



**HAL**  
open science

## A solution to the two-person implementation problem

Jean-François Laslier, Matias Nunez, M Remzi Sanver

► **To cite this version:**

Jean-François Laslier, Matias Nunez, M Remzi Sanver. A solution to the two-person implementation problem. 2019. halshs-02173504v2

**HAL Id: halshs-02173504**

**<https://hal.science/halshs-02173504v2>**

Preprint submitted on 8 Nov 2019 (v2), last revised 5 May 2020 (v3)

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



WORKING PAPER N° 2019 – 32

## A solution to the two-person implementation problem

Jean-François Laslier  
Matías Núñez  
M. Remzi Sanver

JEL Codes:  
Keywords:



# A SOLUTION TO THE TWO-PERSON IMPLEMENTATION PROBLEM<sup>\*</sup>

JEAN-FRANÇOIS LASLIER,<sup>†</sup> MATÍAS NÚÑEZ<sup>‡</sup> AND M. REMZI SANVER<sup>§</sup>

OCTOBER 2019

## Abstract

We propose a solution to the classical problem of Hurwicz and Schmeidler [1978] and Maskin [1999] according to which, in two-person societies, no Pareto efficient rule is Nash-implementable. To this end, we consider implementation through mechanisms that are deterministic-in-equilibrium while lotteries are allowed off-equilibrium. For strict preferences over alternatives and under a very weak condition for extending preferences over lotteries, we build simple veto mechanisms that Nash implement a class of Pareto efficient social choice rules called *Pareto-and-veto* rules. Moreover, under mild richness conditions on the domain of preferences over lotteries, any Pareto efficient Nash-implementable rule is a Pareto-and-veto rule and hence is implementable through one of our simple veto mechanisms.

---

<sup>\*</sup>We would like to thank Lorenzo Bastianello, Chris Chambers, Bhaskar Dutta, Yukio Koriyama, Matthew Jackson, Phil Reny, Larry Samuelson for useful remarks. We are particular indebted to William Thomson for his comments and suggestions. Financial support from the project IDEX ANR-10-IDEX-0001-02 PSL MIFID is gratefully acknowledged.

<sup>†</sup>CNRS, Paris School of Economics, France.

<sup>‡</sup>CREST, Ecole Polytechnique & CNRS, France.

<sup>§</sup>Université Paris Dauphine, PSL Research University, CNRS, UMR [7243], LAMSADE, 75016 Paris, France.

# 1 Introduction

Can one design some protocol that ensures that two players reach a Pareto efficient agreement in equilibrium? The theorems of Hurwicz and Schmeidler [1978] and Maskin [1999], at the outset of implementation theory, provide a negative answer to this question: no deterministic mechanism, except dictatorship, can guarantee that every Nash equilibrium is Pareto efficient. In fact, there is a tension between the conditions for the existence of an equilibrium at every preference profile and those which ensure that each outcome is Pareto efficient. This impossibility, to which we refer as the *two-person implementation problem*, is particularly striking since it is based on a very mild set of assumptions.

We propose a solution to this problem based on a modification of the mechanisms used for implementation. More precisely, we examine the consequences of allowing lotteries off-equilibrium, while still ensuring deterministic outcomes in equilibrium. That is, we consider Nash implementation through deterministic-in-equilibrium mechanisms or simply **DE** mechanisms.<sup>1</sup>

Since we introduce lotteries, the notion of Pareto efficiency needs some qualification (see Bogomolnaia and Moulin [2001] for a discussion). Two classical definitions are *ex-ante* and *ex-post* Pareto efficiency. A lottery is *ex-ante* Pareto efficient if no other lottery Pareto dominates it, whereas it is *ex-post* Pareto efficient if no alternative that can be selected by the lottery is Pareto dominated by some other alternative. While we show that the possibility of *ex-ante* Pareto efficient implementation is severely limited, we establish that *ex-post* Pareto efficient implementation is possible, by **DE** mechanisms, as soon as preferences over alternatives are strict.<sup>2</sup>

Our main result is that a SCR is Pareto efficient and Nash-implementable if and only if it is a *Pareto-and-veto* rule: for some pair of integers  $v = (v_1, v_2)$  with  $v_1 + v_2 + 1$  being the number of alternatives, it selects all Pareto efficient alternatives that are

---

<sup>1</sup>To our knowledge, this paper is the first one to consider this idea with two players. See Özkal-Sanver and Sanver [2006], Bochet [2007] and Benoît and Ok [2008] for related ideas. We could replace lotteries with sets of alternatives and the results presented would be identical.

<sup>2</sup>The current results do not extend to the setting where the players are indifferent among several alternatives. Indeed, as proved by Sanver [2006], no selection of Pareto set is (Maskin) monotonic and hence can be implemented.

not among the  $v_i$  worst alternatives for each player  $i$ .<sup>3</sup> The Pareto-and-veto rule with vector  $v$  is denoted  $\text{PV}_v$ .

The sufficiency part, rather than relying on classical integer games,<sup>4</sup> builds for each Pareto-and-veto rule  $\text{PV}_v$  a simple veto mechanism, which we call a *strike mechanism*, that Nash implements  $\text{PV}_v$ . A strike mechanism endows each player  $i$  with  $v_i$  vetoes to be distributed among the alternatives with, again,  $v_1 + v_2 + 1$  being the number of alternatives. The game is simultaneous and the outcome is a full-support lottery over the non-vetoed alternatives. The best-response reasoning is straightforward: given the vetoes of his opponent, a player can induce any alternative non-vetoed by his opponent as the outcome by adequately casting his vetoes. Thus, his best response amounts to select his best element among the non-vetoed alternatives.

We prove that this game has pure strategy equilibria. Then, the nice feature of best responses has three consequences. First, each veto mechanism is **DE** since a unique alternative remains non-vetoed in equilibrium, otherwise there is a conflict with best responses. Second, an equilibrium outcome is Pareto efficient since otherwise a player can always, by deviating, select a Pareto dominating alternative. Third, the equilibrium strategies have a natural shape: if  $x$  is the implemented alternative and  $v_i$  is the number of vetoes, player  $i$  vetoes all alternatives preferred to  $x$  by his opponent (say  $k$  alternatives) and he vetoes also  $v_i - k$  among the alternatives less preferred than  $x$  by his opponent. If both strategies veto disjoint sets of alternatives, this forces each player to accept his opponent's strategy. In any equilibrium, this is case: the players veto disjoint sets of alternatives and only one alternative, the implemented one, remains non-vetoed.

This result holds under the standard von Neumann and Morgenstern expected utility framework and is even more general than that. It remains true under a mild condition that we term “best-element bias”: for any set of alternatives, a player prefers the (sure) lottery that consists of his most preferred element in the set to

---

<sup>3</sup>It is not the first time that Pareto-and-veto rules are found to be of interest in the literature: Abreu and Sen [1991] (pp. 1016-17) present this class of rules as the main example that is virtually implementable but fails to be Nash-implementable. In a setting where monetary transfers are allowed, Sanver [2018] designs a direct veto mechanism that implements alternatives which are Pareto efficient and preferable to some disagreement outcome by both players.

<sup>4</sup>Jackson [2001] summarizes some views on the limits of these games.

any lottery with support in the same set. Furthermore, we characterize the class of (ex-post) Pareto efficient social choice rules (SCRs) that can be Nash implemented through these mechanisms. Importantly, the implementation result just relies on the existence of a best-element bias for both players. It does not even require completeness<sup>5</sup> or transitivity of the preferences over lotteries.

The necessity part is more involved. Here, the key concept is the veto power generated by a mechanism: a mechanism  $\mu$  endows player  $i$  with veto power over some set  $X$  of alternatives if and only if player  $i$  has some strategy that prevents any alternative in  $X$  to be selected with positive probability whatever his opponent plays. As we show, any mechanism  $\mu$  that ensures Pareto efficient outcomes must endow each player  $i$  with veto power over every set of alternatives whose cardinality does not exceed some integer  $v_i^\mu$  with  $v_1^\mu + v_2^\mu + 1$  being the number of alternatives. This is a strong result which almost directly entails that only sub-correspondences of  $\text{pv}_v$  are Nash-implementable. The necessity is established on a domain of preference extensions over lotteries that is rich enough to include specific extended preferences that we label “priority” extensions. In words, a “priority” extended preference is defined by the property that whenever all the elements of a set  $X$  are preferred to all elements outside  $X$ , any lottery that put some weight (however small) on some element of  $X$  is preferred to any lottery that puts no weight on  $X$ . For instance, the domain of vNM preferences satisfies this requirement.

The structure of the paper is as follows: Section 2 introduces the basic notions. Section 3 presents the strike mechanisms and show that the strike mechanism with parameter  $v$  Nash implements the Pareto-and-veto rule with the same parameter. Section 4 tackles the necessity issue. It shows that if a SCR is Pareto efficient and Nash-implementable, then it is a Pareto-and-veto rule. Section 5 further studies the game-theoretical properties of the proposed mechanisms and shows in particular that their equilibrium are obtained through individual best-response dynamics. Section 6 discusses the limitations of ex-ante implementation through **DE** mechanisms.

---

<sup>5</sup>See Schmeidler [1989] who underlines the importance of weakening completeness and writes: "Out of the seven axioms listed here the completeness of the preferences [...] seems to me the most restrictive and most imposing assumption".

Section 7 presents a review of the literature and Section 8 makes some concluding remarks.

