Soc Choice Welfare (2004) 23: 229–247
DOI: 10.1007/s00355-003-0246-6

Social Choice and Welfare


# Stability and voting by committees with exit

**Dolors Berga**[1], **Gustavo Bergantiños**[2], **Jordi Massó**[3], **Alejandro Neme**[4]

[1] Departament d'Economia, Campus de Montilivi, Universitat de Girona, 17071 Girona, Spain (e-mail: dolors.berga@udg.es)
[2] Departamento de Estadística, Facultade de Economicas, Universidade de Vigo, 36200 Vigo, Pontevedra, Spain (e-mail: gbergant@uvigo.es)
[3] Departament d'Economia i d'Història Econòmica and Code, Universitat Autònoma de Barcelona, 08193 Bellaterra (Barcelona), Spain (e-mail: jordi.masso@uab.es)
[4] 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)

Received: 3 October 2002/ Accepted: 12 May 2003

**Abstract.** We study the problem of a society choosing a subset of new members from a finite set of candidates (as in Barberà et al. 1991). However, we explicitly consider the possibility that initial members of the society (founders) may want to leave it if they do not like the resulting new society. We show that, if founders have separable (or additive) preferences, the unique strategy-proof and stable social choice function satisfying founder's sovereignty (on the set of candidates) is the one where candidates are chosen unanimously and no founder leaves the society.

---

A previous version of this paper was entitled "Voting by Committees with Exit". We are very grateful to the associate editor and two anonymous referees for their detailed suggestions and comments. We also thank Salvador Barberà, Carmen Beviá, Anna Bogomolnaia, Renan Goetz, Matthew Jackson, Howard Petith, Carmelo Rodríguez-Álvarez, James Schummer, and Tayfun Sönmez for their helpful comments and suggestions. The work of D. Berga, G. Bergantiños, and J. Massó is partially supported by Research Grants AGL2001-2333-C02-01, BEC2002–04102-C02-01, and BEC2002-02130 from the Spanish Ministerio de Ciencia y Tecnología, respectively. The work of A. Neme is partially supported by Research Grant 319502 from the Universidad Nacional de San Luis. The work of D. Berga is also partially supported by Research Grant 9101100 from the Universitat de Girona. The work of G. Bergantiños is also partially supported by Research Grant PGIDT00PXI30001PN from the Xunta de Galicia. The work of J. Massó is also partially supported by Research Grant 2001SGR-00162 and the Barcelona Economics Program of CREA from the Generalitat de Catalunya. The paper was partially written while A. Neme was visiting the UAB under a sabbatical fellowship from the Generalitat de Catalunya.

## 1 Introduction

Barberà et al. (1991) considered the problem where a finite set of agents who originally make up a society has to decide which candidates, to be chosen from a given set, will become new members of the society. They analyzed this problem without considering the possibility that current members of the society may want to leave it as a result of its change in composition. In particular, they characterized *voting by committees* as the class of strategy-proof and onto social choice functions whenever founders' preferences over subsets of candidates are either separable or additively representable and founders *cannot* leave the society.

In this paper we are interested in studying the consequences of considering explicitly the possibility that founders have the option to leave the group in case they do not like the resulting composition of the society. In our context, a social choice function is a rule that associates with each founders' preference profile a newly composed society consisting of both candidates and founders. This set up is sufficiently general to include as social choice functions mechanisms which select, given each founders' preference profile, the new composition of the society in a potentially complex procedure. For instance, mechanisms where the subset of admitted candidates is first selected (using a pre-specified voting rule) and then founders decide sequentially to stay or to leave the society after being informed about the chosen candidates.

Notice that our framework is not a particular case of Barberà et al. (1991) model. One of the main consequences of the fact that a founder might leave the society is that each founder's preferences have to be defined on subsets where he is excluded. We will assume that founders are indifferent between any pair of societies to which they do not belong. Moreover, for all societies *containing* a given founder, we will assume, as in Barberà et al. (1991), that this founder has separable preferences. A founder has *separable* preferences if the division between good and bad agents guides the ordering of subsets of agents, in the sense that adding a good agent leads to a better set, while adding a bad agent leads to a worse set. However, when considered as binary relations on the set of all possible societies our separability condition is not the same as Barberà et al.'s (1991).[1]

We are especially interested in social choice functions satisfying the property that no founder ever has an incentive to misrepresent his preferences in order to obtain personal advantages. Functions satisfying this property are called *strategy-proof* social choice functions.[2] In order to capture the main feature of our problem, we will concentrate on social choice functions that are *stable* in the sense that no founder that remains in the final society wants to

[1] At the end of Sect. 3, and after presenting Barberà et al.'s (1991) model, we compare the two preference domains.

[2] See Sprumont (1995), Barberà (1996), and Barberà (2001) for three excellent surveys on strategy-proofness.

leave it (internal stability) and no founder that left the society wants to rejoin it (external stability). Finally, we require that social choice functions satisfy the property of *founder's sovereignty on the set of candidates*. It implies that a function must be sensitive to founders' preferences in two ways: all commonly agreed good candidates have to be elected, and no commonly agreed bad candidates can be elected.

Our main result demonstrates that the unique strategy-proof and stable social choice function satisfying founder's sovereignty on the set of candidates is the one such that, for each profile of separable preferences satisfying the non-initial exit condition,[3] the final chosen society consists of *all* initial founders and the *unanimously* good candidates. In other words, founders do not leave the society, but the existence of such a possibility reduces substantially the number of ways candidates are elected. Stability requires the use of the most qualified majority to get candidates in. But again, this extremely qualified majority makes exit unnecessary since each founder has veto power for all candidates and the original society was originally acceptable for all founders.[4] We also show that not only stability, strategy-proofness, and founder's sovereignty on the set of candidates are independent properties but also that once we relax one of the two stability criteria new social choice functions appear where some founders leave the society at some preference profiles.

However, our model is not limited to the interpretation given so far; i.e., the choice of the composition of the final society. It can be also used to analyze the problem where a society has to define its formal and public positions on a set of issues. One can think of political parties or religious communities deciding on different issues like abortion, death penalty, health reform, and so on. A social choice function should be understood as deciding both on the composition of the new society (as a set of members) and on the set of approved issues. We require that the first decision be stable.

Before concluding this Introduction, we want to comment on two lines of research existing in the literature. The first one is composed of two recent and related papers. Barberà et al. (2001) consider a society that, during a fixed and commonly known number of periods, may admit in each period a subset of new members. Within this dynamic setup, an interesting issue arises: voters, at earlier stages, vote not only according to whether or not they like a candidate but also according to their tastes concerning future candidates. They study

[3] A profile of preferences satisfies the non-initial exit condition if no founder wants to exit the initial society. See condition (C4) in Sect. 2 for a formal statement of this property.

[4] Regional free trade associations and alliances such as NATO seem to generally require unanimous assent, or something close to it, before admitting new members. The importance of stability in connection with such organizations is evident, and this paper gives a strong theoretical connection between stability and conservative standards for admitting new members.

the particular case where agents have dichotomous preferences (candidates are either friends or enemies) and the voting rule used by the society is quota one (it is sufficient to receive one vote to be elected). They identify and study (subgame perfect and trembling-hand perfect) *equilibria* where agents exhibit, due to the dynamics of the game, complex strategic voting behavior.

Granot et al. (2000) study a similar model with expulsion; current members of the society have to decide each period whether to admit new members into the society *and* whether to expel current members of the society for good. They study equilibria for different protocols which depend on whether the expulsion decision has to be taken each period either simultaneously with, before, or after the admission decision.

