Soc Choice Welfare (2001) 18: 555–570

Social Choice and Welfare

© Springer-Verlag 2001

Stable sets and standards of behaviour

Robert Delver, Herman Monsuur

Department of International Security Studies, Royal Netherlands Naval Academy PO Box 10.000, 1780 CA, Den Helder, The Netherlands (e-mail: r.delver@kim.nl)

Received: 4 May 1998/13 March 2000

Abstract. In this paper we present a constructive, behavioural and axiomatic approach to the notion of a stable set as a model of the standard of behaviour of a social organisation. The socially stable set we introduce is a generalisation of the von Neumann-Morgenstern stable set. In contrast with the original version, our stability concept is always solvable. The standard of behaviour, reflecting the established conceptual order of a society or organisation, emerges from a dominance relation on alternative conceptions that are relevant with regard to a certain issue. This common social choice phenomenon, that permeates our societies and organisations, we have tried to clarify. Two axiomatic characterisations as well as a construction algorithm for socially stable sets are presented. These characterisations are based on behavioural postulates regarding the individual or collective strategic behaviour of effective sets. Relations between socially stable sets and other notions of stability are discussed.

1 Introduction

In Sect. 4.5.3 of the Theory of Games and Economic Behaviour, von Neumann and Morgenstern define the *vN-M stable set*. This abstract notion formalises their idea of a *standard of behaviour* of a social economy. While the authors describe this economy as a game of n participants with payoffs in the form of imputations, incidentally, they also refer to a more general setting for their stable sets. In Sects. 4.4.3 and 4.6.1, this more general setting appears to

We acknowledge a number of useful remarks by the referees and an associate editor, which we gladly used to improve our exposition.

be a theory of social phenomena based on effective preferences between various states of society or a organisation.

We follow this direction and primarily view the standard of behaviour as an established conceptual order of an organisation or society. Such a standard of behaviour often functions as the frame of reference for the collective decision making. In a public or private organisation, the mission statement and policy alternatives to realise it are subject to scrutiny from the perspective of the standard of behaviour. In a society, the standard of behaviour reflects the degree of civilisation. It is used to select policies and constitutes the normative base for the political correctness of opinions and public statements and the like.

The structure from which the standard of behaviour emerges are *alternatives* and *dominations* between certain pairs of these alternatives. The alternatives we think of are strategic options, economic doctrines, various possible organisational designs or other lasting intellectual conceptions with regard to a certain issue. A domination between a pair of alternatives is assumed to be generated by at least one *effective coalition*. Such a coalition consists of members of the organisation who are together capable of enforcing their preference of the one alternative over the other, if only these two were considered.

An effective coalition will be inclined to apply its binary dominance, but, in the larger context of the dominance relation, it may have strategic reasons for not exercising its power. Hence, the existence of a domination between two alternatives does not imply that it is enforced. In our approach, this phenomenon of non-enforcement occurs in two instances. A dominance will be not enforced by any of its effective coalitions if the preferred alternative is already suppressed by at least one other effective coalition which does enforce its preference. Such a suppressed alternative we call subdued. Situationally, it can not serve as a viable alternative because it could at once be overturned or it might even be generally considered as discredited. The other reason for non-enforcement is equalisation of dominations along circular patterns within the standard of behaviour. Here the effective coalitions involved may be motivated by mutual interest. This conditional behaviour of effective sets we call the *non-enforcement principle*. We consider it to be part of the organisational or societal culture.

Let the set of alternatives be denoted by X and let A be the set of agents or members of the organisation or society. R is the dominance relation on X that is generated by the effective coalitions: If there is at least one effective coalition generating a domination of x over y, then $(x, y) \in R$. The collection of effective coalitions corresponding to an elementary dominance $(x, y) \in R$ is $E_{x,y} \subseteq 2^A \backslash \emptyset$ while $\varepsilon_{x,y} \in E_{x,y}$ denotes a specific effective coalition for $(x, y) \in R$. A pair of alternatives $\{x, y\}$ between which no effective coalitions exist is *mutually independent*. In case of opposing effective coalitions, $(x, y) \in R$ and $(y, x) \in R$, the two alternatives are *discordant*. If $(x, y) \in R$ and $(y, x) \notin R$, then (x, y) is *asymmetric*. If R contains no discordant pairs of alternatives, then R is *asymmetric*. We will also refer to the elements of X as nodes and to $(x, y) \in R$ as an arc from x to y.

Indeed, the structure of R is that of a directed graph on X. It contains neither loops nor multiple arcs[2], although cycles, including discordant pairs configurations, are allowed. We assume R to be irreflexive and X to be finite.

A distinguishing feature of our theory is that no assumptions involving value functions of the agents over X or individual rational behaviour are made. Another point to notice is that we use the effective sets only to introduce our behavioural postulates, and for interpreting our stability concept and the theory. The formal development of our theory is based only on a given dominance relation R, which may represent different social situations.

Also the vN-M stable set has this special nature. For $S \subseteq X$ to be a vN-M stable set, only two properties are required. Firstly, *inner stability*: No y contained in S is dominated by an x contained in S. Secondly, *external stability*: Every y not contained in S is dominated by some x contained in S, (von Neumann and Morgenstern, Sect. 4.5.3).

The vN-M stable set has been criticised on various grounds. Often mentioned are the possibility of non-existence and that solutions may not be unique. For example, in an odd cyclic pattern of dominations, the vN-M stable set does not exist, while in case of an even cycle there are two solutions. For these and other reasons attempts were made to generalise or alter the definition of a stable set.

Some authors have, in stead of R, used the transitive closure R^{cl} in their definition of either inner- or external stability or both[3]. Vickrey (1959) introduces the *policing property* by looking at R^{cl} for the inner stability. Kreinovich and Kosheleva (1990) introduce *hierarchically stable sets* by using R^{cl} for the external stability. Van Deemen (1991) introduces his *generalised stable sets*, requiring both inner and external stability with respect to R^{cl}. The notions of stability of Vickrey and of van Deemen both exclude many classical vN-M stable sets.

