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

Gustavo Bergantiños[†], Jordi Massó[‡] and Alejandro Neme[§]

February 21, 2016

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 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 non-bossy rules as the family of all serial dictator rules.

Journal of Economic Literature Classification Number: D71.
Keywords: Strategy-proofness; Unanimity; Non-bossiness.

*The work of G. Bergantiños is partially supported by research grant ECO2011-23460 from the Spanish Ministry of Science and Innovation and FEDER. 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.

[†]Research Group in Economic Analysis. Facultade de Económicas, 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

1 Introduction

A classical social choice problem is the following. A society formed by the 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 independent of the chosen outcome. There are many situations for which this assumption is not appropriate because the final composition of the society may depend on the chosen outcome. For instance, membership of a political party 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. A professor in a department may start looking for a position elsewhere if he considers that the recruitment of the department has not being satisfactory to his standards; and this, in turn might trigger further exits. Hence, to be able to deal with such situations the classical social choice model has to be modified to include explicitly the possibility that initial members may leave it as the consequence of the chosen outcome and thus, preferences have to be extended to order pairs formed by the final society and the chosen outcome.

There is a literature that has already considered explicitly the dependence of the final society on its choices in specific settings.[1] Barberà, Maschler and Shalev (2001) consider a dynamic model in which the set of founders and the set of candidates are fixed, and the society repeatedly holds elections (to admit new members) for a fixed number of periods using voting by quota 1 (one vote is sufficient for admission, and voters can support as many candidates as they wish). Chosen candidates at one period become voters, together with remaining members, next period. They show that very interesting strategic behavior may emerge in equilibrium, even when the used voting method is very simple. Giving the right to vote to elected candidates and not allowing non elected candidates to vote at all, are two extreme ways of transferring influence among agents. Barberà and Perea (2002) study a similar model in which the transfer of influence to new members or non elected candidates behaves in a continuous way instead of being binary. They study the (essentially) unique subgame perfect equilibrium of a model with these features and identify its simple dynamic structure. Berga,

[1] See for instance Roberts (1999) for problems related to club formation and Sobel (2000) for the declining of standards in societies that chose their members.

Bergantiños, Massó and Neme (2004) study also the problem of a society choosing a subset of new members, from a finite set of candidates, using voting by committees as in Barberà, Sonnenschein and Zhou (1991). 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. Berga, Bergantiños, Massó and Neme (2006) study equilibria of a finite extensive form game in which, after knowing the chosen alternative, members may reconsider their membership by either staying or exiting. In turn, and as a consequence of the exit of some of its members, other members might now find undesirable to belong to the society as well. For general exit procedures they analyze the exit behavior of members after knowing the chosen alternative. All these papers mentioned above study specific models in terms of the voting methods under which members choose the outcome and the timing under which members reconsider their membership.

In this paper we look at the general setting without being specific about the two issues. 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 N of agents that will remain in the society) and an outcome in X. Then, we assume that agents' preferences are defined over the set of alternatives $2^N \times X$ and satisfy two natural 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.

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). 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, and so on. Strategy-proof rules make all these considerations unnecessary since truthtelling is a weakly dominant strategy of the direct revelation game 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. 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 non-bossy 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 non-bossy 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. This often can be done in a strategy-proof way and satisfying at the same time other desirable properties; for instance, weak notions of efficiency (like unanimity), non arbitrariness (like non-bosiness), or neutrality, consistency, and so on.

Indeed, serial dictator rules have been characterized as the family of strategy-proof rules (satisfying in addition some other properties) in many different settings. For instance, Satterthwaite and Sonnenschein (1981) contains several characterizations of serial dictator rules for restricted domains of preference profiles associated with economic environments. Svensson (1999) shows that, for allocation problems where a finite number of indivisible goods have to be assigned to a finite number of agents, the set of all serial dictator rules coincides with the family of strategy-proof, non-bossy and neutral rules; Papai (2001) extends this result when agents may receive more than one object and she shows that a rule is strategy-proof, totally non-bossy and efficient if and only if it is a serial dictator rule. Bade (2015) shows that, in house allocation problems with costly endogenous information acquisition that influences agents' valuations of objects, a rule is strategy-proof, non-bossy and ex-ante efficient if and only if it is a serial dictator rule.