In contrast to the works cited above, our framework is static. In particular, candidates in our model do not count: they do not have preferences over societies. We are implicitly assuming that they want to become new members of the society regardless of its final composition, and this is restrictive. But this hypothesis allows us to include the interpretation offered earlier where the society has to decide a subset of binary issues which cannot have preferences. Moreover, our paper also differs from the mentioned ones because of the following three features. First, our focus is on *voluntary exit* rather than expulsion; it seems to us that voluntary exit is a relevant and common problem societies face (members often leave a society just by not renewing their annual membership rather than being expelled). Second, we do not restrict ourselves to specific protocols or specific voting rules. Our setup is general and corresponds to the standard framework used in social choice theory: social choice functions mapping agents' preferences into the set of social alternatives. Third, our main interest is in identifying strategy-proof social choice functions instead of analyzing different types of equilibria.

The second line of research started with a work by Dutta et al. (2001) on candidate stability by considering only single-valued voting rules, and continued with the work of Ehlers and Weymark (2003), Eraslan and McLennan (2001), and Rodríguez-Álvarez (2001) on multi-valued voting rules. In these papers, a set of voters and a set of candidates (which may overlap) must select a representative candidate (or a subset of them). The key issue this literature addresses is the incentives of candidates, given a particular voting rule (how voters choose a candidate or a subset of candidates), to enter or exit the election in order to strategically affect the outcome of the rule. By imposing some independence conditions and an "internal stability" condition (the losing candidates must not have an incentive to drop out of the election) they prove that the class of voting rules immune to this strategic manipulation is only composed of dictatorial rules. In contrast to our paper, these articles consider the stability condition to be "strategic" in the sense that, if considering exiting, an agent anticipates the new choice with the smaller set of candidates.

The paper is organized as follows. We introduce preliminary notation and basic definitions in Sect. 2. Section 3 contains the description and characterization of voting by committees due to Barberà et al. (1991) and compares both models (with and without exit). In Sect. 4 we state and prove our main result and in Sect. 5 we show the independence of the axioms characterizing it. Section 6 analyzes the relevancy of the non-initial exit condition for our result. Some concluding remarks are included in Sect. 7 while all omitted proofs of Sect. 5 are in the Appendix.

## 2 Preliminary notation and definitions

Let $N = \{1, ..., n\}$ be the set of *founders* of a society and $K=\{n+1, ..., k\}$ be the set of *candidates* who may become new members of the society. We assume that $n$ and $k$ are finite, $n \geq 2$, and $k \geq 3$. Founders have *preferences* over $2^{N \cup K}$, the set of all possible final societies. We identify the empty set with the situation where the society has no members.[5]

Founder $i$'s *preferences* over $2^{N \cup K}$, denoted by $R_i$, is a complete and transitive binary relation. As usual, let $P_i$ and $I_i$ denote the strict and indifference preference relations induced by $R_i$, respectively. We suppose that founders' preferences satisfy the following conditions:

(C1) *Strictness:*For all $S, S' \subset N \cup K$, $S \neq S'$ such that $i \in N \cap S \cap S'$, either $SP_iS'$ or $S'P_iS$.
(C2) *Indifference:*For all $S$ such that $i \notin S$ , $SI_i\emptyset$.
(C3) *Loneliness:* (a) $\{i\}R_i\emptyset$. (b) If $SI_i\emptyset$ and $i \in S$, then $S = \{i\}$.
(C4) *Non-initial Exit:*For all $i \in N$, $NP_iN\backslash\{i\}$.

*Strictness* means that founder $i$'s preferences over sets containing himself are strict. *Indifference* means that founder $i$ is indifferent between not belonging to the society and the situation where the society has no members. Part (a) of *Loneliness* means that either founder $i$ finds specific benefits to being the only member of the society (in which case $\{i\}P_i\emptyset$) or else, founder $i$ could provide them without being a member of the society (in which case $\{i\}I_i\emptyset$), while part (b) says that the only society containing $i$ that may be indifferent to not being in the society is the society formed by $i$ alone. Finally, the *Non-initial Exit* condition says that no founder wants to exit the initial society.[6]

We denote by $\mathscr{R}_i$ the set of all such preferences for founder $i$, by $\mathscr{R}$ the Cartesian product $\mathscr{R}_1 \times \cdots \times \mathscr{R}_n$, by $\widehat{\mathscr{R}}_i$ a generic subset of $\mathscr{R}_i$, and by $\widehat{\mathscr{R}}$ the Cartesian product $\widehat{\mathscr{R}}_1 \times \cdots \times \widehat{\mathscr{R}}_n$. Notice that conditions (C1), (C2), (C3), and (C4) are founder specific and therefore $\mathscr{R}_i \neq \mathscr{R}_j$ for different founders $i$ and $j$.

---

[5] Remember that, as we already argued in the Introduction, we could interpret the set $K$ as the set of issues that the society has to decide upon. In this case the interpretation of a final society is the subset of approved issues and the subset of members that remain in the society.

[6] In Sect. 6 we will argue that we need condition (C4) for the existence of "stable" social choice functions.

Given $R_i \in \mathscr{R}_i$, denote by $\tau(R_i)$ the best element of $2^{N \cup K}$ according to $R_i$. As a consequence of conditions (C1), (C2), and (C4) this element is unique.

A *preference profile* $R = (R_1, ..., R_n) \in \mathscr{R}$ is a $n$-tuple of preferences. It will be represented by $(R_i, R_{-i})$ to emphasize the role of founder $i$'s preference.

A *social choice function* $f$ is a function $f : \widehat{\mathscr{R}} \longrightarrow 2^{N \cup K}$. Given a social choice function $f$, we will denote by $f_N$ and $f_K$ the functions that specify the subsets of $N$ and $K$, respectively. Namely, $f_N(R) = f(R) \cap N$ and $f_K(R) = f(R) \cap K$ for all $R \in \widehat{\mathscr{R}}$.

Now we define two basic properties that social choice functions may satisfy. The first one is *strategy-proofness*. It says that no founder can gain by lying when reporting his preferences.

**Definition 1.** *A social choice function* $f : \widehat{\mathscr{R}} \to 2^{N \cup K}$ *is* strategy-proof *if for all* $R = (R_1, ..., R_n) \in \widehat{\mathscr{R}}$, $i \in N$, *and* $R_i' \in \widehat{\mathscr{R}}_i$,

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

If $f\left(R_i', R_{-i}\right) P_i f(R)$, we say that *founder* $i$ *manipulates* $f$ *at profile* $R$ *via* $R_i'$.

We are especially interested in social choice functions satisfying the property of *stability* in a double sense: *internal stability* (no founder that remains in the final society wants to leave it) and *external stability* (no founder that left the society wants to rejoin it). Formally,

**Definition 2.** *A social choice function* $f : \widehat{\mathscr{R}} \to 2^{N \cup K}$ *satisfies* internal stability *if for all* $R \in \widehat{\mathscr{R}}$,

$$i \in f(R) \cap N \Longrightarrow f(R) R_i (f(R) \backslash \{i\}).$$

*A social choice function* $f : \widehat{\mathscr{R}} \to 2^{N \cup K}$ *satisfies* external stability *if for all* $R \in \widehat{\mathscr{R}}$,

$$i \in N \text{ and } i \notin f(R) \Longrightarrow f(R) R_i (f(R) \cup \{i\}).$$

*A social choice function* $f$ *is stable if* $f$ *satisfies internal and external stability.*

As in Barberà et al. (1991) we will restrict ourselves to preferences that order subsets of agents (containing agent $i$) according to two basic characteristics of their elements. Consider a preference $R_i \in \mathscr{R}_i$ and an agent $j \in K \cup N \backslash \{i\}$. We say that *$j$ is good for $i$ according to $R_i$* whenever $\{j, i\} P_i \{i\}$; otherwise, we say that *$j$ is bad for $i$ according to $R_i$*. Denote by $G(R_i)$ and $B(R_i)$ the set of good and bad agents for $i$ according to $R_i$, respectively. To simplify notation, let $G_K(R_i) = G(R_i) \cap K$, $B_K(R_i) = B(R_i) \cap K$, and $G_N(R_i) = G(R_i) \cap N$. Now, we are ready to formally define separable preferences.