Another modification, the *largest consistent set*, was introduced by Chwe (1994). Besides existence and uniqueness, this concept also solves an other perceived problem of the theory of stability, that of myopia. Largest consistent sets apply to dominance relations where the alternatives are economic imputations. A difference with our and earlier notions of stability is that with the same dominance relation the solution may depend on the specific imputations. In the largest consistent set the central role is played by the notion of *deterrence* of deviations from the solution. The idea is that any deviation from a stable imputation is deterred because, by an uncontrollable sequence of deviations, some stable imputation might be reached that, by the first deviating coalition, is less preferred than the original stable imputation. Deterrence is based on risk aversion. It may be understood as an extremely farsighted conservative mental disposition of the effective sets.

---

[2] Arcs of the form (x, x) or more than one copy of an arc (x, y).
[3] $(x, y) \in R^{cl}$ if there is a finite sequence of nodes $x = x_0, x_1, \ldots, x_m = y, m \geq 1$, such that $(x_k, x_{k+1}) \in R, k = 0, \ldots, m-1$.

Our *non-enforcement principle* reflects a different attitude. By way of introduction to our notion of socially stable sets and to further illustrate the role of effective coalitions, we first look at two simple examples. A more elaborate application of these ideas is Example 3 in Sect. 2. In Example 1 we particularly demonstrate how in that case non-enforcement makes sense of the concept of a stable set of von Neumann and Morgenstern.

Example 1. Let $X = \{x, y, z\}$ and $R = \{(x, y), (y, z)\}$. Then the vN-M stable set $\{x, z\}$ exists thanks to the effective sets in $E_{y,z}$ not exercising their power to discredit alternative z.

Why would these coalitions behave so obligingly? We suggest the following explanation: The element x is undominated. Therefore an effective coalition $\varepsilon_{x,y} \in E_{x,y}$ will feel no restraint in propagandising x to the detriment of y. This unrestrained manifestation of the effective coalitions $\varepsilon_{x,y}$ subdues y. Hence the effective coalitions in $E_{y,z}$ will not enforce, thus leaving room for alternative z.

In Theorem 5 we prove that the essential elements in this reasoning characterise the vN-M solution for acyclic dominance relations. The above example makes plausible that the direct external domination by R, as was originally required, should not be weakened. This conjecture is further supported by our axiomatic Lemma 4 in Sect. 3. In our next example we give a first indication towards the plausibility that cyclic patterns in a stable set should be allowed.

Example 2. Let $X = \{a, b, c, d\}$ and $R = \{(a, b), (b, c), (c, a), (a, d), (b, d), (c, d)\}$. This dominance relation admits no vN-M stable set. However, the effective sets along the cycle neutralise each other, because, by showing restraint, these effective sets can realise the inner stability of $\{a, b, c\}$. Since the external stability is satisfied $\{a, b, c\}$ then in effect becomes a vN-M stable set. This we consider to be sufficient motivation for the effective sets along the cycle not to enforce their dominations.

From Example 2 and our axiomatic treatment in Proposition 6 and Theorem 7 of Sect. 3, it appears that the notion of inner stability can be relaxed to *generalised inner stability*: $a(R_{|S})^{\text{cl}} = \emptyset$, where $a(R_{|S})^{\text{cl}}$ is the asymmetric part[4] of the transitive closure of $R_{|S}$[5]. This means that for $\{x, y\} \subseteq S$, we allow x to dominate y if y dominates x directly or indirectly within $R_{|S}$. Our generalised notion of stability, introduced in Monsuur (1994), depends on these two observations.

Definition 1. *A* socially stable set *is a subset S of X such that:* (i) $a(R_{|S})^{\text{cl}} = \emptyset$ *and* (ii) *If $y \notin S$, then there is an $x \in S$, such that $(x, y) \in R$.*

A vN-M stable set also is a solution of Definition 1, but, where this first notion may not be solvable, a socially stable set always exists. A detailed

[4] For an arbitrary relation R: $(x, y) \in aR$ if $(x, y) \in R$ and $(y, x) \notin R$.
[5] $R_{|S} = \{(x, y) \in R : \{x, y\} \subseteq S\}$.

description of the way in which the non-enforcement principle works as well as a construction algorithm for socially stable sets, reflecting this principle, is found in Sect. 2. In Sects. 3 and 4 we introduce and explain behavioural postulates and investigate various of their consequences. Our Theorems 8 and 10 are characterisations of socially stable sets by these postulates. While the first of these two characterisations depends more on the behaviour of individual effective sets, the second reflects the collective power of the effective sets that support an established standard of behaviour. In Sect. 5 we make various concluding remarks and suggest lines for further research.

Our research connects with the axiomatic program for the social sciences that has originated with the work of J. F. Nash, (1950). Other contributions to this program may be found in Barnett et al. (1995), see for example Thomson (1995).

2 Non-enforcement

Any concept of stable sets derives its stability from the restraint that certain effective coalitions can summon in exercising their existing binary dominations. Whether or not this restraint applies to a specific dominance depends on the position of it in the dominance relation R, the structure of R and the actual standard of behaviour that has emerged.

A dominance $(x, y) \in R$ is *non-enforced* if no $\varepsilon_{x,y} \in E_{x,y}$ exercises its binary domination of x over y. An alternative y in X is in *primary position* if all $(x, y) \in R$ are non-enforced or if no incoming arcs exist. It is *subdued* if some $\varepsilon_{x,y} \in E_{x,y}$ does exercise its binary dominance. In our view, there are two instances where $(x, y) \in R$ will be non-enforced. The first case is that (x, y) lies on a cycle in the stable set. Then we say that (x, y) is *equalised*. The other case is where x is subdued. Unless (x, y) is equalised, $(x, y) \in R$ not only means that x subdues y if x is in primary position, but also that then, for any $(y, z) \in R$, no $\varepsilon_{y,z} \in E_{y,z}$ will use y to subdue z.

Definition 2. *The* non-enforcement principle *is the general attitude of effective coalitions $\varepsilon_{x,y}$ to refrain from enforcing (x, y) if x is subdued or if x and y lie on a cycle in the stable set.*

In general, given $S \subseteq X$, we may distinguish four types of binary dominations. (1) *Internal dominations* $(x, y) \in R$, with $\{x, y\} \subseteq S$. (2) *Outgoing dominations* $(x, y) \in R$, with $x \in S$ and $y \in X \backslash S$. (3) *Incoming dominations* $(x, y) \in R$, with $x \in X \backslash S$ and $y \in S$. (4) *External dominations* $(x, y) \in R$, with $\{x, y\} \subseteq X \backslash S$.

