Working Paper in Economics

202313

Implementation in vNM stable sets.

Ville Korpela[*], Michele Lombardi[†] and Riccardo D. Saulle[‡]

[*] Turku School of Economics, University of Turku. E-mail: vipeko@utu.fi.
[†] University of Liverpool Management School, Liverpool, UK. Email: michele.lombardi@liverpool.ac.uk. Department of Economics and Statistics, University of Napoli Federico II.
[‡] Department of Economics and Management, University of Padova. E-mail: riccardo.saulle@unipd.it.

Implementation in vNM stable sets*

Ville Korpela[†] Michele Lombardi[‡] Riccardo D. Saulle[§]

November 28, 2023

Abstract

We fully identify the class of social choice functions that are implementable in von Neumann Morgenstern (vNM) stable sets (von Neumann and Morgenstern, 1944) by a rights structure. A rights structure formalizes the idea of power distribution in a society. Following the so-called *Harsanyi's critique* (Harsanyi, 1974), we also study the implementation of social choice correspondences in *strict* vNM stable sets.

Keywords: vNM Stable Set, Implementation, Rights Structures

JEL Codes: C71; D02; D71; D82.

*We greatly appreciate the comments of and discussions with Lars Ehlers, Hannu Vartiainen and the participants at 12^{th} Conference on Economic Design (CED12), 17^{th} European Meeting on Game Theory (SING17) and 23^{th} Annual SAET Conference (SAET2023). Riccardo Saulle acknowledge financial support by the University of Padova under grant StG-497985, the Department of Economics and Management "Marco Fanno" and Unicredit Foundation.

†Turku School of Economics, University of Turku. E-mail: vipeko@utu.fi.

‡University of Liverpool Management School, Liverpool, UK. Email: michele.lombardi@liverpool.ac.uk. Department of Economics and Statistics, University of Napoli Federico II.

§Department of Economics and Management, University of Padova. E-mail: riccardo.saulle@unipd.it

Contents

1 Introduction

The first solution concept for a general model of binding agreements has been introduced by von Neumann and Morgenstern (1944) in their monumental work on game theory. "The solution", so eloquently named by the authors and now widely referred to as the von Neumann Morgenstern (vNM) stable set, builds on a notion of dominance. A social state x dominates a state y if a coalition of agents has the right to move from y to x and each of its members strictly prefers x over y. The vNM stable set satisfies two properties. *Internal stability*: No state in the set is dominated by another state in the set. *External stability*: Every state outside the set is dominated by a state in the set.

Despite its applications in several areas, we still do not have a full understanding of the vNM stable set. We know that it is usually not unique (Lucas, 1968) and may fail to exist (Lucas, 1992). Also, the problem of its computation is undecidable (Deng and Papadimitriou, 1994). As Aumann (1987) aptly noted:

> Finding stable sets involves a new tour de force of mathematical reasoning for each game or class of games that is considered. Other than a small number of very elementary truisms (e.g., that the core is contained in every stable set), there is no theory, no tools, certainly no algorithm.

These facts imposed the Core as the central solution concept for games where coalitions are the fundamental decision units (Gillies, 1959). The Core is the set of undominated states, so its points are immune to coalitional deviations. A limitation of the Core is that it excludes a state just because another state dominates it without requiring that the coalitional deviation itself is credible.[1]

As pointed out by Ray and Vohra (2019), the idea of credibility is circular. A deviation is credible if no other credible deviation challenges it. The vNM stable set embodies this idea since it can be equivalently defined as the set of states that are not dominated by any state in the vNM stable set (von Neumann and

[1]This recently motivated Grabish and Sudhölter (2021) to identify necessary ad sufficient conditions for TU-games under which the Core coincides with the unique vNM stable set.

Morgenstern, 1944). Of course, the vNM stable set includes the Core, but it may also include other elements. This feature is significant because the Core might not be able to fully describe all agents' bargaining possibilities. In contrast, the vNM stable set may offer consistent predictions in several environments such as pure exchange economies (e.g. Einy and Shitovitz, 2003), matching (e.g. Ehlers, 2007; Herings, Mauleon and Vannetelbosch, 2017) and assignment problems (e.g. Núñez and Rafels, 2013). Recently, Ehlers and Morrill (2020) introduce a notion of stability *à la* von Neumann and Morgenstern in public school choice problems.

To make this point clearer, let us consider the following simple trading example. A seller has one indivisible object to be sold to one of two potential buyers. The valuation of the seller is zero, and the valuations of buyers, say v_1 and v_2, lie in the interval $[0, M] \subseteq \mathbb{R}_+$. Suppose that buyer 1's valuation is positive but smaller than buyer 2's valuation—that is, $0 < v_1 < v_2$. Let us assume that buyer $i = 1, 2$ would be willing to pay at most his valuation for the object. The Core of this trading example includes all trades where buyer 2 receives the object and pays a price $p \in [v_1, v_2]$ to the seller. Suppose that the seller and buyer 2 would trade at a price $p \in [0, v_1)$ that is outside the Core. The standard story says that this is impossible because buyer 1 would offer a higher price. However, the offer by buyer 1 is not credible because buyer 2 can credibly make an even larger counteroffer. The vNM stable set embodies this recursive idea: An allocation is stable even when a coalition can improve upon it because the new allocation can be credibly challenged by another deviation. In this example, the vNM stable is larger than the Core because it consists of all trades where buyer 2 receives the object and pays a price $p \in [0, v_2]$ to the seller.

The above trading example is not an exception.[2] Then, the solution concept of the Core, and so mechanisms implementing in Core, may not always function properly. To overcome these shortcomings, the designer may consider implementing in vNM stable instead.

[2]The Core can possibly exclude states that would not be credibly dominated. This, among others, is confirmed by empirical findings in matching markets that suggest that the Core can be small (e.g. Roth and Peranson, 1999).

The importance of the vNM stable set as a solution concept is undoubted. However, its normative investigation is almost entirely unexplored territory.[3] This paper aims to fill this vacuum by studying the vNM stable set in the realm of implementation theory.

Implementation theory offers a framework for the design of institutions, emphasizing the problem of incentives. A common interpretation of an implementation problem is that a hypothetical planner wants to achieve a socially desirable outcome without knowing agents' preferences. The social objective the planner wants to achieve is summarized by a social choice function (SCF), that is, a single-valued function mapping agents' preferences into an outcome. The planner decentralizes the decision-making by designing a mechanism or game form to achieve his goal. Roughly speaking, a mechanism represents the communication and decision aspects of an organization. Formally, it specifies a message space for each agent and an outcome function mapping vectors of messages into decisions. A mechanism implements an SCF if its equilibrium outcome is consistent with the SCF, irrespective of agents' preferences.

Although successful results have been obtained in the last decades in identifying the classes of SCFs that can be implemented with this approach, it is still unclear how to replicate the circular character of the vNM stable set via a game form. To overcome this issue, we follow the approach developed by Sertel (2001) and Koray and Yildiz (2018), who propose a notion of rights structure as a formalization of what a coalition can or cannot do in a society. A rights structure is a flexible tool for designing institutions such as constitutions, legal codes, and rules of conduct. Moreover, it has the merit of formulating rules of behavior in a language that is easily understandable and closer to "real life" rules. With this approach, an implementation problem consists of designing a rights structure such that its equilibrium outcome corresponds to the outcome of the SCF, irrespective of agents' preferences. In solving this problem, the planner describes the available alternatives via a set of possible states and specifies which agents or coalitions

[3] A notable contribution is due by Kimya (2023) who provides an axiomatization of a farsighted version of the vNM stable set.

have the right to move from one state to another.[4] A devised rights structure implements a given SCF in vNM stable sets when the outcome corresponding to its vNM stable states is consistent with the SCF. As reflected by recent contributions (Koray and Yildiz, 2018; Korpela, Lombardi and Vartiainen, 2020, 2021), this "blocking" approach to implementation theory suits the normative investigation of cooperative solution concepts well.

The main contribution of this paper is the characterization of SCFs implementable in vNM stable sets. The class of implementable SCFs is completely characterized by three conditions: NO SIMULTANEUS DOMINATIONS, vNM MONOTONICITY, and TEST CYCLE. These conditions also imply that implementation in the stable set is independent of implementation in (externally stable) core. Indeed, whereas only Pareto efficient SCFs are implementable in vNM stable sets, Pareto-dominated outcomes can be implemented in Core. Furthermore, any implementable SCF in Core is (Maskin) monotonic, whereas monotonicity is not necessary for implementation in vNM stable sets. This asymmetry is recomposed partially when linear orderings represent agents' preferences. In this case, implementability in vNM stable sets implies implementability in Core. The converse implication, however, is false.

We mainly focus on SCFs because there is no other reason than the lack of existence of a single-outcome vNM stable set that led von Neumann and Morgenstern (1944, p.39) to focus on multi-outcomes vNM stable sets. Indeed, by paraphrasing von Neumann and Morgenstern (1944, p. 34, 4.2.2):

> We can see no reason why one should no be satisfied with a solution of this nature, providing it can be found: i.e. a single-imputation which meets reasonable requirements for optimal (rational) behaviour.

[4] Although an alternative can represent a state, they are typically different objects. Indeed, a state reflects a situation that society may end up with, supported by an argument or evidence, and may also include an outcome. For instance, suppose that there are two candidates $\{x, y\}$ and three agents and that preferences are strict. Suppose our goal is to implement the majority solution: $f(R) = a$ if a is preferred to b by at least two agents. In this case, the set of states may consist of an outcome paired with a coalition of size two. The interpretation is that a state $(a, \{1, 2\})$ is a claim that a is preferred to b by agents 1 and 2. And so on. Like in standard mechanism design theory, after we have found an implementing rights structure, we must ask what kind of social organization produces it.

Actually, in the original idea of von Neumann and Morgenstern (1944, p.34), a single-outcome vNM stable set is an ideal solution because it corresponds to what they called "an absolute state of equilibrium" or the "first element" of the orderings induced by the dominance relation—that is, to the outcome that dominates all the others. Since the dominance relation is usually not transitive, such a "first element" is unlikely to exist. However, this is not an issue for us because an SCF is said to be implementable in vNM stable sets whenever each possible social environment admits a single-outcome vNM stable set. In addition, a single-outcome vNM stable set has the property to be a consistent set (Chwe, 1994) and a farsighted stable set (Harsanyi, 1974; Chwe, 1994), and so our conditions also guarantee the partial implementation of an SCF in such solutions.

In Section 4, we also study the case that the planner's goal is represented by a social choice correspondence (SCC, hereafter). A primary difficulty in studying SCCs is the so-called *Harsanyi's critique*. Harsanyi (1974) criticized the vNM stable because it is based on the assumption that coalitions are myopic. Specifically, he argued that the property of internal stability can be violated when agents are farsighted. For this reason, Harsanyi (1974) proposed a farsighted version of the vNM stable set, called *strict* vNM stable. We also characterize a class of SCCs implementable in strict vNM stable sets.

2 Preliminaries

We consider a finite non-empty set of *agents*, denoted by N, and a non-empty set of *alternatives*, denoted by Z. For each agent $i(\in N)$, a *preference relation* over Z is a complete and transitive binary relation $R_i \subseteq Z \times Z$. We denote by P_i the asymmetric part of R_i, i.e., xP_iy if and only if xR_iy and not yR_ix, while the symmetric part of R_i is denoted by I_i, i.e., xI_iy if and only if xR_iy and yR_ix. A *preference profile* $R \equiv (R_i)_{i\in N}$ lists the preferences of all agents in N. Let $\mathcal{R}$ be the collection of all admissible preference profiles. A *coalition* K is any non-empty subset of N. For any preference profile $R(\in \mathcal{R})$ and coalition $K \subseteq N$, we write xR_Ky and xP_Ky for xR_iy for all $i \in K$ and xP_iy for all $i \in K$, respectively. For any $R \in \mathcal{R}$, and any $x, y \in Z$, let $K(R, x, y)$ be a coalition defined by the rule: $i \in K(R, x, y) \Longleftrightarrow xP_i y$. That is, $K(R, x, y)$ is the set of agents that strictly prefer x to y at R. As usual, $L_i(x, R)$ denotes the *lower contour set* of x at R for agent i. The goal of the planner can be represented by a social choice correspondence or a social choice function. A social choice correspondence (SCC) is a correspondence $F : \mathcal{R} \rightrightarrows Z$ such that $\varnothing \neq F(R) \subseteq Z$ for all $R \in \mathcal{R}$. A social choice function (SCF), denoted by $f : \mathcal{R} \to Z$, is a single-valued SCC. We say that x is F-optimal at R if $x \in F(R)$. The *range* of $F : \mathcal{R} \rightrightarrows Z$ is the set $F(\mathcal{R}) \equiv \{x \in Z | x \in F(R) \text{ for some } R \in \mathcal{R}\}$, and the *graph* of $F : \mathcal{R} \rightrightarrows Z$ is the set $Gr(F) \equiv \{(x, R) | x \in F(R), R \in \mathcal{R}\}$. For all $x \in Z$, let $F^{-1}(x) \equiv \{R \in \mathcal{R} | x \in F(R)\}$ be the inverse image of F at x.

For all $R \in \mathcal{R}$ and all $x, z \in Z$, we say that z is *equivalent* to x at R if $xI_N z$, and that z is *welfare equivalent* to x at R if $z \in f(\mathcal{R})$ and x is equivalent to z at R. We write $I^f(x, R) = \{z \in f(\mathcal{R}) \mid zI_N x\}$ for the set of all welfare equivalent outcomes to x at R.

To implement his goal, the planner designs a *rights structure* $\Gamma = (S, h, \gamma)$, where S is the *state space*, $h : S \to Z$ the *outcome function*, and $\gamma : S \times S \rightrightarrows N$ a *code of rights*, which specifies, for each pair of distinct states (s, t), the collection of coalitions $\gamma(s, t) \subseteq 2^N$ that is entitled to move from state s to t. If $\gamma(s, t) = \varnothing$, then no coalition is entitled to move from s to t. To save notation, we denote $S^x = \{s \in S | h(s) = x\}$, with a typical element s^x, the set of states where the outcome is x. A rights structure and a preference profile return a *social environment*, a general framework to model strategic interaction among agents or groups.[5]

DEFINITION **1** (Social Environment)**.** A *social environment* is a pair $\langle \Gamma, R \rangle$ consisting of a rights structure Γ and a preference profile R.

Agents' behavior is described by the solution concept Σ which select a subset of S for each social environment $\langle \Gamma, R \rangle$. A right structure Γ implements $F : \mathcal{R} \rightrightarrows Z$ in the solution concept Σ if the outcomes corresponding to the states selected by

[5]When S is the set of outcomes and h is the identity function, then our social environment coincides with the social environment of Chwe (1994).

Σ coincide with F at any preference profile. That is, if $F(R) = h \circ \Sigma(\Gamma, R)$ for all $R \in \mathcal{R}$. Therefore, from an economic design perspective, the rights structure is the planner's design variable and corresponds to a "mechanism" in the economic theory jargon. FIGURE 1 adapts the well-known Mount-Reiter diagram (Mount and Reiter, 1974) to illustrate the implementation via rights structure.

$$\begin{tikzcd}
\mathcal{R} \arrow[rr, "F"] \arrow[dr, dash, "\Sigma"] & & Z \\
& S, \gamma \arrow[ur, "h"] &
\end{tikzcd}$$

FIGURE 1: A Mount-Reiter type diagram illustrating the implementation via rights structure

DEFINITION 2 establishes the *dominance* relation: a state $y \in S$ dominates another state $x \in S$ if there is a coalition such that (i) it can move from x to y and (ii) each of its members strictly prefer to do so.

DEFINITION **2** (Dominance)**.** Given a social environment $\langle \Gamma, R \rangle$ and states $s, s' \in S$, the state $s \in S$ dominates $s' \in S$ under γ at $R \in \mathcal{R}$, if there is a coalition $K \subseteq N$ such that: (i) $K \in \gamma(s', s)$; and (ii) $h(s) P_K h(s')$.

Given $\langle \Gamma, R \rangle$, if s dominates s' under Γ at R, then we write $s >_{(\Gamma,R)} s'$. DEFINITION 3 introduces the notion of a vNM stable set for any social environment $\langle \Gamma, R \rangle$.

DEFINITION **3.** Let $\langle \Gamma, R \rangle$ be a social environment. The set $V(\Gamma, R) \subseteq S$ is a vNM stable set of Γ at R if it satisfies the following conditions:

Internal Stability: for all $s, s' \in V(\Gamma, R)$, not $s' >_{(\Gamma,R)} s$.

External Stability: for all $s \notin V(\Gamma, R)$, there exists $s' \in V(\Gamma, R)$ such that $s' >_{(\Gamma,R)} s$.

Internal Stability requires that no two states inside the set dominate each other. External Stability requires that each state outside the set is dominated by a state inside the set. Internal and external stability work together: No two allocations threaten each other, and jointly, the stable allocations dominate all non-stable allocations. As von Neumann and Morgenstern (1944) pointed out, the notion of the vNM stable set can be stated as a single condition. For a given social environment (Γ, R) and any subset $A \subseteq S$ define $Dom_{(\Gamma,R,>)}(A)$, the *dominion* of A, as the subset of states that are dominated by some element of A, formally, $Dom_{(\Gamma,R,>)}(A) \equiv \{s \in S | \exists s' \in A : s' >_{(\Gamma,R)} s\}$. Then, any vNM stable set at (Γ, R) is $V(\Gamma, R) \equiv S - Dom_{(\Gamma,R,>)}(V(\Gamma, R))$ that is the set of states that are not dominated by any state in the vNM stable set. We denote by $vNM(\Gamma, R)$ the union of all vNM stable sets at $\langle \Gamma, R \rangle$. On the other hand, the Core of Γ at R can be defined as $C(\Gamma, R) = S - Dom_{(\Gamma,R,>)}(S)$, which consists of the set of states that are not dominated by any coalition.