**Definition 3.** *A preference* $R_i \in \mathscr{R}_i$ *is* separable *if for all* $j \in K \cup N \backslash \{i\}$ *and* $S \subset N \cup K \backslash \{j\}$ *such that* $i \in S$,

$$[\{j\} \cup S] P_i S \text{ if and only if } j \in G(R_i).$$

Let $\mathscr{S}_i \subset \mathscr{R}_i$ denote the set of separable preferences for founder $i$ that satisfy (C1)–(C4) and let $\mathscr{S}$ denote the Cartesian product $\mathscr{S}_1 \times ... \times \mathscr{S}_n$.

A careful examination of all preferences used in all proofs below shows that the statements of our results still hold if we consider social choice functions defined on the subdomain of additive preferences, where a preference $R_i \in \mathscr{R}_i$ is said to be *additive* if there exists a function $u_i : N \cup K \cup \{\emptyset\} \to \boldsymbol{R}$ such that for all $S$ and $S'$ with $i \in S \cap S'$,

$$SP_iS' \text{ if and only if } \sum\nolimits_{x\in S} u_i(x) > \sum\nolimits_{y\in S'} u_i(y)$$

and

$$SP_i\emptyset \text{ if and only if } \sum\nolimits_{x\in S} u_i(x) > u_i(\emptyset).$$

Note that additivity implies separability but the converse is false for $k > 3$, since a separable ordering $R_1$ could simultaneously have $\{1,3\}P_1\{1,4\}$ and $\{1,2,4\}P_1\{1,2,3\}$. However, if $R_1$ were additive, $\{1,3\}P_1\{1,4\}$ would imply $\{1,2,3\}P_1\{1,2,4\}$, but this would seem too restrictive, though, to capture some degree of complementarity among agents, which can be very natural in our setting.

We are also interested in social choice functions satisfying the property of *founder's sovereignty on K* in a double sense. Namely, candidates that are good for all founders have to be admitted to the society. On the contrary, candidates that are bad for all founders cannot be admitted. Formally,

**Definition 4.** *A social choice function* $f : \widehat{\mathscr{R}} \to 2^{N\cup K}$ *satisfies* founder's sovereignty on $K$ *if for all* $R \in \widehat{\mathscr{R}}$,

$$\bigcap_{i\in N} G_K(R_i) \subseteq f_K(R) \subseteq \bigcup_{i\in N} G_K(R_i).$$

Barberà et al. (1991) characterized the class of strategy-proof and onto social choice functions without exit (see Proposition 1 in Sect. 3). They used the phrase *voters' sovereignty* to indicate the onto condition (for all $K' \subseteq K$, there exists $R \in \widehat{\mathscr{R}}$ such that $f_K(R) = K'$). Our founder's sovereignty (on $K$) condition is stronger. However, our condition is reasonable because, in addition to ontoness, it only requires the natural coherence between the preference profile and its corresponding subset of elected candidates.

## 3 Voting by committees

In this section we first present the main ingredients of Barberà et al.'s (1991) model in order to state their characterization of voting by committees, on which part of our proof is built. We finish the section with a discussion of the differences between the two models.

Since in the problem considered by Barberà et al. (1991) founders cannot leave the society, the social alternatives are subsets of candidates. Therefore, founder $i$'s preferences, denoted by $\succsim_i$, is a complete, asymmetric and transitive binary relation over $2^K$. As usual, let $\succ_i$ denote the strict preference relation induced by $\succsim_i$. Let $\tau(\succsim_i)$ denote the best element of $2^K$ according to $\succsim_i$ and let $\succsim = (\succsim_1, ..., \succsim_n)$ be a preference profile.

**Definition 5.** *A preference* $\succsim_i$ *is* BSZ- separable *if for all* $S \subseteq K$ *and all* $x \notin S$,

$$S \cup \{x\} \succ_i S \text{ if and only if } \{x\} \succ_i \emptyset.$$

Let $\mathcal{S}_i^{BSZ}$ be the set of all BSZ-separable preferences on $2^K$ (note that this set is the same for all founders) and let $\mathcal{S}^{BSZ} = \mathcal{S}_1^{BSZ} \times ... \times \mathcal{S}_n^{BSZ}$.

A *voting scheme* $g$ is a function from $\mathcal{S}^{BSZ}$ to $2^K$. A voting scheme $g$ is *strategy-proof* if it satisfies the natural translation of Definition 1 to this setup.

We now turn to defining voting by committees. Rules in this class are defined by a collection of families of winning coalitions (committees), one for each candidate. Founders vote for sets of candidates. To be elected, a candidate must get the vote of all members of some coalition among those that are winning for that candidate. Formally,

**Definition 6.** *A committee* $\mathcal{W}$ *is a nonempty family of nonempty coalitions of* $N$, *which satisfies coalition monotonicity in the sense that if* $I \in \mathcal{W}$ *and* $I' \supseteq I$, *then* $I' \in \mathcal{W}$.

Coalition $I \in \mathcal{W}$ is a *minimal winning coalition* if, for all $I' \subsetneq I$, $I' \notin \mathcal{W}$. Given a committee $\mathcal{W}$ we denote by $\mathcal{W}^m$ the set of minimal winning coalitions and call it the *minimal committee*.

**Definition 7.** *A voting scheme* $g : \mathcal{S}^{BSZ} \to 2^K$ *is* voting by committees *if for each* $x \in K$, *there exists a committee* $\mathcal{W}_x$ *such that for all* $\succsim \in \mathcal{S}^{BSZ}$

$$x \in g(\succsim) \text{ if and only if } \{i \in N \mid x \in \tau(\succsim_i)\} \in \mathcal{W}_x.$$

**Proposition 1.** *(Theorem 1 in Barberà et al. 1991) A voting scheme* $g : \mathcal{S}^{BSZ} \to 2^K$ *is strategy-proof and onto if and only if* $g$ *is voting by committees.*

We could now extend voting by committees to our context by saying that a social choice function $f : \mathcal{S} \to 2^{N \cup K}$ is *voting by committees* if for each agent $x$ (founder and candidate) there exists a committee $\mathcal{W}_x$ such that for all $R \in \mathcal{S}$,

$$x \in f(R) \text{ if and only if } \{i \in N \mid x \in \tau(R_i)\} \in \mathcal{W}_x.$$

We now argue that the two models are different due to fundamental differences of the two preference domains. The following three are crucial.

First, to deal with voluntary exit and voluntary membership we allow a founder's preference of joining a society to depend on the other members in the society; that is, founder $i$ may prefer joining a society $S$ to not joining it, i.e., $S \cup \{i\} P_i S$ and *at the same time*, prefer not joining another society $S'$ to joining it, i.e., $S' P_i S' \cup \{i\}$ (so BSZ-separability is violated). Second, each founder is indifferent to any two societies to which he does not belong. Third, each founder belongs to his best society; that is, $i \in \tau(R_i)$ for all $R_i \in \mathscr{S}_i$ and $i \in N$ (this holds by transitivity and (C2) since $\tau(R_i) R_i N$ and $N P_i \emptyset$ by (C4)). We think that these three aspects are meaningful and necessary to deal with the social choice problem we want to study here. We want to emphasize that, due to these domain differences, Barberà et al.'s (1991) model cannot be applied directly here, although we will use their main result after showing that no founder ever wants to leave the society.