## 2 Basic notions and notation

A set  $N = \{1, 2\}$  of two players faces a finite set  $A$  of  $n + 1 \geq 3$  alternatives. We write  $\mathcal{A} = 2^A$  for the power set of  $A$  and  $\overline{\mathcal{A}} = \mathcal{A} \setminus \{\emptyset, A\}$ . The set of probability distributions (or “lotteries”) over  $A$  is denoted  $\Delta = \{p : A \rightarrow [0, 1] \mid \sum_{x \in A} p(x) = 1\}$ . For each lottery  $p \in \Delta$ , we let  $\text{SUPP}(p) = \{x \in A \mid p(x) > 0\}$  denote the support of  $p$ . For each  $X \in \mathcal{A}$ ,  $p[X] = \sum_{x \in X} p(x)$  stands for the probability that  $p$  selects an alternative in  $X$ . Let  $\Delta^{\text{uni}}$  denotes the set of uniform probability distributions over the non-empty subsets of  $A$ . Slightly abusing notation, we let  $\{x\}$  denote both the singleton set consisting of alternative  $x$  and the lottery that selects  $x$  with probability one.

We define a “strike mechanism” as follows. Each player  $i \in N$  is endowed with a non-negative number  $v_i$  of vetoes, with  $v_1 + v_2 = n$ . The set

$$\mathcal{M}_i = \{X \subseteq A \mid \#X = v_i\}$$

represents the sets of alternatives  $i$  can veto, and  $\mathcal{M} = \mathcal{M}_1 \times \mathcal{M}_2$  is the joint message space. The mechanism  $\mu_v : \mathcal{M} \rightarrow \Delta^{\text{uni}}$  associates to each pair of messages  $m = (m_1, m_2)$ , the lottery  $\mu_v(m)$  that is uniform<sup>6</sup> over the set

$$\text{SUPP}(\mu_v(m)) = A \setminus (m_1 \cup m_2).$$

In other words, an alternative is uniformly drawn from the non-vetoed alternatives.

Note that, as  $v_1 + v_2 = n$ , the set  $m_1 \cup m_2$  contains at most  $n$  elements, so that  $\text{SUPP}(\mu_v(m))$  is always non-empty.

The set of linear orders over  $A$  is denoted by  $\mathcal{L}_A$  and its generic element  $\succ_i$  is the preference of  $i \in N$ .<sup>7</sup> The set of (strict) preference profiles over  $A$  is  $\mathcal{L}_A^2 = \mathcal{L}_A \times \mathcal{L}_A$

<sup>6</sup>Theorem 1 holds replacing the uniform distribution by any probability distribution with full support over the non-vetoed alternatives.

<sup>7</sup>More precisely, one of  $x \succ_i y$  and  $y \succ_i x$  holds for any distinct  $x, y \in A$  while  $x \succ_i x$  fails for all

with  $\succ = (\succ_1, \succ_2)$  denoting a generic preference profile. We write

$$\text{PE}(\succ) = \{x \in A \mid \nexists y \in A : \forall i \in N, y \succ_i x\}$$

for the set of **Pareto efficient** alternatives at  $\succ \in \mathcal{L}_A^2$ . Let  $L(x, \succ_i) = \{y \in A : x \succ_i y\}$  be the (strict) lower contour set and  $U(x, \succ_i) = \{y \in A : y \succ_i x\}$  be the (strict) upper contour set of  $x \in A$  at  $\succ_i \in \mathcal{L}_A$ .

A **social choice rule** (SCR) is a mapping  $f : \mathcal{L}_A^2 \rightarrow \mathcal{A}$ . A SCR is Pareto efficient iff  $f(\succ) \subseteq \text{PE}(\succ)$  for all  $\succ \in \mathcal{L}_A^2$ . We say that  $f$  is a sub-correspondence of  $g$  and write  $f \subseteq g$  whenever  $f(\succ) \subseteq g(\succ)$  for all  $\succ \in \mathcal{L}_A^2$ .

In general, a **mechanism** is a function  $\mu : \mathcal{M} \rightarrow \Delta$  with  $\mathcal{M} = \mathcal{M}_1 \times \mathcal{M}_2$  where  $\mathcal{M}_i \neq \emptyset$  is the message space of  $i \in N$ . In order to properly define the game associated to  $\mu$ , we do not need to extend preferences over the whole  $\Delta$  but simply over  $\mu(\mathcal{M}) := \{p \in \Delta \mid p = \mu(m) \text{ for some } m \in \mathcal{M}\}$ , the range of  $\mu$ . In this paper, we only consider mechanisms with finite ranges.<sup>8</sup> For example, the set of uniform lotteries over  $A$ , denoted  $\Delta^{\text{uni}} = \{p \in \Delta \mid p(x) = p(y) \text{ for any } x, y \in \text{supp}(p)\}$  is finite. The strike mechanisms, which play a central role in this work, have  $\Delta^{\text{uni}}$  as their range.

We let  $\mathcal{P}_{\mu(\mathcal{M})}$  stand for the set of binary relations over  $\mu(\mathcal{M})$ . A typical element of  $\mathcal{P}_{\mu(\mathcal{M})}$  is denoted  $\succeq_i^*$  with  $\succ_i^*$  being its strict part. We say that  $\succeq_i^*$  is an **extension** of  $\succ_i$  when  $x \succ_i y \implies \{x\} \succeq_i^* \{y\}, \forall x, y \in A$ .

For a mechanism  $\mu : \mathcal{M} \rightarrow \Delta$  and a preference profile over lotteries  $\succeq^* = (\succeq_1^*, \succeq_2^*)$ , a **Nash equilibrium** is a pair of messages  $(m_1, m_2) \in \mathcal{M}$  such that, for all  $m'_1 \in \mathcal{M}_1$  and all  $m'_2 \in \mathcal{M}_2$ ,  $\mu(m_1, m_2) \succeq_1^* \mu(m'_1, m_2)$  and  $\mu(m_1, m_2) \succeq_2^* \mu(m_1, m'_2)$ . Let  $\mathcal{N}^\mu(\succeq^*)$  denote the set of Nash equilibria of the mechanism  $\mu$  at the profile  $\succeq^*$ .

We now turn to the question of the domain of preferences to be considered. As already mentioned we work under the assumption that preferences over alternatives are strict, but we are flexible as to the way preferences are extended from alternatives to lotteries. Since there are many ways to do so, we use a notion of admissible extended preferences. Let  $\kappa(\succ_i) \subseteq \mathcal{P}_\Delta$  be a set of admissible preferences over lot-

---

$x \in A$ . Moreover,  $x \succ_i y$  and  $y \succ_i z$  implies  $x \succ_i z$  for all  $x, y, z \in A$ .

<sup>8</sup>While our results still hold extending over the whole simplex, the richness condition **PREX** becomes harder to satisfy. We would like to thank Bhaskar Dutta for pointing this out.



teries of player  $i$  associated with  $\succ_i \in \mathcal{L}_A$ . Abusing notation, let  $\kappa(\succ) \subseteq \mathcal{P}_\Delta^2$  be the set of admissible preference profiles over  $\Delta$  associated with the preference profile  $\succ = (\succ_1, \succ_2)$ . Such a correspondence  $\kappa$  that associates to each preference a set of extended preferences (and to each profile of preference a set of profiles of extended preferences) is called a **domain of preference extensions**. Throughout the paper we use the property of Best-element bias: a player with a best-element bias prefers the (sure) lottery that selects his best element in  $X$  to any (considered) lottery with support in  $X$ .

**Best-element bias:** Let  $\succ_i \in \mathcal{L}_A$  be a strict preference on  $A$ , and let  $\bar{\Delta} \subseteq \Delta$  be a set of lotteries. An extension  $\succeq_i^*$  of  $\succ_i$  **exhibits the best element bias in  $\bar{\Delta}$**  when for any  $X \in \mathcal{A}$  with  $\#X > 1$  and any  $x \in X$ , if  $x \succ_i y$  for any  $y \in X \setminus \{x\}$ , then  $\{x\} \succeq_i^* p$  for all  $p \in \bar{\Delta}$  with  $\text{SUPP}(p) \subseteq X$  and  $p \neq \{x\}$ .

A domain  $\kappa$  is said to satisfy the best element bias (in short:  $\kappa$  satisfies **BEB**) in  $\bar{\Delta}$  if, for any strict preference  $\succ \in \mathcal{L}_A$ , any extension  $\succeq_i^* \in \kappa(\succ)$  exhibit the best element bias in  $\bar{\Delta}$ . Note that **BEB** is satisfied by virtually all domain of preference extensions that are considered in the literature, including the von Neumann and Morgenstern domain.

Given a domain  $\kappa$ , a mechanism  $\mu$  is **admissible** iff for all  $\succ \in \mathcal{L}_A^2$  and all  $\succeq^* \in \kappa(\succ)$ ,  $\mathcal{N}^\mu(\succeq^*) \neq \emptyset$ . It is **deterministic-in-equilibrium (DE)** iff for all  $\succ \in \mathcal{L}_A^2$ , all  $\succeq^* \in \kappa(\succ)$ , and all  $m \in \mathcal{N}^\mu(\succeq^*)$ ,  $\#\text{supp}(\mu(m)) = 1$ . It **Nash-implements** the SCR  $f : \mathcal{L}_A^2 \rightarrow \mathcal{A}$  iff for all  $\succ \in \mathcal{L}_A^2$  and all  $\succeq^* \in \kappa(\succ)$ ,  $f(\succ) = \bigcup_{m \in \mathcal{N}^\mu(\succeq^*)} \text{SUPP}(\mu(m))$ . Note that if  $\mu$  Nash-implements some SCR  $f$ , then  $\mu$  is admissible.