Definition **4** (Implementation in vNM stable sets)**.** A rights structure Γ implements $F : \mathcal{R} \rightrightarrows Z$ in vNM stable sets if $F(R) = h \circ vNM(\Gamma, R)$ for all $R \in \mathcal{R}$. If such a rights structure exists, $F : \mathcal{R} \rightrightarrows Z$ is implementable in vNM stable sets by a rights structure.

The following axiom, introduced by Korpela, Lombardi and Vartiainen (2020) to characterize the class of functions implementable in externally stable Core, will be used thereafter.

Definition **5** (no simultaneous domination)**.** $F : \mathcal{R} \rightrightarrows Z$ satisfies no simultaneous domination (NSD, henceforth) provided that there exists Y, with $F(\mathcal{R}) \subseteq Y$, such that for all $R \in \mathcal{R}$ and all $x \in Y \backslash F(R)$, there exist $i \in N$ and $x' \in F(R)$ such that $x' P_i x$.

NSD simply states that if outcome x is not F-optimal at R, it cannot be that this x dominates every outcome in the range of F at R, in the sense that x is at least as good as every F-optimal outcome at R for every agent $i \in N$. When preference domain $\mathcal{R}$ is the domain of linear orderings, the condition implies that an outcome x that is not F-optimal at R cannot Pareto dominate every F-optimal outcome at R.[6]

It turns out that NSD is necessary for the implementation of SCCs in vNM stable set.

THEOREM **1.** *If $F : \mathcal{R} \rightrightarrows Z$ is implementable in vNM stable sets by a rights structure, then it satisfies NSD.*

3 A full characterization for SCFs

In coalition theory, a vNM stable set $V(\Gamma, R)$ is said single-payoff if for all $s, s' \in V(\Gamma, R)$, $h(s) = h(s')$. Single-payoff cooperative solutions are widely studied in coalition theory.[7] Their relevance is also motivated by von Neumann and Morgenstern (1944, p.37). From an implementation point of view, this implies focussing on SCFs. In this section, we identify the class of SCFs that are implementable in vNM stable sets.[8]

3.1 Necessary and Sufficient Conditions

Koray and Yildiz (2018) and Korpela, Lombardi and Vartiainen (2020) show that (Maskin) monotonicity is necessary for implementation in Core via a rights structure. Monotonicity requires that if an outcome x is f-optimal at R,[9] preferences change from R to R', and the outcome x does not fall in any agent's preference ordering relative to any other alternative, then x remains f-optimal at R'[10]. The following example shows that monotonicity is not necessary for implementation in vNM stable sets via a rights structure. This supports Serrano's

[6]For any profile R, we say that outcome x *Pareto dominates* y if xP_iy for all $i \in N$.

[7]Prominent examples are the *Shapley value* (Shapley, 1951) and the *nucleolus* (Schmeidler, 1969). Furthermore, single-payoff analysis is pervasive in the farsighted coalition formation literature: Béal, Durieu, and Solal (2008); Mauleon, Vannetelbosch, and Vergote (2011); Ray and Vohra (2015); Dutta and Vohra (2017); Ray and Vohra (2019); Herings, Mauleon and Vannetelbosch (2020); Bloch and van den Nouweland (2021); Karos and Robles (2021).

[8]It may be worth noting that the single-payoff vNM stable set is unique when it exists. Moreover, it is a consistent set (Chwe, 1994) and a farsighted stable set (Harsanyi, 1974; Chwe, 1994). Consistent sets and farsighted sets are formally introduced in Section 4.

[9]x is f-optimal at R means that $f(R) = x$.

[10]Formally, for all $R, R' \in \mathcal{R},\ \ L_i(f(R), R) \subseteq L_i(f(R), R')\ \forall i \in N\ \ \rightarrow f(R) = f(R')$

conjecture (Serrano, 1997), according to which only the Core satisfies monotonicity among the major cooperative solutions.

Example 1. There are two agents $\{1, 2\}$, three outcomes $\{x, y, z\}$ and two preference profiles $R, R' \in \mathcal{R}$. The table below specifies agents' preferences.[11] The SCF $f : \{R, R'\} \to \{x, y, z\}$ is such that $f(R) = x$ and $f(R') = y$. Note that $f : \mathcal{R} \to Z$ is not monotonic: x is f-optimal at R. No agent experiences a preference reversal around x when the state changes from R then R', but x is not f-optimal at R'. However, the SCF is implementable in vNM stable sets. The right-hand side of Figure 2 is an example of implementing rights structure. First, we impose that states are outcomes. An oriented graph represents the rights structure. The vertices are the states. The edges represent the code of rights: Agent 1 can move from x to y and from z to x. Agent 2 can move from y to x and from z to y and *vice versa*. According to this rights structure, the unique vNM stable set at

R		R'	
1	*2*	*1*	*2*
y	x	y	x, y
x	z	x, z	z
z	y		

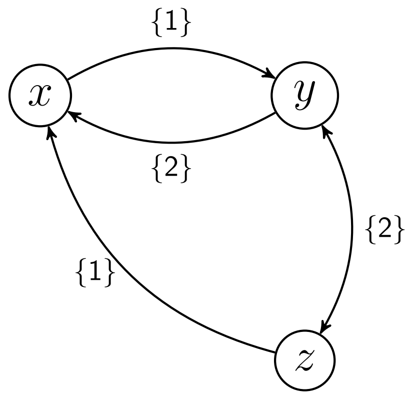

Figure 2: An example of a non-monotonic SCF and an implementing rights structure.

R and R' are, respectively, $vNM(\Gamma, R) = \{x\}$ and $vNM(\Gamma, R') = \{y\}$. To see this, take, as an example, the preference profile R. Then, $\{x\}$ trivially satisfies internal stability. External stability is also satisfied since z and y are dominated by x. Note that $\{x\}$ is the unique vNM stable set at R. Indeed, one can check that at R, any subset of $\{x, y, z\}$ different from $\{x\}$ violates either internal or external stability. A similar argument applies to R'.

[11] We allow for indifferences because implementation in vNM stable sets implies implementation in Core when preferences are strict. On this point, see Corollary 4

REMARK **1.** Note that in the above example, the Core is empty at the profile R. This is the reason that the above example violates monotonicity.

To guarantee the property of "external stability" of $f : \mathcal{R} \to Z$ at the profile R, for every outcome x different from $f(R)$, one agent needs to prefer strictly $f(R)$ to x. This property is captured by the axiom discussed at the end of SECTION 2 and named NO SIMULTANEOUS DOMINATION.

Note that if $f(R)$ is Pareto dominated at R by an outcome $x \in Y$, then f violates NO SIMULTANEOUS DOMINATION. Moreover, if f satisfies NO SIMULTANEOUS DOMINATION, then f is a sub-solution of the Pareto correspondence where the set of feasible outcomes is Y. This implies that only efficient SCFs (relative to Y) can be implemented in vNM stable sets. This contrasts with the implementation in Core, which does not require NSD (Koray and Yildiz, 2018; Korpela, Lombardi and Vartiainen, 2020)

It is straightforward to see that the SCF described in EXAMPLE 1 satisfies NSD. However, NSD is not sufficient for the implementability in vNM stable sets. We show this in the example below.

EXAMPLE **2.** There are two agents $\{1, 2\}$, three outcomes $\{x, y, z\}$ and two preference profiles $R, R' \in \mathcal{R}$. The table below specifies agents' preferences.

R		R'	
1	*2*	*1*	*2*
y	x	y	x
x, z	z	x	y, z
	y	z	

The SCF $f : \{R, R'\} \to \{x, y, z\}$ is such that $f(R) = x$ and $f(R') = y$. Note that the SCF satisfies NSD: Agent 2 strictly prefers $f(R)$ to y and to z at R and agent 1 strictly prefers $f(R')$ to x and to z at R'.

However, $f : \mathcal{R} \to Z$ is not implementable in vNM stable sets. Indeed, if S^x is a vNM stable set at R and S^y a vNM stable set at R', then it has to be that S^x is also a vNM stable set at R'. To see it, note that S^x satisfies internal stability at any preference profile, including R'. Also, any rights structure implementing S^x

at R must satisfy the following property. Agent 2 must be allowed to move from $s' \in S^y$ to $s \in S^x$ and from $s'' \in S^z$ to $s \in S^x$. Otherwise, external stability is not satisfied. Since this guarantees external stability at R' for the set S^x, it follows that S^x is a vNM stable set at R'.

EXAMPLE 2 suggests that another property is required to rule out undesirable outcomes. In the particular case of EXAMPLE 2, the planner wants to achieve S^x as the unique vNM stable set at R and S^y as the unique vNM stable set at R'. However, x happens to be a vNM stable set at R' because the agents strictly preferring x to y and x to z at R, namely agent 2, and the same happens at R'. In other words, from one side, agent 2 guarantees external stability of S^x at R; from the other side, no other agent is breaking the external stability of S^x at R'. To break this, we need a preference reversal when we move from R to R'.

An implementable SCF satisfies the von Neumann Morgenstern monotonicity. We abbreviate this condition as VNM MONOTONICITY. It requires that for any x in the range of $f : \mathcal{R} \to Z$ that is not f-optimal at some R', an outcome z acting as a breaking-point of the vNM stability of x at R' exists. In particular, VNM MONOTONICITY requires that an outcome z exists such that, for every preference profile R at which x is f-optimal, one agent who strictly prefers x to z experiences a preference reversal when we move from R to R'. Moreover, this preference reversal over the pair $\{x, z\}$ must also hold over the pair $\{x^*, z\}$, where x^* is any outcome that is welfare equivalent to x at R'.

DEFINITION **6** (VNM MONOTONICITY). $f : \mathcal{R} \to Z$ satisfies VNM MONOTONICITY if there exists $Y \subseteq Z$ such that $f(\mathcal{R}) \subseteq Y$, and that for all $(x, R') \in Z \times \mathcal{R}$ with $x \in f(\mathcal{R}) \backslash f(R')$, there exists $z \in Y$ such that $K(R, x^*, z) \nsubseteq K(R', x^*, z)$ for all $x^* \in I^f(x, R')$ and all $R \in f^{-1}(x^*)$.

Henceforth, we denote by $\mathcal{M}^f(x, R')$ the set of outcomes satisfying VNM MONOTONICITY at (x, R'). That is, $\mathcal{M}^f(x, R')$ contains all attainable outcomes z such that for all $x^* \in I^f(x, R')$ and all $R \in f^{-1}(x^*)$, the set of agents strictly preferring x^* to z at R' differs from the set of agents strictly preferring x^* to z at R.

THEOREM **2.** *If $f : \mathcal{R} \to Z$ is implementable in vNM stable sets by a rights structure, then it satisfies* vNM MONOTONICITY.

The reader can check that, in EXAMPLE 2, the SCF violates vNM MONOTONICITY. To this end, note that $x \in f(R)\backslash f(R')$, $\{2\} = K(R, x, y) = K(R', x, y)$ and $K(R, x, z) = K(R', x, z) = \{1, 2\}$. Therefore, $\mathcal{M}^f(x, R') = \varnothing$.

Next, we show that NSD and vNM MONOTONICITY are not sufficient for implementing $f : \mathcal{R} \to Z$ in vNM stable sets. The following example makes the point.

EXAMPLE **3.** There are two agents $\{1, 2\}$, three outcomes $\{x, y, z\}$ and two preference profiles $R, R' \in \mathcal{R}$. The table below specifies agents' preferences.

R		R'	
1	*2*	*1*	*2*
y, z	x	y	x, z
x	y	x, z	y
	z		

Again, $f : \{R, R'\} \to \{x, y, z\}$ is such that $f(R) = x$ and $f(R') = y$. Note that this non-monotonic SCF satisfies NSD: Agent 2 strictly prefers $x = f(R)$ to y and z at R and agent 1 strictly prefers $y = f(R')$ to x and z at R'. It also satisfies vNM MONOTONICITY because $\mathcal{M}^f(x, R') = \{z\}$ and $\mathcal{M}^f(y, R) = \{z\}$. However, this $f : \mathcal{R} \to Z$ is not implementable. Indeed, if S^x were a vNM stable set at R and S^y a vNM stable set at R', then it would have to be that $S^x \cup S^z$ is a vNM stable set at R'. To see the latter point, note that any implementing rights structure where S^x is a vNM stable set at R must allow agent 2 to move from each $s^y \in S^y$ to an s^x and from each $s^z \in S^z$ to an s^x. Otherwise, external stability would not be satisfied for S^x at R. Since agent 2 can move from a state in S^y to a state in S^x and he has incentives to do so at R' and since agent 2 can move from a state in S^z to a state in S^x but he is indifferent between x and z at R', it follows that the set $S^x \cup S^z$ satisfies external stability at R'. Since agents are indifferent between x and z at R', it follows that $S^x \cup S^z$ is a vNM stable set at R'.

EXAMPLE 3 suggests that vNM MONOTONICITY is too weak for ruling out all undesirable vNM stable sets. How can we rule them out?

A clue comes from the studies of Richardson (1946, 1953) and Harary et al. (1966), which show that an *odd cycle* exists when no vNM stable set exists. An odd cycle at P' is a sequence of outcomes $z^1, z^2, ..., z^k$ where $k \in \mathbb{N}$ is odd and such that $z^k P'_{i_h} z^1 P'_{i_1} z^2 P'_{i_2} ... P'_{i_{h-1}} z^k$ holds for $i_1, i_2, ..., i_h \in N$.[12] Although their result is undoubtedly relevant from a positive point of view,[13] it is helpful for our purposes as well. Indeed, to be sure that in Example 3 the set $S^x \cup S^z$ is not a vNM stable set at R', we need to make sure that when we move from R to R' and $f(R) \neq f(R')$, either an odd cycle at R' exists among the states in S^z—that is, among the outcomes in $\mathcal{M}^f(x, R')$—or an odd cycle at R' exists among a state in S^x, the f-optimal outcome at R' and a state in S^z—that is, among x, $f(R')$ and z. These requirements allow us to violate the internal stability of $S^x \cup S^z$ at R'. This is the core idea of our following necessary condition, which is called TEST CYCLE and builds over the notion of an odd cycle.

Definition **7** (TEST CYCLE)**.** $f : \mathcal{R} \to Z$ satisfies TEST CYCLE if for all $R' \in \mathcal{R}$ and all $x \in f(\mathcal{R}) \backslash f(R')$ such that $x P'_i f(R')$ for some $i \in N$, one of the following requirements holds:

(i) There exists $z \in \mathcal{M}^f(x, R')$ such that for all $x^* \in I^f(x, R')$, $x^* P'_i f(R') P'_j z P'_k x^*$ holds for some $j, k \in N$.

(ii) There exists an odd cycle at R' with outcomes in $\mathcal{M}^f(x, R') \cup I^f(x, R')$.

(iii) $f(R') \in \mathcal{M}^f(x, R')$.

Theorem **3.** *If $f : \mathcal{R} \to Z$ is implementable in vNM stable set by a rights structure, then it satisfies* TEST CYCLE*.*

Remark **2.** In Example 3, f is not implementable because it violates TEST CYCLE. To see it, observe that $x = f(R) \neq f(R') = y$ and that $x P'_2 f(R')$. Parts (ii)-(iii) of the condition are violated because $f(R') \notin \mathcal{M}^f(x, R') = \{z\}$ and $I^f(x, R') = \{x\}$. Part (i) is also violated because $\mathcal{M}^f(x, R') = \{z\}$, $I^f(x, R') = \{x\}$ and no agent k exists who strictly prefers z to x at R'.

[12] From definition, it follows that $k \geqslant 3$.

[13] The equivalent statement of Richardson's result is that if there are no odd cycles, then a vNM stable set exists.

Theorem 1, Theorem 2 and Theorem 3 prove the following corollary.

Corollary 1 (Necessity). *If $f : \mathcal{R} \to Z$ is implementable in vNM stable sets, then there exists $X \subseteq Z$ such that $f(\mathcal{R}) \subseteq X$ and that $f : \mathcal{R} \to Z$ satisfies NSD, vNM MONOTONICITY and TEST CYCLE with respect to X.*

Next, we show that NSD, vNM MONOTONICITY, and TEST CYCLE are also sufficient for implementation in vNM stable sets.

Theorem 4 (Sufficiency). *Let $X \subseteq Z$ be such that $f(\mathcal{R}) \subseteq X$. If $f : \mathcal{R} \to Z$ satisfies NSD, vNM MONOTONICITY and TEST CYCLE with respect to X, then $f : \mathcal{R} \to Z$ is implementable in vNM stable sets by a rights structure.*

3.2 Environments with a simpler characterization

This subsection shows that the TEST CYCLE condition is redundant in environments with transfers and when linear orderings represent agents' preferences.

3.2.1 Transfers

Let D be a set of potential social decisions with typical element $d \in D$. A transfer of agent i is any real number $t_i \in \mathbb{R}$. As usual, we write $t_{-i} \equiv (t_i)_{i \in N \setminus \{i\}} \in \mathbb{R}^{n-1}$. In this environment, an outcome $z \in Z \equiv D \times \mathbb{R}^n$ consists of a social decision d together with a profile of transfers $t = (t_1, ..., t_n)$. For any $i \in N$, agent i's preference relation R_i is defined over Z. An *environment with transfers* is a triplet $\langle N, Z, (R_i)_{i \in N} \rangle$. We impose over R_i the following requirement:

Definition 8 (Money Monotonicity). Agent i 's preference relation R_i is *money monotonic* if for all $d \in D$, all $t_{-i} \in \mathbb{R}^{n-1}$, and all $t_i, t'_i \in \mathbb{R}$, $t_i > t'_i \quad \Rightarrow (d, (t_i, t_{-i})) P_i (d, (t'_i, t_{-i}))$.

The next result shows that in an environment with transfers where preferences satisfy some requirements, vNM MONOTONICITY implies TEST CYCLE. In light of this result and of Theorem 4, we obtain that in an environment with transfers where preferences are continuous and money monotonic and where the domain $\mathcal{R}$ is finite, NSD and vNM MONOTONICITY fully characterize the class of implementable functions.