Furthermore, and as a consequence of the fact that each $i$ belongs to $\tau(R_i)$ (each founder always votes for himself) we have now an insubstantial multiplicity of voting by committees inducing the same social choice function. To see that, consider the following two possibilities. On the one hand, consider any pair of committees $\mathscr{W}$ and $\mathscr{W}'$ such that $\mathscr{W}_x = \mathscr{W}'_x$ for all $x \in K$ and for any founder $i$, $\mathscr{W}'_i = \left\{ \{S \cup \{i\}\}_{S \in \mathscr{W}_i} \right\}$. Since $i \in \tau(R_i)$ for all $i \in N$ and all $R_i \in \mathscr{S}_i$, we conclude that both committee structures ($\mathscr{W}$ and $\mathscr{W}'$) induce the same social choice function. On the other hand, if $\mathscr{W}$ and $\mathscr{W}'$ are such that $\{i\} \in \mathscr{W}_i$ and $\{i\} \in \mathscr{W}'_i$ for all $i \in N$, and $\mathscr{W}_x = \mathscr{W}'_x$ for all $x \in K$, then both committee structures induce the same social choice function. Therefore, because of these two situations, from now on and in order to state our results more compactly, we will assume that a committee for founder $i$ is a nonempty family of subsets containing $i$. Formally,

**Definition 8.** *A social choice function $f : \mathscr{S} \to 2^{N \cup K}$ is voting by committees if for each $x \in N \cup K$ there exists a committee $\mathscr{W}_x$ such that for all $R \in \mathscr{S}$,*

$$x \in f(R) \text{ if and only if } \{i \in N \mid x \in \tau(R_i)\} \in \mathscr{W}_x,$$

*where for all $i \in N$ and all $I \in \mathscr{W}_i$, $i \in I$.*

## 4 The characterization result

Theorem 1 below characterizes the class of strategy-proof and stable social choice functions satisfying founder's sovereignty on $K$ as the voting by committees social choice function satisfying the properties that the minimal committee of each founder is himself and the minimal committee for each candidate is the set of all founders. That is, it is the single rule which chooses, for each preference profile, the final society consisting of *all initial founders and all unanimously good candidates*. Formally,

**Theorem 1.** *Let $f : \mathscr{S} \longrightarrow 2^{N \cup K}$ be a social choice function. Then, $f$ is strategy-proof, stable, and satisfies founder's sovereignty on $K$ if and only if $f$ is voting by committees with the following two properties:*

(*Founders*) *For all $i \in N$, $\mathscr{W}_i^m = \{\{i\}\}$.*
(*Candidates*) *For all $x \in K$, $\mathscr{W}_x^m = \{N\}$.*

Remember that the assumptions about the domain, (C1)-(C4), have been incorporated into the definition of $\mathscr{S}$.

**Remark 1.** *Alternatively, we can write the social choice function characterized above as follows: for all $R \in \mathscr{S}$, $f(R) = N \cup (\bigcap_{i \in N} G_K(R_i))$.*

*Proof of Theorem 1.* To prove sufficiency, assume that for all $R \in \mathscr{S}$, $f(R) = N \cup (\bigcap_{i \in N} G_K(R_i))$. Clearly, $f$ satisfies external stability and founder's sovereignty on $K$. Since $f_N(R) = N$, $f_K(R) \subset G_K(R_i)$ for all $i \in N$, and preferences are separable and satisfy (C2) and (C4), we have that $f(R) R_i N P_i N \backslash \{i\} I_i (f(R) \backslash \{i\})$ for all $i \in N$ which shows that $f$ satisfies internal stability.

To show that $f$ is strategy-proof, let $i \in N$, $R \in \mathscr{S}$, and $R'_i \in \mathscr{S}_i$ be arbitrary and suppose that $f(R) \neq f(R'_i, R_{-i})$ (otherwise, the proof is trivial). Since $f_N(R) = f_N(R'_i, R_{-i}) = N$, there must exist $x \in K$ such that either $x \in f_K(R)$ and $x \notin f_K(R'_i, R_{-i})$ or else $x \notin f_K(R)$ and $x \in f_K(R'_i, R_{-i})$. Note that for both cases, $f_K(R) = \bigcap_{j \in N} G_K(R_j)$ and $f_K(R'_i, R_{-i}) = G' \cup B'$ where $G' \subset G_K(R_i)$, $B' \subset B_K(R_i)$, and $G' \subset \bigcap_{j \in N} G_K(R_j)$. Then, since $R_i$ is a separable preference we obtain $(N \cup \bigcap_{j \in N} G_K(R_j)) P_i (N \cup G' \cup B')$; that is, $f(R) P_i f(R'_i, R_{-i})$ which shows that $f$ is strategy-proof.

To prove necessity, let $f$ be a strategy-proof and stable social choice function satisfying founder's sovereignty on $K$. First note that the following claim holds.

*Claim 1. If $R \in \mathscr{S}$ is such that $G_K(R_i) = A$ for all $i \in N$, then $f(R) = N \cup A$.*

*Proof of Claim 1.* Let $R \in \mathscr{S}$ be such that $G_K(R_i) = A$ for all $i \in N$. By founder's sovereignty on $K$, $f_K(R) = A$. To prove that $f_N(R) = N$ we use an induction argument. First observe that $f_N(R) \neq N \backslash \{i\}$ for all $i \in N$; otherwise, if $f_N(R) = N \backslash \{i\}$ for some $i \in N$, $f$ would not be externally stable since, by separability of $R_i$, $(N \cup f_K(R)) R_i N$, by (C4), $N P_i N \backslash \{i\}$, by (C2), $N \backslash \{i\} I_i (N \cup f_K(R)) \backslash \{i\}$, and by transitivity, $(N \cup f_K(R)) P_i (N \cup f_K(R)) \backslash \{i\}$.

*Induction hypothesis.* Suppose that for all $R \in \mathscr{S}$ such that $G_K(R_i) = A$ for all $i \in N$ and for all $S \subset N$ such that $1 \leq \#S \leq s < n$, $f_N(R) \neq N \setminus S$.[7]

[7] The symbol # stands for the cardinality of a set. Observe that $f_N(R) \neq N \backslash S$ means that $f_N(R)$ either equals $N$ or has less than $n - s$ elements.

We will show that for all $R \in \mathscr{S}$ such that $G_K(R_i) = A$ for all $i \in N$ and for all $T \subset N$ with $\#T = s+1$, $f_N(R) \neq N \backslash T$. Suppose there exists $R \in \mathscr{S}$ and $T \subset N$ with $\#T = s+1$ such that $f_N(R) = N \setminus T$.

Consider $i_1 \in T$ and $R'_{i_1} \in \mathscr{S}_{i_1}$ such that $G(R'_{i_1}) = (N \setminus \{i_1\}) \cup A$ and $\{i_1\} P'_{i_1} \emptyset$. We define $R^{(1)} = (R'_{i_1}, R_{-i_1})$. By founder's sovereignty on $K$, $f_K(R^{(1)}) = A$. Note that $f(R^{(1)}) \subset G(R'_{i_1}) \cup \{i_1\}$. Then, by separability $\left(f(R^{(1)}) \cup \{i_1\}\right) P'_{i_1} \emptyset$ and by external stability, $i_1 \in f_N(R^{(1)})$. The induction hypothesis implies that we can write $f_N(R^{(1)}) = N \backslash T^{(1)}$ for some $T^{(1)}$ such that $\#T^{(1)} \in [s+1, n-1]$ or $\#T^{(1)} = 0$. If $\#T^{(1)} = 0$, that is, $f_N(R^{(1)}) = N$, we have that $f(R^{(1)}) = (N \cup A) P_{i_1} \emptyset I_{i_1} f(R)$, which means that $i_1$ manipulates $f$ at $R$ via $R'_{i_1}$ contradicting strategy-proofness of $f$. Thus, $\#T^{(1)} \in [s+1, n-1]$.