### 3 Pareto efficient implementation

For any  $v = (v_1, v_2) \in \{0, 1, \dots, n\}^2$  with  $v_1 + v_2 = n$ , we define the Pareto-and-veto rule  $\text{PV}_v : \mathcal{L}_A^2 \rightarrow \mathcal{A}$  as the SCR:

$$\text{PV}_v(\succ) = \overbrace{\text{PE}(\succ)}^{\text{Pareto}} \cap \underbrace{\{x \in A \mid \#L(x, \succ_1) \geq v_1\}}_{\text{Best } n-v_1 \text{ alternatives for 1}} \cap \underbrace{\{x \in A \mid \#L(x, \succ_2) \geq v_2\}}_{\text{Best } n-v_2 \text{ alternatives for 2}}.$$

The Pareto-and-veto rule  $\text{PV}_v$  picks all Pareto efficient alternatives with a lower-contour set at least as large as  $v_i$  for every player  $i$ .

Our first observation is that  $\text{PV}_v$  is non empty when  $v_1 + v_2 \leq n$ . To see this, just observe that eliminating  $n$  alternatives at most, out of  $n + 1$ , leaves at least one, say  $a$ . If  $a$  is Pareto efficient, we are done. If not,  $a$  is Pareto-dominated by some  $a' \in \text{PE}_v$ , but since  $a'$  is at least as good as  $a$  for player  $i$ ,  $a$  is still among his  $n - v_i$  best alternatives. As soon as  $v_1 + v_2$  is at least  $n + 1$ , the example of completely opposed preferences shows that  $\text{PV}_v$  can be empty.

We now turn to the implementation  $\text{PV}_v$  by a “strike” mechanism. For a strike mechanism, given a strategy  $m_j$  that vetoes some set of  $v_j$  alternatives, the objective for player  $i$  is to select the support of the lottery that determines the outcome. Let  $g_v(\mathcal{M}_i, m_j) = \{X \in \mathcal{A} \mid \text{SUPP}(\mu_v(m_i, m_j)) = X \text{ for some } m_i \in \mathcal{M}_i\}$  be the attainable set of player  $i$  at  $m_j$  under the strike mechanism  $\mu_v$ . So the set  $g_v(\mathcal{M}_i, m_j)$  contains the different supports of the uniform lotteries that player  $i$  can induce when player  $j$  selects  $m_j$  under the strike mechanism  $\mu_v$ . Because of the number of vetoes at his disposal, player  $i$  can choose the support of the outcome by adequately casting his vetoes as described by the following result:

**Proposition 1.** *For each player  $i$  and each strategy  $m_j \in \mathcal{M}_j$ , the attainable set equals:*

$$g_v(\mathcal{M}_i, m_j) = \{X \subseteq A \setminus m_j \mid 1 \leq \#X \leq \min\{n + 1 - v_i, n + 1 - v_j\}\}.$$

*Proof.* Take some player  $i$  and some strategy  $m_j \in \mathcal{M}_j$ . Take first the case with  $v_i < v_j$  so that  $n + 1 - v_j < n + 1 - v_i$ . We want to prove that for each non-empty  $X \subseteq A \setminus m_j$  (hence with  $\#X \leq n + 1 - v_j$ ), there is some  $m_i \in \mathcal{M}_i$  with  $\text{SUPP}(\mu_v(m_i, m_j)) = X$ . Note that each non-empty subset of  $A \setminus m_j$  is of the form  $A \setminus (m_j \cup C)$  with  $0 \leq \#C \leq v_i$ . Thus, it suffices to pick  $m_i$  such that  $m_i \setminus m_j = C$  which ensures that  $\text{SUPP}(\mu_v(m_i, m_j)) = A \setminus (m_i \cup m_j) = A \setminus (m_j \cup C)$ , as required. In the case  $v_i \geq v_j$ , take  $m_i$  with  $m_i \setminus m_j = C$ .

Since  $v_i \geq v_j$ , it follows that  $\#C \geq v_i - v_j$  and hence for each non-empty  $X \subseteq A \setminus m_j$  with  $\#X \leq n+1 - v_j - (v_i - v_j) = n+1 - v_i$ , there is some  $m_i \in \mathcal{M}_i$  with  $\text{SUPP}(\mu_v(m_i, m_j)) = X$ .  $\square$

Proposition 1's main implication is that player  $i$  can induce any singleton in  $A \setminus m_j$  as the support of the outcome: formally, for any player  $i$  and any alternative  $x \in A$ :

$$x \in A \setminus m_j \implies \{x\} \in g_v(\mathcal{M}_i, m_j).$$

Building on this key property of the attainable set, the rest of the section proves that the strike mechanism  $\mu_v$  Nash implements the Pareto-and-veto rule  $\text{pv}_v$ . The first consequence of this property is that strike mechanisms are deterministic in equilibrium as long as the domain satisfies **BEB**.

**Proposition 2.** *For any strike mechanism  $\mu_v$ , if the domain  $\kappa$  satisfies **BEB** in the range of  $\mu_v$ , then  $\mu_v$  is **DE**.*

*Proof.* Assume that there is some equilibrium  $m = (m_1, m_2)$  with  $\#\text{SUPP}(\mu_v(m)) > 1$ . Consider some player  $i$  and some alternative  $x \in \text{SUPP}(\mu_v(m))$  with  $x \succ_i y$  for all  $y \in \text{SUPP}(\mu_v(m)) \setminus \{x\}$ . Since  $x \in A \setminus m_j$ , Proposition 1 shows that  $\{x\} \in g_v(\mathcal{M}_i, m_j)$ . Thus, there is some  $m'_i \in \mathcal{M}_i$  with  $\mu_v(m'_i, m_j) = \{x\}$ . Furthermore,  $\{x\} \succ_i^* \mu_v(m)$  due to **BEB**, which contradicts that  $m$  is an equilibrium.  $\square$

Since a strike mechanism is **DE**, no uncertainty remains in equilibrium: players veto disjoint sets of alternatives and a unique alternative is selected. The next result shows that any alternative selected by a Pareto-and-veto rule with veto vector  $v$  is selected by some strict equilibrium of the strike mechanism  $\mu_v$ . In a strict equilibrium, the best response of each player is unique. Few games admit this sort of equilibria and when they exist, strict equilibria have all the usual desiderata that the theory of refinements requires<sup>9</sup>. As a by-product of our proof, we also obtain the admissibility of each strike mechanism  $\mu_v$ .

---

<sup>9</sup>In particular, a strict equilibrium is proper and hence perfect (see van Damme [1991] for a classic treatment). Note that equilibrium refinements rely on expected utility (to derive utility from perturbed profiles) whereas most of our results do not depend on this assumption.

**Proposition 3.** *For any veto vector  $v$ , let the domain  $\kappa$  satisfy **BEB** in the range of  $\mu_v$ , then any Pareto-and-veto alternative  $x \in \text{PV}_v(\succ)$  is the unique outcome of a strict equilibrium of the strike mechanism  $\mu_v$ .*

*Proof.* Take any  $v \in \{0, \dots, n\}^2$  with  $v_1 + v_2 = n$ , any  $\succ \in \mathcal{L}_A^2$ , and any  $\succeq^* \in \kappa(\succ)$  with  $\kappa$  satisfying **BEB**. Take  $x \in \text{PV}_v(\succ)$ . Because  $x$  is Pareto-optimal, any of the  $n$  other alternatives is either strictly better than  $x$  for one (and only one) player or strictly worse for both. So counting these  $n = v_1 + v_2$  alternatives we obtain:

$$v_1 + v_2 = \#U(x, \succ_1) + \#U(x, \succ_2) + \#(L(x, \succ_1) \cap L(x, \succ_2)). \quad (1)$$

By definition of  $\text{PV}_v$ ,  $v_1 \leq \#L(x, \succ_1) = n - \#U(x, \succ_1)$ . Therefore  $v_2 \geq \#U(x, \succ_1)$ , which means that player 2 has enough vetoes to block all the alternatives that player 1 strictly prefer to  $x$ . The same holds for player 1 with respect to player 2. Writing Equation (1) as:

$$[v_1 - \#U(x, \succ_2)] + [v_2 - \#U(x, \succ_1)] = \#(L(x, \succ_1) \cap L(x, \succ_2)),$$

one can see that it is possible to have players 1 and 2 respectively veto  $v_1 - \#U(x, \succ_2)$  and  $v_2 - \#U(x, \succ_1)$  different alternatives in  $L(x, \succ_1) \cap L(x, \succ_2)$ , so that all  $n$  alternatives are vetoed by one player or the other.