THEOREM **5.** *Assume that preferences in $\mathcal{R}$ are continuous, and money monotonic and that the cardinality of $\mathcal{R}$ is finite. If $f : \mathcal{R} \to Z$ satisfies NSD and vNM MONOTONICITY, then it satisfies TEST CYCLE.*

COROLLARY **2.** *Assume that preferences in $\mathcal{R}$ are continuous, and money monotonic and that the cardinality of $\mathcal{R}$ is finite. Then, $f : \mathcal{R} \to Z$ is implementable in vNM stable sets by a rights structure if and only if it satisfies NSD and vNM MONOTONICITY in some set $Y \subseteq Z$ such that $f(\mathcal{R}) \subseteq Y$.*

3.2.2 Linear Orderings

A binary relation $R_i \subseteq Z \times Z$ is a linear order if it is reflexive, transitive, and anti-symmetric. Let $\mathcal{L}$ be a domain of profiles of linear orderings. The next result states that TEST CYCLE is redundant under the domain restriction.

THEOREM **6.** *If $f : \mathcal{L} \to Z$ satisfies* NO SIMULTANEOUS DOMINATION *and vNM MONOTONICITY, then it satisfies TEST CYCLE.*

Therefore, NSD and vNM MONOTONICITY fully characterize the class of implementable functions in vNM stable sets when agents' preferences are linear orderings.

COROLLARY **3.** *$f : \mathcal{L} \to Z$ is implementable in vNM stable sets via a rights stricture if and only if there exists a set $X \subseteq Z$ such that $f : \mathcal{L} \to Z$ satisfies NSD and vNM MONOTONICITY with respect to X.*

In SECTION 3.4, we show that a similar result applies in environments with transfers under mild conditions on agents' preferences.

3.3 Connections with the (Externally Stable) Core

The fact that an SCF is implementable in vNM stable sets does not imply that it is implementable in Core. Indeed, monotonicity, which is necessary for implementation in Core (Koray and Yildiz, 2018; Korpela, Lombardi and Vartiainen, 2020), is not a necessary condition for implementation in vNM stable sets (see EXAMPLE 1 above). Recall that, since NO SIMULTANEOUS DOMINATION, only Pareto efficient SCFs are implementable in vNM stable sets, whereas this is not the case for implementation in Core. EXAMPLE 4 below makes this point.

EXAMPLE **4.** Suppose that $N = \{1,2,3\}$, $Z = \{x,y,z\}$, and $\mathcal{R} = \{R, R', R''\}$. Preferences of agents at different profiles are defined in the table below.

R			R'			R''		
1	*2*	*3*	*1*	*2*	*3*	*1*	*2*	*3*
x	y	x	z	y	y	x	x	x
z	x	z	x	z	z	y	y	y
y	z	y	y	x	x	z	z	z

Let SCF f be such that $f(R) = z$, $f(R') = y$, and $f(R'') = x$. This SCF is monotonic and unanimous. Therefore, f is implementable in the core (Korpela, Lombardi and Vartiainen, 2020). However, it does not satisfy NSD; hence, it is not implementable in vNM stable sets.

The asymmetry between implementation in vNM stable sets and implementation in Core is partially recomposed when preferences are linear orderings. First, one can prove that under this domain restriction, NSD, together with vNM MONOTONICITY, implies monotonicity and unanimity, which are the necessary and sufficient conditions for implementation in Core.

THEOREM **7.** *If $f : \mathcal{L} \to Z$ is implementable in vNM stable sets, then it satisfies monotonicity and unanimity in some set $Y \subseteq Z$ such that $f(\mathcal{R}) \subseteq Y$.*

Then, since monotonicity and unanimity w.r.t. Y fully characterize the implementation in Core (Korpela, Lombardi and Vartiainen, 2020), the following result directly follows.

COROLLARY **4.** *If SCF $f : \mathcal{L} \to Z$ is implementable in vNM stable sets, then it is also implementable in Core.*

Nevertheless, the implementation in vNM stable sets and implementation in Core are very different design exercises. Indeed, the following example shows that implementation in Core does not imply implementation in vNM stable sets even when $f : \mathcal{L} \to Z$ is Pareto efficient. The example below makes it crystal clear.

EXAMPLE **5.** Suppose that $N = \{1, 2\}$, $Z = \{x, y, z, u\}$, and $\mathcal{R} = \{R, R', R''\}$. Preferences of agents at different profiles are defined in the table below.

R		R'		R''	
1	*2*	*1*	*2*	*1*	*2*
u	z	u	z	x	u
x	y	z	x	z	z
z	x	x	y	u	y
y	u	y	u	y	x

Let f be such that $f(R) = x$, $f(R') = z$, and $f(R'') = z$. This SCF is monotonic and unanimous. Therefore, f is implementable in the core (Korpela, Lombardi and Vartiainen, 2020). However, it does not satisfy vNM MONOTONICITY; hence, it is not implementable in vNM stable sets. To see this, notice that z belongs to $f(\mathcal{R})\backslash\{f(R)\}$, $I^f(z, R) = \{z\}$, and $f^{-1}(z) = \{R', R''\}$. Since $\{2\} = K(R', z, u) \subseteq K(R, z, u) = \{2\}$, $\{1, 2\} = K(R', z, y) \subseteq K(R, z, y) = \{1, 2\}$, and $\{2\} = K(R'', z, x) \subseteq K(R, z, x) = \{2\}$, f does not satisfy vNM MONOTONICITY.

The relationship between implementation in vNM stable sets and implementation in Core via rights structure goes behind COROLLARY 4. Indeed, it is well known the Core is contained in every stable set and moreover, if the Core satisfies the property of external stability, then it becomes the unique vNM stable set. This observation led to the following remark.

REMARK **3.** If $f : \mathcal{R} \to Z$ is implementable in externally stable Core, then it is also implementable in vNM stable sets.

In what follows, we argue that the converse of Remark 3 is false. Koray and Yildiz (2018) introduce a notion of Winner monotonicity, which strengthens

(Maskin) monotonicity, and they show that it is necessary for the implementation of SCCs in externally stable Core.[14]

DEFINITION **9** (Winner monotonicity (Koray and Yildiz, 2018))**.** An SCC $F : \mathcal{R}Z$ satisfies Winner monotonicity provided that, for all $R, R' \in \mathcal{R}$, if $x \in F(R)$, and $L_i(x, R) \cap F(R') \subseteq L_i(x, R')$ for all $i \in N$, then $x \in F(R')$.

The example below shows the existence of an SCF that is not implementable in externally stable core, though it is implementable in vNM stable set. Another SCF with these features is *minimum distance SCF* for facility location problems, which is introduced in SECTION 3.4.

EXAMPLE **6.** Suppose that $N = \{1, 2\}$, $Z = \{x, y, z\}$, and $\mathcal{R} = \{R, R'\}$. Preferences of agents at different profiles are defined in the left-hand side of FIGURE 3. The SCF f is such that $f(R) = x$ and $f(R') = y$. The SCF is implementable in vNM stable sets and the right-hand side of FIGURE 3 represents an implementing rights structure. However, f violates Winner monotonicity. To see this, suppose that preferences move from R to R'. The premises of Winner monotonicity are satisfied in this case. Indeed, $x = f(R)$, $L_1(x, R) \cap f(R') = \{y\} \subseteq \{x, y, z\} = L_1(x, R')$ and $L_2(x, R) \cap f(R') = \{\varnothing\} \subseteq \{x\} = L_2(x, R')$. However $x \neq f(R')$.

R		R'	
1	*2*	*1*	*2*
x	y	x	y
z	x	y	z
y	z	z	x

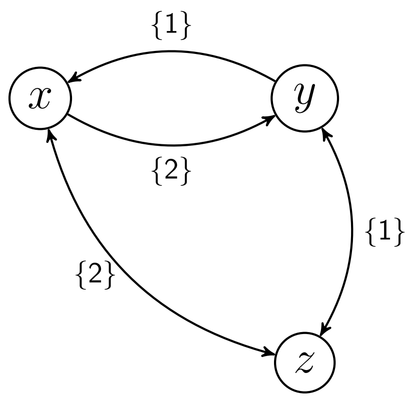

FIGURE 3: An example of a non-Winner Maskin monotonic SCF and an implementing rights structure.

[14]When agents' preferences are described by linear orderings and the domain is full, Winner monotonicity is also a sufficient condition for implementation.

Section 3.4 provides an example of implementable SCF in the realm of bilateral trading. In a previous preliminary draft of the manuscript (Korpela, Lombardi and Saulle, 2023), we identified further SCFs that are implementable in vNM stable sets. For instance, it turns out the minimum distance rule in facility location problems although non monotonic, and thus non-implementable in Core, it is implementable in vNM stable set. An another example of implementable SCF is the Condorcet rule in voting environment when the preference domain is Condorcet. Furthermore, we provide an example of non-implementable SCF. For instance, the Vickrey auction rule is not implementable because it violates vNM monotonicity. We refer the interested reader to (Korpela, Lombardi and Saulle, 2023).

3.4 An Application to Bilateral Trading

A basic model of bilateral trading (Myerson and Satterthwaite, 1983; Chatterjee and Samuelson, 1983) consists of one indivisible object to be traded between agent 1 (the seller) and agent 2 (the buyer). The value of agent i is denoted by v_i. Both values lie in the interval $[a, b]$ and all value profiles $(v_1, v_2) \in [a, b]^2$ are admissible. The set of outcomes Z is the set of all possible trading prices $p \in \{0\} \cup [a, b]$ where 0 means that there is no trade and $p \in [a, b]$ means that agents trade with price p. Agents' utility functions are $u_1(p) = p - v_1$ and $u_2(p) = v_2 - p$.

f maps any profile of valuations (v_1, v_2) to a trading price $p \in [a, b]$, or to 0 if there is no trade. We require f to be individually rational – both agents must benefit from trade when it takes place.

Fix any $p \in [a, b]$. f_p is a fixed-price rule if and only if $f_p(v_1, v_2) = p$ for $v_1 < p < v_2$, and $f_p(v_1, v_2) = 0$, otherwise.[15]

Theorem 8. *Let us consider a bilateral trading environment and fix any $p \in [a, b]$. An SCF f is implementable in vNM stable sets by a rights structure if and only if $f = f_p$.*

The rights structure employed in the proof is such that $S = Z$ and trade occurs

[15] This rule is not efficient. Sometimes trade would be Pareto improving but will not take place at the pre-specified price. This no-trade situation also happens under incomplete information (Myerson and Satterthwaite, 1983).

according to γ only if both parties agree. In such an environment, the vNM stable set outcome equals the one-price Bayesian Nash equilibrium, as defined by Gibbons (1992).

4 Correspondences and Farsighted Rationality

So far, we have focused on implementing SCFs in vNM stable sets. In this section, we study the implementation of SCCs in vNM stable sets, which may consist of multiple outcomes. Thus, the objective of the planner is to ensure that the set of vNM stable outcomes coincides with $F(R)$ for every R.

4.1 Implementation in vNM Stable Sets

This section shows that NSD and an auxiliary condition, INDEPENDENCE OF IRRELEVANT ALTERNATIVES (IIA, henceforth), are sufficient for implementation in vNM stable sets. However, since Harsanyi (1974) criticized the vNM stable set for being myopic, this sufficiency result is an intermediate step for addressing the so-called Harsanyi (1974)'s critique in our framework. Section 4.3 shows that NSD and a strengthening of IIA are sufficient for implementation in strict vNM stable sets as put forward by Harsanyi (1974).

A condition that is easy to check can be stated as follows. Before stating it, it is worth mentioning that this condition is a necessary condition for implementation in some preference domains, such as the universal domain and single-crossing domain.[16]

Definition **10.** (IIA) $F : \mathcal{R} \rightrightarrows Z$ satisfies INDEPENDENCE OF IRRELEVANT ALTERNATIVES (IIA, henceforth) provided that for all $R, R' \in \mathcal{R}$ and all $x, x' \in F(\mathcal{R})$,

$$x \in F(R),\, x' \notin F(R),\ \text{and}\ \varnothing \neq K(R, x, x') \subseteq K(R', x, x') \implies x' \notin F(R').$$

IIA simply requires that if those agents who strictly prefer x to x' at R when x is F-optimal at R but x' is not also strictly prefer x to x' at R', then x' cannot

[16]Details available from authors upon request.

be F-optimal at R'. Note that the condition does not require that x remains F-optimal at R'. IIA recasts the well-known *independence of irrelevant alternatives* condition, introduced by Kenneth Arrow in his seminal paper on Arrow's impossibility theorem (Arrow, 1950). One can show[17] that the fixed price rule in bilateral trading, as defined in SECTION 3.4, satisfies IIA.

THEOREM **9.** *If $F : \mathcal{R} \rightrightarrows Z$ satisfies IIA and NSD w.r.t. $Y = Z$, then it is implementable in vNM stable by a rights structure.*

The next section provides examples of SCCs implementable in vNM stable sets in light of THEOREM 9.

To study implementation in vNM stable sets, one can focus on the implementation in externally stable Core. The reason is that the externally stable Core is the unique vNM stable of any social environment. However, we do not follow this route here because our main objective is to address the Harsanyi (1974)'s critique of myopia, and the externally stable Core is subject to this criticism. However, it may be worth noting that IIA and NSD implies Winner Monotonicity, which is a necessary condition for implementation in externally stable Core.[18] Moreover, when agents' preferences are described by linear orderings and the domain of preferences is full, denoted by $\mathcal{L}^*$, Winner monotonicity fully characterizes the class of SCCs that are implementable in externally stable Core (see SECTION 3.3). Therefore:

COROLLARY **5.** *If $F : \mathcal{L}^* \rightrightarrows Z$ satisfies IIA and NSD w.r.t. $Y = Z$, then it is implementable in externally stable Core by a rights structure.*

4.2 Applications

4.2.1 Strict Majority Rule

Suppose that there are only two outcomes x and x', so that $Z = \{x, x'\}$. Let us also assume that each agent (voter) $i\,(\in N)$'s preferences over Z are represented

[17] Details available from authors upon request.

[18] Details available from authors upon request.

by a transitive and complete preference relation R_i. The preference domain is denoted by $\mathcal{R}$ with R as a typical preference profile. Given R, let $q_R(x, x') = \{i \in N | xP_ix'\}$. Then, $q_R(x, x')$ is the fraction of voters striclty preferring x to x' at R. $F^{SM} : \mathcal{R} \rightrightarrows Z$ is the *strict majority rule* if

$$F^{SM}(R) = \begin{cases} \{x\} & \text{if } q_R(x, x') \geqslant \frac{|N|}{2} + 1 \\ \{x'\} & \text{if } q_R(x', x) \geqslant \frac{|N|}{2} + 1 \\ Z & \text{otherwise.} \end{cases}$$

for all $R \in \mathcal{R}$.

THEOREM **10.** *The strict majority rule $F^{SM} : \mathcal{R} \rightrightarrows Z$ is implementable in vNM stable set by a rights structure.*

4.2.2 Unanimity with Status-Quo

Let $x^* \in Z$ denote the status-quo outcome. Our preference domain places a restriction on agents' preferences over outcomes in Z. Fix an agent $i \in N$. We assume that agent i's preference relation over Z has a form of trichotomy. Formally, $R_i \in \mathcal{R}_i$ is *trichotomous* for agent i if there exists a set of desirable outcomes $D_i(R_i) \subseteq Z$ with $x^* \notin D_i(R_i)$ such that (i) aI_ib for all $a, b \in D_i(R_i)$, (ii) aP_ix^* for all $a \in D_i(R_i)$, (iii) x^*P_ia for all $a \in Z \backslash (D_i(R_i) \cup \{x^*\})$, and (iv) aI_ib for all $a, b \in Z \backslash (D_i(R_i) \cup \{x^*\})$. In other words, R_i is *trichotomous* for i if there exists a set of desirable outcomes $D_i(R_i)$ for i such that R_i ranks every desirable outcome in its top indifference class, the status-quo in the second indifference class, and all remaining outcomes in its third indifference class. Let $\mathcal{R}(x^*)$ denote the set of profiles of *trichotomous preferences where* x^* is the status-quo outcome. $F : \mathcal{R}(x^*) \rightrightarrows Z$ is the unanimity with status-quo rule if there exists an outcome x^* such that for all $R \in \mathcal{R}(x^*)$,

$$F(R) = \begin{cases} \bigcap_{i \in N} D_i(R_i) & \text{if } \bigcap_{i \in N} D_i(R_i) \neq \varnothing \\ \{x^*\} & \text{otherwise.} \end{cases}$$

This SCC is implementable in vNM stable set because it satisfies NSD and IIA.

Theorem **11.** *The unanimity with status-quo rule $F : \mathcal{R}(x^*) \rightrightarrows Z$ is implementable in vNM stable set by a rights structure.*

4.2.3 Collusion-Proof Stable Rule

A *matching problem* is a quadruplet $(M, W, P, \mathcal{M})$ where M and W are non empty sets of men and women respectively; $P \in \mathcal{L}$ is a profile of linear orderings so that (i) every man $m \in M$'s preference relation is represented by a linear ordering P_m over $W \cup \{m\}$ and (ii) every woman $w \in W$'s preference relation is represented by a linear ordering P_w over $M \cup \{w\}$; $\mathcal{M}$ is a collection of all matchings, with μ as a typical element. $\mu : M \cup W \to M \cup W$ is a bijective function, matching every agent $i \in M \cup W$ either to a partner of the opposite sex or with himself/herself. If an agent i is matched with himself/herself, we say that this i is *single* under μ.

We refer to $(M, W, \mathcal{L}, \mathcal{M})$ as a *class of matching problems*, with $(M, W, P, \mathcal{M})$ as a typical matching problem. Note that $M \cup W = N$ and $Z = \mathcal{M}$.

To apply the above partial characterization to matching problems, we extend agent i's linear ordering $P_i \in \mathcal{L}_i$ to the preference ordering $\succsim_i$ on $\mathcal{M}$ as follows: for all $\mu, \mu' \in \mathcal{M}$ and all $P_i \in \mathcal{L}_i$, $\mu \succsim_{P_i} \mu' \Leftrightarrow$ either $\mu(i) P_i \mu'(i)$ or $\mu(i) = \mu'(i)$.

Let $\mathcal{R}$ denote the preference domain over $\mathcal{M}$ derived from $\mathcal{L}$ with $\succsim$ as a typical element.