In a companion note (Bergantiños, Massó and Neme (2016)) we consider the same setting but assume that the 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 doble 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 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 non-bossy rules as the family of all serial dictator rules. Section 5 contains the proof of Theorem 1.

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.[2] 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

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

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 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*.

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.*,

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

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 $aR_i[\varnothing]_i$ to represent that aR_ia' for all $a' \in [\varnothing]_i$.

To clarify the model, we relate it with the two 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 $f_X(R) \in X$.

3 Properties of rules

We present several properties that a rule $f : \mathcal{R} \to A$ may satisfy, and that we will consider later on. The first two impose conditions at each profile.

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 $j \in N$.

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 three 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'.

A rule is monotonic if when the chosen alternative at a profile improves in the ordering of the preferences of an agent, the rule selects the same alternative in the new profile.

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 $U(f(R), R_i) \supset U(f(R), R_i')$, $f(R) = f(R_i', R_{-i})$.

A rule is non-bossy if an agent that is not a member of the final society at a profile changes his preferences and remains a nonmember, the rule chooses the same alternative at the two profiles.

Non-bossiness 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 non-bossiness was introduced by Satterthwaite and Sonnenschein (1981) and different variants of it have been intensively used in the literature. Thomson (2014) contains a systematic analysis of this property by giving alternative definitions and interpretations of non-bosiness, and by relating them to a large family of allocation problems. Since our definition imposes conditions to the rule only after a change of preferences of an agent that is not a member of the two final societies, but at the same time agents are only indifferent among alternatives for which they do not belong to their corresponding final societies, our notion requires that the alternative does not change at all. We found that this is the natural requirement of non-bossiness in our setting; otherwise, the agent could remain indifferent and still his change in preferences could induce the rule to select different alternatives.

4 Characterization result

In this section we state the result of the paper characterizing the class of all strategy-proof, unanimous and non-bossy rules.[3] This class coincides with the family of all serial dictator rules. To define a serial dictator rule in our setting we need some preliminaries. Let $\pi : N \to \{1, \ldots, n\}$ be a permutation (one-to-one mapping) of the set of agents. Given $i \in N$, $\pi(i)$ is the position assigned to i after applying the permutation π to N. The set of all permutations $\pi : N \to \{1, \ldots, n\}$ will be denoted by Π. For $\pi \in \Pi$ and $1 \leq k \leq n$, we write π_k to denote the agent $\pi^{-1}(k)$.

A serial dictator rule induced by $\pi \in \Pi$ and $x \in X$, denoted by $f^{\pi,x}$, proceeds as follows. Fix a profile $R \in \mathcal{R}$ and look for the best alternative (S_1, x_1) of agent π_1, the first in the ordering induced by π. If $\pi_1 \in S_1$, set $f^{\pi,x}(R) = (S_1, x_1)$. Otherwise, look for the best alternative (S_2, x_2) of agent π_2, the second in the ordering induced by π, with the property that $\pi_1 \notin S_2$. If $\pi_2 \in S_2$, set $f^{\pi,x}(R) = (S_2, x_2)$. Otherwise, look for the best alternative (S_3, x_3) of agent π_3, the third in the ordering induced by π, provided that $\pi_1, \pi_2 \notin S_3$, and so on. At the end, look for the best alternative (S_n, x_n) of agent π_n, the last in the ordering induced by π, with 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.

Stage 1. 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 Stage 2.

We now define Stage k $(1 < k < n)$, assuming that the stage $k-1$ has been reached and A_{k-1} was defined on it.

[3] 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 (see Gibbard (1973) and Satterthwaite (1975)).