Consider $i_2 \in T^{(1)}$ and $R'_{i_2} \in \mathscr{S}_{i_2}$ such that $G(R'_{i_2}) = (N \setminus \{i_2\}) \cup A$ and $\{i_2\} P'_{i_2} \emptyset$. Define $R^{(2)} = (R'_{i_2}, R^{(1)}_{-i_2})$. Using similar arguments to those used above for $i_1$ we can conclude that $\{i_1, i_2\} \subset f_N(R^{(2)}) = N \setminus T^{(2)}$ where $\#T^{(2)} \geq s+1$. Repeating this process we obtain that there exists $V \subset N$ such that $V \subset f_N(R^{(n-s)}) = N \setminus T^{(n-s)}$ where $\#T^{(n-s)} \geq s+1$ and $\#V = n-s$, which is a contradiction. ■

We decompose the necessity part of the proof into two Lemmata.

**Lemma 1.** *For all $R \in \mathscr{S}$, $f_N(R) = N$.*

*Proof of Lemma 1.* We use an induction argument over all good candidates. Let $R \in \mathscr{S}$ be arbitrary and define $m = \sum_{i \in N} \#G_K(R_i)$. If $m = 0$, we get that $f_N(R) = N$ by Claim 1.

*Induction hypothesis.* Suppose that $f_N(R) = N$ holds for all $R \in \mathscr{S}$ such that $m \leq l$.

To prove that $i \in f(R)$ for all $i \in N$ and all $R \in \mathscr{S}$ such that $m = l+1$, we distinguish the following two cases:

- $G_K(R_i) \neq \emptyset$.

Consider any $R'_i \in \mathscr{S}_i$ with the properties that $G_K(R'_i) = \emptyset$, $G(R'_i) = G_N(R_i)$, and

$$\text{if } S R'_i \emptyset, \text{ then } S R_i \emptyset. \tag{1}$$

The reader can check that, by making all candidates in $G_K(R_i)$ extremely bad (i.e., $\emptyset P'_i S$ whenever $S \cap K \neq \emptyset$), such a preference exists. By the induction hypothesis, $f_N(R'_i, R_{-i}) = N$. By strategy-proofness, $f(R) R_i f(R'_i, R_{-i})$, and by internal stability, $f(R'_i, R_{-i}) R'_i \emptyset$. By condition (1) in the construction of $R'_i$, $f(R'_i, R_{-i}) R_i \emptyset$. Since $N \subseteq f(R'_i, R_{-i})$ and part (b) of the loneliness condition (C3), $f(R'_i, R_{-i}) P_i \emptyset$. Therefore, by transitivity of $R_i$, $f(R) P_i \emptyset$ holds. Moreover, by the indifference condition (C2), $i \in f(R)$.

- $G_K(R_i) = \emptyset$.

Suppose that $i \notin f(R)$. Since $m \geq 1$, there exists $j \in N$ such that $G_K(R_j) \neq \emptyset$. By the previous case, $j \in f(R)$. Consider any $R'_j \in \mathcal{S}_j$ with the property that $G_K(R'_j) = \emptyset$ and $SP'_jS'$ for all $S, S'$ such that $i \in S' \backslash S$ and $j \in S \cap S'$. By the induction hypothesis, $f_N(R'_j, R_{-j}) = N$. Since $i \in f(R'_j, R_{-j})$ and $i \notin f(R)$, by definition of $R'_j$, $f(R)P'_jf(R'_j, R_{-j})$, which contradicts strategy-proofness.

Hence, for all $R \in \mathcal{S}$, $f_N(R) = N$. ■

**Lemma 2.** *For all* $R \in \mathcal{S}$, $f_K(R) = \bigcap_{i \in N} G_K(R_i)$.

*Proof of Lemma 2.* We will now use the result of Barberà, et al. (1991) stated in Proposition 1 above. In order to do so, we will identify our $f_K : \mathcal{S} \to 2^K$ with a voting scheme over $\mathcal{S}^{BSZ}$, $g : \mathcal{S}^{BSZ} \to 2^K$ as follows: Given $\succsim_i \in \mathcal{S}_i^{BSZ}$ choose any $R_i \in \mathcal{S}_i$ such that $(N \cup S)P_i(N \cup S')$ if and only if $S \succ_i S'$ for all distinct $S, S' \in 2^K$. Therefore, we have defined a mapping $p : \mathcal{S}^{BSZ} \to \mathcal{S}$; notice that there are many $p$'s. Define $g : \mathcal{S}^{BSZ} \to 2^K$ as follows: $g(\succsim) := f_K(p(\succsim))$ for all $\succsim \in \mathcal{S}^{BSZ}$. We want to show that $g$ is well-defined, strategy-proof, and onto.

- $g$ is well-defined.

It is sufficient to show that, for all $\succsim \in \mathcal{S}^{BSZ}$, $f_K(p^1(\succsim)) = f_K(p^2(\succsim))$ for any pair of functions $p^1$ and $p^2$. Assume otherwise; that is, there exist $\succsim \in \mathcal{S}^{BSZ}$, $p^1$ and $p^2$ such that $f_K(p^1(\succsim)) = S^1 \neq S^2 = f_K(p^2(\succsim))$. Hence, $p^1(\succsim) \neq p^2(\succsim)$. Let $p^1(\succsim) = (R_1^1, ..., R_n^1)$ and $p^2(\succsim) = (R_1^2, ..., R_n^2)$ be the two different preference profiles. By Lemma 1, all $f(R)$ are of the form $N \cup S$; going from $f(R^1) = N \cup S^1$ to $f(R^2) = N \cup S^2$, there exist $M \subseteq N$ and $i \in M$ such that $f(R_M^1, R_{-M}^2) = N \cup S^1$ and $f(R_{M\backslash\{i\}}^1, R_{-(M\backslash\{i\})}^2) = N \cup T$ with $T \neq S^1$ (eventually, $T$ could be equal to $S^2$). By the strictness condition (C1) either $(N \cup T)P_i^1(N \cup S^1)$ or $(N \cup S^1)P_i^1(N \cup T)$ . If $(N \cup T)P_i^1(N \cup S^1)$, then $i$ manipulates $f$ at profile $(R_M^1, R_{-M}^2)$ with $R_i^2$. If $(N \cup S^1)P_i^1(N \cup T)$, and hence $(N \cup S^1)P_i^2(N \cup T)$, then $i$ manipulates $f$ at profile $(R_{M\backslash\{i\}}^1, R_{-(M\backslash\{i\})}^2)$ with $R_i^1$.

- $g$ is strategy-proof.

Assume otherwise; that is, there exist $\succsim \in \mathcal{S}^{BSZ}$, $i \in N$, and $\succsim'_i \in \mathcal{S}_i^{BSZ}$ such that $g(\succsim'_i, \succsim_{-i}) \succ_i g(\succsim)$. Since $g$ is well-defined, we can find $R \in \mathcal{S}$, $R'_i \in \mathcal{S}_i$, and $p$ such that $p(\succsim) = R$ and $p(\succsim'_i, \succsim_{-i}) = (R'_i, R_{-i})$. Therefore, by Lemma 1 and the definition of $g$ and $p$, $f(R'_i, R_{-i}) = (N \cup g(\succsim'_i, \succsim_{-i}))P_i(N \cup g(\succsim)) = f(R)$, which implies that $f$ is not strategy-proof.

- $g$ is onto $2^K$.

This is an immediate consequence of Claim 1, using the definitions of $g$ and $p$.

Then by Proposition 1, $g$ is voting by committees. Let $\{\mathcal{W}_x\}_{x \in K}$ be its associated family of committees. We next show that $f$ is voting by committees. Given $R \in \mathcal{S}$ let $p$ and $\succsim \in \mathcal{S}^{BSZ}$ be such that $p(\succsim) = R$ (the strictness condition (C1) guarantees the existence of a unique preference profile $\succsim$).