A matching μ is *blocked by agent i at $\succsim_P \in \mathcal{R}$* if $iP_i\mu(i)$. A matching μ is *blocked by a pair $(m, w) \in M \times W$ at $\succsim^P \in \mathcal{R}$* if $mP_w\mu(w)$ and $wP_m\mu(m)$. A matching μ is *stable at $\succsim \in \mathcal{R}$* if it is not blocked by any agent or any pair of a man and a woman at $\succsim$. Given a matching problem, the *stable solution*, denoted by St, can be defined, for each $\succsim \in \mathcal{R}$, by $St(\succsim) \equiv \{\mu \in \mathcal{M} | \mu \text{ is stable at } \succsim\}$.

Following Kimya (2022b), let us define when a matching μ is a collusion-proof matching. For all $\mu \in \mathcal{M}$ and all $K \in \mathcal{N}_0$, let $\mu(K) = \{i \in N : \mu(j) = i \text{ for some } j \in K\}$.

Definition **11.** A matching μ' can be obtained from a matching μ through collusion by $K \in \mathcal{N}_0$ at $\succsim$ if the following requirements hold: (1) Either $K \subseteq M$ or $K \subseteq W$; (2) $K \subseteq K(\succsim, \mu', \mu)$; (3) $\mu'(i) \in \mu(K)$ for all $i \in K$; (4) $\mu'(i) = \mu(i)$ if $i \notin K$ and $\mu(i) \notin \mu(K)$.

If μ' can be obtained from μ through collusion by some coalition K at $\succsim$, then we say that μ' is obtainable from μ through collusion at $\succsim$. A matching μ is *collusion-proof* at $\succsim$ if there does not exist any matching μ' that is obtainable from μ through collusion at $\succsim$. In other words, a matching μ' can be obtained from a matching μ through collusion at $\succsim$ if a same-sex coalition can reallocate their mates obtained under μ among themselves so that each of its member is strictly better off under μ'.

Given a matching problem, the *collusion-proof solution*, denoted by CP, can be defined, for each $\succsim \in \mathcal{R}$, by $CP(\succsim) \equiv \{\mu \in \mathcal{M} | \mu \text{ is collusion-proof at } \succsim\}$.

A matching μ is a collusion-proof stable matching at $\succsim$ if $\mu \in CP(\succsim) \cap St(\succsim)$. Kimya (2022b, Lemma 2) has shown that if $CP(\succsim)$ is not empty at $\succsim$, then there exists a unique stable matching at $\succsim$. Kimya (2022b, Lemma 5) shows that the preference domain satisfying the so-called top coalition property of Banerjee, Konishi and Sonmez (2001) , denoted by $\mathcal{R}^{TC}$, guarantees the existence of the collusion-proof stable matching—for all $\succsim \in \mathcal{R}^{TC}$, $CP(\succsim) \cap St(\succsim) \neq \varnothing$, which is the top-coalition matching of the matching problem. Given $P \in \mathcal{L}$, its extension $\succsim_P \in \mathcal{R}^{TC}$ satisfies the *top coalition property* if for all $K \in \mathcal{N}_0$, either a) there exists $(m, w) \in M \times W$ with $m, w \in K$ such that $m \in \arg\max_{P_w} K$ and $w \in \arg\max_{P_m} K$, or b) there exists $i \in N$ such that $i \in \arg\max_{P_i} K$. This pair or singleton is called the top coalition in K at $\succsim_P$. The top coalition matching at $\succsim_P$ is $\mu^*(\succsim_P) = \{S_1, S_2, ..., S_n\}$, where S_1 is the top coalition in N at $\succsim_P$, S_2 is the top coalition in $N \backslash S_1$ at $\succsim_P$, and so on.

For the class of matching problems $(M, W, \mathcal{R}^{TC}, \mathcal{M})$, the *collusion-proof stable rule* $f^{CP-St} : \mathcal{R}^{TC} \rightrightarrows \mathcal{M}$ is the *collusion-proof stable rule* if $f^{CP-St}(\succsim) = CP(\succsim) \cap St(\succsim)$ for all $\succsim \in \mathcal{R}^{TC}$.

We show below that for the class $(M, W, \mathcal{R}^{TC}, \mathcal{M})$, the collusion-proof stable rule is implementable in vNM stable sets.

Theorem **12.** *For the class of matching problems $(M, W, \mathcal{R}^{TC}, \mathcal{M})$, the collusion-proof stable rule $f^{CP-St} : \mathcal{R}^{TC} \rightarrow \mathcal{M}$ is implementable in vNM stable set*

4.3 Implementation in Strict vNM Stable Sets

Harsanyi (1974) criticized the vNM stable set for being myopic. He argued that when agents are farsighted, an alternative might be "unstable" even if it belongs to a vNM stable set. The following example reproduces the *Harsanyi's critique* in our environment.

Example **7.** There are two agents $\{1, 2\}$, three outcomes $\{x, y, z\}$ and two preference profiles $R, R' \in \mathcal{R}$. The table in Figure 4 specifies agents' preferences.

R		R'	
1	*2*	*1*	*2*
x	z	y	x, y
y	x	x	z
z	y	z	

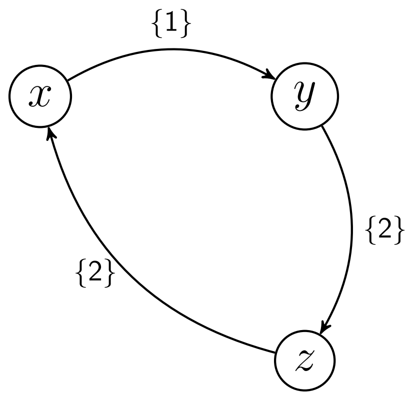

Figure 4: An implementing rights structure suffering the *Harsanyi's critique*.

The SCC $F : \{R, R'\} \to \{x, y, z\}$ is such that $F(R) = \{x, z\}$ and $F(R') = \{y, z\}$. The SCC satisfies IIA and NSD w.r.t. $Y = Z$, and hence according to Theorem 9 is implementable in vNM stable sets. The right-hand side of Figure 4 illustrates the implementing rights structure employed by the planner. Thus, $\{x, z\}$ and $\{y, z\}$ are respectively vNM stables sets at R and R'. Consider the profile R' and note that z is dominated by x. However, in the idea of von Neumann and Morgenstern (1944), a deviation from z to x is not credible since the alternative x is, in turn, dominated by y, which belongs to the vNM stable set. Harsanyi (1974) pointed out that this argument works only when y is not preferred to z by the coalition that moves from z to x, which is not the case. Indeed, if agent 2 is farsighted, then he deviates to x only to reach y and be better off.

Example 7 shows that when the planner ignores the *Harsanyi's critique*, he may fail in his economic deign exercise since the vNM stable set is no longer a valid prediction when agents are farsighted. A natural question then arises: Can the planner design a rights structure such that agents select the socially desirable alternatives, irrespective of whether they are farsighted? In the case of Example 7, the goal can be achieved by giving only to agent 1 the power to move from x to y, and back to x.

To find a general solution to this problem, we borrow from Harsanyi (1974) the notions of *indirect dominance* and *strict vNM stable sets*.

Indirect dominance is a way to incorporate farsightedness in models with binding agreements. Given a social environment $\langle \Gamma, R \rangle$, a state s' indirectly dominates s if there exists a path from s to s' such that every coalition effective on this path prefers the final state of the path s' to the state they replace.

Definition **12** (Indirect Dominance[19])**.** For a given social environment $\langle \Gamma, R \rangle$, a state s is indirectly dominated by s' under γ, denoted by $s' \gg_{(\Gamma,R)} s$, if there are states $s^0, s^1, ..., s^m$ and corresponding coalitions $K^1, ..., K^m$ where $s = s^0$ and $s' = s^m$ such that for all $\ell = 1, ..., m$, (1) $K^\ell \in \gamma(s^{\ell-1}, s^\ell)$, and (2) $h(s')P_{K^\ell}h(s^{\ell-1})$.

Following Harsanyi (1974), we define the so called *strict vNM stable set.*

Definition **13.** Let $\langle \Gamma, R \rangle$ be a social environment. The set $V(\Gamma, R) \subseteq S$ is a strict vNM stable set of Γ at R if it satisfies the following conditions:

Indirect Internal Stability: for all $s, s' \in V(\Gamma, R)$, not $s' \gg_{(\Gamma,R)} s$

External Stability: for all $s \notin V(\Gamma, R)$, there exists $s' \in V(\Gamma, R)$ such that $s' >_{(\Gamma,R)} s$.

Note that a strict stable set is immune to the *Harsanyi's critique*. To define implementation in strict vNM stable set, let us denote by $SvNM(\Gamma, R)$ the union of all strict vNM stable sets at $\langle \Gamma, R \rangle$.

[19] Harsanyi (1974) introduced two notions of indirect dominance. One is based on the idea of a monotone chain: x indirectly dominates y if there is a sequence connecting x and y such that the deviating agents do not only prefer x to the status-quo, but in addition, their deviation must also be preferred to the status-quo. Another definition, later formalized by Chwe (1994), is the one that we employ here: Alternatives along the sequence are not required to directly dominate each other.

DEFINITION 14 (Implementation in strict vNM stable sets). A rights structure Γ implements $F : \mathcal{R} \rightrightarrows Z$ in strict vNM stable sets if $F(R) = h \circ SvNM(\Gamma, R)$ for all $R \in \mathcal{R}$. If such a rights structure exists, $F : \mathcal{R} \rightrightarrows Z$ is implementable in strict vNM stable sets by a rights structure.

Our implementing condition relies on the following notion of *indirect independence of irrelevant alternatives*, which strengthens the notion of IIA.

DEFINITION 15. (SIIA) $F : \mathcal{R} \rightrightarrows Z$ satisfies STRONG INDEPENDENCE OF IRRELEVANT ALTERNATIVES (SIIA, henceforth) provided that for all $R, R' \in \mathcal{R}$ and all $x, x', x'' \in F(\mathcal{R})$ such that $x' \neq x''$,

$$x \in F(R),\ x' \notin F(R) \text{ and } \varnothing \neq K\left(R, x, x'\right) \subseteq K\left(R', x'', x'\right) \implies x' \notin F\left(R'\right).$$

REMARK 4. SIIA implies IIA. It is plain by setting $x = x''$.

SIIA simply states that if those agents who strictly prefer x to x' at R when x is F-optimal at R but x' is not F-optimal at R and they also strictly prefer to x'' to x' at R', then x' cannot be F-optimal R'. The condition does not require that x must be F-optimal at R'. In IIA, the comparison between x and x' is direct, while in SIIA the comparison can happen indirectly through a third outcome x''. Examples[20] of social choice rules satisfying SIIA are the one in EXAMPLE 7, the fixed price rule in a bilateral trading environment, the strict majority rule (SECTION 4.2.1), and the unanimity with status-quo rule (SECTION 4.2.2). The main result of this section can be stated as follows.

THEOREM 13. *If $F : \mathcal{R} \rightrightarrows Z$ satisfies SIIA and No Simultaneous Domination w.r.t. $Y = Z$, then it is implementable in strict vNM stable sets.*

We conclude this section by pointing out that the *Harsanyi's critique* does not exhaust the problem of farsightedness in coalition theory. The notion of indirect dominance provided the input for a plethora of solution concepts, each of them captures different aspects of farsighted rationality. In light of this observation, one can show that the implementing rights structure in the proof of THEOREM 13

[20] Details available from authors upon request.

also implements in the following solutions as put forward by Chwe (1994): the farsighted Core[21], the largest consistent set[22], and the farsighted stable set[23].

Corollary **6.** *If $F : \mathcal{R} \rightrightarrows Z$ satisfies NSD and SIIA, then there exists a rights structure implementing $F : \mathcal{R} \rightrightarrows Z$ in strict vNM stable sets, farsighted Core, largest consistent set and farsighted stable set.*

Corollary 6 is in line with recent contributions studying dominance invariance in coalitional games (Mauleon, Molis, Vannetelbosch and Vergote, 2014; Kimya, 2022a). A social environment satisfies dominance invariance if direct and indirect dominance are equivalent. Kimya (2022a) shows that dominance invariance plays a fundamental role in eliminating differences among various farsighted solutions. Our result shows that SIIA and NSD are sufficient for designing a rights structure that exhibits dominance invariance when it is restricted to the set of F-optimal states. This fact sheds new light on the role played by dominance invariance to harmonize different solutions, myopic and farsighted.

5 Concluding Remarks

Social Choice Correspondences: In this paper, we fully identified the class of SCFs that are implementable in vNM stable sets. This characterization extends immediately to SCCs when the planner views the outcomes selected by the SCC F as equally good (Abreu and Sen, 1991). Under this interpretation that multiple outcomes express the planner's indifference or neutrality, implementing an SCC can be formulated as implementing an SCF. It also extends immediately when a social choice set represents the planner's goal.[24] Indeed, say that a social choice set $F = \{f | f : \mathcal{R} \to Z\}$ is implementable in vNM stable sets if and only if for

[21] The farsighted Core of $\langle \Gamma, R \rangle$ is the set of states $C(\Gamma, R) \subseteq S$ that are indirectly undominated by any states, that is $C(\Gamma, R) = S - Dom_{(\Gamma, R, \gg)}(S)$.

[22] Given $\langle \Gamma, R \rangle$, a set of states $Y \subseteq S$ is consistent if the following statement holds: $s \in Y \iff \forall s' \in S, K \in \gamma(s, s')$ there is an $s'' \in Y$ such that either $(s'' = s')$ or $(s'' \gg_{(\Gamma, R)} s')$ and not $h(s'') P_K h(s)$. The largest consistent set is the maximal consistent set with respect to set inclusion.

[23] The farsighted stable set of $\langle \Gamma, R \rangle$ is the set of states $F(\Gamma, R) \subseteq S$ that are indirectly undominated by any states in $F(\Gamma, R)$, that is $F(\Gamma, R) = S - Dom_{(\Gamma, R, \gg)}(F(\Gamma, R))$.

[24] The concept of social choice set is prevalent in the literature of incomplete information; see, for instance, Jackson (1991).

each $f \in F$, a rights structure implements f in vNM stable sets.[25] Then, in these situations, the necessary and sufficient conditions for SCFs can be directly applied to provide necessary and sufficient conditions for fully implementing both an SCC and a social choice set.

In contrast, the analysis would change significantly if we wanted to derive general necessary and sufficient conditions for SCCs. The crux of the matter is how to build connections between states so that external stability breaks down when we move from R to R' with $F(R) \neq F(R')$. Guidance is given by vNM MONOTONICITY for SCFs. Roughly speaking, this condition suggests that we can allocate power to coalitions of the type $K(R, f(R), z)$, with $z \neq f(R) = x$, in designing the implementing rights structure to guarantee external stability at R. Indeed, in our implementing rights structure, the set $S^x = \{s \in S | h(s) = x\}$ is the unique vNM stable at every profile R'' such that $f(R'') = x$. In each of these profiles, external stability is guaranteed by allocating to coalition $K(R, x, h(s))$ the power to move from s to a state in S^x. Moreover, vNM MONOTONICITY says that S^x cannot be a vNM stable at R' by breaking down its external stability at R' via a preference reversal. Indeed, to make S^x externally unstable at R'', the condition requires the existence of an outcome z such that $K(R'', x, z) \nsubseteq K(R', x, z)$ for all profile R'' such that S^x is a vNM stable set at R''. When the goal of the planner is represented by an SCC, we lose this guidance and it remains unclear how to design a rights structure that breaks down the external stability of the vNM stable set at R when agents' preferences are represented by R'.

Game Forms: The design of a rights structure is more flexible than the design of a game form: It is always possible to represent a game form as a rights structure, but the converse is not always true.[26] However, our full characterization result relies on constructing a rights structure that cannot be represented as a game form. A rights structure to represent a game form must be individually transitive. Individual transitivity requires that when agent i has the power to move from a

[25] In this notion of implementation, the implementing rights structure depends on the function selected from the social choice set. A similar notion of implementation has been proposed by Bergemann, Morris and Tercieux (2011) by using game forms.

[26] See Koray and Yildiz (2018) for a discussion.

state s to another state s' and the power to move from s' to a third state s'', then agent i must have the power to move from s to s'' directly. Thus, any state agent i can obtain via a chain of movements among states should be able to obtain it directly. Our implementing rights structure is not individually transitive. Let us clarify this point.

As discussed in the previous paragraph, when the profile changes from R to R' where $f(R) = x \neq f(R')$, vNM MONOTONICITY allows us to make the set S^x externally unstable at R' via a preference reversal. Specifically, to make S^x externally unstable at R', the condition requires the existence of an outcome z such that $K(R'', x, z) \nsubseteq K(R', x, z)$ for every profile R'' such that S^x is a vNM stable set at R''. This outcome z allows us to define a state s such that its outcome is z and s is not dominated by any state in S^x at R', though it is dominated by a state in S^x when agents' preferences are $R'' \in f^{-1}(x)$. To rule out the possibility that s becomes an unwanted stable state at R', we follow a result due to Richardson (1946, 1953) and Harary et al. (1966), according to which a vNM stable set exists when no odd cycle exists. By using the TEST CYCLE condition, we embed s into an odd cycle of states, $s_1, ..., s_{2k+1}$ with $s = s_j$ for some j, where the same agent i can have a right and an incentive to move between any two consecutive states. Therefore, the TEST CYCLE condition allows us to insert s as a feasible state of our rights structure without generating an unwanted stable state. This, however, comes at the expense of individually intransitivity of our rights structure. It is still an open question on how to devise an individually transitive rights structure implementing in vNM stable set.

Appendix

Proof of Theorem 1 Suppose that $\Gamma = (S, g, \gamma)$ implements F in vNM stable sets. Let $Y = \{h(s) | s \in S\}$. Clearly, $F(\mathcal{R}) \subseteq Y$. Since the proof is an immediate consequence of the external stability of the vNM stable set, we omit it here. ■

The following lemmata has been used in the proofs of Theorem 2 and Theorem 3.