Let  $m_1$  and  $m_2$  be such strategies. We now prove that, under **BEB**,  $m_1$  is a strict best response to  $m_2$ . To this end, recall that  $U(x, \succ_1) \subseteq m_2$ : any alternative strictly preferred by player 1 to  $x$  is vetoed by player 2. So when player 1 deviates to  $m'_1 \in \mathcal{M}_1$ , the support  $A \setminus (m'_1 \cup m_2)$  of the outcome lottery excludes  $U(x, \succ_1)$ . Because of the constraints on the number of vetoes,  $\mu(m'_1, m_2) = \{x\}$  is impossible for  $m'_1 \neq m_1$ . Therefore, for player 1, the support of  $\mu(m'_1, m_2)$  either contains only alternatives that are strictly worse than  $x$ , or contains  $x$  and some other alternatives that all are worse than  $x$ . By **BEB**, player 1 strictly prefers  $\{x\}$  to such outcomes, so  $m_1$  is the unique best response to  $m_2$ . The same holds for the other player, hence the result.  $\square$

Equipped with this result, we are now ready to prove Theorem 1, according to which the strike mechanism  $\mu$  with veto vector  $v$  Nash-implements the Pareto-and-

veto rule with the same veto vector  $v$ .

**Theorem 1.** *Let the domain  $\kappa$  satisfy **BEB** in  $\Delta^{\text{uni}}$ . The strike mechanism  $\mu_v$  Nash-implements the Pareto-and-veto rule  $\text{PV}_v$  for any  $v \in \{0, \dots, n\}^2$  with  $v_1 + v_2 = n$ .*

*Proof.* Take any  $v \in \{0, \dots, n\}^2$  with  $v_1 + v_2 = n$ , any  $\succ \in \mathcal{L}_A^2$ , and any  $\succeq^* \in \kappa(\succ)$ . Proposition 3 establishes that  $\text{PV}_v(\succ) \subseteq \bigcup_{m \in \mathcal{N}^{\mu_v(\succeq^*)}} \mu_v(m)$  provided that the domain  $\kappa$  satisfies **BEB** in the range of  $\mu_v$ . Moreover, it is easy to check that the range of  $\mu_v$  is precisely  $\Delta^{\text{uni}}$ .

We now show  $\text{PV}_v(\succ) \supseteq \bigcup_{m \in \mathcal{N}^{\mu_v(\succeq^*)}} \mu_v(m)$ . Take some  $x$  with  $\mu_v(m) = \{x\}$  for some  $m \in \mathcal{N}^{\mu_v(\succeq^*)}$ . We first show that  $x \in \text{PE}(\succ)$ . Suppose not, i.e., there exists  $y \in A$  with  $y \succ_i x$  for all  $i \in N$ . Since  $\mu_v(m) = \{x\}$ , we have  $m_1 \cap m_2 = \emptyset$ . Thus,  $y \in m_i$  for some  $i \in N$ , say  $i = 1$ , without loss of generality. It follows that  $y \in A \setminus m_2$  and thus  $\{y\} \in g_v(\mathcal{M}_1, m_2)$ . Therefore,  $\mu_v(m'_1, m_2) = \{y\}$  for some  $m'_1$  and as  $\{y\} \succ_1^* \mu_v(m) = \{x\}$ , we contradict  $m \in \mathcal{N}^{\mu_v(\succeq^*)}$ .

We now show  $\#L(x, \succ_i) \geq v_i \forall i \in N$ . Suppose, without loss of generality, that  $v_1 > \#L(x, \succ_1)$ . For any  $m_2 \in \mathcal{M}_2$ , there is some  $y \in A$  with  $y \in A \setminus m_2$ ,  $y \succ_1 z$  for any  $z \in A \setminus m_2$  and  $\#L(y, \succ_1) \geq n + 1 - v_2$  since  $m_2$  contains  $v_2$  vetoes. Remark that  $n + 1 - v_2 = v_1 + 1$  and thus  $y \succ_1 x$  since  $\#L(y, \succ_1) > \#L(x, \succ_1)$  so  $\{y\} \succ_1^* \{x\}$ . Moreover,  $\mu_v(m'_1, m_2) = \{y\}$  for some  $m'_1 \in \mathcal{M}_1$  since  $y \in A \setminus m_2$  and thus  $\{y\} \in g_v(\mathcal{M}_1, m_2)$ . Finally  $\{y\} \succ_1^* \{x\} = \mu_v(m)$ , contradicting that  $m$  is an equilibrium.  $\square$

## 4 On the necessity of vetoes and the uniqueness of the mechanism

We now turn to the uniqueness question. This section shows in which sense the strike mechanism is the only possibility for Pareto-efficient Nash implementation.

### 4.1 Domain of extended preferences

We now define some conditions on the domain  $\kappa$  to be used throughout. The first one restricts admissible extensions in the same spirit as the **BEB** condition. A player

with a worst-element bias (or simply **WEB**) prefers any lottery with support in  $X$  to the (sure) one that selects his worst element in  $X$ .

**Worst-element bias:** Let  $\succ_i \in \mathcal{L}_A$  be a strict preference on  $A$ , and let  $\bar{\Delta} \subseteq \Delta$  be a set of lotteries. An extension  $\succeq_i^*$  of  $\succ_i$  **exhibits the worst element bias in  $\bar{\Delta}$**  when for any  $X \in \mathcal{A}$  with  $\#X > 1$  and any  $x \in X$ , if  $y \succ_i x$  for any  $y \in X \setminus \{x\}$ , then  $p \succ_i^* \{x\}$  for all  $p \in \bar{\Delta}$  with  $\text{SUPP}(p) \subseteq X$  and  $p \neq \{x\}$ .

As in the case of **BEB**, **WEB** is satisfied by virtually all preference extensions over lotteries.<sup>10</sup> We say that a domain  $\kappa$  satisfies **WEB** in  $\bar{\Delta}$  iff **WEB** is satisfied in  $\bar{\Delta}$  for all  $\succeq^* \in \kappa(\succ)$ , for all  $\succ \in \mathcal{L}_A^2$ .

The next condition, Priority Extension, deals with the richness of the domain of preference extensions. For any lottery  $p \in \Delta$ , we write  $p[\cdot \succeq x] = \sum_{y: y \succeq x} p(y)$  to refer to the probability, according to  $p$ , of obtaining an alternative weakly preferred to  $x$  according to  $\succ$ .

**Priority extension:** Let  $\succeq_i^*$  extend  $\succ_i$  and let  $x \in A$ , the extension  $\succeq_i^*$  is a (**PREX**) of  $\succ_i$  for  $x$  in  $\bar{\Delta}$  iff given any two lotteries  $p, q \in \bar{\Delta}$ , if  $p[\cdot \succeq x] > 0$  and  $q[\cdot \succeq x] = 0$ , then  $p \succ^* q$ .

The interpretation of this property is clear: under a priority extension, each alternative is used as a grading benchmark: The individual prefers to reach the benchmark  $x$ , even with a tiny probability, than not reaching it. The argument “What is the best alternative I have some chance to obtain with that lottery?” has priority over the precise values of the probabilities. We say that a domain  $\kappa$  satisfies **PREX** in  $\bar{\Delta}$  iff for all  $\succ \in \mathcal{L}_A^2$ , there is some  $\succeq^* \in \kappa(\succ)$  that is a priority extension of  $\succ$  in  $\bar{\Delta}$  for all  $x \in A$ .<sup>11</sup>

Here is an example of a domain of extension that satisfies the condition in the set  $\Delta^{\text{uni}}$  of uniform lotteries. Similar examples can be found for any finite set of lotteries. Consider the correspondence  $\kappa^{\text{vNM}} : \mathcal{L}_A \rightarrow \mathcal{P}_{\Delta^{\text{uni}}}$  that allows any von Neumann and

<sup>10</sup>In fact, **BEB** and **WEB** are satisfied if one considers the well-known preference extension axioms of the literature (such as Gärdenfors [1976] or Kelly [1977]) and deduces preferences over lotteries through the preferences over their supports. If  $\kappa$  satisfies **BEB** and **WEB** (which are universally quantified), every sub-correspondence of  $\kappa$  satisfies them as well.

<sup>11</sup>Note that if  $x$  is bottom-ranked in  $\succ$ , there is no lottery  $q$  with  $q[\cdot \succeq x] = 0$ , so that any extension is (vacuously) a priority extension for  $x$ .

Morgenstern extension of  $\succ$ . In other words, for  $\succ \in \mathcal{L}_A$ ,  $\kappa^{\text{vNM}}(\succ)$  is the set of all  $\succ^* \in \mathcal{P}_{\Delta^{\text{uni}}}$  such that there exists a vector  $u \in \mathbb{R}^A$  with  $a \succ b \iff u_a > u_b$  for all  $a, b \in A$  and:

$$\forall p, q \in \Delta^{\text{uni}}, p \succ^* q \iff \sum_{a \in A} p(a)u_a > \sum_{a \in A} q(a)u_a.$$

The domain  $\kappa^{\text{vNM}}(\succ)$  contains priority extensions of  $\succ$  to  $\Delta^{\text{uni}}$ . To see this, label the alternatives in  $A$  according to the preference:  $a_{n+1} \succ a_n \succ \dots \succ a_1$  and let  $u_{a_k} = (n+1)^k$  for any  $a_k \in A$ . Take any pair  $p, q \in \Delta^{\text{uni}}$  with  $p[\cdot \geq a_k] > 0$  and  $q[\cdot \geq a_k] = 0$  for some  $a_k$ . The expected value of  $p$ , that is  $\sum_{a \in A} p(a)u_a$ , reaches its minimum when the lottery contains in its support  $a_k$  but no better alternative according to  $\geq$  (and hence has  $k$  alternatives in its support). The expected value  $\sum_{a \in A} p(a)u_a$  is at least

$$\frac{u_{a_k}}{k} > \frac{u_{a_k}}{k+1} = \frac{(n+1)^k}{k+1} \geq (n+1)^{k-1}.$$

The expected value of  $q$ ,  $\sum_{a \in A} q(a)u_a$ , reaches its maximum when  $q = \{a_{k-1}\}$  and hence its value is at most  $(n+1)^{k-1}$ . Therefore, for any  $a_k \in A$ ,  $p[\cdot \geq a_k] > 0$  and  $q[\cdot \geq a_k] = 0$  implies that  $p \succ^* q$ . Thus, uniform lotteries are ordered following the priority rule.