In case of the vN-M stable set internal dominations do not exist, all outgoing dominations are enforced, while all incoming and external dominations are non-enforced.

With our socially stable sets all outgoing dominations are enforced, incoming and external dominations are non-enforced. Cyclic internal dominations we assume to be equalised. As in Example 2, an explanation that often applies is that effective sets involved in cyclic patterns realise that by collectively not enforcing their dominations their favourite alternatives will end up in a vN-M stable set. An axiomatic rationalisation of equalising cycles in the stable set is given by Proposition 6 and Theorem 7 of Sect. 3.

Definition 3. *Let R be a dominance relation on X. Then the core of R is defined as* $C(R) = \{x \in X : \text{there exists no } y \in X \text{ such that } (y,x) \in R\}$. *Further, the set of* top elements *of R,* $T(R)$, *is defined as* $C(a(R^{\mathrm{cl}}))$.

According to this definition, $x \in T(R)$ if $(y,x) \in R^{\mathrm{cl}}$ implies that $(x,y) \in R^{\mathrm{cl}}$. To further explicate the set of top elements, we firstly observe that the R-connected component[6] of $x \in T(R)$ is entirely contained in $T(R)$. Secondly, we note that if $x \in T(R)$ and $(y,x) \in R$, then also $y \in T(R)$, meaning that $T(R)$ has no incoming arcs. So $T(R)$ consists of R-connected components without incoming arcs.

Given a subset S of X, we may divide X into S, the S-dominated set $D(S)$ and the S-undominated set $U(S) : D(S) = \{x \in X \backslash S$: there is a node $s \in S$ such that $(s,x) \in R\}$, $U(S) = X \backslash (S \cup D(S))$. Note that the pairwise intersections of these sets are empty. Further, we define $N(S)$ by $\{(y,z) \in R : y \in D(S)\}$. Assuming that outgoing dominations are enforced, $N(S)$ consists of non-enforced external and incoming dominations. We may now construct all socially stable sets of X with a given dominance relation R. The algorithm precisely reflects the non-enforcement idea of Definition 2 and is illustrated in Example 3.

Construction algorithm

Initial step.	*Determine a socially stable set S for* $T(R)$.
Step 1.	*If S is externally stable for X, then go to the final step, else go to Step 2.*
Step 2.	*Extend S with* S' *in such a way that* $S \cup S'$ *is a socially stable set for* $T(R \backslash N(S))$. *Replace* S by $S \cup S'$ *and return to Step 1.*
Final Step.	*Observe that S is a socially stable set for X.*

Theorem 1. *Each run of the construction algorithm gives a socially stable set.*

Proof. Let R be an arbitrary dominance relation on X. We show that each step of the construction algorithm is practicable. Moreover, we show that before (and after) each cycle of Steps 1 and 2 we have (α) : S is generalised inner stable and (β) : $U(S)$ has no arcs to S (there are no arcs $(u,s) \in R$ with $u \in U(S), s \in S$).

At the initial step, we may take $S = T(R)$ as a socially stable set for $T(R)$. Any choice of socially stable set for $T(R)$ at this step gives a generalised inner stable set. Furthermore, $U(S) \subset X \backslash T(R)$ because S is externally stable for $T(R)$. So $U(S)$ has no arcs to $S \subseteq T(R)$. We now have (α) and (β).

---

6 A node y is in the R-connected component of x if $y = x$ or if (y,x) and $(x,y) \in R^{\mathrm{cl}}$

Suppose that at some Step 1, we observe that S is not externally stable for X. Since there are no arcs from S to $U(S)$ and vice versa, we have the following equality:

$$R \backslash N(S) = R_{|S} \cup R_{|U(S)} \cup \{(x,t) \in R : t \in D(S), x \in S \text{ or } U(S)\}.$$

This implies that $T(R \backslash N(S)) = T(R_{|S}) \cup T(R_{|U(S)})$, and since S is generalised inner stable, we obtain $T(R \backslash N(S)) = S \cup T(R_{|U(S)})$. Altogether, this shows that $S \cup S'$, where S' is any socially stable set for $T(R_{|U(S)})$, is socially stable for $T(R \backslash N(S))$. Notice that we may always take $S' = T(R_{|U(S)})$. Furthermore, $S \cup S'$ is generalised inner stable, showing (α). We finally prove that $U(S \cup S')$ has no arcs to $S \cup S'$, so $S \cup S'$ also satisfies (β). To this end, suppose that there would exist $(u,s) \in R$ with $u \in U(S \cup S')$ and $s \in S \cup S'$. Since $u \notin D(S)$, we have $(u,s) \in R \backslash N(S)$. Since $s \in T(R \backslash N(S))$, we obtain that u also is in $T(R \backslash N(S))$. But then, by the external stability of $S \cup S'$ for $T(R \backslash N(S))$, there is a node $z \in S \cup S'$ with $(z,u) \in R$, contradicting $u \in U(S \cup S')$.

This shows that all the steps are practicable. Moreover, S is extended each time we observe that it is not externally stable. Since X is finite, we eventually detect external stability at Step 1. Then, because of (α), S is socially stable. □

To prove Theorem 3, we need the following lemma.

Lemma 2. *Let R be a dominance relation and let M be a socially stable set for a set of alternatives X. Then* (i) *$S = M \cap T(R)$ is socially stable for $T(R)$ and* (ii) *$M \backslash S \subseteq U(S)$.*

Proof. (i) To prove the external stability of $M \cap T(R)$, let $x \in T(R) \backslash M$. As M is externally stable, there exists a node $s \in M$ such that $(s,x) \in R$. Because $x \in T(R)$, we have $s \in T(R)$, so $s \in M \cap T(R)$. Next consider the generalised inner stability. Suppose $(x,y) \in R$ for $x,y \in M \cap T(R)$. Then, M being socially stable, x and y are on a cycle in $R_{|M}$. This cycle is part of the R-connected component of x (and y) and since $x \in T(R)$, it is a subset of $T(R)$. So this cycle is also in $M \cap T(R)$. (ii) Suppose there would exist a node $m \in M \cap D(S)$. So there is a node $s \in S$ with $(s,m) \in R$. By the generalised inner stability of M, we have $(m,s) \in (R_{|M})^{\mathrm{cl}}$. Because $s \in T(R)$, we have $m \in T(R)$, so $m \in T(R) \cap M = S$, contradicting $m \in D(S)$. □

Theorem 3. *The construction algorithm produces all socially stable sets.*

Proof. Let M be a socially stable set. We show that we may reconstruct M by a series of Steps 1 and 2. As follows from Lemma 2, we may take $M \cap T(R)$ at the initial step of the construction algorithm. Next, suppose that at some Step 1, the set $S \subseteq M$ constructed so far, is not externally stable. As follows from the proof of Theorem 1, S has to be extended with a socially stable set S' for $T(R_{|U(S)})$. We therefore show that $S' = (M \backslash S) \cap T(R_{|U(S)}) \subseteq M$ is socially stable for $T(R_{|U(S)})$, so it can be used at Step 2 of the algorithm.