Notice that, for all $i \in N$, separability implies that $G_K(R_i) = \tau(\succsim_i)$. Therefore, for each $x \in K$,

$$x \in f_K(R) \Longleftrightarrow x \in g(\succsim)$$
$$\Longleftrightarrow \{i \in N \mid x \in \tau(\succsim_i)\} \in \mathscr{W}_x$$
$$\Longleftrightarrow \{i \in N \mid x \in G_K(R_i)\} \in \mathscr{W}_x.$$

To show that all minimal committees coincide with $\{N\}$, assume that there exist $x \in K$ and $S \subsetneq N$ such that $S \in \mathscr{W}_x^m$. Take $i \in N \backslash S$ and $R \in \mathscr{S}$ where for all $j \in S$, $x \in G_K(R_j)$, and

$$\emptyset P_i T \text{ whenever } x \in T. \tag{2}$$

Then, $x \in f(R)$. By Lemma 1, $i \in f(R)$. But this and conditions (C2) and (2) contradict internal stability of $f$. This ends the proof of Lemma 2. ■

By Remark 1, the statement of Theorem 1 follows from Lemmata 1 and 2. ■

## 5 Independence of the axioms

In this section we show the independence of all properties used in the characterization of Theorem 1.

Note first that the constant function $f(R) = N$ for all $R \in \mathscr{S}$ is strategy-proof and stable but it does not satisfy founder's sovereignty on $K$.

Second, there exist social choice functions satisfying founder's sovereignty on $K$ and stability but not strategy-proofness. For any $R \in \mathscr{S}$ define

$$T(R) = \{S \subset \bigcup_{j \in N} G_K(R_j) \mid (N \cup S) R_i (N \cup (\bigcap_{j \in N} G_K(R_j)) \text{ for all } i \in N\}.$$

Consider now the social choice function $f : \mathscr{S} \to 2^{N \cup K}$ such that $f(R) = N \cup B$ where $B \in T(R)$ and $(N \cup B) P_1 (N \cup S)$ for any $S \in T(R) \backslash \{B\}$. Of course $f$ satisfies founder's sovereignty on $K$ and stability. Because $f_K(R)$ is not equal to $\bigcap_{i \in N} G_K(R_i)$ for all $R \in \mathscr{S}$, Theorem 1 implies that $f$ is not strategy-proof.

Third, there exist strategy-proof social choice functions satisfying founder's sovereignty on $K$ that, although they are not stable, satisfy either internal or external stability. Propositions 2 and 3 below identify, among the class of voting by committees, those that are internal and external stable, respectively.

To state Proposition 2 we need the following definitions. We say that a committee $\mathscr{W}_i$ is *unanimous* if $\mathscr{W}_i^m = \{N\}$; *decisive* if $\mathscr{W}_i^m = \{\{i\}\}$; and *bipersonal* if $\mathscr{W}_i^m = \{\{i,j\}_{j \in N \backslash \{i\}}\}$. When $n = 3$ we say that the committees $\mathscr{W}_i$, $\mathscr{W}_j$, and $\mathscr{W}_l$ are *cyclical* if $\mathscr{W}_i^m = \{\{i,j\}\}$, $\mathscr{W}_j^m = \{\{j,l\}\}$, and $\mathscr{W}_l^m = \{\{l,i\}\}$.

**Proposition 2.** *Assume $f : \mathscr{S} \longrightarrow 2^{N \cup K}$ is voting by committees. Then, $f$ satisfies internal stability if and only if:*

(*Candidates*) $\mathscr{W}_x^m = \{N\}$, *for all* $x \in K$.

(*Founders*) *When* $n \geq 4$, *either* (i) $\mathscr{W}_i$ *is unanimous for all* $i \in N$ *or (ii) for all* $i \in N$, $\mathscr{W}_i$ *is either decisive or bipersonal. When* $n = 3$, *either (i)* $\mathscr{W}_i$ *is unanimous for all* $i \in N$, (ii) $\{\mathscr{W}_1, \mathscr{W}_2, \mathscr{W}_3\}$ *are cyclical, or (iii) for all* $i \in N$, $\mathscr{W}_i$ *is either decisive or bipersonal.*

*Proof.* See the Appendix.

We now characterize the set of voting by committees satisfying external stability.[8]

**Propostion 3.** *Assume* $f : \mathscr{S} \longrightarrow 2^{N \cup K}$ *is voting by committees. Then,* $f$ *satisfies external stability if and only if for all* $i \in N$, $\mathscr{W}_i^m = \{\{i\}\}$.

*Proof.* See the Appendix.

Fourth, all voting by committees social choice functions satisfy strategy-proofness and founder's sovereignty on $K$. Other rules satisfying both properties can be defined by dropping the non-emptiness condition for committees. For instance, those where a subset of founders $N_1$ is *always* in the society and another subset $N_2$ is *never* in the society. These can be expressed as *generalized* voting by committees by allowing that the committees of all founders in $N_2$ be empty (that is, without any winning coalition) and for all $i \in N_1$ the committees of founder $i$ have the singleton $\{i\}$ as minimal winning coalition.

## 6 Necessity of the non-initial exit condition

In this section we argue that the non-initial exit condition (C4) is indispensable for the existence of stable social choice functions; that is, there might not exist social choice functions satisfying stability if (C1), (C2), and (C3) hold but (C4) fails. Examples 1 and 2 below illustrate this fact for separable and non-separable preferences, respectively. Observe that the constant function $f(R_1, ..., R_n) = N$ for all $(R_1, ..., R_n)$ in any domain is stable as long as $R_i$ satisfies (C4) for all $i \in N$ .

*Example 1.* Assume that $N = \{1, 2, 3\}$ ($K$ could be any set of candidates). Let $R$ be the additive preference profile induced by the following utility functions:

| | $u_1$ | $u_2$ | $u_3$ |
|---|---|---|---|
| 1 | 1 | 10 | −5 |
| 2 | −5 | 1 | 2 |
| 3 | 10 | −5 | 1 |
| $x \in K$ | $\pm\varepsilon_x$ | $\pm\varepsilon_x$ | $\pm\varepsilon_x$ |
| $\emptyset$ | 0 | 0 | 0 |

[8] Observe that the two characterizations in Propositions 2 and 3 are established assuming voting by committees. We conjecture that they will still be valid, if instead, we assume that $f$ is a strategy-proof social choice function respecting founder's sovereignty on $K$.

where the absolute value of all $\varepsilon_x$'s are sufficiently small such that $\sum_{x \in K} |\varepsilon_x| < 1$.

Notice that $\emptyset P_3 N$ and by (C2) (Indifference) $N \backslash \{3\} I_3 \emptyset$. Thus, (C4) (Non-initial exit) fails since $N \backslash \{3\} I_3 \emptyset P_3 N$. We now check that there is no social choice function satisfying stability. Let $X$ denote any arbitrary subset of $K$.

- If $f(R) = X$, then $f$ does not satisfy external stability because $1 \notin f(R)$ and $(X \cup \{1\}) P_1 X$.
- If $f(R) = \{1\} \cup X$, then $f$ does not satisfy external stability because $2 \notin f(R)$ and $(f(R) \cup \{2\}) P_2 f(R)$.
- If $f(R) = \{2\} \cup X$, then $f$ does not satisfy external stability because $3 \notin f(R)$ and $(f(R) \cup \{3\}) P_3 f(R)$.
- If $f(R) = \{3\} \cup X$, then $f$ does not satisfy external stability because $1 \notin f(R)$ and $(f(R) \cup \{1\}) P_1 f(R)$.
- If $f(R) = \{1, 2\} \cup X$, then $f$ does not satisfy internal stability because $1 \in f(R)$ and $(f(R) \setminus \{1\}) P_1 f(R)$.
- If $f(R) = \{1, 3\} \cup X$, then $f$ does not satisfy internal stability because $3 \in f(R)$ and $(f(R) \setminus \{3\}) P_3 f(R)$.
- If $f(R) = \{2, 3\} \cup X$, then $f$ does not satisfy internal stability because $2 \in f(R)$ and $(f(R) \setminus \{2\}) P_2 f(R)$.
- If $f(R) = N \cup X$, then $f$ does not satisfy internal stability because $3 \in f(R)$ and $(f(R) \setminus \{3\}) P_3 f(R)$.