## 4.2 Implementable rules

We are now ready to state the counterpart to Theorem 1, according to which, if one wants to implement Pareto efficient SCRs through a **DE** mechanism, the SCR must be a Pareto-and-veto rule. Precisely we prove the following:

**Theorem 2.** *Let  $f$  be a Pareto efficient SCR that is Nash-implementable by a **DE** mechanism  $\mu$  on a domain  $\kappa$ . Let the domain  $\kappa$  satisfy **BEB**, **WEB** and **PREX** in the range of  $\mu$ . Then  $f = \text{PV}_v$  for some  $v \in \{0, \dots, n\}^2$  with  $v_1 + v_2 = n$ .*

To prepare for the proof we provide lemmas showing that we can restrict attention to “veto-neutral” mechanisms. For each player  $i$ , let

$$\text{veto}(\mathcal{M}_i) = \{X \in \mathcal{A} \mid \exists m_i \in \mathcal{M}_i \text{ s.t. } \text{SUPP}(\mu(m_i, m_j)) \cap X = \emptyset \text{ for all } m_j \in \mathcal{M}_j\},$$

denote the veto set for player  $i$ . When  $X \in \text{veto}(\mathcal{M}_i)$ , we say that player  $i$  has veto

power over the set  $X$  of alternatives, i.e., he has a strategy that ensures that no alternative in this set belongs to the support of the outcome independently of the strategy of his opponent. We first state a result on the structure of the veto power that **DE** mechanisms generate.

**Lemma 1.** *Under the hypothesis of Theorem 2, for any partition  $\{X, Y\}$  of  $A$  with  $X \in \overline{\mathcal{A}}$ , either  $Y \in \text{veto}(\mathcal{M}_1)$  or  $X \in \text{veto}(\mathcal{M}_2)$  but not both.*

*Proof.* Let  $\mu : \mathcal{M} \rightarrow \Delta$  be admissible and **DE** and let  $X \in \overline{\mathcal{A}}$ . Write  $Y = A \setminus X$ . Pick some  $\succ \in \mathcal{L}_A^2$  such that  $\forall x \in X, \forall y \in Y, x \succ_1 y$  and  $y \succ_2 x$ . The existence of such preference  $\succ$  is ensured by our assumption that the domain contains all strict preferences on alternatives. Take also  $\succeq^* \in \kappa(\succ)$  such that  $p \succ_1^* q$  for all  $p, q \in \mu(\mathcal{M})$  with  $p[X] > 0$  and  $q[X] = 0$ , and such that  $p \succ_2^* q$  for all  $p, q \in \mu(\mathcal{M})$  with  $p[Y] > 0$  and  $q[Y] = 0$ . The existence of such extended preference  $\succeq^*$  is ensured by **PREX**. Now suppose, for a contradiction, that  $Y \notin \text{veto}(\mathcal{M}_1)$  and  $X \notin \text{veto}(\mathcal{M}_2)$ . Because  $\mu$  is admissible and **DE**, there exists an equilibrium  $m = (m_1, m_2) \in \mathcal{N}^\mu(\succeq^*)$  with  $\mu(m) = \{a\}$  for some  $a \in A$ . Two cases are possible:

- If  $a \in X$ . As  $Y \notin \text{veto}(\mathcal{M}_1)$ ,  $\exists m'_2 \in \mathcal{M}_2$  such that  $\text{supp}(\mu(m_1, m'_2)) \cap Y \neq \emptyset$ , hence  $\mu(m_1, m'_2) \succeq_2^* \{a\}$  due to **WEB**, contradicting  $m \in \mathcal{N}^\mu(\succeq^*)$ .
- If  $a \in Y$ . As  $X \notin \text{veto}(\mathcal{M}_2)$ ,  $\exists m'_1 \in \mathcal{M}_1$  such that  $\text{supp}(\mu(m'_1, m_2)) \cap X \neq \emptyset$ , hence  $\mu(m'_1, m_2) \succeq_1^* \{a\}$ , again contradicting  $m \in \mathcal{N}^\mu(\succeq^*)$ .

Thus,  $Y \in \text{veto}(\mathcal{M}_1)$  or  $X \in \text{veto}(\mathcal{M}_2)$ . Because the mechanism is well-defined, it is impossible that  $Y$  belongs to  $\text{veto}(\mathcal{M}_1)$  and its complement  $X$  belongs to  $\text{veto}(\mathcal{M}_2)$ . Therefore either  $Y \in \text{veto}(\mathcal{M}_1)$  or  $X \in \text{veto}(\mathcal{M}_2)$  but not both. □

A mechanism  $\mu$  is **veto neutral** for player  $i$  iff for any  $X \in \mathcal{A}$  and any permutation  $\rho : A \rightarrow A$ ,  $X \in \text{veto}(\mathcal{M}_i) \iff \rho(X) \in \text{veto}(\mathcal{M}_i)$ . When  $\mu$  is veto neutral for some player  $i$ , if a set with a given cardinality belongs to  $\text{veto}(\mathcal{M}_i)$  then any other set with the same cardinality belongs to  $\text{veto}(\mathcal{M}_i)$  as well. Note that a player with veto power



over  $X$  has also veto power over any  $X' \subset X$ . Hence, the veto set for player  $i$  can be written as:

$$\text{veto}(\mathcal{M}_i) = \{X \in \mathcal{A} \mid \#X \leq v_i\},$$

where the integer  $v_i$  stands for the cardinality of the highest cardinality set over which  $i$  has veto power.

**Lemma 2.** *Under the hypothesis of Theorem 2,  $\mu$  is veto neutral for both players.*

*Proof.* Let  $X \in \text{veto}(\mathcal{M}_1)$ ,  $x \in X$  and  $x' \in A \setminus X$ .<sup>12</sup> Thus, there exists  $m_1 \in \mathcal{M}_1$  that vetoes  $X$ . The set  $X' = X \setminus \{x\} \cup \{x'\}$  has the same cardinal as  $X$ . Write  $Y = A \setminus (X \cup \{x'\})$ , so that we have a partition

$$A = (X \setminus \{x\}) \cup \{x\} \cup \{x'\} \cup Y.$$

Suppose, for a contradiction, that  $X' \notin \text{veto}(\mathcal{M}_1)$ . Lemma 1 then implies that  $Y \cup \{x\} \in \text{veto}(\mathcal{M}_2)$ . Therefore there exists  $m_2 \in \mathcal{M}_2$  that vetoes  $Y \cup \{x\}$ . Since  $x'$  is neither vetoed by  $m_1$  nor by  $m_2$ ,  $\mu(m_1, m_2) = \{x'\}$ . Now consider a unanimous preference profile  $\succ = (\succ_1, \succ_2)$  such that  $x \succ_i x' \succ_i y$  for all  $y \neq x, x'$  and for  $i = 1, 2$ . For this preference profile, the second-best alternative  $x'$  is Pareto-dominated by  $x$  but, at  $(m_1, m_2)$ , both players veto  $x$ . Thus, no unilateral deviation can obtain, with any probability, a better outcome than  $x'$ . Thanks to **BEB**, that implies that  $(m_1, m_2)$  is a Nash equilibrium, in contradiction with the Pareto efficiency assumption.

The proof of the proposition is established by noting that given any  $X, X' \in \overline{\mathcal{A}}$  with  $\#X = \#X'$ , there is a finite sequence of sets  $X = X_1, \dots, X_s = X'$  with  $\#(X_i \cap X'_{i+1}) = \#X - 1$  for each  $i \in \{1, \dots, s-1\}$  and applying repeatedly the argument above.  $\square$

We can now complete the proof of the Theorem.

*Proof of Theorem 2.* We first establish the existence of  $v$  such that  $f \subseteq \text{PV}_v$ . Let  $\mathcal{M}$  be the joint message space of  $\mu$  that **DE**-implements  $f$ . Take any preference profile  $\succ = (\succ_1, \succ_2)$  and any  $x \in f(\succ)$ . For each  $\succ_i$  with  $i = 1, 2$ , let  $\succ_i^*$  denote its associated **PREX** for  $x$ . Thus, for all  $p, q \in \mu(\mathcal{M})$ , for each  $z \in A$  and for  $i = 1, 2$ , if  $p[\cdot \succ_i z] > 0$

<sup>12</sup>The two extreme cases  $\text{veto}(\mathcal{M}_1) = \{\{\emptyset\}\}$  and  $A \in \text{veto}(\mathcal{M}_1)$  are trivial.

and  $q[\cdot \succ_i z] = 0$  then  $p \succ_i^* q$ . Since  $x \in f(\succ)$ ,  $\mu$  admits a Nash equilibrium  $(m_1, m_2)$  with  $\mu(m_1, m_2) = \{x\}$ . By the definition of an equilibrium, player 2 has no better response to  $m_1$  than  $m_2$ . However, under  $\succ_2^*$ , a deviation  $m'_2$  is profitable for player 2 iff  $\text{SUPP}(\mu(m_1, m'_2)) \cap U(\succ_2, x) \neq \emptyset$ . Therefore:

$$\forall m_2 \in \mathcal{M}_2, \text{SUPP}(\mu(m_1, m_2)) \cap U(\succ_2, x) = \emptyset$$

and likewise for player 1. In other words,  $m_1$  makes the set  $U(\succ_2, x)$  unattainable for player 2 under  $\mu$ . We say that  $m_1$  **gives player 1 veto power on** the set  $U(\succ_2, x)$ , and likewise for player 2.