Using Lemma 2(i), we only have to prove that $M \backslash S$ is socially stable for

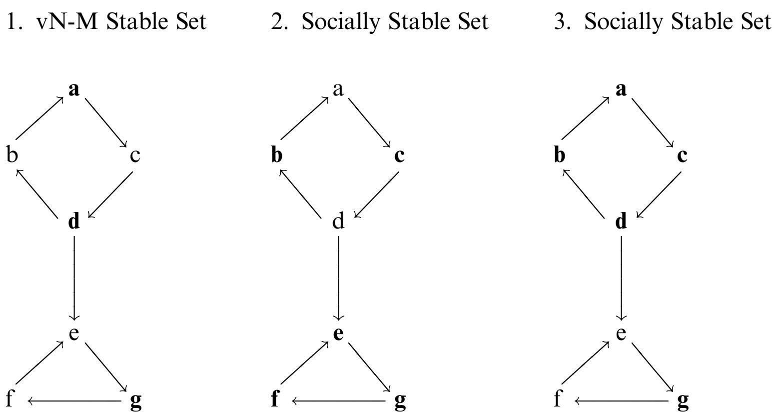
$U(S)$. To this end, we show that before each cycle of Steps 1 and 2 we have $(\gamma) : M \backslash S \subseteq U(S)$. Since M is socially stable for X, (γ) then implies that $M \backslash S$ is socially stable for $U(S)$.

As follows from Lemma 2, the choice above at the initial step gives (γ), thus making possible our choice of S'. To prove that (γ) still holds for $S \cup S'$ constructed at some Step 2, observe that $M \backslash S$ is socially stable for $U(S)$. So, by making use of Lemma 2(ii), we deduce that the set $M \backslash (S \cup S') = (M \backslash S) \backslash S'$ is included in $U(S')$. Since we already have $M \backslash (S \cup S') \subseteq M \backslash S \subseteq U(S)$, we obtain (γ).

Finally, note that (γ), together with $S \subseteq M$, shows that if at Step 1 we observe that S is externally stable, we have $S = M$. □

We conclude this section with an example.

Example 3. Let $X = \{a, b, c, d, e, f, g\}$ and $R = \{(a, c), (b, a), (c, d), (d, b), (d, e), (e, g), (g, f), (f, e)\}$. We observe that there are three socially stable sets: $S_1 = \{a, d, g\}$, $S_2 = \{b, c, e, f, g\}$ and $S_3 = \{a, b, c, d, g\}$. We first discuss S_1, which also is vN-M-stable. Due to the positive feedback in the cycle $\{a, c, d, b\} = T(R)$, once established, the solution $S = \{a, d\}$ for this cycle, created at the initial step of the construction algorithm, may easily stabilise: the dominations (a, c), (d, b) and (d, e) are enforced, therefore (c, d), (b, a) and (e, g) are non-enforced. Removing these non-enforced arcs at Step 2, g comes in primary position: $g \in T(R \backslash N(S)) = \{a, d, g\}$. Next (g, f) is also enforced and (f, e) therefore is non-enforced. Altogether (a, c), (d, b), (d, e) and (g, f) are enforced, while (c, d), (b, a), (e, g) and (f, e) are non-enforced. The set S_2 is constructed similarly. In this case, the solution $\{b, c\}$ for $T(R)$ gives rise to a cycle in the lower part of the solution. The socially stable set S_3 was constructed by choosing $S = T(R)$ at the initial step. In Step 2, we remove the non-enforced domination (e, g) and obtain the node g as a new core element.



The socially stable sets of X. Elements of S are in bold, $x \rightarrow y$ denotes $(x, y) \in R$.

3 Behavioural characterisation of the socially stable set

In this section we introduce some behavioural postulates and show that these elementary properties characterise the vN-M stable set for acyclic relations. If we want universal domain, we show that it is unavoidable to allow for cyclic patterns of domination within the stable set. Thus we argue that socially stable sets are a natural extension of the vN-M stable set. Further, we present our first characterisation of socially stable sets.

Let Φ be a solution concept assigning to relations R on X in its domain a subset of $2^X \backslash \emptyset$, where 2^X denotes the set of subsets of X. The elements of $\Phi(R)$ are to be interpreted as stable sets. Examples are $\Phi_{\text{vN-M}}$, determining the classical vN-M stable sets, Φ_C, assigning to R its core $C(R)$ and Φ_{soc}, giving the socially stable sets: $\Phi_{\text{soc}}(R) = \{S : S \text{ is a socially stable set for } R\}$.

Neither $\Phi_{\text{vN-M}}(R)$ nor $\Phi_C(R)$ are defined for all relations on X. Von Neumann and Morgenstern already remark that, as regards existence, no concessions can be made, (1944) 4.6.3. The concept of socially stable sets satisfies this *universal domain* requirement as is shown in Theorem 1 of Sect. 2.