*Example 2.* Assume that $N = \{1, 2, 3\}$ and $K = \emptyset$. Consider the non-separable preference profile $R = (R_1, R_2, R_3)$ where

$$\{1,2\} P_1 \{1\} P_1 \emptyset P_1 \{1,3\} P_1 \{1,2,3\}$$
$$\{2,3\} P_2 \{2\} P_2 \emptyset P_2 \{1,2\} P_2 \{1,2,3\}$$
$$\{1,3\} P_3 \{3\} P_3 \emptyset P_3 \{2,3\} P_3 \{1,2,3\}.$$

Observe that for any $i \in N$ the non-initial exit condition (C4) does not hold since $N \backslash \{i\} I_i \emptyset P_i N$. There is no "stable" set of members; that is, for any $S \subseteq N$ either (1) there exists $i \in S$ such that $S \backslash \{i\} P_i S$ or (2) there exists $j \notin S$ such that $S \cup \{j\} P_j S$. Thus, it is not possible to define a stable social choice function on any domain of preferences containing $R$.

## 7 Concluding remarks

In this paper we have expanded Barberà et al.'s (1991) framework of a society choosing new members to allow the possibility of voluntary exit from the society. The voluntary nature of exit is modelled by requiring that outcomes be stable in the sense that each founder prefers the outcome to the one that would result if his membership status was reversed. We have shown that strategy-proofness, stability, and founder's sovereignty on the set of candidates are equivalent to a particular form of voting by committees: founders decide for themselves whether to stay or leave, and new members are admitted if and only if they are unanimously approved by the founders.

There are a number of questions not answered in this paper. First, and since our primary interest is on voluntary exit, we have not fully characterized the class of strategy-proof social choice functions, although we think that they are also voting by committees. In addition, we conjecture that our founder's sovereignty on $K$ condition (defined as an unanimity condition over candidate decisions) could be weakened to a full-range condition over candidate decisions. Second, we do not know if there is a meaningful characterization of the maximal domain of preferences under which stable and strategy-proof social choice functions exist. Third, we have not analyzed here the (subgame perfect) equilibrium voting behavior of founders who take into account the effect of their votes not only on the chosen candidates, but also on the final composition of the society (see Berga et al. (2003) for a subgame perfect equilibrium analysis of exiting after voting in a general set up).

## Appendix

*Proof of Proposition 2.* First, we define the set of *vetoers* of $\mathscr{W}_i$ as the set $V_i = \{j \in N \backslash \{i\} \mid j \in S \text{ for all } S \in \mathscr{W}_i\}$. We will use the following result.

**Lemma 3.** *Assume that* $f : \mathscr{S} \longrightarrow 2^{N \cup K}$ *is voting by committees and that* $f$ *satisfies internal stability:*

*(a) Suppose that* $n \geq 3$ *and there exist* $i, j \in N$ *such that* $i \neq j$ *and* $\{i, j\} \notin \mathscr{W}_j$. *Then, for all* $l \in N \backslash \{i, j\}$, $i \in V_l$.
*(b) Suppose that* $n \geq 3$ *and there exist* $i, j \in N$ *such that* $i \neq j$ *and* $i \in V_j$. *Then for all* $l \in N \setminus \{i, j\}$, $l \in V_q$ *for all* $q \in N \setminus \{l, j\}$.
*(c) Suppose that* $n \geq 4$ *and there exist* $i, j \in N$ *such that* $i \neq j$ *and* $i \in V_j$. *Then for all* $l \in N$, $l \in V_q$ *for all* $q \in N \setminus \{l\}$.

*Proof.* (a) We prove it by contradiction. Assume that there exists $l \in N \backslash \{i, j\}$ such that $i \notin V_l$. Let $R$ be a preference profile satisfying:

- $\tau(R_i) = (N \setminus \{l\}) \cup K$. Given $S \subset N \cup K$ such that $i \in S$, if $j \notin S$ and $l \in S$, then $\emptyset P_i S$.
- $\tau(R_j) = N \cup K$.
- $\tau(R_r) = (N \setminus \{j\}) \cup K$ for all $r \in N \setminus \{i, j\}$.

Since $\{i, j\} \notin \mathscr{W}_j$ we conclude that $j \notin f(R)$. Moreover, $l \in f(R)$ because $i \notin V_l$. Agents of $(N \setminus \{j, l\}) \cup K$ belong to $f(R)$ because they are unanimously good. But this contradicts internal stability since $i \in f(R) = (N \setminus \{j\}) \cup K$ and $\emptyset P_i f(R)$.

(b) We prove it by contradiction. Assume that there exist $l \in N \backslash \{i, j\}$ and $q \in N \setminus \{l, j\}$ such that $l \notin V_q$ . Let $R$ be a preference profile satisfying:

- $\tau(R_i) = (N \setminus \{j\}) \cup K$.
- $\tau(R_l) = (N \setminus \{q\}) \cup K$. Given $S \subset N \cup K$ such that $l \in S$, if $j \notin S$ and $q \in S$, then $\emptyset P_l S$.
- $\tau(R_r) = (N \setminus \{j\}) \cup K$ for all $r \in N \setminus \{i, l\}$.

Since $i \in V_j$ we conclude that $j \notin f(R)$. Moreover, $q \in f(R)$ because $l \notin V_q$. Agents of $(N \setminus \{j, q\}) \cup K$ belong to $f(R)$ because they are unanimously good. But this contradicts internal stability since $l \in f(R) = (N \setminus \{j\}) \cup K$ and $\emptyset P_l f(R)$.

(c) Without loss of generality assume that $2 \in V_1$. By part (b) we conclude that for all $i \geq 3$, $i \in V_q$ for all $q \in N \setminus \{1, i\}$. Since $3 \in V_4$, by part (b), $2 \in V_i$ for all $i \in N \setminus \{2, 4\}$ and $1 \in V_i$ for all $i \in N \setminus \{1, 4\}$. Since $4 \in V_3$, by part (b), $2 \in V_i$ for all $i \in N \setminus \{2, 3\}$ and $1 \in V_i$ for all $i \in N \setminus \{1, 3\}$. Then, $2 \in V_i$ for all $i \in N \setminus \{2\}$ and $1 \in V_i$ for all $i \in N \setminus \{1\}$. Since $1 \in V_2$, by part (b), $3 \in V_i$ for all $i \in N \setminus \{2, 3\}$. Then, $3 \in V_i$ for all $i \in N \setminus \{3\}$. Similarly, since $1 \in V_2$, by part (b), for all $l \geq 4$, $l \in V_q$ for all $q \in N \setminus \{2, l\}$. Then, $l \in V_q$ for all $q \in N \setminus \{l\}$. ■

Assume first that $f : \mathscr{S} \to 2^{N \cup K}$ is voting by committees and satisfies internal stability. To show that all committees for the candidates are unanimous, assume that there exist $x \in K$ and $S \subsetneq N$ such that $S \in \mathscr{W}_x^m$. Take $i \in N \setminus S$ and $R \in \mathscr{S}$ such that $x \in G_K(R_j)$ for all $j \in S$, $i \in G_N(R_j)$ for all $j \in N$, and

$$\emptyset P_i T \text{ whenever } x \in T. \tag{3}$$

Then, $x \in f(R)$ and $i \in f(R)$. But this and condition (3 ) contradict internal stability of $f$.

We now prove the statement for founders distinguishing two cases: (1) $n \geq 4$ and (2) $n = 3$. No restriction has to be imposed on committees for $n = 2$, since for this case we can check that any committee structure defines voting by committees satisfying internal stability, by (C4) and because $i \in \tau(R_i)$ for all $i \in N$ and for all $R_i \in \mathscr{S}_i$.

*Case 1.* $n \geq 4$. Again, we consider two cases:

(a) There exist $i, j \in N$, $i \neq j$, such that $\{i, j\} \notin \mathscr{W}_j$.

