consistent to these two strict orders (and so, the preference would violate (P.1) as well). Most likely then, separability of the rule (as in Jackson and Nicolò (2004)) would be a necessary condition of strategy-proofness.[8]

Third, the class of rules that satisfy only strategy-proofness and unanimity but fail outsider independence is extremely large, even inside the class of rules based on serial dictators (which guarantee that strategy-proofness and unanimity are also satisfied). Its richness comes from the fact that, without outsider independence, the sequence of dictators may not be given from the very beginning, and independently of the preference profile $R \in \mathcal{R}$. Instead, at each step, the dictator may be selected as a function of the entire subprofile of the set of agents, say $S$, that have already declared that $\tau(R_i) = [\varnothing]_i$, for $i \in S$. Then, the choice of the agent selected as dictator in the current step could depend, in so many and arbitrary ways, on any conceivable characteristic of the subprofile $R_S$, that its full and systematic description seems unworkable.

[8] We thank a referee of this journal for suggesting us these two domain restrictions. However, to obtain full characterizations of interesting classes of non-trivial and strategy-proof rules, on the two corresponding domains, seems to require a complete analysis, which is outside the scope of the present paper.

We think that outsider independence structures in a simple way this arbitrariness by requiring that the sequence of dictators is not profile dependent. In addition, together with strategy-proofness and unanimity, insider independence imply group strategy-proofness, because serial dictators are trivially group strategy-proof; however, it is easy to see that the rules that are not outsider independent just described above, are not group strategy-proof because a change in a preference of an agent may produce a strictly improvement of the welfare of another agent, without affecting the welfare of the agent that has changed the preference, and hence inducing a group manipulation of the rule.

Fourth, our proof of Theorem 1 requires explicitly that $|X| \geq 3$. In particular, in the proof of Lemma 2, we show that the rule $g$, defined from $f$ on the universal domain of preferences over $X$, is onto and dictatorial. This follows from the Gibbard-Satterthwaite theorem and the unanimity and strategy-proofness of $f$, which can be applied because $|X| \geq 3$. And this is a key step in the full proof. Of course, the case $|X| \leq 2$ is interesting since $|A|$ may still be large, whenever the set of agents is large; in particular, if $|X| = 1$ the problem resembles the "Who is a J?" problem, where a set of $N$ agents has to decide who, among them, fulfills a specific binary characteristic, although we are not aware of any strategic analysis of this problem based explicitly on agents' preferences over families of subsets of agents. But we have not attempted to perform this analysis; it remains open for further research.

# References


[1] Bade, S. (2015): "Serial dictatorship: The unique optimal allocation rule when information is endogenous," *Theoretical Economics* 10, 385–410.

[2] Barberà, S., M. Maschler and J. Shalev (2001): "Voting for voters: a model of electoral evolution," *Games and Economic Behavior* 37, 40–78.

[3] Barberà, S., H. Sonnenschein and L. Zhou (1991): "Voting by committees," *Econometrica* 59, 595–609.

[4] Berga, D., G. Bergantiños, J. Massó and A. Neme (2004): "Stability and voting by committees with exit," *Social Choice and Welfare* 23, 229–247.

[5] Berga, D., G. Bergantiños, J. Massó and A. Neme (2006): "On exiting after voting," *International Journal of Game Theory* 34: 33–54.

[6] Berga, D., G. Bergantiños, J. Massó and A. Neme (2007): "An undominated Nash equilibrium for voting by committees with exit," *Mathematical Social Sciences* 54, 152–175.

[7] Bergantiños, G., J. Massó and A. Neme (2016): "On societies choosing social outcomes, and their memberships: internal stability and consistency," mimeo.

[8] Deb, R. and L. Razzolini (1999): "Auction-like mechanisms for pricing excludable public goods," *Journal of Economic Theory* 88, 340–368.

[9] Demange, G. and M. Wooders [editors] (2005): *Group Formation in Economics: Networks, Clubs, and Coalitions*, Cambridge University Press, New York.

[10] Gibbard, A. (1973): "Manipulation of voting schemes: a general result," *Econometrica* 41, 587–601.

[11] Jackson, M. and A. Nicolò (2004): "The strategy-proof provision of public goods under congestion and crowding preferences," *Journal of Economic Theory* 115, 278–308.

[12] Papai, S. (2001): "Strategy-proofness and nonbossy multiple assignments," *Journal of Public Economic Theory* 3, 257–271.

[13] Roberts, K. (1999): "Dynamic voting in clubs," mimeo, London School of Economics.

[14] Satterthwaite, M. (1975): "Strategy-proofness and Arrow's conditions: existence and correspondence theorems for voting procedures and social welfare functions," *Journal of Economic Theory* 10, 187–217.

[15] Satterthwaite, M. and H. Sonnenschein (1981): "Strategy-proof allocation mechanisms at differentiable points," *Review of Economic Studies* 48, 587–597.

[16] Svensson, L.-G. (1999): "Strategy-proof allocation of indivisible goods," *Social Choice and Welfare* 16, 557–567.

[17] Thomson, W. (2016): "Non-bossiness," *Social Choice and Welfare* 47, 665–696.