Lemma **1.** *Suppose that f is implementable in vNM stable sets via a rights structure. If $x \in f(\mathcal{R})\backslash f(R')$ for some $R' \in \mathcal{R}$, then for all $x^* \in I^f(x, R')$, $x^* \neq f(R')$.*

Proof of Lemma 1. Fix any $x^* \in I^f(x, R')$. Suppose toward a contradiction that $x^* = f(R')$. Since $x \in f(\mathcal{R})$, it holds that $x \in I^f(x, R')$. Then, by definition of $I^f(x, R')$ we have that $x I'_N x^*$ or, in other terms, $K(R', x^*, x) = \varnothing$. Since f is implementable and since $x \neq f(R') = x^*$, we have that $vNM(\Gamma, R')) = S^{x^*} \neq S^x$. Then, by external stability of S^{x^*}, for any $s \in S^x$ (where $h(s) = x$) and some $s^* \in S^{x^*}$ (where $h(s^*) = x^*$), it must be the case that $K(R', h(s^*), h(s)) \neq \varnothing$, which is a contradiction. ■

Proof of Theorem 2. Suppose that $\varGamma = (S, h, \gamma)$ implements f in vNM stable sets. Let $h(S) = Y \subseteq Z$, where $h(S) = \{h(s) \in Z | s \in S\}$. Recall that, for all $y \in Y$, $S^y = \{s \in S | h(s) = y\}$ denotes the set of states where the outcome is y. Fix any $(x, R') \in Z \times \mathcal{R}$ with $x \in f(\mathcal{R})\backslash f(R')$. Let $S^{I^f(x,R')}$ be defined by $S^{I^f(x,R')} = \{s \in S \mid h(s) \in I^f(x, R')\}$. For all $x^* \in I^f(x, R')$ and all $R \in f^{-1}(x^*)$, since $\varGamma = (S, h, \gamma)$ implements f in vNM stable sets, it follows that $vNM(\varGamma, R) = S^{x^*}$. Moreover, for all $x^* \in I^f(x, R')$, $S^{x^*} \subseteq S^{I^f(x,R')}$. Since $x \in f(\mathcal{R})\backslash f(R')$, Lemma 1 implies that $f(R') \notin I^f(x, R')$. It follows from the implementability of f that $S^{I^f(x,R')}$ is not a vNM stable set at R'. Note that at R', since $S^{I^f(x,R')}$ is internally stable, then $S^{I^f(x,R')}$ must violate external stability. Then, there exists $s \in S\backslash S^{I^f(x,R')}$ such that for all $s' \in S^{I^f(x,R')}$ and all $K \subseteq K(R', h(s'), h(s))$, it holds that $K \notin \gamma(s, s')$. Since for all $x^* \in I^f(x, R')$ and all $R \in f^{-1}(x^*)$, $vNM(\varGamma, R) = S^{x^*} \subseteq S^{I^f(x,R')}$, and since for all $x^* \in I^f(x, R')$ and all $R \in f^{-1}(x^*)$, there exists $\bar{s} \in S^{x^*}$ and a coalition K such that $K \in \gamma(s, \bar{s})$, it follows that $K(R, x^*, h(s)) \nsubseteq K(R', x^*, h(s))$ for all $x^* \in I^f(x, R')$ and all $R \in f^{-1}(x^*)$. Thus, f satisfies vNM monotonicity.
■

The following lemma will be used in the proof of Theorem 3.

Lemma **2.** *(Richardson, 1953) If a vNM stable set does not exist, then there is an odd cycle.*

Proof of Theorem 3. Suppose that Γ implements f in vNM stable sets. Fix any $R' \in \mathcal{R}$ and any $x \in f(\mathcal{R}) \backslash f(R')$. Suppose that $xP'_i f(R')$ for some $i \in N$. Since $x \in f(\mathcal{R})$, it holds that $x \in I^f(x, R')$.

Let $S^{I^f(x,R')} = \left\{s \in S | h(s) \in I^f(x, R')\right\}$. Since $f(R') \neq x$, LEMMA 1 implies that $vNM(\Gamma, R') = S^{f(R')} \neq S^{I^f(x,R')}$. Moreover, let

$$S' = \left\{s' \in S \backslash S^{I^f(x,R')} | s \not>_{(\Gamma,R')} s' \text{ for all } s \in S^{I^f(x,R')}\right\}$$

Since $vNM(\Gamma, R') \neq S^{I^f(x,R')}$, and since $S^{I^f(x,R')}$ is internally stable at R', it must be that $S^{I^f(x,R')}$ violates external stability at R', and so $S' \neq \emptyset$. By construction, $h(S') \subseteq \mathcal{M}^f(x, R')$ and $S^{I^f(x,R')} \cup S'$ is externally stable at R'. However, since by implementability of f, $vNM(\Gamma, R') \neq S^{I^f(x,R')} \cup S'$, it follows that $S^{I^f(x,R')} \cup S'$ is not internally stable at R'.

Suppose that there exists $s' \in S'$ such that $s' >_{(\Gamma,R)} s$ for some $s \in S^{I^f(x,R')}$. Then, there exists $K \in \gamma(s, s')$ such that $h(s') P'_K h(s)$. Since $s \in S^{I^f(x,R')}$ and preferences are transitive, it holds that $h(s') P'_K x^*$ for all $x^* \in I^f(x, R')$. Fix any $l \in K$, so that $h(s') P'_l x^*$ for all $x^* \in I^f(x, R')$. Let us proceed according to whether $f(R') = h(s')$ or not.

Suppose that $f(R') = h(s')$. Since $s' \in S'$ and since $h(s') \in \mathcal{M}^f(x, R')$, it follows that $f(R') \in \mathcal{M}^f(x, R')$. This shows that part (iii) of the TEST CYCLE condition is satisfied.

Suppose that $f(R') \neq h(s')$. Since f satisfies NSD, there exists $j \in N$ such that $f(R') P'_j h(s')$. Since, by our initial assumption, there exists an agent $i \in N$ such that $xP'_i f(R')$ and agent i's preferences are transitive, it follows that $x^* P'_i f(R')$ for all $x^* \in I^f(x, R')$. Since $f(R') P'_j h(s')$ and since $h(s') P'_l x^*$ for all $x^* \in I^f(x, R')$, we have that for all $x^* \in I^f(x, R')$, $x^* P'_i f(R') P'_j h(s') P'_l x^*$ some $i, j, l \in N$. This shows that part (i) of the TEST CYCLE condition is satisfied.

Otherwise, suppose that there does not exist any $s' \in S'$ such that $s' >_{(\Gamma,R)} s$ for some $s \in S^{I^f(x,R')}$. Then it has to be that S' is not internally stable at R'. Hence, by definition of S', $S^{I^f(x,R')} \cup S'$ is not internally stable at R' because S' is not internally stable at R'. Given a rights structure Γ, a restriction of Γ to $S' \subseteq S$, denoted by $\Gamma_{|S'} = \left(S', h_{|S'}, \gamma_{|S'}\right)$, is a rights structure such that for all $s \in S'$, $h_{|S'}(s) = h(s)$, and for all $s, s' \in S'$, $\gamma_{|S'}(s, s) = \gamma(s, s')$. Suppose that $vNM\left(\Gamma_{|S'}, R'\right) \neq \varnothing$. Then, by construction, $S^{I^f(x,R')} \cup vNM\left(\Gamma_{|S'}, R'\right) = vNM(\Gamma, R')$, which is a contradiction. Then it must be that $vNM\left(\Gamma_{|S'}, R'\right) = \varnothing$. LEMMA 2 implies that there exists a sequence of states $(s^1, ..., s^k)$ in S' yielding an odd cycle at R'. Since $h(S')$ is contained in $\mathcal{M}^f(x, R')$, this shows that part (ii) of the TEST CYCLE condition is satisfied. ■

PROOF OF THEOREM 4. Let us construct a rights structure that implements f under the given conditions. We will denote outcome z in condition (i) of TEST CYCLE by $z(x, R)$, and outcome z^h in condition (ii) of TEST CYCLE by $z^h(x, R)$. Thus, whenever we speak of $z(x, R)$, we mean that for the pair (x, R) it is condition (i) of TEST CYCLE that is satisfied. Furthermore, we will denote the agent who prefer $z^k(x, R)$ to $z^{k+1}(x, R)$ at R in condition (ii) by $j(x, R, k, k+1)$ modulo k.

Let f satisfy conditions (i)-(iii) with respect to $Y \subseteq Z$ such that $f(\mathcal{R}) \subseteq Y$. In what follows, we construct an implementing Γ. Let $\bar{S}$ be defined by

$$\bar{S} = \bigcup_{R \in \mathcal{R}} \bigcup_{x \in f(\mathcal{R})} \left\{ \left(y, I^f(x, R)\right) | y \in I^f(x, R) \text{ and } f(R) \neq x \right\} \cup Gr(f)$$

Furthermore, fix any $x \in f(\mathcal{R})$ and any $R \in \mathcal{R}$ such that $f(R) \neq x$. If either (i), (ii), or (iii) holds, then we say that there exists a test cycle for (x, R).

Suppose that there exists a test cycle for (x, R). Let us define the following sets of states according to whether condition (i), condition (ii), or condition (iii) applies:

$$S((x, R), \mathrm{i}) = \left\{ (u, (x, R), \mathrm{i}) \;\middle|\; \begin{array}{l} \text{the test cycle for } (x, R) \\ \text{satisfies condition (i) and} \\ u \in \{f(R), z(x, R)\} \cup I^f(x, R) \end{array} \right\}$$

$$S((x, R), \mathrm{ii}) = \left\{ \left(z^h(x, R), (x, R), \mathrm{ii}\right) \;\middle|\; \begin{array}{l} \text{the test cycle for } (x, R) \\ \text{satisfies condition (ii) and} \\ h = 1, ..., k \end{array} \right\}$$

$$S\left((x,R),\text{iii}\right)=\left\{\left(f(R),(x,R),\text{iii}\right)\;\middle|\;\begin{array}{l}\text{the test cycle for } (x,R)\\ \text{satisfies condition (iii).}\end{array}\right\}$$

Let us define the set of states S by

$$S=\bar{S}\cup\left\{\bigcup_{R\in\mathcal{R}}\bigcup_{x\in f(\mathcal{R})\backslash\{f(R)\}}\left(S\left((x,R),\text{i}\right)\cup S\left((x,R),\text{ii}\right)\cup S\left((x,R),\text{iii}\right)\right)\right\}.$$

Then, for all $s\in S$, let us defined the outcome function h by $h(s)=s_1$, where s_1 is the outcome of the first entry of the tuple s. Finally, let γ be defined, for all $s,s'\in S$ and all $i\in N$, by the following rules.

RULE 1: If $s,s'\in\bar{S}$, then:

(a) if $s,s'\in Gr(f)$, then $\{i\}\in\gamma(s,s')$.

(b) if $s\in\bar{S}\backslash Gr(f)$ and $s'\in Gr(f)$, then $\{i\}\in\gamma(s,s')$.

RULE 2: If $s,s'\in S((x,R),\text{i})$, then:

(a) if $s=(f(R),(x,R),\text{i})$ and $s'=(x^*,(x,R),\text{i})$, then $\{i\}\in\gamma(s,s')$.

(b) if $s=(x^*,(x,R),\text{i})$ and $s'=(z(x,R),(x,R),\text{i})$, then $\{i\}\in\gamma(s,s')$.

(c) if $s=(z(x,R),(x,R),\text{i})$ and $s'=(f(R),(x,R),\text{i})$, then $\{i\}\in\gamma(s,s')$.

RULE 3: If $s\in S((x,R),\text{i})$ and $s'=(R',y)\in Gr(f)$, then:

(a) if $s=(x^*,(x,R),\text{i})$, then $K(R',y,x^*)\in\gamma(s,s')$.

(b) if $s=(z(x,R),(x,R),\text{i})$, then $K(R',y,z(x,R))\in\gamma(s,s')$.

RULE 4: If $s=(f(R),(x,R),\text{i})\in S((x,R),\text{i})$, $s'=(R',y)\in Gr(f)$ and $y\notin I^f(x,R)$, then $K(R',y,f(R))\in\gamma(s,s')$.

RULE 5: If $s,s'\in S((x,R),\text{ii})$, $s=\left(z^{h+1}(x,R),(x,R),\text{ii}\right)$ and $s'=\left(z^h(x,R),(x,R),\text{ii}\right)$ for some $h=1,...,k$ and $z^h(x,R)\,P_i z^{h+1}(x,R)$, then $\{i\}\in\gamma(s,s')$, where $z^{k+1}(x,R)=z^1(x,R)$.

RULE 6: If $s \in S((x, R), \text{ii})$ and $s' = (R', y) \in Gr(f)$, then $K(R', y, h(s)) \in \gamma(s, s')$.

RULE 7: If $s \in S((x, R), \text{iii})$ and $s' = (R', y) \in Gr(f)$, then $K(R', y, h(s)) \in \gamma(s, s')$.

RULE 8: If $s \in \{(y, I^f(x, R)) \,|\, y \in I^f(x, R) \text{ and } f(R) \neq x\}$ and $s' \in S((x, R), \text{iii})$, then $\{i\} \in \gamma(s, s')$.

RULE 9: Otherwise, $\gamma(s, s') = \varnothing$.

By construction, Γ is a rights structure. Let us show that Γ implements f in single-payoff vNM stable sets. To this end, suppose that R is the true preference profile, and let $f(R) = x$. We show that $S^x \equiv \{s \in S \mid h(s) = x\}$ is the unique vNM stable set of (Γ, R).

Clearly, S^x satisfies internal stability. Then, let us show that S^x satisfies external stability.

To this end, note that NSD implies that for all $z \in Y$, xP_iz for some $i \in N$. Thus, by construction of Γ, (x, R) dominates all states in $\bar{S} \backslash S^x$ by **RULE 1**, all states in $S((y, R'), \text{ii}) \backslash S^x$ by RULE 6, and all states in $S((y, R'), \text{iii}) \backslash S^x$ by **RULE 7**. The set $S(R', y, i)$ needs a more careful examination.

Suppose that $S((y, R'), \text{i}) \neq \varnothing$. We proceed according to whether $f(R) \neq y$ or not.

Suppose that $f(R) \neq y$. By NSD, we have that $f(R)P_iy$ for some $i \in N$. **RULE 3** implies that (x, R) dominates all states $s \in S((y, R'), \text{i}) \backslash S^x$ such that $h(s) \in \{z(R', y)\} \cup I(y, R')$. Suppose that $f(R) = f(R')$. Then, $(f(R'), (y, R'), \text{i}) \in S^x$. Suppose that $f(R) \neq f(R')$. Suppose that $f(R) \subseteq I(y, R')$. Then, (x, R) dominates $(f(R'), (y, R'), \text{i})$ via **RULE 2**. Thus, let $f(R) \neq f(R')$ and $f(R) \cap I(y, R') = \varnothing$. Since $K(R, f(R), f(R')) \neq \varnothing$ and since $f(R) \cap I(y, R') = \varnothing$, we have that (x, R) dominates $(f(R'), (u, R'), \text{i})$ via **RULE 4**.

Suppose that $f(R) = y$. Then, $f(R) \neq f(R')$, $f(R) \cap I(y, R') \neq \varnothing$ and $(f(R), (y, R'), \text{i}) \in S^x$. By NSD, we have that $f(R)P_iw$ for some $i \in N$ if $w \neq f(R)$. Since for all $s \in S((y, R'), \text{i})$ such that $h(s) = f(R)$, it holds that $s \in S^x$, we need to focus only on the cases that both $z(R', y) \neq y$ and $y^* \neq y$. Since $K(R, f(R), z(R')) \neq \varnothing$ and since $K(R, f(R), y^*)) \neq \varnothing$, it follows that (x, R)

dominates any state $s \in S((y, R'), \mathrm{i})$ such that either $h(s) = z(R', y)$ or $h(s) = y^*$ via **RULE 3**. Thus, we are left to show that $(f(R'), (y, R'), \mathrm{i}))$ is dominated by a state in S^x. To this end, note that $f(R) \neq f(R')$, and so NSD implies that $f(R) P_i f(R')$ for some $i \in N$. Since $(f(R), (y, R'), \mathrm{i}) \in S^x$ and since $f(R) P_i f(R')$ for some $i \in N$, it follows from **RULE 2**(a) that agent i has the power and incentive to move from $(f(R'), (y, R'), \mathrm{i})$ to $(f(R), (y, R'), \mathrm{i})$. Thus, a state in S^x dominates $(f(R'), (y, R'), \mathrm{i})$.

We conclude that S^x is externally stable, and so S^x is a vNM stable set of (Γ, R). Next, we show that this is the only stable set at R. Assume, to the contrary, that there exists a nonempty set $S^* \subseteq S$ that is a vNM stable set of (Γ, R) such that $S^x \neq S^*$. Note that at least one state of $\bar{S}$ must be in S^* by external stability. The reason is that the rights structure Γ does not allow any move from states inside the $Gr(f)$ to states outside of $\bar{S}$. Moreover, **RULE 1** implies that if $s \in S^* \cap Gr(f)$ and $h(s) = z$, then $\big\{(z, R') \mid R' \in \mathcal{R},\ z = f(R')\big\} \subseteq S^*$. Given that S^* is externally stable and since $S^* \cap \bar{S} \neq \varnothing$, it follows from Rule 1 that $s \in S^* \cap Gr(f)$. Fix any $s \in S^* \cap Gr(f)$. We proceed according to whether $h(s) = x$ or not.

Suppose that $h(s) = x$. Thus, $(x, R) \in S^*$, and so $S^x \subseteq S^*$. Since we have already shown that S^x is a vNM stable set of (Γ, R) and since S^* is a vNM stable set of (Γ, R), it follows that $S^x = S^*$, yielding a contradiction.

Suppose that $h(s) = y \neq x$. Since S^* is internally stable and since f is vNM efficient, it follows from **RULE 1** that $\{(z, R') \big| R' \in \mathcal{R},\ f(R') = z \in I^f(y, R)\} = S^* \cap Gr(f)$. We proceed according to whether $f(R) R_N y$ or not.