From Lemma 2, if a player has veto power on some set, she has also veto power on any set of the same cardinality. Let  $v_i$  be the largest number of outcomes that  $i$  can veto. For the mechanism to be well-defined, one needs  $v_1 + v_2 \leq n$ , so that not all the  $n + 1$  alternatives can be vetoed simultaneously. The existence of a deterministic equilibrium (an equilibrium with a singleton outcome) shows that  $v_1 + v_2 \geq n$ .

Clearly, an outcome that would be among the  $v_i$  worse alternatives for a player  $i$  cannot be an equilibrium outcome under  $\mu$  because  $i$  could then veto her  $v_i$  worse alternatives. Due to **WEB**, a player prefers any lottery with support not included in the  $v_i$  worst alternatives to any lottery that selects (surely) one of the worst  $v_i$  alternatives. Hence  $f$  being implementable imposes the required veto conditions on the ranks of the implemented alternatives in the individual preferences. Since we assumed that  $f$  is also efficient, we obtain  $f \subseteq \text{PV}_v$ .

For this  $v$ , we now prove the reverse inclusion. Given  $\succ = (\succ_1, \succ_2)$ , let  $x \in \text{PV}_v(\succ)$ . Consider the profile  $\succ'$  defined as follows.

Label the  $n + 1$  alternatives in two ways:  $a_{n+1} \succ_1 a_n \succ_1 \dots \succ_1 a_1$  and  $b_{n+1} \succ_2 b_n \succ_2 \dots \succ_2 b_1$ . Write  $a_{w_1} = b_{w_2} = x$ . The veto conditions in the definition of  $\text{PV}_v$  are that  $w_1 > v_1$  and  $w_2 > v_2$ , which implies that:

$$\begin{aligned} a_{n+1} \succ_1 \dots \succ_1 a_{w_1} = x \succ_1 \dots \succ_1 a_{v_1} \succ_1 \dots \succ_1 a_1, \\ b_{n+1} \succ_2 \dots \succ_2 b_{w_2} = x \succ_2 \dots \succ_2 b_{v_2} \succ_2 \dots \succ_2 b_1. \end{aligned}$$

The preference  $\succ'_1$  is obtained by lowering the ranks of all those, among the alternatives  $a_{v_1+1}, \dots, a_{w_1-1}$ , which are preferred to  $x$  by the other player, player 2. If  $w_1 = v_1 + 1$  we simply let  $\succ'_1 = \succ_1$ . If  $w_1 \geq v_1 + 2$ , consider the set

$$H_1 = \{a_{v_1+1}, \dots, a_{w_1-1}\} \cap \{b_{w_2+1}, \dots, b_{n+1}\}$$

and observe that

$$\#H_1 \leq n - w_2 \leq n - v_2 = v_1.$$

Starting from  $\succ_1$ , we define  $\succ'_1$  by switching in the ranking the first elements  $a_1, \dots, a_{\#H_1}$  with the elements of  $H_1$ , where  $a_1$  is switched with the most preferred element of  $H_1$  of player 1,  $a_2$  is switched with the second most preferred element of  $H_1$  of player 1 and so on...

We now claim that if  $x \in f(\succ'_1, \succ_2)$  then  $x \in f(\succ)$ . Let  $\mu$  **DE**-implement  $f$ . If  $x \in f(\succ'_1, \succ_2)$ , there exists a pure strategy equilibrium  $(m'_1, m'_2)$  for the game with preferences  $(\succ'_1, \succ_2)$  with  $\{x\} = \mu(m'_1, m'_2)$ . With the initial preferences  $(\succ_1, \succ_2)$ ,  $m_2$  is also a best response since player 2 does not change her preference, and  $m'_1$  is also a best response for player 1 because her preferences differ only below  $x$ . As previously argued,  $m_2$  gives player 2 veto power on the set  $U(\succ_1, x)$ . Since  $U(\succ'_1, x) = U(\succ_1, x)$  by construction, it follows that the support of any lottery that player 1 can attain given  $m_2$  is included in  $A \setminus U(\succ_1, x)$ . Hence, due to **BEB**,  $m_1$  is a best response for player 1 since  $\mu(m_1, m_2) = \{x\}$ . Therefore this equilibrium for  $(\succ'_1, \succ_2)$  is also an equilibrium for  $(\succ_1, \succ_2)$ , that is:  $x \in f(\succ'_1, \succ_2) \implies x \in f(\succ)$ .

The same construction for player 2 yields the preference profile  $\succ'' = (\succ'_1, \succ'_2)$  with the property:

$$x \in f(\succ'') \implies x \in f(\succ). \quad (2)$$

But notice that, by construction of  $\succ'_1$ , all the alternatives  $y$  such that  $y \succ_2 x$  are now among the  $v_1$  worse alternatives according to  $\succ'_1$ . Therefore  $x$  is the preferred alternative, according to  $\succ'_2$ , among the alternatives in the intersection of the top  $n - v_1$  alternatives for player 1 and  $n - v_2$  alternatives for player 2 in  $\succ'$ . Since the same is true for the other player, we find that  $x$  is the unique Pareto optimum in the alter-

natives among the top  $n - v_1$  alternatives for player 1 and the top  $n - v_2$  alternatives for player 2 in  $\succ''$ . Since  $f$  itself is assumed to be efficient and is selecting in  $\text{PV}_v$ , we obtain that  $f(\succ'') = \{x\}$ . From (2) it follows that  $x \in f(v)$  as requested.  $\square$

Theorem 2 shows the existence of a strong link between implementation through **DE** mechanisms and veto power. Indeed, it shows that under the conditions **BEB**, **WEB**, and **PREX**, a SCR has to admit some veto structure in order to be both Pareto efficient and implementable. This theorem is related to the impossibility result by Hurwicz and Schmeidler [1978] in the following sense. Hurwicz and Schmeidler [1978] show that the only SCRs which are both Pareto efficient and implementable (through a deterministic mechanism) are the dictatorial ones. Note that a dictatorial SCR corresponds to  $\text{PV}_v$  with  $v = (n, 0)$  (if player 1 is the dictator) or  $v = (0, n)$  (if player 2 is the dictator). Our theorem shows that by allowing lotteries as off-equilibrium punishments, the Pareto-and-veto rules appear as a class of intermediate and, interestingly, non dictatorial SCRs.

Note that  $\text{PV}_v$  is neutral for any  $v \in \{0, \dots, n\}^2$  and that it is anonymous if and only if  $v_1 = v_2$ . Thus, under the assumptions of Theorem 2, the following observations trivially follow. With an odd number of alternatives, an anonymous, neutral and Pareto efficient SCR  $f$  is Nash-implementable by a **DE** mechanism iff  $f$  is a Pareto-and-veto rule with  $v_1 = v_2$ . On the contrary, with an even number of alternatives, there exist no anonymous, neutral and Pareto efficient SCR that is Nash-implementable by a **DE** mechanism.

### 4.3 Maskin Monotonicity

Maskin Monotonicity has played a key role in the development of implementation theory. It stands as the necessary condition for implementation through deterministic mechanisms: if a SCR is implementable, then it satisfies this condition. As we show, this statement also applies to the current setting. Formally, a SCR  $f$  is **Maskin Monotonic** iff for any  $x \in A$  and any  $\succ, \succ' \in \mathcal{L}_A^2$  with  $L(x, \succ_i) \subseteq L(x, \succ'_i) \forall i \in N$ ,  $x \in f(\succ) \implies x \in f(\succ')$ . Maskin monotonic is satisfied by each Pareto-and-veto rule as shown by the next lemma.

**Lemma 3.** *For any veto vector  $v$ , the Pareto-and-veto rule  $\text{PV}_v$  is Maskin monotonic.*

*Proof.* For any veto vector  $v$ , take any  $\succ \in \mathcal{L}_A^2$  and any  $x \in \text{PV}_v(\succ)$ . Let  $\succ' \in \mathcal{L}_A^2$  be some profile with  $L(x, \succ_i) \subseteq L(x, \succ'_i) \forall i \in N$ . Note that  $x \in \text{PE}(\succ)$  implies that  $x \in \text{PE}(\succ')$ . Moreover,  $\#L(x, \succ'_i) \geq \#L(x, \succ_i)$  for each  $i \in N$  (by construction of  $\succ'$ ) and  $\#L(x, \succ_i) \geq v_i \forall i \in N$  (by the definition of  $\text{PV}_v$ ). Thus  $x \in \text{PV}_v(\succ')$ , as desired.  $\square$

If the domain satisfies **BEB**, **WEB** and **PREX**, a Pareto efficient SCR that is Nash-implementable by a **DE** mechanism is a Pareto-and-veto rule (as stated by Theorem 2). Since any such rule is Maskin monotonic, we conclude that Maskin monotonicity is still necessary for implementation with **DE** mechanisms.