As we will show next, socially stable sets may be axiomatically characterised by four behavioural postulates. Each of these we will first introduce and try to elucidate. The elements of $C(R)$ are undominated. Therefore, for each $x \in C(R)$ and $(x, y) \in R$, the effective coalitions $\varepsilon_{x,y} \in E_{x,y}$ will feel no restraint in propagandising x to the detriment of y. We therefore require both ***core primacy***:

$$\text{If } S \in \Phi(R), \text{ then } C(R) \subseteq S \tag{1}$$

and ***core subduction***:

$$\text{If } S \in \Phi(R), x \in C(R) \text{ and } (x, y) \in R, \text{ then } y \notin S \tag{2}$$

We consider it undesirable if a stable set S would change when the effective coalitions in $E_{x,y}$, finding that the alternative x they propagandise does not lie in S, give up or dissolve. This property we call ***independence of non-enforced dominations***:

$$\text{If } S \in \Phi(R) \text{ and } (x, y) \in R, x \notin S \text{ then } S \in \Phi(R'),$$
$$\text{where } R' = R \backslash \{(x, y)\} \tag{3}$$

For completeness we remark that we assume the domain of Φ to be closed under to the transformation of R into R'. Also in what follows, when introducing independence conditions, we will implicitly assume that the domain is closed under its operations.

In fact, for the purpose of Lemma 4, we use a weaker version: *independence of non-enforced external dominations*

$$\text{If } S \in \Phi(R) \text{ and } (x, y) \in R, x \notin S, y \notin S \text{ then } S \in \Phi(R'),$$
$$\text{where } R' = R \backslash \{(x, y)\} \tag{4}$$

The following lemma is our formal argument for not weakening the direct external domination of the vN-M stable set. Modifications of the vN-M stable set, such as those of van Deemen and Kreinovich and Kosheleva, that use the transitive closure in the definition of external domination, do not simultaneously satisfy fairly minimal properties as core primacy and independence of non-enforced dominations.

Lemma 4. *Let Φ satisfy* core primacy *and* independence of non-enforced external dominations. *Then for all relations R in its domain, all $S \in \Phi(R)$ and all $x \in X \backslash S$, there is an $s \in S : (s, x) \in R$.*

Proof. Assume there exists $S \in \Phi(R)$ and $x \in X \backslash S$, such that for all $s \in S$: $(s, x) \notin R$. Then we get a contradiction. Consider $R' = R \backslash \{(y, x) : (y, x) \in R\}$, where we remove all non-enforced external dominations involving x. By repeated application of the independence of non-enforced external dominations, we have $S \in \Phi(R')$. But $x \in C(R')$, so by core primacy we have $x \in S'$, for all $S' \in \Phi(R')$. So $x \in S$, contradicting $x \in X \backslash S$. □

If R is acyclic, then there is a unique vN-M stable set $\Phi_{\text{vN-M}}(R)$, see Sect. 5. The following theorem characterises $\Phi_{\text{vN-M}}$ for acyclic relations.

Theorem 5. *Let the domain of a solution concept Φ be the set of acyclical relations. Then Φ satisfies* core primacy, core subduction *and* independence of non-enforced dominations, *if and only if $\Phi(R) = \Phi_{\text{vN-M}}(R)$ for all acyclical relations R.*

Proof. The 'if-part' being a straightforward verification of the three postulates, we only prove the 'only if'-part. By Lemma 4, $S \in \Phi(R)$ directly dominates $X \backslash S$: every y not contained in S is dominated by some x in S. If there exists $\{a, b\} \subseteq S$ with $(a, b) \in R$ then we may derive a contradiction. Since $(a, b) \in R_{|S}$ and $R_{|S}$ is acyclic then, by going in opposite direction along arcs of $R_{|S}$, we deduce that there exist nodes $\{x, y\} \subseteq S$ with $(x, y) \in R_{|S}$ and $x \in C(R_{|S}) = \{v \in S$: there is no $u \in S$ such that $(u, v) \in R\}$. Let R' be obtained from R by deleting all non-enforced arcs, amongst them all $(z, x) \in R, z \in X \backslash S$. Then $x \in C(R')$ and by independence of non-enforced dominations we have $S \in \Phi(R')$. But now, by core subduction, we obtain $y \notin S$ contradicting $\{x, y\} \subseteq S$. Summarising, we have shown that $\{S\} = \Phi_{\text{vN-M}}(R)$. □

As the following proposition regarding the cyclic triple shows, if we want universal domain, it is unavoidable to allow for the possibility of cyclic dominations in the stable set.

Proposition 6. *Let $\# X = 3$. Let the domain of Φ contain all* asymmetric *relations on X. If Φ satisfies* core primacy, core subduction *and* independence of non-enforced dominations, *then $\Phi(R) = \Phi_{\text{soc}}(R)$ for all asymmetric dominance relations R on X.*

Proof. Let $X = \{a, b, c\}$. If R is an acyclic asymmetric relation, it is straightforward to verify that $\Phi = \Phi_{\text{soc}}$. Now consider a cycle. Without loss of generality, we may assume that this cycle is $R = \{(a, b), (b, c), (c, a)\}$ in which case $\Phi_{\text{soc}}(R) = X$. Since $S \neq \emptyset$ for each $S \in \Phi(R)$, we only need to show that $\#S = 1$ or 2 are impossible. The case $\#S = 1$ is impossible because of the direct external domination, as follows from Lemma 4. Now consider for example $S = \{a, b\}$. Then $(c, a) \in R$ is non-enforced and by independence of non-enforced dominations $S \in \Phi(R')$, where $R' = R \backslash \{(c, a)\}$. But by core subduction and $a \in C(R')$ we deduce that $b \notin S$, a contradiction. □

In socially stable sets we allowed for the possibility of cyclic patterns of domination in S. To present a more formal support of this notion of equalisation then was given in Sect. 2, in Theorem 7 we show that the socially stable set is a natural generalisation of the vN-M stable set. To prepare this result we introduce the *S-equalised dominance relation*, which we denote by $R_{\otimes S}$.

$$R_{\otimes S} = \{(x, y) \in R\colon (x, y) \text{ does not lie on a cycle of } R_{|S}\} \tag{5}$$

Theorem 7. *Let R be in the domain of a solution concept Φ. (i) Suppose that $S \in \Phi(R)$ implies $S \in \Phi_{\text{vN-M}}(R_{\otimes S})$. Then $S \in \Phi_{\text{soc}}(R)$. (ii) Suppose that $S \in \Phi(R)$ if and only if $S \in \Phi_{\text{vN-M}}(R_{\otimes S})$. Then $\Phi(R) = \Phi_{\text{soc}}(R)$.*