Suppose that $f(R) R_N y$. Since f is vNM efficient, there exists $i \in N$ such that $f(R) P_i y$. Since agent i has the power to move from s to (x, R) via **RULE 1**(a) and since S^* is internally stable, it follows that $(x, R) \notin S^*$. Since $f(R) R_N y$, it follows that no agent has incentive to move from (x, R) to any state $\bar{s} \in S^* \cap Gr(f)$, though they have the power to do so via **RULE 1**(a). Therefore, S^* is not externally stable, which is a contradiction.

Suppose that $y P_i f(R)$ for some agent $i \in N$. Since $y \in f(\mathcal{R}) \backslash f(R)$, since

$yP_if(R)$ for some agent $i \in N$ and since, moreover, f satisfies the TEST CYCLE property, it follows that a test cycle for (y, R) exists. There are three cases to be considered according to whether the test cycle for (y, R) is given either by condition (i), or by (ii), or by (iii).

Case 1: The test cycle is given by condition (i). Then, for some $i, j, k \in K$, it holds that $y^* P_i f(R) P_j z(y, R) P_k y^*$ for some $z(y, R) \in \mathcal{M}^f(y, R)$ and all $y^* \in I^f(y, R)$. Moreover, vNM MONOTONICITY implies that for all $y^* \in I^f(y, R)$ and all $R'' \in f^{-1}(y^*)$, it holds that $K(R'', y^*, z(y, R)) \subsetneq K(R, y^*, z(y, R))$. Suppose that $(z(y, R), (y^*, R), \mathrm{i}) \notin S^*$ for some $y^* \in I^f(y, R)$. Since S^* satisfies external stability, it follows that there exists $K \in \gamma((z(y, R), (y^*, R), \mathrm{i}), t)$ for some $t \in S^*$. By construction of Γ, since K can move only to a state in $S^* \cap Gr(f)$ via **RULE 3**, we have that $t = (R'', z) \in S^*$ for some $z \in Y$ and $R'' \in f^{-1}(z)$ and $K = K(R'', z, z(y, R))$.

Since $S^* \cap Gr(f) = \{(R', z) \,|\, f(R') = z \in I^f(y, R)\}$, it follows that $t = (R'', z)$ is such that $z \in I^f(y, R)$. Since $z \in I^f(y, R)$ and since, for all $y^* \in I^f(y, R)$ and all $R'' \in f^{-1}(y^*)$, $K(R'', y^*, z(y, R)) \subsetneq K(R, y^*, z(y, R))$, it follows that S^* violates external stability at R, which is a contradiction. Therefore, it must be the case that $(z(y, R), (y^*, R), \mathrm{i}) \in S^*$ for all $y^* \in I^f(y, R)$.

Suppose that $(y^*, (y, R), \mathrm{i}) \notin S^*$ for some $y^* \in I^f(y, R)$. Again, since S^* satisfies external stability, there exists a coalition K such that $K \in \gamma((y^*, (y, R), \mathrm{i}), t)$ for some $t \in S^*$. Since $y^* \in I^f(y, R)$, it follows from Γ that K can move only to a state in $S^* \cap Gr(f)$ via **RULE 3**. This implies that $t = (R'', z) \in S^*$ for some $z \in Y$ and $R'' \in f^{-1}(z)$ and that $K = K(R'', z, y^*)$. Again, since $S^* \cap Gr(f) = \{(R', z) \,|\, f(R') = z \in I^f(y, R)\}$, it follows that $t = (R'', z)$ is such that $z \in I^f(y, R)$. Since $z \in I^f(y, R)$, we have that the state $t = (R'', z) \in S^*$ cannot dominate at R the state $(y^*, (y, R), \mathrm{i})$, in violation of the external stability of S^*. We conclude that $(y^*, (y, R), \mathrm{i}) \in S^*$ for all $y^* \in I^f(y, R)$. Fix any $y^* \in I^f(y, R)$. Then, $(y^*, (y, R), \mathrm{i}) \in S^*$ and $(z(y, R), (y^*, R), \mathrm{i}) \in S^*$. Since, by condition (i) of TEST CYCLE, there exists $k \in N$ such that $z(y, R) P_k y^*$ and since **RULE 2**(b) implies that $k \in \gamma((y^*, (y, R), \mathrm{i}), (z(y, R), (y^*, R), \mathrm{i}))$, it follows that

S^* violates internal stability at R, which is a contradiction.

Case 2: The test cycle is given by condition (ii). The states that are designed as a test cycle for (y, R) are $(z^1(y, R), (y, R), \text{ii}), (z^2(y, R), (y, R), \text{ii}), \ldots, (z^k(y, R), (y, R), \text{ii})$.

Note that, by construction, if $(z^h(y, R), (y, R), \text{ii}) \notin S^*$, then we can move only to states of the type $(z, R') \in S^*$ with $z \in I^f(y, R)$. Fix any $h = 1, ..., k$. Suppose that $(z^h(y, R), (y, R), \text{ii}) \notin S^*$. Then, vNM MONOTONICITY implies that $K(R'', y^*, z^h(y, R)) \nsubseteq K(R, y^*, z^h(y, R)$ for all $y^* \in I^f(y, R)$ and all $R'' \in f^{-1}(y^*)$. This implies that S^* is not externally stable, which is a contradiction. Therefore, it must be the case that $(z^1(y, R), (y, R), \text{ii}), (z^2(y, R), (y, R), \text{ii}), \ldots, (z^k(y, R), (y, R), \text{ii}) \in S^*$. Since condition (ii) of the TEST CYCLE implies that there is a cycle at R of odd length among the outcomes $z^1(y, R), ...z^k(y, R)$, it follows from **RULE 5** that S^* is not internally stable, which is a contradiction.

Case 3: The test cycle is given by condition (iii). Then, $f(R) \in \mathcal{M}^f(y, R')$. By definition of the rights structure Γ, only states in $\bar{S}$ can dominate the state $(f(R), (y, R), \text{iii})$ (via **RULE 7**). Since $f(R) \in \mathcal{M}^f(y, R')$, no state in $S^* \cap \bar{S}$ dominates the state $(f(R), (y, R), \text{iii})$ by vNM MONOTONICITY; the reason is that $f(R) \in \mathcal{M}^f(y, R')$, and so $K(R'', y^*, f(R)) \nsubseteq K(R, y^*, f(R))$ for all $y^* \in I^f(y, R)$ and all $R'' \in f^{-1}(y^*)$. Thus, it must be the case that $(f(R), (y, R), \text{iii}) \in S^*$. Since $s \in S^* \cap Gr(f)$ and $h(s) = y$, it follows that $(y, I^f(y, R)) \in S^*$. Since $(y, I^f(y, R)) \in S^*$ and since, by NSD, there exists an agent i such that $f(R) P_i y$, it follows that the internal stability of S^* is violated because agent i has the incentive and the power (via **RULE 8**) to move from $(y, I^f(y, R))$ to $(f(R), (y, R), \text{iii})$.

Since the choice of state $s \in S^* \cap Gr(f)$ was arbitrary, we conclude that S^* is not a vNM stable set of (Γ, R), which is a contradiction. ■

Proof of Theorem 5. Suppose preferences are continuous, money monotone, and that the preference domain $\mathcal{R}$ is finite. Suppose that f satisfies NSD and vNM MONOTONICITY with respect to Y. We show that f satisfies TEST CYCLE.

Fix any $R' \in \mathcal{R}$ and any $x \in Y$. Suppose that $x \in f(\mathcal{R}) \backslash f(R')$ and that $x P'_i f(R')$ for some $i \in N$. Since f satisfies vNM MONOTONICITY, it follows that there exists a $z \in Y$ such that for all $R \in f^{-1}(x)$ and for all $x^* \in I^f(x, R')$,

$K(R, x^*, z) \nsubseteq K(R', x^*, z)$. Thus, $z \in \mathcal{M}^f(x, R')$. We proceed according to whether $z = f(R')$ or not.

Suppose that $z = f(R')$. Then, the requirement (iii) of the TEST CYCLE property is satisfied.

Suppose that $z \neq f(R')$. Let $z = (d, t)$. Since agents' preferences are continuous, it follows that there exists $\hat{\varepsilon} > 0$ such that for all $R \in f^{-1}(x)$, all $x^* \in I^f(x, R')$ and all $i \in K(R, x^*, z) \backslash K(R', x^*, z)$, it holds that $x_i^* P_i (d, t_i + \hat{\varepsilon})$. Moreover, since preferences are money monotonic and transitive, we have that for all $R \in f^{-1}(x)$, all $x^* \in I^f(x, R')$ and all $i \in K(R, x^*, z) \backslash K(R', x^*, z)$, $(d, t_i + \hat{\varepsilon}) P_i' x_i^*$. Therefore, for all $R \in f^{-1}(x)$, all $x^* \in I^f(x, R')$ and all $i \in K(R, x^*, z) \backslash K(R', x^*, z)$, it holds that $x_i^* P_i (d, t_i + \hat{\varepsilon})$ and $(d, t_i + \hat{\varepsilon}) P_i' x_i^*$.

Since $\mathcal{R}$ is finite and since, moreover, f is vNM efficient and agents' preferences are continuous, it follows that there exists $\varepsilon' > 0$ such that for all $\bar{R} \in \mathcal{R}$ and all $i \in N$ such that $f_i(\bar{R}) \bar{P}_i z_i$, it holds that $f_i(\bar{R}) \bar{P}_i (d, t_i + \varepsilon')$.

Let $\varepsilon = \frac{\min\{\hat{\varepsilon}, \varepsilon'\}}{2}$. By construction, we have that:

1. For all $R \in f^{-1}(x)$, all $x^* \in I^f(x, R')$ and all $i \in K(R, x^*, z) \backslash K(R', x^*, z)$, it holds that $x_i^* P_i (d, t_i + \varepsilon)$ and $(d, t_i + \varepsilon) P_i' x_i^*$.

2. For all $\bar{R} \in \mathcal{R}$ and all $i \in N$ such that $f_i(\bar{R}) \bar{P}_i (d, t_i)$, it holds that $f_i(\bar{R}) \bar{P}_i (d, t_i + \varepsilon)$.

Let us define z' by

$$z_i' = \begin{cases} (d, t_i + \varepsilon_i) & \text{if } i \in \cup_{R \in f^{-1}(x)} \cup_{x^* \in I^f(x,R')} K(R, x^*, z) \\ z_i & \text{otherwise.} \end{cases}$$

By construction of z', we have that if $i \in \cup_{R \in f^{-1}(x)} \cup_{x^* \in I^f(x,R')} K(R, x^*, z)$, then $i \in K(R, x^*, z') \cap K(R', z', x^*)$ and that for all $\bar{R} \in \mathcal{R}$ and all $i \in N$ such that $f_i(\bar{R}) \bar{P}_i z_i$, it holds that $f_i(\bar{R}) \bar{P}_i z_i'$. Moreover, by construction, we also have that f satisfies NSD and vNM MONOTONICITY with respect to $Y \cup \{z'\}$, that $z' \in \mathcal{M}^f(x, R')$ and that $z' \neq f(R')$.

Since $z' \neq f(R')$ and since f is vNM efficient with respect to $Y \cup \{z'\}$, it follows that $f(R') P_k' z'$ for some $k \in N$. Since, by our initial supposition,

$xP'_i f(R')$ for some $i \in N$, and since R'_i is transitive, we have that for all $x^* \in I^f(x, R')$, $x^* P'_i f(R')$ for some $i \in N$. Thus, we have that for all $x^* \in I^f(x, R')$, $x^* P'_i f(R') P'_k z'$ for some $i, k \in N$, with $z' \in \mathcal{M}^f(x, R')$. Fix any $j \in K(R, x^*, z) \backslash K(R', x^*, z)$ for some $R \in f^{-1}(x)$ and some $x^* \in I^f(x, R')$. Then, by construction, $j \in K(R, x^*, z') \cap K(R', z', x^*)$, and so $z' P'_j x^*$. Therefore, we have that $x^* P'_i f(R') P'_k z' P'_j x^*$ for some $i, j, k \in N$, with $z' \in \mathcal{M}^f(x, R')$. Since the previous argument holds for all $x^* \in I^f(x, R')$, we have that there exists $z' \in \mathcal{M}^f(x, R')$ such that for all $x^* \in I^f(x, R')$, there exists $i, j, k \in N$ such that $x^* P'_i f(R') P'_k z' P'_j x^*$. Thus, f satisfies requirement (i) of TEST CYCLE.

Since the above arguments hold for any $(R', x) \in \mathcal{R} \times F(\mathcal{R})$ such that $f(R') \neq \{x\}$ and $xP'_i f(R')$ for some $i \in N$, it follows that we can construct a set Y', with $Y \subseteq Y'$, such that f satisfies NSD and vNM MONOTONICITY with respect to Y', and so f satisfies TEST CYCLE with respect to Y'. ■

PROOF OF THEOREM 6. Fix any $R' \in \mathcal{L}$ and suppose that $x \in f(\mathcal{L}) \backslash f(R')$ and that $xP'_i f(R')$ for some $i \in N$. Since preferences are linear orders, $I^f(x, R') = \{x\}$. By vNM MONOTONICITY, for all $R \in f^{-1}(x)$, we have that $K(R, x, z) \nsubseteq K(R', x, z)$ for some outcome $z \in \mathcal{M}^f(x, R')$. If $z = f(R')$, then requirement (iii) of the TEST CYCLE condition is satisfied. In what follows, let $z \neq f(R')$. Since $K(R, x, z)$ is non empty, take any $j \in K(R, x, z) \backslash K(R', x, z)$. Then, $xP_j z$ and $zR'_j x$. Since $xP_j z$, it follows that $x \neq z$. Since R'_j is a linear order, it follows that $zP'_j x$. Since, by our initial supposition, $xP'_i f(R')$ for some $i \in N$, we have that $zP'_j xP'_i f(R')$ for some $i \in N$, and some $j \in K(R, x, z) \backslash K(R', x, z)$. Since $f(R') \neq z$ and since f satisfies NSD, it follows that there exists $k \in N$ such that $f(R') P'_k z$. We have established that $zP'_j xP'_i f(R') P'_k z$ for some $i, j, k \in N$ and some $z \in \mathcal{M}^f(x, R')$. Thus, f satisfies requirement (i) of TEST CYCLE. ■

PROOF OF THEOREM 7. Suppose that $f : \mathcal{L} \to Z$ satisfies NSD and vNM MONOTONICITY w.r.t $Y \subseteq Z$. First, we show that f satisfies monotonicity w.r.t. Y. Take any $R, R' \in \mathcal{R}$, denote $f(R) = x$, and assume that $L_i(x, R) \subseteq L_i(x, R')$ holds for all $i \in N$. We need to show that x is selected at R' to verify the claim. Suppose, toward a contradiction, that $x \neq f(R')$. Since $x = f(R)$ and $x \neq f(R')$, then by vNM MONOTONICITY there exists a $z \in Y$ such that $K(R, x, z) \nsubseteq K(R', x, z)$. Thus, there exists an agent $j \in N$ who experiences a preference reversal when preferences move from R to R', that is xR_jz and zR'_jx. Therefore, $L_j(x, R) \nsubseteq L_j(x, R)$, which contradicts the premises. We conclude that f must satisfy monotonicity w.r.t. Y.

Next, we show that f satisfies NSD w.r.t. Y. Take any $x \in Y$ and let assume that for any $R \in \mathcal{L}$, it holds that xR_iy for all $i \in N$ and all $y \in Y$. Suppose toward a contradiction that $x \neq f(R)$ for some $R \in \mathcal{L}$. Let $y = f(R) \in Y$. Then, by NSD of f there exists an agent $j \in N$ such that yR_jx, a contradiction. We conclude that that f must satisfy unanimity w.r.t. Y. ■

PROOF OF THEOREM 8. First, we show that if f is implementable in vNM stable sets, then f is a fixed price rule. For any $v_1 \in [a, b]$, let f be an SCF implementable in vNM stable sets. Suppose toward a contradiction that f is not a fixed price rule, that is, for some v_2 and v'_2 with $v_2 > v'_2$, f is such that $f(v_1, v_2) \neq f(v_1, v'_2) \equiv p' > 0$. Take any $z \in Z$. Note that if trading with price p' is more profitable than z to a buyer of type v'_2, then it is more profitable to a buyer of type v_2 too. Since the argument holds for any $z \in Z$, we have that $K((v_1, v'_2), p', z) \subseteq K((v_1, v_2), p', z)$ holds for all $z \in Z$ which contradicts vNM MONOTONICITY of f. Therefore, if f is implementable in vNM stable sets, then for any $v_1 \in [a, b]$, the function $f_{v_1} \equiv f(v_1, \cdot)$ must be a fixed price rule, conditionally on v_1. To complete the proof, it remains to show that f is a fixed price rule unconditionally on v_1.

Fix any $v_1 \in [a, b)$,[27] such that the price p of the fixed price rule $f(v_1, x)$ satisfies $b > p > 0$. Notice that by individual rationality, this implies $p > v_1$. If v_1 does not exist, then f must be the zero price rule f_0, a particular case of a fixed price rule. To show that f is the fixed price rule f_p (unconditionally on v_1), we must verify that $f_{v'_1}$ is the fixed price rule with a price p for any $v'_1 < p$, and the zero price rule for any $v'_1 \geqslant p$. We study the two cases separately.

Suppose towards a contradiction that $v'_1 < p$ but $f_{v'_1}$ is not the fixed price rule

[27]The rule f_b is equivalent to the zero price rule since trade never takes place.

with a price p. Take any value of the buyer $v_2 \in [p, b)$. Recall that f is implementable in vNM; hence it must satisfy vNM MONOTONICITY by THEOREM 2. However, it is straightforward to see that $K((v_1, v_2), p, z) \subseteq K((v_1', v_2), p, z)$ holds for all $z \in Z$- a contradiction. Hence $f_{v_1'}$ is indeed the fixed price rule with a price p for any $v_1' < p$ as claimed.