Stage k. Let $A_k = C(A_{k-1}, R_{\pi_{k-1}})$. Consider two cases.

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. $|C(A_k, R_{\pi_k})| > 1$. 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 Stage $k+1$.

We now define Stage n, the last stage of the procedure, assuming that the stage $n-1$ has been reached and A_{n-1} was defined on it.

Stage n. Let $A_n = C(A_{n-1}, R_{\pi_{n-1}})$. Consider two cases.

1. $|C(A_n, R_{\pi_n})| = 1$. Then, $C(A_n, R_{\pi_n}) = \tau(R_{\pi_n})$. Set $(S_n, x_n) = C(A_n, R_{\pi_n})$ and observe that $\pi_n \in S_n$. Define
$$f^{\pi,x}(R) = (S_n, x_n).$$
2. $|C(A_n, R_{\pi_n})| > 1$. 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 the set of agents and the set of outcomes, and consider the identity permutation π, 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)	$\{(S, y) \in A \mid 1 \notin S, y \in A\}$	(N, a)	(N, a)
		(N, b)	(N, b)
		$(\{2\}, c)$	$\{(S, y) \in A \mid 2 \notin S, y \in A\}$

Then,

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

$\Diamond$

We are now ready to state Theorem 1, the characterization of the class of all strategy-proof, unanimous and non-bossy 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 non-bossy if and only if f is a serial dictator rule for some permutation $\pi \in \Pi$ and alternative $x \in X$.*

5 Proof of Theorem 1

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

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

Notice that the option set of i at R does not depend on R_i. We use the full profile R just for notational convenience.

The proof that, for any π and x, the serial dictator rule $f^{\pi,x}$ is strategy-proof, unanimous and non-bossy is easy. The main idea of the proof of the other implication is as follows. We first show that any strategy-proof, unanimous and non-bossy 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 at R. 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 non-bossy rule. Then, the following hold.*

(1) *f satisfies monotonicity.*

(2) *f satisfies efficiency.*

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

Proof Assume $f : \mathcal{R} \to A$ is strategy-proof, unanimous and non-bossy. 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 non-bossiness, $f(R'_i, R_{-i}) = f(R)$ which is a contradiction.

(2) Suppose f is not efficient. Namely, there exist $R \in \mathcal{R}$ and $a \in A$ such that $a R_i f(R)$ for all $i \in N$ and $a P_j f(R)$ 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 non-bossiness. 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, $f(R') \in \bigcap_{k \in N} \tau(R'_k)$. 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 $a = (S, x) \in C(o_i(R), R_i)$. Then, $a R_i f(R)$. Assume $f(R) \neq a$. Two cases are possible:

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

2. $i \notin S$. By non-bossiness, $i \in f_N(R)$ and hence, $aP_i f(R)$. 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 $x,y \in X$.[4]

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 social choice function $g : \mathcal{F}^{N^*} \to X$ as follows. For each subprofile $(\succ_i)_{i\in N^*} \in \mathcal{F}^{N^*}$ 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 non-bossy, 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 $y \in X \backslash \{\tau(\succ_j)\}$.

Lemma 2 *Let $f : \mathcal{R} \to A$ be strategy-proof, unanimous and non-bossy. Then, for all $N^* \subseteq N$, the social choice function g is dictatorial.*

Proof Fix $N^* \subseteq N$. Since g is defined on the universal domain of strict preference profiles on 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.

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

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 $(N^*, x) \in \tau\left(R_{N^*,\succ_i}\right)$ if $i \notin N^*$. Since f is unanimous and $\bigcap\limits_{i\in N} \tau\left(R_{N^*,\succ_i}\right) = (N^*, x)$, $f((R_{N^*,\succ_i})_{i\in N}) = (N^*, x)$. 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 $S \subset N$, $S \neq N^*$, 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,

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

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

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

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

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

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

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 R_{-ij} means $R_{N\setminus\{i,j\}}$.