Proof. (i) To show the direct external domination by S, let $x \in X \backslash S$. Since S is a vN-M stable set for $R_{\otimes S}$, there is an element $s \in S$ with $(s, x) \in R_{\otimes S} \subseteq R$. To establish the generalised inner stability, let $(x, y) \in R$ and $\{x, y\} \subseteq S$. The solution S being a vN-M stable set of $R_{\otimes S}$, it follows that $(x, y) \notin R_{\otimes S}$, showing that $(y, x) \in (R_{|S})^{\text{cl}}$. (ii) Because of part (i), we only show that $\Phi_{\text{soc}}(R) \subseteq \Phi(R)$. If $S \in \Phi_{\text{soc}}(R)$ then $S \in \Phi_{\text{vN-M}}(R_{\otimes S})$, so $S \in \Phi(R)$. □

In line with Theorem 7, we introduce our fourth behavioural postulate, the ***independence of stable cycles***:

$$\text{If } S \in \Phi(R), \text{ then } S \in \Phi(R_{\otimes S}) \tag{6}$$

Our characterisation of socially stable sets in terms of the above properties is as follows:

Theorem 8. *(i) Φ_{soc} satisfies core primacy, core subduction, independence of non-enforced dominations and independence of stable cycles. (ii) If a solution concept Φ satisfies these same four postulates, then $\Phi(R) \subseteq \Phi_{\text{soc}}(R)$, for all R in the domain of Φ.*

Proof. Part (i) being straightforward, we only prove (ii). In Lemma 4, we showed that for all $S \in \Phi(R)$ and all $x \in X \backslash S$, there is an $s \in S$: $(s, x) \in R$. This property is direct external domination. Next suppose there would exist nodes $\{x, y\} \subseteq S$ with $(x, y) \in R_{\otimes S}$, contradicting the generalised inner stability. Then let R' be obtained from R by deleting all arcs (u, v) with $u \notin S$. Using independence of non-enforced dominations, we obtain $S \in \Phi(R')$. Next, by independence of stable cycles, $S \in \Phi(R'_{\otimes S})$, where $R'_{\otimes S}$ is an acyclic relation. By repeating the 'only if'-part of the proof of Theorem 5, we deduce that $S = \Phi_{\text{vN-M}}(R'_{\otimes S})$. But we also have $(x, y) \in R'_{\otimes S}$ with $\{x, y\} \subseteq S$, contradicting the inner stability of this vN-M stable set for $R'_{\otimes S}$. □

Below we show that the postulates (1), (2), (3) and (6) are logically independent, where we only consider Φ defined on the set of all asymmetric relations.

In each example three postulates are satisfied while the fourth is violated. Moreover, in each case $\Phi \not\subset \Phi_{\text{soc}}$, showing non-redundancy of each postulate with respect to the other three. Let $X = \{a, b, c\}$.

(1). Let $\Phi(\{(a,b),(b,c)\}) = \Phi(\{(a,b)\}) = \{\{a\}\}$. For the other relations R on X (not being isomorphic to a previous relation R) we take $\Phi(R) = \Phi_{\text{soc}}(R)$. Then Φ does not satisfy core primacy since $\Phi(\{(a,b)\}) \neq \{\{a,c\}\}$. (2). Let $\Phi(R) = \{X\}$ for all R. Then Φ does not satisfy core subduction. (3). Let $\Phi(\{(a,b),(b,c)\}) = \{\{a\}\}$ while for other relations R (not being isomorphic to a previous relation R), let $\Phi(R) = \Phi_{\text{soc}}(R)$. Then Φ is not independent of non-enforced dominations: $\Phi(\{(a,b),(b,c)\}) = \{\{a\}\}$ while $\Phi(\{(a,b)\}) = \{\{a,c\}\}$. (4). Let $X = \{a,b,c,d\}$; $\Phi(\{(a,b),(b,c),(c,a),(a,d),(b,d),(c,d)\}) = \{\{a,b,c,d\}\}$. For other relations R we let $\Phi(R) = \Phi_{\text{soc}}(R)$. Then Φ is not independent of stable cycles since $\{a,b,c,d\} \notin \Phi(\{(a,d),(b,d),(c,d)\}) = \{\{a,b,c\}\}$.

4 Collective power and non-frontal opposition

In this section we examine the collective power of the effective coalitions supporting a standard of behaviour. This approach leads to an alternative axiomatic treatment of socially stable sets. We also present additional results on the extension of the classical vN-M solution concept. The main result of this section is the characterisation by inclusion of Φ_{soc} in Theorem 10. In Theorem 12, we also obtain an equality characterisation of Φ_{soc}.

Let the *effective inner grand coalition* I_S of $S \in \Phi_{\text{soc}}(R)$ be defined by $I_S = \bigcup_{x \in S, y \in S}(\varepsilon_{x,y})$. Analogously, the *effective external grand coalition* is given by $E_S = \bigcup_{x \in S, y \in X \setminus S}(\varepsilon_{x,y})$. The corresponding *effective grand coalition* G_S we define by $G_S = E_S \cup I_S$. While the effective inner grand coalition of a socially stable set maintains the inner stability, the effective grand external coalition affirms the external stability.

An effective external grand coalition may assert its external stability by enforcing out-going dominations, while ignoring incoming arcs. External stability may even be maintained if incoming arcs from subdued alternatives are added. A socially stable set may only be upset by the elimination or reversion of one or more of its outgoing arcs.

To make the situation regarding additional incoming arcs more precise, we introduce the property of ***independence of non-frontal opposition***:

$$\text{If } S \in \Phi(R), s \in S \text{ and } x \notin S, \text{ then } S \in \Phi(R \cup \{(x,s)\}) \tag{7}$$

It is easy to verify that Φ_{soc} and $\Phi_{\text{vN-M}}$ satisfy this axiom. We also have the following characterisation of $\Phi_{\text{vN-M}}$.