Next, suppose towards a contradiction that $v_1' \geqslant p$, but $f_{v_1'}$ is not the zero price rule. One can easily see that this case follows directly from the previous case. Let $p' > 0$ be the fixed price of the rule $f_{v_1'}$. By individual rationality $p' > v_1' \geqslant p$. However, by the previous argumentation, f_{v_1} must be a fixed price rule with a price p' too - a contradiction.

Finally, it is easy to see that any fixed price rule f_p is implementable in vNM stable sets. A simple rights structure (code of rights) $\mathit{\Gamma} = (S, \gamma)$, where $S = \{0, p\}$, $\gamma(0, p) = \{\{1, 2\}\}$ (trade must be accepted by both), and $\gamma(p, 0) = \{\{1\}, \{2\}\}$ (trade can be rejected by either), implements it. ■

PROOF OF THEOREM 9. Suppose that F satisfies IIA and NSD w.r.t. $Y = Z$. Let $\Gamma = (S, h, \gamma)$ be defined as follows. Let

$$S = \{(x, R) \in Z \times \mathcal{R} | x \in F(R)\} = Gr(F).$$

Let $h : S \to Z$ be defined by $h(x, R) = x$ for all $(x, R) \in S$. Finally, let $\gamma : S \times S \rightrightarrows \mathcal{N}$ be defined by the following two rules. For all $(x, R), (x', R') \in S$,

1. if $x' \notin F(R)$, then $K(R, x, x') \in \gamma((x', R'), (x, R))$;
2. otherwise, $\gamma((x', R'), (x, R)) = \varnothing$.

By construction, we have that $\gamma((x', R), (x, R)) = \varnothing$ for all $x, x' \in F(R)$, and so the states (x', R) and (x, R) are not connected.

Fix any $R \in \mathcal{R}$. Let us show that $F(R) = h \circ vNM(\Gamma, R)$. To this end, let

$$S(F(R)) = \{s \in S | h(s) \in F(R)\}. \tag{1}$$

Let us first show that $F(R) \subseteq h \circ vNM(\Gamma, R)$. To this end, it suffices to show that $S(F(R))$ is a vNM stable set of (Γ, R); that is, $S(F(R))$ is externally and internally stable at (Γ, R).

$S(F(R))$ *is externally stable.* To see it, take any $(x', R') \in S$ such that $(x', R') \notin S(F(R))$. Since F satisfies No Simultaneous Domination, it follows that there exist $x \in F(R)$ and $i \in N$ such that xP_ix'. Since xP_ix', it holds that $K(R, x, x') \neq \varnothing$. Since $(x, R) \in S(F(R))$ and $(x', R') \in S$, and since $x' \notin F(R)$, it also follows that $K(R, x, x') \in \gamma((x', R'), (x, R))$, and so $(x, R) >_{(\Gamma,R)} (x', R')$. Since the choice of (x', R') was arbitrary, it follows that $S(F(R))$ is externally stable.

$S(F(R))$ *is internally stable.* Assume, to the contrary, that it is not internally stable; that is, there are $(x', R'), (x'', R'') \in S(F(R))$ such that $(x'', R'') >_{(\Gamma,R)} (x', R')$. This implies that there exists $K \in \mathcal{N}_0$ such that $K \in \gamma((x', R'), (x'', R''))$ and $K \subseteq K(R, x'', x')$. By definition of γ, it follows that $K = K(R'', x'', x')$ and $x' \notin F(R'')$. Since $(x'', R'') \in S$, it holds that $x'' \in F(R'')$. Since $(x'', R'') >_{(\Gamma,R)} (x', R')$, it also holds that $K(R'', x'', x') \subseteq K(R, x'', x')$. IIA implies that $x' \notin F(R)$, which is a contradiction. Thus, $S(F(R))$ is internally stable.

Finally, let us show that $h \circ vNM(\Gamma, R) \subseteq F(R)$. To this end, take any $\bar{S} \subseteq S$ and suppose that it is a vNM stable set of (Γ, R). To show that $S(F(R)) = \bar{S}$, we proceed in two steps.

Step 1: $S(F(R)) \subseteq \bar{S}$. Take any $(x', R') \in S(F(R))$. Assume, to the contrary, that $(x', R') \notin \bar{S}$. Since $\bar{S}$ is externally stable, it follows that exists $(x'', R'') \in \bar{S}$ such that $(x'', R'') >_{(\Gamma,R)} (x', R')$. This implies that there exists $K \in \mathcal{N}_0$ such that $K \in \gamma((x', R'), (x'', R''))$ and $K \subseteq K(R, x'', x')$. By definition of γ, it follows that $K = K(R'', x'', x')$ and $x' \notin F(R'')$. Since $(x'', R'') \in S$, it holds that $x'' \in F(R'')$. Since $K \subseteq K(R, x'', x')$, it also holds that $K(R'', x'', x') \subseteq K(R, x'', x')$. IIA implies that $x' \notin F(R)$, and so $(x', R') \notin S(F(R))$, which is a contradiction.

Step 2: $\bar{S} \subseteq S(F(R))$. Assume, to the contrary, that $\bar{S} \nsubseteq S(F(R))$. Step 1 implies that $S(F(R)) \subseteq \bar{S}$. Then, $S(F(R))$ is a proper subset of $\bar{S}$; that is, there exists $(x', R') \in \bar{S}$ such that $(x', R') \notin S(F(R))$. Since $S(F(R))$ is a vNM stable set of (Γ, R) and since $(x', R') \notin S(F(R))$, the external stability of $S(F(R))$ implies that there exists $(x'', R'') \in S(F(R))$ such that $(x'', R'') >_{(\Gamma,R)} (x', R')$.

Then, $\bar{S}$ is not internally stable, which is a contradiction.

Since the choice of $\bar{S} \subseteq S$ was arbitrary, we conclude that any vNM stable set $\bar{S}$ of (Γ, R) coincides with $S(F(R))$. In other words, $S(F(R))$ is the unique vNM stable set of (Γ, R). Since $S(F(R)) = vNM(\Gamma, R)$, it follows from definition of $S(F(R))$ that $h \circ vNM(\Gamma, R) \subseteq F(R)$. ■

Proof of Theorem 10. It suffices to show that the strict majority rule F^{SM} satisfies IIA and NSD under the specification that $Y = Z$. Since it is plain that F^{SM} satisfies NSD, we only show that it satisfies IIA. To this end, take any $R, R' \in \mathcal{R}$. Suppose that $x \in F^{SM}(R)$ and $x' \notin F^{SM}(R)$. Then, it must the case that $q_R(x, x') \geqslant \frac{|N|}{2} + 1$. Moreover, suppose that $K(R, x, x') \subseteq K(R', x, x')$. Since $K(R, x, x') \subseteq K(R', x, x')$, it follows that $q_{R'}(x, x') \geqslant \frac{|N|}{2} + 1$. It follows from definition of F^{SM} that $F^{SM}(R') = \{x\}$, and so $x' \notin F^{SM}(R')$. Thus, F^{SM} satisfies IIA. Lemma 9 implies that F^{SM} is implementable in vNM stable set by a rights structure. ■

Proof of Theorem 11. It suffices to show that the unanimity with status-quo rule satisfies IIA and NSD under the specification that $Y = Z$. To see that it satisfies NSD, fix any $R \in \mathcal{R}(x^*)$ and suppose that $x' \in Z \backslash F(R)$. Then, $x' \notin \bigcap_{i \in N} D_i(R_i)$. We proceed according to whether $\bigcap_{i \in N} D_i(R_i) \neq \varnothing$ or not.

Suppose that $\bigcap_{i \in N} D_i(R_i) \neq \varnothing$. Since $x' \notin \bigcap_{i \in N} D_i(R_i)$ and $\bigcap_{i \in N} D_i(R_i) \neq \varnothing$, there exist $x \in \bigcap_{i \in N} D_i(R_i)$ and $i \in N$ such that $x P_i x'$.

Suppose that $\bigcap_{i \in N} D_i(R_i) = \varnothing$. Then, $F(R) = \{x^*\}$. Since $\bigcap_{i \in N} D_i(R_i) = \varnothing$, there exists $i \in N$ such that $x' \notin D_i(R_i)$. Since $x' \neq x^*$, it follows that x' belongs to the third indifference class for i at R_i, and so $x^* P_i x'$.

Let us show that F satisfies IIA. To this end, fix any $R, R' \in \mathcal{R}(x^*)$ such that $x \in F(R)$, $x' \notin F(R)$ and $K(R, x, x') \subseteq K(R', x, x')$. Since $x' \notin F(R)$ and $x \in F(R)$, it follows that $K(R, x, x') \neq \varnothing$. We proceed according to whether $\bigcap_{i \in N} D_i(R_i) \neq \varnothing$ or not.

Suppose that $\bigcap_{i \in N} D_i(R_i) \neq \varnothing$. Then, $x \neq x^*$. Fix any $i \in K(R, x, x')$. Then, $x P_i x'$, and so $x P'_i x'$. It follows that $x' \notin D_i(R'_i)$. Clearly, $x' \notin F(R')$ if $\bigcap_{i \in N} D_i(R'_i) \neq \varnothing$. Otherwise, suppose that $\bigcap_{i \in N} D_i(R'_i) = \varnothing$, so that $F(R') =$

$\{x^*\}$. If $x' \neq x^*$, then $x' \notin F(R')$. Suppose that $x' = x^*$. Since $\bigcap_{i \in N} D_i(R_i) \neq \varnothing$ and x^* belongs to the middle indifference class, it follows that $N = K(R, x, x^*) \subseteq K(R, x, x^*)$, and so $x \in \bigcap_{i \in N} D_i(R'_i)$, which is a contradiction.

Suppose that $\bigcap_{i \in N} D_i(R_i) = \varnothing$. Then, $F(R) = \{x^*\}$. Fix any $i \in K(R, x^*, x')$, so that $x^* P_i x'$ and so $x^* P'_i x'$. It follows that $x' \notin D_i(R'_i)$. Since $x' \neq x^*$, it follows that $x' \notin F(R')$. Thus, F satisfied both NSD and IIA. ■

Proof of Theorem 12. Let $(M, W, \mathcal{R}^{TC}, \mathcal{M})$ be given. We show that the collusion-proof stable rule $f^{CP-St} : \mathcal{R}^{TC} \rightarrow \mathcal{M}$ satisfied NSD and IIA. Let us first show that it satisfies NSD. Fix any $\succsim \in \mathcal{R}^{TC}$ and any $\mu, \mu' \in \mathcal{M}$. Suppose that $f^{CP-St}(\succsim) = \mu$ and that $\mu' \neq f^{CP-St}(\succsim) = \mu$. Assume, to the contrary, that $\mu' \succsim_j \mu$ for all $j \in N$. Since $\mu' \neq f^{CP-St}(\succsim) = \mu$, and so $\mu'(i) \neq \mu(i)$ for some $i \in N$, it follows that $\mu' \succ_i \mu$ for some $i \in N$. Fix any $i \in N$ such that $\mu' \succ_i \mu$. Since μ is stable *at* $\succsim$, it cannot be blocked by i at $\succsim$, and so $\mu'(i) \neq i$. Since the pair $(i, \mu'(i))$ cannot block μ at $\succsim$ and since $\mu' \succ_i \mu$ and $\mu' \succsim_{\mu'(i)} \mu$, it follows that $\mu'(\mu'(i)) = \mu(\mu'(i))$. This implies that $i = \mu'(\mu'(i)) = \mu(\mu'(i))$, and so $\mu'(i) = \mu(i)$, which is a contradiction. Thus, the collusion-proof stable rule f^{CP-St} satisfies NSD.

Let us show that f^{CP-St} satisfies IIA. Fix any $\succsim, \succsim' \in \mathcal{R}^{TC}$ and any $\mu, \mu' \in \mathcal{M}$. Suppose that $\mu' \neq f^{CP-St}(\succsim) = \mu$ and that $K(\succsim, \mu, \mu') \subseteq K(\succsim', \mu, \mu')$. We show that $\mu' \neq f^{CP-St}(\succsim')$. Since f^{CP-St} satisfies NSD, it follows that $K(\succsim, \mu, \mu') \neq \varnothing$. By definition of f^{CP-St}, it follows that $\mu \in CP(\succsim) \cap St(\succsim)$. We proceed according to whether $\mu(i) \in K(\succsim, \mu, \mu')$ for some $i \in K(\succsim, \mu, \mu')$ or not.

Case 1: $\mu(i) \in K(\succsim, \mu, \mu')$ for some $i \in K(\succsim, \mu, \mu')$. Suppose $\mu(i) \in K(\succsim, \mu, \mu')$ for some $i \in K(\succsim, \mu, \mu')$. We proceed according to whether $i = \mu(i)$ or not. Suppose that $i = \mu(i)$. Since $i \in K(\succsim', \mu, \mu')$, it holds that $\mu \succ'_i \mu'$. Then, i blocks μ' at $\succsim'$. Suppose that $i \neq \mu(i)$. Since $i, \mu(i) \in K(\succsim', \mu, \mu')$, it follows that the pair $(i, \mu(i))$ blocks μ' at $\succsim'$. In both cases, $\mu' \notin St(\succsim')$, and so $\mu' \neq f^{CP-St}(\succsim')$.

Case 2: $\mu(i) \notin K(\succsim, \mu, \mu')$ for all $i \in K(\succsim, \mu, \mu')$. The proof of this case relies on the following result. Without loss of generality, let us assume that $m_1 \in K(\succsim, \mu, \mu')$—exactly the same reasoning holds if we assume that $w_1 \in K(\succsim, \mu, \mu')$.

Claim 1. *For all $m_i \in M$, if $m_i \in K(\succsim, \mu, \mu')$, then $\mu(m_i) = w_i$, $\mu'(\mu(m_i)) \in K(\succsim, \mu, \mu')$ and $\mu'(\mu(m_i)) \in M \setminus \{m_i\}$.*

Proof. Since $m_i \in K(\succsim, \mu, \mu')$, it follows from our initial supposition that $\mu(m_i) \notin K(\succsim, \mu, \mu')$. So, $m_i \neq \mu(m_i) = w_i$ and $\mu' \succsim_{w_i} \mu$. Let us show that $\mu'(w_i) \in K(\succsim, \mu, \mu')$. Assume, to the contrary, that $\mu'(w_i) \notin K(\succsim, \mu, \mu')$, and so $\mu' \succsim_{\mu'(w_i)} \mu$. We proceed according to whether $\mu'(w_i) = w_i$ or not.

Suppose that $\mu'(w_i) = w_i$. Since $\mu(w_i) \neq w_i$ and $\mu' \succsim_{w_i} \mu$, it follows that $\mu' \succ_{w_i} \mu$. Thus, w_i blocks μ at $\succsim$.

Suppose that $\mu'(w_i) \neq w_i$. Since $\mu'(w_i) \notin K(\succsim, \mu, \mu')$ and $m_i = \mu(w_i) \in K(\succsim, \mu, \mu')$, it holds that $\mu'(w_i) \neq \mu(w_i)$. Since $\mu'(w_i) \neq \mu(w_i)$ and $\mu' \succsim_{w_i} \mu$, it follows that $\mu' \succ_{w_i} \mu$. Moreover, since $\mu'(w_i) \neq m_i$ and $\mu(m_i) = w_i$, it holds that $w_i = \mu'(\mu'(w_i)) \neq \mu(\mu'(w_i))$. Since $\mu' \succsim_{\mu'(w_i)} \mu$ and $\mu'(\mu'(w_i)) \neq \mu(\mu'(w_i))$, it follows that $\mu' \succ_{\mu'(w_i)} \mu$. Thus, the pair $(w_i, \mu'(w_i))$ blocks μ at $\succsim$.

In both cases, $\mu \notin St(\succsim)$, which is a contradiction. Thus, $\mu'(w_i) \in K(\succsim, \mu, \mu')$.

Finally, let us show that $\mu'(w_i) \neq m_i$. Since $w_i \notin K(\succsim, \mu, \mu')$ and $\mu'(w_i) \in K(\succsim, \mu, \mu')$, it holds that $\mu'(w_i) \neq w_i$, and so $\mu'(w_i) \in M$. Suppose that $\mu'(w_i) = m_i$. Then, $\mu'(w_i) = m_i = \mu(w_i)$, and so $\mu'(m_i) = \mu(m_i)$. However, since $m_i \in K(\succsim, \mu, \mu')$, it follows that $\mu'(m_i) \neq \mu(m_i)$, which is a contradiction. Thus, $\mu'(w_i) \in K(\succsim, \mu, \mu')$ and $\mu'(w_i) \in M \setminus \{m_i\}$. □

Suppose that $m_1 \in K(\succsim, \mu, \mu')$. Then, by our initial supposition, $\mu(m_1) \notin K(\succsim, \mu, \mu')$. Claim 1 implies that $\mu(m_1) = w_1$, $\mu'(w_1) \in K(\succsim, \mu, \mu')$ and $\mu'(w_1) = m_2 \in M \setminus \{m_1\}$. Then, by our initial supposition, $\mu(m_2) \notin K(\succsim, \mu, \mu')$, and Claim 1 implies that $\mu(m_2) = w_2$, $\mu'(w_2) \in K(\succsim, \mu, \mu')$ and $\mu'(w_2) \in M \setminus \{m_2\}$. We proceed according to whether $\mu'(w_2) = m_1$ or not.

Suppose that $\mu'(w_2) = m_1$. Then, $\mu(m_1) = w_1 \neq \mu'(m_1) = w_2$ and $\mu(m_2) = w_2 \neq \mu'(m_2) = w_1$. Let $K = \{m_1, m_2\}$, and so $\mu'(K) = \{w_1, w_2\}$. Since $K \subseteq K(\succsim, \mu, \mu')$ and since $K(\succsim, \mu, \mu') \subseteq K(\succsim', \mu, \mu')$, we have that $K \subseteq K(\succsim', \mu, \mu')$.