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

Proof: We only prove that $o_i\left(R\right) = o_i(R'_i, R'_j, R_{-i,j})$ holds (the proof of the case $o_j\left(R\right) = o_j(R'_i, R'_j, R_{-i,j})$ is similar and we omit it). Suppose otherwise and assume $o_i\left(R\right) \backslash o_i(R'_i, R'_j, R_{-i,j}) \neq \varnothing$ (the proof of the other case $o_i(R'_i, R'_j, R_{-i,j}) \backslash o_i\left(R\right) \neq \varnothing$ is similar and we omit it). Take any $\widetilde{a} \in o_i\left(R\right) \backslash o_i(R'_i, R'_j, R_{-i,j})$. Since $\widetilde{a} \in \tau(R_i^{\widetilde{a}})$, $\widetilde{a} \in o_i\left(R\right) = o_i(R_i^{\widetilde{a}}, R_{-i})$ and, by (3) of Lemma 1, $f(R_i^{\widetilde{a}}, R_{-i}) = C(o_i(R_i^{\widetilde{a}}, R_{-i}), R_i^{\widetilde{a}})$. Hence, $f(R_i^{\widetilde{a}}, R_{-i}) = \widetilde{a}$. Since $\widetilde{a} \notin o_i(R'_i, R'_j, R_{-i,j})$, $f(R_i^{\widetilde{a}}, R'_j, R_{-i,j}) \neq \widetilde{a}$. 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, $f(R_i^{\widetilde{a}}, R'_j, R_{-i,j}) = f(R_i, R'_j, R_{-ij}) = a$. We now distinguish between two cases.

<u>Case 1</u>: $\widetilde{a}P'_j a$. 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.

<u>Case 2</u>: $aP'_j\widetilde{a}$. By definition of R'_j, $aP_j\widetilde{a}$. Then,

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

Thus, j manipulates f at $(R_i^{\widetilde{a}}, R_j, R_{-i,j})$ via R'_j, which is also a contradiction. Since $\widetilde{a}I'_j a$ is not possible because $j \in S$ and $a = (S, x)$ we have finished the proof of Claim 1. $\square$

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

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

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

Proof:

(i) $f(\tilde{R}_i, R'_j, R_{-ij}) = a'$. Since $a' \in o_i(R)$, by Claim 1, $a' \in o_i(R'_i, R'_j, R_{-ij})$. If aP_ia'', $C(o_i(R'_i, R'_j, R_{-ij}), \tilde{R}_i) = a'$. Since f is strategy-proof, $f(\tilde{R}_i, R'_j, R_{-ij}) = a'$. If $a''P_ia$ then $a'' \in \tau(\tilde{R}_i)$. Assume first that $a'' \in o_i(R'_i, R'_j, R_{-ij})$. Since f is strategy-proof, $f(\tilde{R}_i, R'_j, R_{-ij}) = a''$. Since $f(R'_i, R'_j, R_{-ij}) = a$, i manipulates f at (R'_i, R'_j, R_{-ij}) via $\tilde{R}_i$, a contradiction. Hence, $a'' \notin o_i(R'_i, R'_j, R_{-ij})$. Since $a' \in o_i(R'_i, R'_j, R_{-ij})$ and f is strategy-proof, $f(\tilde{R}_i, R'_j, R_{-ij}) = a'$ because $a' = C(A\backslash\{a''\}, \tilde{R}_i)$.