## 5 Iterative best responses

As mentioned above the equilibria of the considered game are pure and strict. This ensures that the usual game-theoretical refinement criteria are satisfied. However, what does this imply concerning the use of veto mechanisms in laboratory experiments or in real-life applications? Fudenberg and Levine [2016] argue that an equilibrium often fails to arise from introspection, but rather from some non-equilibrium learning dynamics. Moreover, as they write, "in laboratory games do not usually resemble Nash equilibrium (except in some special cases); instead, there is abundant experimental evidence that play in many games moves toward equilibrium as subjects play the game repeatedly and receive feedback" (see Van Huyck et al. [1990] for a classic treatment and Goeree and Yariv [2011] and Chan et al. [2017] for recent treatments).

In this section we consider the simplest learning dynamics (best responses). Since there may multiple equilibria in the game-form associated to a strike mechanism, there is no hope that the synchronous best response dynamics converge necessarily. If  $(m_1, m_2)$  and  $(m'_1, m'_2)$  are two different equilibria then the sequence  $(m_1, m'_2), (m'_1, m_2), (m_1, m'_2), (m'_1, m_2), \dots$  is such that each player best-responds to her opponent's previous moves, but they never coordinate (This remark is very general: it holds for any two player game with multiple equilibria). We thus consider alternate

best response dynamics and show that these processes lead to our equilibria. That point underlines the relevance of our mechanisms in applied settings.

**Alternate best responses dynamics.** Let  $m^0, m^1, m^2, m^3, \dots$  be a sequence of messages from alternating players. Say, for instance, and without loss of generality, that player 2 plays  $m^0, m^2$ , etc. Suppose that for any  $t \geq 1$ ,  $m^t$  is a best response (for player 1 if  $t$  is odd and for player 2 if  $t$  is even) to  $m^{t-1}$ . So  $\#m^t$  is equal to  $v_1$  for  $t$  odd and to  $v_2$  for  $t$  even.

First notice that, thanks to our strict preferences assumption (**BEB**), best responses are unique. Precisely, when player  $i$  is facing a veto on the  $v_j$  alternatives  $m^{t-1}$ , her best response is to pick her unique preferred alternative among the remaining set  $A \setminus m^{t-1}$  and to veto the other  $v_i$  alternatives. Thus the whole sequence is uniquely defined by its first element  $m^0$ , and we have, for any  $t \geq 1$ :

$$m^{t-1} \cap m^t = \emptyset. \quad (3)$$

Let  $r_t$  for  $t \geq 1$  denotes the outcome at date  $t$ ; this is the unique alternative such that:

$$m^{t-1} \cup m^t \cup \{r_t\} = A.$$

By definition, both  $m^t$  and  $m^{t+2}$  contain  $v_j$  alternatives. However, as previously mentioned,  $m^t$  and  $m^{t+1}$  are disjoint, and so are  $m^{t+1}$  and  $m^{t+2}$ . Therefore, both  $m^t$  and  $m^{t+2}$  contain  $v_j$  alternatives from the set  $A \setminus m^{t+1}$ , which contains  $n - v_i$  alternatives. Thus, since  $v_i + v_j = n - 1$ ,  $m^t$  and  $m^{t+2}$  differ on at most one alternative. If  $m^t = m^{t+2}$ , an equilibrium is reached. If  $m^t \neq m^{t+2}$  then  $m^t$  and  $m^{t+2}$  differ on one alternative exactly.

The following property of the best response correspondence is used in our proof of convergence. Suppose that one alternative, say  $a$ , is erased from the set  $A$ . In case  $a \in m^{t-1}$ , the best response to  $\tilde{m}^{t-1} = m^{t-1} \setminus \{a\}$  is the same  $m^t$ . In case  $a \notin m^{t-1}$  and  $a \in m^{t+1}$ , the best response to  $\tilde{m}^{t-1} = m^{t-1}$  is  $\tilde{m}^t = m^t \setminus \{a\}$ , and is the best response of the same player in the modified game where  $a$  is not available and the player has one veto less.

We now prove that the sequence of best responses leads to an equilibrium in at most  $n$  iterations. Let  $a_n$  denote the worst alternative for player 1. If for some  $k \geq 0$ ,  $a_n \notin m^{2k}$ , then the best response for player 1 implies to veto  $a_n$ , that is:  $a_n \in m^{2k+1}$ . This in turn implies (because of property 3) that player 2 does not veto  $a_n$  at date  $2k + 2$ . It follows that the following chain holds: for all  $t \geq 0$

$$a_n \notin m^t \implies a_n \in m^{t+1} \implies a_n \notin m^{t+2} \implies \dots$$

Consequently,  $a_n$  belongs either to all  $m^t$  for  $t$  odd and starting at 1 (call this case 1), or to all  $m^t$  for  $t$  even and starting at 2 (call this case 2).

Now consider the sequence of sets  $\tilde{m}^t = m^t \setminus \{a_n\}$  for all  $t \geq 1$ . We claim that this new sequence is again a sequence of alternating best responses in the game where the set of alternatives is  $A \setminus \{a_n\}$  and the numbers of vetoes are, in case 1,  $v' = (v_1 - 1, v_2)$  and in case 2,  $v' = (v_1, v_2 - 1)$ . This is true in case 1 because, in the original sequence, player 1 always had to veto  $a_n$  that is her worst alternative and player 2 never had to block  $a_1$  that is never available to her. This is also true in case 2 because, in the original sequence, player 1 never had to veto  $a_n$  that was never available to her, and player 2 always had to veto  $a_n$ .

The same logic applies to the worst element for the other player as well. The argument can be repeated for player 1 or for player 2 until all vetoes are exhausted and about the sequences starting at  $m^1$  then at  $\tilde{m}^2$ , then at  $\tilde{\tilde{m}}^3$ , etc. It follows that in the original sequence, for all  $t \geq n$ ,  $m^t = m^{t+2}$ . We conclude that the iterative process of alternate best responses converges to an equilibrium in at most  $n$  iterations.

## 6 Ex-ante Pareto efficiency

This section shows that ensuring ex-ante Pareto efficient equilibria through **DE** mechanisms is in general not possible. It presents two separate results for two notions: ex-ante efficiency for mechanisms (Section 6.1) and for SCRs (Section 6.2). The first one shows that no ex-ante Pareto efficient admissible mechanism ensures minimal veto rights to each player. The second one proves that any ex-ante Pareto efficient

and implementable SCR is a dictatorship.

## 6.1 Ex-ante efficient mechanisms

Ex-ante efficiency means that efficiency is observed at the level of lotteries, before their realization. Received knowledge on this issue (see, for instance, Börgers and Postl [2009]) highlights that ex-ante efficiency is difficult to obtain. An example is published (Núñez and Laslier [2015]) of an ex-ante Pareto efficient two-player mechanism for three alternatives; this mechanism, called Approval mechanism, is not **DE** and fails to be efficient for four alternatives or more. The existence of a non-**DE** efficient mechanism for many alternatives remains an open problem.

The difficulty can be described by the following argument. Let  $A = \{a, b, c\}$  with  $a \succ_1 b \succ_1 c$  and  $c \succ_2 b \succ_2 a$ . Consider the strike mechanism that gives one veto to each player. If the domain  $\kappa$  satisfies **BEB**, the unique equilibrium outcome is  $b$ . Now, assume that both players prefer a non degenerate lottery with support  $\{a, c\}$  to the pure outcome  $b$ . This is the case when both players extend their preference over alternatives to uniform lotteries through expected utility and their intensity of preference for  $b$  is low. In this case, the unique equilibrium outcome is Pareto dominated by a lottery, that is a possible outcome of the mechanism, therefore non dictatorial ex-ante Pareto efficiency cannot be reached with deterministic outcomes. Our first result is a negative result, that generalizes this observation to veto rules, as studied in this paper.

Instead of social choice rules, defined on profiles of preferences over pure alternatives, we are here dealing with **social lottery rules** (SLR), defined on profiles of preferences over lotteries. For such a preference profile,  $\succeq^*$ , the SLR  $F$  defines a set of lotteries  $F(\succeq^*) \subseteq \Delta$ . We will consider SLRs that are defined on the same domains that were used in the previous sections: preferences over pure alternatives are strict, and all strict preferences are admitted, and the preferences on lotteries are described by a product correspondence  $\kappa$ .

For a mechanism  $\mu$  and a profile of preferences over alternatives  $\succeq^*$ , let  $F_\mu(\succeq^*)$  denote the set of **Nash outcomes**:  $F_\mu(\succeq^*) = \{\mu(m) : m \in \mathcal{N}_\mu(\succeq^*)\}$ . This is a subset of



$\mu(\mathcal{M})$ , the range of  $\mu$ . A mechanism  $\mu$  is **ex-ante Pareto efficient** on the domain  $\kappa$  if for any  $\succ \in \mathcal{L}_A^2$  and any  $\succeq^* \in \kappa(\succ)$  there is no  $p \in \mu(\mathcal{M})$  and  $q \in F_\mu(\succeq^*)$  such that  $p \succeq_i^* q$  for all  $i$  with at least one strict preference. Say that  $\mu$  is **DE** at  $\succeq^*$  if all its Nash outcomes are deterministic, that is, with our loose notation:  $F_\mu(\succeq^*) \subseteq A$ . A mechanism  $\mu$  is a **dictatorship** iff there is some  $i \in N$  such that for each  $x \in A$ , there exists  $m_i \in \mathcal{M}_i$  such that  $\mu(m_i, m_j) = \{x\}$  for all  $m_j \in \mathcal{M}_j$ .