Let us define μ'' by:

$$\begin{aligned} \mu''(m_i) &= \mu(m_i) = w_i \text{ and } \mu''(w_i) = m_i \text{ for all } i = 1,2, \\ \mu''(i) &= \mu'(i) \text{ for all } i \in N \backslash (K \cup \mu'(K)). \end{aligned}$$

It can be checked that parts 1)-4) of Definition 11 are satisfied, and so μ'' is obtainable from μ' through collusion at $\succsim$. This implies that $\mu' \notin CP(\succsim)$, and so $\mu' \notin f^{CP-St}(\succsim)$. Thus, suppose that $\mu'(w_2) = m_3 \in M \backslash \{m_1, m_2\}$. Since the cardinality of M is finite, by repeating the above reasoning, we can see that there exists $K = \{m_1, ..., m_\ell\} \subseteq K(\succsim, \mu, \mu')$ such that

$$\begin{aligned} \mu(m_i) &= w_i \notin K(\succsim, \mu, \mu') \text{ for all } i = 1, .., \ell, \\ \mu'(w_i) &\in K(\succsim, \mu, \mu') \text{ and } \mu'(w_i) = m_{i+1} \text{ for } i = 1, ..., \ell - 1, \\ \mu'(w_\ell) &\in K(\succsim, \mu, \mu') \text{ and } \mu'(w_\ell) = m_1. \end{aligned}$$

Moreover, it also holds that $\mu(m_1) = w_1 \neq \mu'(m_1) = w_\ell$ and $\mu(m_i) = w_i \neq \mu'(m_i) = w_{i-1}$ for $i = 2, .., \ell$. Since $K \subseteq K(\succsim, \mu, \mu')$ and since $K(\succsim, \mu, \mu') \subseteq K(\succsim', \mu, \mu')$, we have that $K \subseteq K(\succsim', \mu, \mu')$. Let us define μ'' by:

$$\begin{aligned} \mu''(m_i) &= \mu(m_i) = w_i \text{ and } \mu''(w_i) = m_i \text{ for all } i = 1, ..., \ell, \\ \mu''(i) &= \mu'(i) \text{ for all } i \in N \backslash (K \cup \mu'(K)). \end{aligned}$$

It can be checked that parts 1)-4) of Definition 11 are satisfied, and so μ'' is obtainable from μ' through collusion at $\succsim$. This implies that $\mu' \notin CP(\succsim)$, and so $\mu' \notin f^{CP-St}(\succsim)$. Thus, f^{CP-St} satisfies IIA. ■

Proof of Theorem 13. Let the premises hold. Let $\Gamma = (S, h, \gamma)$ be defined as in the proof of Theorem 9. Fix any $R \in \mathcal{R}$. Let us show that $F(R) = h \circ SvNM(\Gamma, R)$. Recall the definition of $S(F(R))$ provided in (1). Also, recall that in the proof of Theorem 9, we showed that $S(F(R))$ is the unique vNM stable set of (Γ, R); hence, $S(F(R))$ satisfies External Stability. It remains to show that $S(F(R))$ satisfies Indirect Internal Stability. Its proof is based on the following claim.

Claim 2. *For all $s' \in S(F(R))$ and all $s'' \in S$, not $s'' \gg_{(\Gamma,R)} s'$.*

Proof of Claim 2. Fix any $(x', R') \in S(F(R))$ and any $(x'', R'') \in S$. Assume, to the contrary, that $(x'', R'') \gg_{(\Gamma,R)} (x', R')$. Let us proceed according to whether $(x'', R'') >_{(\Gamma,R)} (x', R')$ or not.

Case 1: $(x'', R'') >_{(\Gamma,R)} (x', R')$. Since $S(F(R))$ is a vNM stable set, it must be the case that $(x'', R'') \notin S(F(R))$. Since $(x'', R'') >_{(\Gamma,R)} (x', R')$, it follows that there exists $K \in \mathcal{N}_0$ such that $K \in \gamma((x', R'), (x'', R''))$ and $K \subseteq K(R, x'', x')$. By definition of Γ, Rule 1 implies that $K = K(R'', x'', x') \in \gamma((x', R'), (x'', R''))$ and $x' \notin F(R'')$. Since $x'' \in F(R'')$, $x' \notin F(R'')$ and $K(R'', x'', x') \subseteq K(R, x'', x')$, IIA implies that $x' \notin F(R)$, and so $(x', R') \notin S(F(R))$, which is a contradiction.

Case 2: *not* $(x'', R'') >_{(\Gamma,R)} (x', R')$. Since $(x'', R'') \gg_{(\Gamma,R)} (x', R')$ but *not* $(x'', R'') >_{(\Gamma,R)} (x', R')$, it follows that there exist $s_0, s_1, ..., s_n \in S$ and $K_1, ..., K_n \in \mathcal{N}_0$, with $s_0 = (x', R')$ and $s_n = (x'', R'')$, such that $K_j \in \gamma(s_{j-1}, s_j)$ and $h(s_n) P_{K_j} h(s_{j-1})$ for all $j = 1, ..., n$. Let $s_1 = (x''', R''')$. Rule 1 implies that $K_1 = K(R''', x''', x') \in \gamma((x', R'), (x''', R'''))$ and $x' \notin F(R''')$. Since $x'' P_{K_1} x'$, it follows that $K_1 = K(R''', x''', x') \subseteq K(R, x'', x')$. Since $x', x'', x''' \in F(\mathcal{R})$ and $x''' \neq x'$, since $x''' \in F(R''')$, $x' \notin F(R''')$ and since $K(R''', x''', x') \subseteq K(R, x'', x')$, SIIA implies that $x' \notin F(R)$. It follows that $(x', R') \notin S(F(R))$, which is a contradiction. □

By Claim 2 it descends that $S(F(R))$ satisfies the property of Indirect Internal Stability. Therefore, $S(F(R)) = SvNM(\Gamma, R)$. ■

Proof of Corollary 6. Let the premises hold. Let $\Gamma = (S, h, \gamma)$ be defined as in the proof of Theorem 9. Fix any $R \in \mathcal{R}$. Recall that by Theorem 9, the rights structure under consideration is implementing in vNM stable set.

Next, recall that, by Claim 2, $S(F(R))$ is the set of states that are not indirectly dominated by any state in S, i.e., $S(F(R)) \equiv S - Dom_{(\Gamma,R,\gg)}(S)$, where $Dom_{(\Gamma,R,\gg)}(A) \equiv \{s \in S | \exists s' \in A : s' \gg_{(\Gamma,R)} s\}$ for all $A \subseteq S$. Therefore, $S(F(R))$ is the farsighted Core of (Γ, R) which is unique by definition. It follows that Γ implements F in farsighted Core.

We now show that $S(F(R))$ is the largest consistent set (LCS) of $\langle \Gamma, R \rangle$. We proceed in two steps.

Step 1: $S(F(R))$ is a consistent set of (Γ, R). Take any $(x', R') \in S(F(R))$ and any $(x'', R'') \in S$. Suppose that there exists $K \in \mathcal{N}_0$ such that $K \in \gamma((x', R'), (x'', R''))$. Then, by definition of Rule 1, it must be the case that $K = K(R'', x'', x')$ and $x' \notin F(R'')$. We proceed according to whether $(x'', R'') \in S(F(R))$ or not.

Suppose that $(x'', R'') \notin S(F(R))$. Since $S(F(R))$ is externally stable, it follows that there exists $(x''', R''') \in S(F(R))$ such that $(x''', R''') >_{(\Gamma,R)} (x'', R'')$. The proof is complete if it holds that *not* $x''' P_K x'$. Assume, to the contrary, that $x''' P_K x'$. It follows that $(x''', R''') \gg_{(\Gamma,R)} (x', R')$. Since $(x', R') \in S(F(R))$, $(x''', R''') \gg_{(\Gamma,R)} (x', R')$ contradicts Claim 2.

Suppose that $(x'', R'') \in S(F(R))$. Then, we are left to show that *not* $x'' P_{K(R'',x'',x')} x'$. Assume, to the contrary, that $x'' P_{K(R'',x'',x')} x'$. It follows that $K = K(R'', x'', x') \in \gamma((x', R'), (x'', R''))$ and $K(R'', x'', x') \subseteq K(R, x'', x')$. This implies that $(x'', R'') >_{(\Gamma,R)} (x', R')$, which contradicts the fact that $S(F(R))$ is internally stable at (Γ, R).

Step 2: $S(F(R)) = LCS(\Gamma, R)$. Assume, to the contrary, that $S(F(R)) \neq LCS(\Gamma, R)$. Since $S(F(R))$ is a consistent set of (Γ, R) by step 1, it follows that there exists a consistent set $\bar{S}$ of (Γ, R) such that $S(F(R))$ is a proper set of $\bar{S}$. Then, there exists $(x', R') \in \bar{S}$ such that $(x', R') \notin S(F(R))$. Since $(x', R') \notin S(F(R))$, it follows that $x' \notin F(R)$. No Simultaneous Domination implies that there exist $x \in F(R)$ and $i \in N$ such that $x P_i x'$. It follows that $i \in K(R, x, x')$. Moreover, by definition of $S(F(R))$, $(x, R) \in S(F(R))$. Since $x' \notin F(R)$, it follows that $K(R, x, x') \in \gamma((x', R'), (x, R))$, by Rule 1.Then, we have that $(x', R') \in \bar{S}$, $(x, R) \in S(F(R))$ and $K(R, x, x') \in \gamma((x', R'), (x, R))$. Since $\bar{S}$ is a consistent set of (Γ, R), there exists $(x'', R'') \in \bar{S}$ such that [either $(x, R) = (x'', R'')$ or $(x'', R'') \gg_{(\Gamma,R)} (x, R)$] and *not* $x'' P_{K(R,x,x')} x'$. Since $(x, R) \in S(F(R))$, Claim 2 implies *not* $(x'', R'') \gg_{(\Gamma,R)} (x, R)$. Then, it must be the case that $(x, R) = (x'', R'')$. Since $K(R, x, x') \neq \varnothing$, it also follows that $x P_{K(R,x,x')} x'$, which is a contradiction. Therefore, Γ implements F in farsighted Core.

Finally, let us show that Γ implements F in farsighted stable set. Recall that the farsighted stable set extends the vNM to the indirect dominance relation. Hence, a farsighted stable set at (Γ, R), denoted by $F(\Gamma, R)$, is defined as the set of states that are not indirectly dominated by any states in the farsighted stable set, i.e., $F(\Gamma, R) \equiv F(\Gamma, R) - Dom_{(\Gamma,R,\gg)}(F(\Gamma, R))$. Let $FF(\Gamma, R)$ be the union of all farsighted stable set at $\langle\Gamma, R\rangle$. It is well known (e.g. Chwe, 1994, Proposition 3) that if a farsighted stable set exists, then it is weakly contained in the LCS (Chwe, 1994) which is just proved to be equal to $S(F(R))$. Then, $FF(\Gamma, R) \subseteq S(F(R))$. To prove the claim it remains to show that $S(F(R)) \subseteq FF(\Gamma, R)$ which must hold since any farsighted stable set at $F(\Gamma, R)$ cannot be a strict subset of $S(F(R))$, otherwise the indirect external stability of the farsighted stable set would be violated. Then $FF(\Gamma, R) = S(F(R))$ and Γ, implements F in farsighted stable set. ■

References

Abreu D, Sen, A. (1991), *Virtual implementation in Nash equilibrium*, Econometrica 59, 997–1021; 29

Arrow, K.J. (1950), *A Difficulty in the Concept of Social Welfare*, Journal of Political Economy, 58, 328–346; 22

Aumann, R. J. (1987), *What is game theory trying to accomplish?*, in Frontiers of Economics, ed. by K. J. Arrow, and S. Honkapohja, Blackwell, Oxford; 1

Banerjee, S., Konishi, H., Sönmez, T. (2001), *Core in a simple coalition formation game*, Social Choice and Welfare, 18(1), 135-153; 25

Béal, S., Durieu, J., Solal, P. (2008), *Farsighted coalitional stability in TU-games*, Mathematical Social Sciences, 56, 303-313; 9

Bergemann, D., Morris, S., and Tercieux, O. (2011), *Rationalizable Implementation*, Journal of Economic Theory, 146, 1253–1274. 30

Bloch, F., van den Nouweland, A. (2020), *Farsighted stability with heterogeneous expectations*, Games and Economic Behavior, 121, 32-54; 9

Chatterjee, K., and W. Samuelson (1983), *Bargaining under Incomplete Information*, Operations Research, 31, 5, 835–51; 20

Chwe, M. Suk-Young (1994), *Farsighted Coalitional Stability*, Journal of Economic Theory, 63, 299–325; 5, 6, 9, 27, 29, 51

Deng, X., and C. H. Papadimitriou (1994), *On the complexity of cooperative solution concepts*, Mathematics of Operations Research, 19, 257-266; 1

Dutta, B., Vohra, R. (2017), *Rational expectations and farsighted stability*, Theoretical Economics, 12, 1191-122; 9

Einy, E., Shitovitz, B. (2003), *Symmetric von Neumann–Morgenstern stable sets in pure exchange economies*, Games and Economic Behavior, 43, 1, 28-43; 2

Ehlers, L. (2007), *Von Neumann–Morgenstern stable sets in matching problems*, Journal of Economic Theory, 134, 1, 537-547; 2

Ehlers, L., Morrill, T. (2020), *(Il)legal Assignments in School Choice*, The Review of Economic Studies, 87, 4, 1837–1875; 2

Jackson, M.O. (1991), *Bayesian Implementation*, Econometrica, 59, 461–477. 29

Karos, D., Robles, L. (2021), *Full farsighted rationality*, Games and Economic Behavior, 130, 409-424; 9

Kimya, M. (2022), *Coalition Formation Under Dominance Invariance*, Dynamic Games and Applications; 29

Kimya, M. (2022b), *Farsighted Objections and Maximality in One-to-one Matching Problems*, Journal of Economic Theory, 204, 105499; 24, 25

Kimya, M. (2023), *Axiomatic Approach to Farsighted Coalition Formation*, mimeo; 3

Koray S, Yildiz K. (2018), *Implementation via a rights structures*, Journal of Economic Theory, 176, 479–502; 3, 4, 9, 11, 16, 18, 19, 30

Korpela V, Lombardi M, Vartiainen H. (2020). *Do Coalitions Matter in Designing Institutions?*, Journal of Economic Theory, 185; 4, 8, 9, 11, 16, 17, 18

Korpela, V., Lombardi, M., Vartiainen, H. (2021), *Implementation in largest consistent set via a rights structures*, Games and Economic Behavior, 128, 202-212; 4

Korpela, V., Lombardi, M. and Saulle, R. (2023), *Implementation in vNM stable sets*, available at SSRN. 20

Gibbons, R. (1992), *Game Theory for Applied Economists*, Princeton University Press; 21

Gillies, D.B. (1959), *Solutions to general non-zero-sum games*, in: A.W. Tucker, R.D. Luce (Eds.), Contributions to the Theory of Games IV, in: Ann. of Math. Stud., 40, 47–85; 1

Grabisch, M. and P. Sudhölter (2021), *Characterization of TU games with stable core by nested balancedness*, Mathematical Programming; 1

Hararay, F., Norman, R.Z. and Cartwright, D.(1966), *Structural Models: An Introduction to the Theory of Directed Graphs*; 14, 31

Harsanyi, J. (1974), *An Equilibrium-Point Interpretation of Stable Sets and a Proposed Alternative Definition*, Management Science, 20, 11, 1472-1495; i, 5, 9, 21, 22, 26, 27

Herings, J.J., Mauleon, A., Vannetelbosch, V. (2017), *Stable sets in matching problems with coalitional sovereignty and path dominance*, Journal of Mathematical Economics, 71, 14-19; 2

Herings, J.J., Mauleon, A., Vannetelbosch, V. (2020), Matching with myopic and farsighted players, Journal of Economic Theory, 190; 9

Lucas, W. (1968), *A game with no solution*, Bull. Amer. Math. Soc. 74 237–239; 1

Lucas, W. (1992), *Von Neumann-Morgenstern stable sets*, in Handbook of Game Theory with Economic Applications, Vol. 1, Chapter 17, 543-590; 1

Mauleon, A., Vannetelbosch, V.J. and Vergote, W. (2011), *von Neumann–Morgenstern farsightedly stable sets in two-sided matching*, Theoretical Economics, 6, 499-52; 9

Mauleon, A., Molis, E., Vannetelbosch, V. J., Vergote, W. (2014), *Dominance invariant one-to-one matching problems*, International Journal of Game Theory, 43(4), 925-943; 29

Myerson R., Satterthwaite M. (1983), *Efficient Mechanisms for Bilateral Trading*, Journal of Economic Theory, 29,2, 265-281; 20

Mount, K., and S. Reiter (1974), *The informational size of message spaces*, Journal of Economic Theory, 8, 161-192; 7

Núñez, M. and Rafels, C. (2013), *Von Neumann–Morgenstern solutions in the assignment market*, Journal of Economic Theory, 148(3), 1282-1291; 2

Ray, D., Vohra, R. (2015), *The Farsighted Stable Set*, Econometrica, 83,3, 977-1011; 9

Ray, D., Vohra, R. (2019), *Maximality in the Farsighted Stable Set*, Econometrica, 87,5, 1763-1779; 1, 9

Richardson, M. (1946), *On weakly ordered systems*, Bulletin of the American Mathematical Society, 52(2), 113-116; 14, 31

Richardson, M. (1953), *Solutions of irreflexive relations*, Annals of Mathematics, 58,573-590; 14, 31, 32

Roth, A. E., Peranson, E. (1999), *The Redesign of the Matching Market for American Physicians: Some Engineering Aspects of Economic Design*, American Economic Review, 89, 4: 748-780; 2

Serrano, R. (1997), *A comment on the Nash program and the theory of implementation*, Economics Letters, 55, 2, 203-208; 10

von Neumann J., and Morgenstern, O. (1944), *Theory of games and economic behavior*, Princeton Univ. Press, Princeton, NJ; i, 1, 4, 5, 8, 9, 26

Schmeidler, D. (1969), *The Nucleolus of a Characteristic Function Game*, SIAM Journal of Applied Mathematics, 17, 1163-1170; 9

Sertel, M. (2001), *Designing rights: Invisible hand theorems, covering and membership*, mimeo, Bogazici University; 3

Shapley, L. S. (1951), *Notes on the N-Person Game — II: The Value of an N-Person Game*, RAND Corporation; 9