(ii) $f(R'_i, \tilde{R}_j, R_{-ij}) = a''$. Since $a'' \in o_j(R)$, by Claim 1, $a'' \in o_j(R'_i, R'_j, R_{-ij})$. If aP_ja', $C(o_j(R'_i, R'_j, R_{-ij}), \tilde{R}_j) = a''$. Since f is strategy-proof, $f(R'_i, \tilde{R}_j, R_{-ij}) = a''$. If $a'P_ja$ then $a' \in \tau(\tilde{R}_j)$. Assume first that $a' \in o_j(R'_i, R'_j, R_{-ij})$. Since f is strategy-proof, $f(R'_i, \tilde{R}_j, R_{-ij}) = a'$. Since $f(R'_i, R'_j, R_{-ij}) = a$, j manipulates f at (R'_i, R'_j, R_{-ij}) via $\tilde{R}_j$, a contradiction. Hence, $a' \notin o_j(R'_i, R'_j, R_{-ij})$. Since $a'' \in o_j(R'_i, R'_j, R_{-ij})$ and f is strategy-proof, $f(R'_i, \tilde{R}_j, R_{-ij}) = a''$ because $a'' = C(A\backslash\{a'\}, \tilde{R}_j)$. And this finishes the proof of Claim 2. □

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

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

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

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

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

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

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

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

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

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 $f(\tilde{R}_i, \tilde{R}_j, R_{-ij}) = a''$ and $f(\tilde{R}_i, \tilde{R}_j, R_{-ij}) = a'$, which is a contradiction with $a' \neq a''$. ■

Given $N^* \subseteq N$ we know, by Lemma 2, that the function 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 $(N^*, \succ_i)$ 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 $\tau(R_{d(N^*)}) = (N^*, x)$ for some $x \in X$ and $(N^*, x) \in \bigcap_{j\in N\setminus N^*} \tau(R_j)$. Then, $f(R) = (N^*, x)$.*

Proof Assume the hypothesis of Lemma 4 holds. By Lemma 2, $f_X(R) = g\left((\succ_i)_{i\in N^*}\right) = x$. By efficiency of f and the definition of $(R_{N^*,\succ_i})_{i\in N}$, $f_N(R) = N^*$. 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) $\tau(R_{d(N^*)}) = (N^*, x)$, (ii) there exists $i \in N^*\setminus\{d(N^*)\}$ such that $R_i \in \mathcal{R}_i$, and for all $j \in N\setminus\{d(N^*), i\}$, $R_j \in \mathcal{R}_j^{N^*}$. We want to show that $f(R) = (N^*, x)$ holds. Consider any $R_i' \in \mathcal{R}_i^{N^*}$. By (2), for all $y \in X$, $(N^*, y) \in o_{d(N^*)}(R_i', R_{-i})$. Since $|X| \geq 3$, $\left|o_{d(N^*)}(R_i', R_{-i})\right| \geq 3$,. By Lemma 3, $|o_i(R_i', R_{-i})| = 1$. Since $(N^*, x) \in o_i(R_i', R_{-i})$ and $|o_i(R_i', R_{-i})| = 1$, $o_i(R_i', R_{-i}) = (N^*, x)$. Since the option set of i at R does not depend on R_i, $o_i(R_i', R_{-i}) = o_i(R)$. This implies that $o_i(R) = (N^*, x)$. Thus, $f(R) = (N^*, x)$.

Applying successively the argument above we obtain that for all $R \in \mathcal{R}$ satisfying (i) $\tau(R_{d(N^*)}) = (N^*, x)$, (ii) for all $i \in N^*\setminus\{d(N^*)\}$, $R_i \in \mathcal{R}_i$, and (iii) for all $j \in N\setminus N^*$, $R_j \in \mathcal{R}_j^{N^*}$, we have that $f(R) = (N^*, x)$. ■

Lemma 5 *Assume $N' \subsetneq N'' \subset N$ are 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) $R_{d(N'')}$ satisfies $\tau(R_{d(N'')}) = (N'', x)$, (ii) $R_{d(N')}$ satisfies $\tau(R_{d(N')}) = (N', x)$ and (iii) for each $i \in N\setminus\{d(N'), d(N'')\}$, R_i is any preference in the subdomain $\mathcal{R}_i^{N^*}$ for $N^* = \{d(N'), d(N'')\}$. By Lemma 4, with $N^* = N''$, $f(R) = (N'', x)$. By Lemma 4 again, with $N^* = N'$, $f(R) = (N', x)$, which is a contradiction. ■

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 non-bossy. To prove the other implication, assume $f : \mathcal{R} \to A$ is strategy-proof, unanimous and non-bossy. We will identify from f a permutation of agents $\pi \in \Pi$ and an alternative $x \in X$ such that $f = f^{\pi,x}$. We first define π recursively 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. Two cases are possible.