By parts (a) and (c) of Lemma 3, we conclude that for all $l \in N$, $l \in V_q$ for all $q \in N \setminus \{l\}$. Now it is easy to conclude that all committees are unanimous.

(b) For all $i, j \in N$, $i \neq j$, $\{i, j\} \in \mathscr{W}_j$.

Let $\{N_1, N_2\}$ be the partition of $N$ where $N_1 = \left\{i \in N \mid \{\{i\}\} = \mathscr{W}_i^m\right\}$ and $N_2 = \left\{i \in N \mid \{\{i\}\} \neq \mathscr{W}_i^m\right\}$. Note that one of $N_1$ or $N_2$ could be empty. Now it is immediate to conclude that all committees for founders in $N_1$ are decisive and all committees for founders in $N_2$ are bipersonal.

*Case 2.* $n = 3$. We now distinguish three cases:

(a) There exist $i, j, l \in N$, $j \in N \setminus \{i\}$, and $l \in N \setminus \{i, j\}$, such that $\{i, j\} \notin \mathscr{W}_i$ and $\{i, l\} \notin \mathscr{W}_i$.

Then $\mathscr{W}_i^m = N$, which means that $j \in V_i$ and $l \in V_i$. Since $j \in V_i$ $(l \in V_i)$, by part (b) of Lemma 3, $l \in V_j$ $(j \in V_l)$. Applying again part (b) of Lemma 3 we conclude that $i \in V_l$ $\left(i \in V_j\right)$.

Hence, for all $q \in N$, $q \in V_r$ for all $r \in N \setminus \{q\}$. Now it is easy to conclude that all committees are unanimous.

(b) There exist $i, j, l \in N$, $j \in N \setminus \{i\}$, and $l \in N \setminus \{i, j\}$, such that $\{i, j\} \notin \mathscr{W}_i$ but $\{i, l\} \in \mathscr{W}_i$.

Then $\mathscr{W}_i^m = \{\{i, l\}\}$ and thus $l \in V_i$. Applying twice part (b) of Lemma 3 we conclude that $j \in V_l$ and $i \in V_j$. For $n = 3$ this implies that $\mathscr{W}_l^m = \{\{l, j\}\}$ and $\mathscr{W}_j^m = \{\{j, i\}\}$. That is, the committees $\mathscr{W}_i$, $\mathscr{W}_j$, and $\mathscr{W}_l$ are cyclical.

(c) For all $i, j \in N$, $i \neq j$, $\{i, j\} \in \mathscr{W}_i$.

Arguing as in Case 1(b) we obtain that some committees are decisive and some are bipersonal.

We now prove the converse. Assume $n \geq 3$ and let $f : \mathscr{S} \to 2^{N \cup K}$ be a voting by committees as defined in the statement of Proposition 2. Let $R \in \mathscr{S}$ and suppose that $i \in f_N(R)$. Note that since $\mathscr{W}_x^m = \{N\}$ for all $x \in K$, $f_K(R) \subset G_K(R_i)$. Consider first that $\mathscr{W}_j^m = \{N\}$ for all $j \in N$. Then, by separability of $R_i$ and (C3), $f(R) R_i \{i\} R_i \emptyset$. Hence, $f$ satisfies internal stability. Consider now that there exist $N_1, N_2 \subset N$ such that $N_1 \cap N_2 = \emptyset$, $N_1 \cup N_2 = N$, and $\mathscr{W}_j^m = \{\{j\}\}$ for all $j \in N_1$ while $\mathscr{W}_{j'}^m = \{\{j', l\}_{l \in N \setminus \{j'\}}\}$ for all $j' \in N_2$. Then, $G_N(R_i) \subset f_N(R)$ which implies that $N \setminus f_N(R) \subset B_N(R_i)$. Therefore, by separability of $R_i$ and by condition (C4), $f(R) R_i f_N(R) R_i N P_i \emptyset$. Hence, by (C2), $f$ satisfies internal stability.

Assume now that $n = 3$. The only case remaining to be considered is $\{\mathscr{W}_1, \mathscr{W}_2, \mathscr{W}_3\}$ is cyclic. Without loss of generality suppose that $\mathscr{W}_1^m = \{1, 2\}$, $\mathscr{W}_2^m = \{2, 3\}$, $\mathscr{W}_3^m = \{3, 1\}$, and $i = 1$. Then, by (C2), $f$ satisfies internal stability since the following four conditions hold:

- If $3 \in G(R_1)$ (and hence $3 \in f_N(R)$) and $f_N(R) = \{1, 3\}$, then $f(R) R_1 f_N(R) P_1 \{1\} R_1 \emptyset$ by separability of $R_1$ and condition (C3).
- If $3 \in G(R_1)$ (and hence $3 \in f_N(R)$) and $f_N(R) = N$, then $f(R) R_1 N P_1 \emptyset$ by separability of $R_1$ and condition (C4).
- If $3 \in B(R_1)$ (and hence $3 \notin f_N(R)$) and $f_N(R) = \{1\}$, then $f(R) R_1 \{1\} R_1 \emptyset$ by separability of $R_1$ and condition (C3).
- If $3 \in B(R_1)$ (and hence $3 \notin f_N(R)$) and $f_N(R) = \{1, 2\}$, then $f(R) R_1 f_N(R) P_1 N P_1 \emptyset$ by separability of $R_1$ and condition (C4). ■

*Proof of Proposition 3.* Since $i \in \tau(R_i)$ for all $i \in N$ and all $R_i \in \mathscr{S}_i$, we conclude that

$$[\mathscr{W}_i^m = \{\{i\}\}, \text{ for all } i \in N] \Longleftrightarrow [N \subset f(R), \text{ for all } R \in \mathscr{S}].$$

Suppose that $f$ is voting by committees and for all $i \in N$, $\mathscr{W}_i^m = \{\{i\}\}$. Then $f$ satisfies external stability because $N \subset f(R)$ for all $R \in \mathscr{S}$.

We now prove the converse by contradiction. Let $R \in \mathscr{S}$ and $i \in N$ be such that $i \notin f(R)$. Consider $R_i' \in \mathscr{S}_i$ such that $\tau(R_i) = \tau(R_i')$ and $SP_i'\emptyset$ when $i \in S$. Since $f$ is voting by committees we conclude that $f(R) = f(R_i', R_{-i})$. But this contradicts external stability because $i \notin f(R_i', R_{-i})$ and $(f(R_i', R_{-i}) \cup \{i\})P_i'\emptyset$. ■

## References

Barberà S (1996) Notes on strategy-proof social choice functions. In: Arrow K, Sen A, Suzumura K (eds) Social choice re-examined. Macmillan, London. French version (1996) Sur les fonctions de choix non manipulables. Rev d'Écon Polit 106 (1)

Barberà S (2001) An introduction to strategy-proof social choice functions. Soc Choice Welfare 18: 619–653

Barberà S, Maschler M, Shalev J (2001) Voting for voters: a model of electoral evolution. Games Econ Behav 37: 40–78

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

Berga D, Bergantiños G, Massó J, Neme A (2003) On exiting after voting. Working paper of the Departament d'Economia, Universitat de Girona

Dutta B, Jackson M, Le Breton M (2001) Strategic candidacy and voting procedures. Econometrica 69: 1013–1037

Ehlers L, Weymark J (2003) Candidate stability and nonbinary social choice. Econ Theory 22: 233–243

Eraslan H, McLennan A (2001) Strategic candidacy for multivalued voting procedures. mimeo, University of Minnesota

Granot D, Maschler M, Shalev J (2000) Unanimity in voting for voters. Inter Game Theory Rev 2: 117–127

Rodríguez-Álvarez C (2001) Candidate stability and voting correspondences. mimeo, Universitat Autònoma de Barcelona

Sprumont Y (1995) Strategy-proof collective choice in economic and political environments. Can J Econ 28: 68–107