Theorem 9. *Let Φ satisfy core subduction and independence of non-frontal opposition. If the domain of Φ contains that of $\Phi_{\text{vN-M}}$, then $\Phi_{\text{vN-M}}(R) \subseteq \Phi(R)$.*

Proof. We first show that core primacy is implied by core subduction and independence of non-frontal opposition. To this end we let $S \in \Phi(R)$ and $x \in C(R)$. Suppose it is possible that $x \notin S$. Using the independence of non-frontal opposition, $S \in \Phi(R \cup \{(x,s)\})$, where s is taken from S. But, since $x \in C(R \cup \{(x,s)\})$ and $(x,s) \in R \cup \{(x,s)\}$, core subduction forces s to be member of $X \backslash S$, a contradiction. Now let R be a relation in the domain of $\Phi_{\text{vN-M}}$. We prove that $S \in \Phi(R)$ if $S \in \Phi_{\text{vN-M}}(R)$. Remove all incoming arcs $(x,s) \in R$, $x \notin S$, $s \in S$, resulting in R'. Since S is a vN-M stable set of R, we have $S = C(R')$. By core primacy and core subduction, $S = \Phi(R')$. Next we add the original arcs (x,s). By independence of non-frontal opposition we obtain $S \in \Phi(R)$. □

Theorem 9 is essential for the observation that $\Phi_{\text{vN-M}}$ is the smallest solution concept, with respect to inclusion, satisfying core subduction and independence of non-frontal opposition.

In the proof of Theorem 9, we showed that core primacy is implied by core subduction and independence of non-frontal opposition. So, our next characterisation of socially stable sets is a corollary of Theorem 8.

Theorem 10. *(i) Φ_{soc} satisfies independence of non-frontal opposition. (ii) If Φ satisfies core subduction, independence of non-enforced dominations, independence of non-frontal opposition and independence of stable cycles, then $\Phi(R) \subseteq \Phi_{\text{soc}}(R)$, for all R in the domain of Φ.* □

Below, we prove the independence of the axioms of Theorem 10 and show the non-redundancy of each postulate with respect to the other three. Again we take as domain the set of all asymmetric relations. Let $X = \{a,b,c\}$.

(1). Let $\Phi(R) = \{X\}$ for all R. Then Φ does not satisfy core subduction. (2). If R is strongly connected, $\Phi(R) = \{\{a\},\{b\},\{c\}\}$. For other relations, let $\Phi(R) = \Phi_{\text{soc}}(R)$. Then Φ does not satisfy independence of non-enforced dominations: if R is a 3-cycle, then $\{a\} \in \Phi(R)$, so (b,c) is non-enforced. However, $\{\{c,b\}\} = \Phi(R \backslash \{(b,c)\})$. (3). Let $\Phi(\{(a,b),(b,c)\}) = \{\{a\}\}$, $\Phi(R) = \Phi_{\text{soc}}(R)$ for other R. Then Φ is not independent of non-frontal oppositions, because $\{a\} \notin \Phi(\{(a,b),(b,c),(c,a)\})$. (4). Let $X = \{a,b,c,d\}$; $\Phi(\{(a,b),(b,c),(c,a),(a,d),(b,d),(c,d)\}) = \{\{a,b,c,d\}\}$. For other relations R, let $\Phi(R) = \Phi_{\text{soc}}(R)$. Then Φ is not independent of stable cycles: $\{a,b,c,d\} \notin \Phi(\{(a,d),(b,d),(c,d)\})$.

Combining Theorems 9 and 10 we obtain a more accurate position of Φ with respect to $\Phi_{\text{vN-M}}$ and Φ_{soc}.

Corollary 11. *Let Φ satisfy core subduction, independence of non-enforced dominations, independence of non-frontal opposition and independence of stable cycles. If we assume that the domain of Φ contains that of $\Phi_{\text{vN-M}}$, then $\Phi_{\text{vN-M}}(R) \subseteq \Phi(R) \subseteq \Phi_{\text{soc}}(R)$.* □

If in addition Φ satisfies the *universal domain* requirement, the first inclusion is strict: there are relations R with $\Phi_{\text{vN-M}}(R) \neq \Phi(R)$. We prove this by showing that there is no solution concept $\Phi \subseteq \Phi_{\text{soc}}$ with universal domain and

$\Phi = \Phi_{\text{vN-M}}$ on the domain of $\Phi_{\text{vN-M}}$, that satisfies the independence of non-frontal opposition. To this end let $X = \{a, b, c, d, e, f, g\}$, $R = \{(a, b), (b, c), (c, a), (b, d), (d, e), (e, f), (f, g), (g, d)\}$ and $R' = R \cup \{(d, c)\}$. Since $\Phi \subseteq \Phi_{\text{soc}}$, we have $\Phi(R) = \{a, b, c, e, g\}$, the unique socially stable set. Further, if $\Phi = \Phi_{\text{vN-M}}$ on the domain of $\Phi_{\text{vN-M}}$, we obtain $\Phi(R') = \{a, d, f\}$. This shows that Φ is not independent of non-frontal oppositions since (d, c) is a non-frontal opposition for $\Phi(R)$.

A solution concept $\Phi_{\max}$, giving a unique stable set, may be constructed by taking $S = T(R)$ at the initial step and $S' = T(R_{|U(S)})$ at each Step 2 of the construction algorithm. This solution concept does not satisfy the independence of non-frontal opposition. To show this let $X = \{a, b, c, d\}$ and $R = \{(a, b), (b, c), (c, d)\}$. Then $\Phi_{\max}(R) = \{a, c\}$. If we add (d, a) we have $\Phi_{\max}(R \cup \{(d, a)\}) = \{a, b, c, d\}$, showing that $\Phi_{\max}$ does not satisfy this independence condition. If we try to extend $\Phi_{\max}$ in order to attain a concept Φ that fulfils this postulate, this example shows that we have to demand that $\{\{a, c\}, \{a, b, c, d\}\} \subseteq \Phi(R \cup \{(d, a)\})$. In fact we have:

Theorem 12. *Let a solution concept Φ have universal domain. (i) If Φ satisfies independence of non-frontal opposition and, in addition, extends $\Phi_{\max}$, then $\Phi_{\text{soc}}(R) \subseteq \Phi(R)$ for all R. (ii) If Φ also satisfies core subduction, independence of non-enforced dominations and independence of stable cycles, we have $\Phi = \Phi_{\text{soc}}$.*

Proof. (i) We use induction to the number of dominations, $\# R$, to show that all socially stable sets S are in $\Phi(R)$. If $\# R = 0$, then obviously there is just one socially stable set: $S = \Phi_{\max}(R)$ which is supposed to be element of $\Phi(R)$. Now suppose the assertion is true for $\# R = k$, we prove it also holds for $\# R = k + 1$. Take a socially stable set S for R with $\# R = k + 1$. If S does not have incoming dominations, then the socially stable set S equals $\Phi_{\max}(R)$ showing that $S \in \Phi(R)$. Now suppose that (x, s) is an incoming arc: $x \notin S$ and $s \in S$. Then S is a socially stable set for $R \backslash \{(x, s)\}$. Using the induction step we obtain $S \in \Phi(R \backslash \{(x, s)\})$. But, since (x, s) is a non-frontal opposition, we obtain $S \in \Phi(R)$, proving the theorem. (ii) Combine theorem 10 and (i) above. □

This theorem shows that in the class of concepts Φ satisfying the independence of non-frontal opposition, Φ_{soc} is a minimal extension of $\Phi_{\max}$. Also requiring Φ to satisfy the other postulates, we finally obtain $\Phi = \Phi_{\text{soc}}$.