<u>Case 1</u>. $|\tau(R_{\pi_1})| = 1$ (*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 $\mathcal{R}_j^{N^*}$ induced when $N^* = S_1$. Observe that $\tau(R'_j) \notin [\varnothing]_j$. Since $\pi_1 \in S_1$, by Lemmata 4 and 5, $f(R_{S_1}, R'_{-S_1}) = (S_1, x_1)$. Let $i \in N \backslash S_1$. By Lemma 4, $(S_1 \cup \{i\}, y) \in o_{\pi_1}(R_{S_1 \cup \{i\}}, R'_{-(S_1 \cup \{i\})})$ for all $y \in X$. By Lemma 3, $\left| o_i(R_{S_1 \cup \{i\}}, R'_{-(S_1 \cup \{i\})}) \right| = 1$. Since $o_i(R_{S_1 \cup \{i\}}, R'_{-(S_1 \cup \{i\})}) = o_i(R_{S_1}, R'_{-S_1})$, because in both problems π_1 determines the outcome and R_{π_1} is the same in both profiles, we deduce that $\left| o_i(R_{S_1}, R'_{-S_1}) \right| = 1$. Since $f(R_{S_1}, R'_{-S_1}) = (S_1, x_1)$, $o_i(R_{S_1}, R'_{-S_1}) = (S_1, x_1)$. 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, $f(R) = f^{\pi,x}(R)$.

<u>Case 2</u>. $|\tau(R_{\pi_1})| > 1$. Thus, $\tau(R_{\pi_1}) = [\varnothing]_{\pi_1}$. We consider two subcases separately.

<u>Case 2.1</u>. $|C(\tau(R_{\pi_1}), R_{\pi_2})| = 1$ (*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 $\pi_2 \in S_2$. It is immediate to see that $f^{\pi,x}(R) = (S_2, x_2)$. We now prove that $f(R) = (S_2, x_2)$. For each $j \in N \backslash S_2$, let R'_j be any preference in the subdomain $\mathcal{R}_j^{N^*}$ induced when $N^* = S_2$. Note that R_{π_1} belongs to the subdomain $\mathcal{R}_{\pi_i}^{N^*}$ and take $R'_{\pi_1} = R_{\pi_1}$. Using arguments similar to those used in Case 1 above, we can show that $f(R) = (S_2, x_2)$.

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

Continuing with this procedure, at the end we would reach agent n and we would need to consider two subcases separately depending on whether $|C(A_n, R_{\pi_n})|$ is equal to one or strictly larger, where

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

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

If $|C(A_n, R_{\pi_n})| > 1$ then $C(A_n, R_{\pi_n}) = \{(\varnothing, y) \in A \mid y \in X\}$. 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 ρ 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 ρ, $f^{AV,\rho}$ is unanimous and non-bossy but it is not strategy-proof.

Any constant rule satisfies strategy-proofness and non-bossiness but fails unanimity.

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 is not non-bossy.

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. and A. Perea (2002): "Supporting others and the evolution of influence," *Journal of Economic Dynamics & Control* 26, 2051–2092.

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

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

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

[7] Bergantiños, G., J. Massó and A. Neme (2015): "The division problem under constraints," *Games and Economic Behavior* 89, 56–77.

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

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

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

[11] 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.

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

[13] Sobel, J. (2000): "A model of declining standard," *International Economic Review* 41, 295–303.

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

[15] Thomson, W. (2014): "Non-bossiness," Rochester Center for Economic Research, Working paper 586.