5 Concluding remarks

Various lines of research on domination have originated from von Neumann and Morgenstern's Theory of Games and Economic Behaviour. Studies in economic and co-operative game theory often presuppose that the numerical value of each coalition is known so that the characteristic function is defined. In this paper, such detailed assumptions are not made. Our results may however apply to the case of a social economy with a finite number of imputations, where there is social interaction between the effective coalitions. For an example see Delver and Monsuur (1997).

In political theory, generally, the dominance relation is the point of departure. An example of this approach is the bipartisan set of a majority tournament, see Laffond et al. (1993).

Another line of research dating back to the Theory of Games and Economic Behaviour concentrating on structural conditions for the existence of von Neumann Morgenstern stable sets, leads into graph theory. Here the object of study often is the reversal R^{-1} of a dominance relation R. A vN-M stable set of R is a *kernel* of R^{-1} and conversely.[7] According to C. Berge and P. Duchet, (1990), the main question in graph theory is: Which structural properties of a graph imply the existence of a kernel? A graph such that all it subgraphs have kernels is said to be *kernel perfect*. Von Neumann showed that a graph without circuits is kernel perfect and has a unique kernel, von Neumann and Morgenstern (1944). Richardson extended this result, requiring only the non-existence of odd circuits, see Richardson (1946) or Ghoshal et al. (1998).

A graph theoretic generalisation of the vN-M stable set, on the reversal of R, satisfying the universal domain requirement is the semikernel. Let d be the distance function in R between subsets or elements of X. Then $S \subseteq X$ is a *semikernel* or *(2,2) kernel* if for any distinct pair $\{x, y\} \subseteq S$ it is true that $d(x, y) \geq 2$, while for any $x \in X \setminus S$, $d(x, S) \leq 2$. Further generalisations of this idea are (k, l) kernels. References and a discussion may be found in Ghoshal et al. (1998).

There are two fashionable notions in public policy discussions for which socially stable sets may provide a theoretical foundation. In a society, public and private statements are tested against the established standard of behaviour. As mentioned briefly in the introduction, this verification functions as an often immediate trial on the *political correctness* of any proclamation or remark. The other notion is *repressive tolerance*. Its meaning comes close to the way in which, in Sect. 4, the effective grand inner coalition is able to keep up the external stability.

Socially stable sets are a model for situations where the combination of characterising postulates of one of our theorems is valid. In for example majority graphs for large elections and many imputations based economic applications these conditions may easily be not fulfilled. In particular the *non-enforcement principle* will in such circumstances probably not apply. On the other hand, in situations where the effective coalitions interact, such as in the political debate, the selection of alternatives in an organisation or society at large and decision making in small committees our behavioural assumptions and characterisations may provide useful insights in both the underlying social processes and the emerging standard of behaviour.

[7] R^{-1} has the same vertices as X while $(x, y) \in R^{-1}$ if and only if $(y, x) \in R$.

References

Arce DG (1994) Stability criteria for social norms with applications to the prisoner's dilemma. J Conflict Resol: 749–765

Barnett WA, Moulin H, Salles M, Schofield NJ (eds) (1995) Social choice welfare and ethics. Cambridge University Press, Cambridge

Berge C, Duchet P (1990) Recent problems and results about kernels in directed graphs. In: Hedetniemi ST, Laskar RC (eds) Topics on domination. North Holland, Amsterdam

Chwe MS-Y (1994) Farsighted coalitional stability. J Econ Theory 63: 299–325

Chvatal V, Lovasz L (1974) Every directed graph has a semi-kernel. (Lecture Notes in Mathematics 411) Springer, Berlin Heidelberg New York

Delver R, Monsuur H (1997) Towards a theory of domination. Scientific publications of the Royal Netherlands Naval Academy

Delver R, Monsuur H (1998) Echelons in incomplete relations. Theory Decision 44: 279–292

Ghoshal J, Laskar R, Pillone D (1998) Topics on domination in directed graphs. In: Haynes TW, Hedetniemi ST, Slater PJ. (eds.) Domination in graphs. Advanced topics. Marcel Dekker Inc

Greenberg J (1990) The theory of social situations. Cambridge University Press, Cambridge

Kreinovich VYa, Kosheleva O (1990) A hierarchic analogon of von Neumann-Morgenstern solution always exists. Technical report UTEP-CS-90-26, Computer Science Department, University of Texas at El Paso

Laffond G, Laslier JF, Le Breton M (1993) The bipartisan set of a tournament game. Games Econ Behav 5: 182–201

Monsuur H (1994) Choice, ranking and circularity in asymmetric relations. Dissertation Tilburg University

Nash JF (1950) The bargaining problem. Econometrica 18: 155–162

Richardson M (1946) On weakly ordered systems. Bull Amer Math Soc 52: 551–584

Thomson W (1995) Population monotonic allocation rules. In: Barnett WA, Moulin H, Salles M, Schofield NJ (eds.). Social choice, welfare and ethics. Cambridge University Press, Cambridge

van Deemen AMA (1991) A note on generalized stable sets. Soc Choice Welfare 8: 255–260

Vickrey W (1959) Self-policing properties of certain imputation sets. In: Tucker AW, Luce RD (ed.). Contributions to the theory of games, vol 4. Princeton University Press, Princeton, NJ pp. 213–246

von Neumann J, Morgenstern O (1944) Theory of games and economic behavior. Princeton University Press, Princeton, NJ