**Theorem 3.** *Let the domain  $\kappa$  satisfy **PREX** and **WEB**. On  $\kappa$ , any admissible **DE** mechanism that is ex-ante Pareto-efficient is a dictatorship.*

*Proof.* Suppose first that the mechanism  $\mu$  is not purely deterministic, that is there exists a strategy combination  $m^* \in \mathcal{M}$  and two distinct alternatives  $a_1, a_{n+1} \in A$  such that  $\{a_1, a_{n+1}\} \subseteq \text{SUPP}(\mu(m^*))$ . Since  $\mu$  is **DE**, it follows that  $F^\mu(\succeq^*) \subseteq A$ . Write  $A = \{a_1, a_2, \dots, a_{n+1}\}$  (recall that  $n+1 \geq 3$ ) and consider the opposed preferences  $\succ = (\succ_1, \succ_2)$  with  $a_1 \succ_1 a_2 \succ_1 \dots \succ_1 a_{n+1}$  and  $a_{n+1} \succ_2 a_n \succ_2 \dots \succ_2 a_1$ . Let the players' preferences over lotteries,  $\succ_1^*$  and  $\succ_2^*$ , be such that, for any  $p, q \in \mu(\mathcal{M})$ , if  $p[\cdot \succeq_1 a_1] > 0$  and  $q[\cdot \succeq_1 a_1] = 0$  then  $p \succ_1^* q$ , and if  $p[\cdot \succeq_2 a_{n+1}] > 0$  and  $q[\cdot \succeq_2 a_{n+1}] = 0$  then  $p \succ_2^* q$ . Such a profile exists because the domain  $\kappa$  satisfies **PREX**. Therefore, since  $\{a_1, a_{n+1}\} \subseteq \text{SUPP}(\mu(m^*))$ ,  $\mu(m^*) \succ_i^* \{x\}$  for  $i = 1, 2$  and any  $x \neq a_1, a_{n+1}$ . Since  $\mu$  is ex-ante Pareto efficient, it follows that  $F^\mu(\succeq^*) \subseteq \{a_1, a_{n+1}\}$ .

Therefore, at this profile, the mechanism  $\mu$  admits either  $\{a_1\}$  or  $\{a_{n+1}\}$  or both as equilibrium outcome. Assume w.l.o.g. that  $\mu$  admits some equilibrium  $\tilde{m}$  with  $\mu(\tilde{m}) = \{a_1\}$ . By definition of equilibrium,  $\{a_1\} = \mu(\tilde{m}) \succeq_2^* \mu(\tilde{m}_1, m'_2)$  for any  $m'_2 \in \mathcal{M}_2$ . Yet, since  $\kappa$  satisfies **WEB**, then  $p \succ_2^* \{a_1\}$  for all  $p \in \mu(\mathcal{M})$  with  $p \neq \{x\}$ . Therefore,  $\mu(\tilde{m}_1, m'_2) = \{a_1\}$  for any  $m'_2 \in \mathcal{M}_2$ . It follows that for every  $a \in A$ , there is either some  $m_1 \in \mathcal{M}_1$  such that  $\mu(m_1, m'_2) = \{a\}$  for any  $m'_2 \in \mathcal{M}_2$  or some  $m_2 \in \mathcal{M}_2$  such that  $\mu(m'_1, m_2) = \{a\}$  for any  $m'_1 \in \mathcal{M}_1$ . It follows that either for every  $a \in A$  there is some  $m_1 \in \mathcal{M}_1$  such that  $\mu(m_1, m'_2) = \{a\}$  for any  $m'_2 \in \mathcal{M}_2$  or for every  $a \in A$  there is some  $m_2 \in \mathcal{M}_2$  such that  $\mu(m'_1, m_2) = \{a\}$  for any  $m'_1 \in \mathcal{M}_1$ . In the first case, player 1 is the dictator and in the second case player 2 is the dictator.

Suppose now that there is no  $m^* \in \mathcal{M}$  such that  $\text{SUPP}(\mu(m^*)) \supset \{x, y\}$  for some pair of alternatives  $\{x, y\} \in A$ . Then, for any  $m \in \mathcal{M}$ ,  $\mu(m) \in A$  so that  $\mu$  is a determinis-

tic mechanism and  $\mu(\mathcal{M}) \subseteq A$ . Thus, for any  $\mathcal{N}^\mu(\geq^*) = \mathcal{N}^\mu(\geq)$  for any  $\geq^* \in \kappa(\succ)$  and any  $\succ \in \mathcal{L}_A^2$ . Hence, ex-ante Pareto efficient is equivalent to Pareto efficiency w.r.t.  $\succ$ . Thus, the two-person implementation problem (as stated by Hurwicz and Schmeidler [1978] and Maskin [1999]) applies: the only mechanisms that are admissible and Pareto efficient are dictatorships.  $\square$

## 6.2 Ex-ante efficiency of implementable social choice rules

The literature on implementation has concentrated on social choice rules (SCRs) which, by definition use only cardinal information: a preference profile  $\succ$  on alternatives, and not a preference profile  $\geq^*$  over lotteries. Since we consider mechanisms that can outcome lotteries, some definitions are useful in order to make the link with this literature.

So consider a SCR  $f$ : for all  $\succ \in \mathcal{L}(A)$ ,  $f(\succ) \subseteq A$ . A mechanism  $\mu$  that is DE on a domain  $\kappa$  is said to implement the SCR  $f$  on  $\kappa$  iff:

$$\forall \succ \in \mathcal{L}(A), \forall \geq^* \in \kappa(\succ), F_\mu(\geq^*) = f(\succ).$$

Note that, for a mechanism to implement a social choice rule, it is required that the outcomes of the mechanism not only are deterministic, but also are independent of the precise preferences over lotteries. The following definition presents a concept of ex-ante Pareto efficient SCR that is suitable for the study of the implementation of SCRs by mechanisms that can output lotteries. It should not be confused with the concept of an ex-ante Pareto efficient mechanism defined above.

Given a set of lotteries  $\bar{\Delta} \subseteq \Delta$  a SCR  $f$  is **ex-ante Pareto efficient in the range**  $\bar{\Delta}$  iff given any  $\succ \in \mathcal{L}_A^2$  and any  $\geq^*$  in  $\kappa(\succ)$ , any  $X \in f(\succ)$  and any  $x \in X$ , there is no  $p \in \bar{\Delta}$  such that  $p \succeq_i^* x$  for all  $i \in N$  with at least one strict preference.

We show that the notions of ex-ante Pareto efficiency and admissibility clash, hence extending the two-player implementation problem to the setting with lotteries and **DE** mechanisms. This shows that ex-ante Pareto efficiency is too restrictive in our setting.

**Theorem 4.** *Let  $f$  be a SCR that is Nash-implementable by a **DE** mechanism  $\mu$  on a domain  $\kappa$ . Suppose that  $\kappa$  satisfies **PREX** and **WEB** in the range of  $\mu$ . If  $f$  is ex-ante Pareto efficient in the range of  $\mu$ , then  $\mu$  is a dictatorship.*

*Proof.* Let  $f$  be an ex-ante Pareto efficient SCR that is Nash-implementable by a **DE** mechanism  $\mu$  on a domain  $\kappa$ . Borrowing the vocabulary of Hurwicz and Schmeidler [1978], we think of the mechanism  $\mu$  as a matrix where player 1 controls rows and player 2 controls columns. We hence write, for every  $x \in X$ , an  $\{x\}$ -row is a row that contains only  $\{x\}$  as an outcome and similarly for an  $\{x\}$ -column.

Take any profile  $\succ \in \mathcal{L}_A^2$ . Let  $a$  and  $b$  respectively denote the best outcomes for player 1 and 2 at  $\succ$ . Take  $\succ^* \in \kappa(\succ)$  such that  $p \succ_1^* q$  for all  $p, q \in \mu(\mathcal{M})$  with  $p(a) > 0$  and  $q(a) = 0$  and such that  $p \succ_2^* q$  for all  $p, q \in \mu(\mathcal{M})$  with  $p(b) > 0$  and  $q(b) = 0$ . The existence of  $\succ^*$  is ensured by **PREX**. Take any alternative  $x \neq a, b$ . According to  $\succ^*$  both players strictly prefer a lottery with support  $\{a, b\}$  to the pure alternative  $x$ . Ex-ante Pareto efficiency thus implies that  $x \notin f(\succ)$ . Indeed, if  $x \in f(\succ)$ , then  $\mu$  admits an equilibrium  $m^*$  with  $\mu(m^*) = \{x\}$  (since  $\mu$  is **DE**). However, both players prefer the lottery  $\{a, b\}$  to  $x$ , contradicting ex-ante Pareto efficiency. So  $f(\succ) \subseteq \{a, b\}$ . Thus, an ex-ante Pareto optimal and admissible **DE** mechanism gives equilibrium outcomes from the union of tops.

Now consider a preference profile  $\succ$  where the players' preferences are completely opposed. Relabel the alternatives as  $a_1, a_2, \dots, a_m$ . Take a preference profile where  $a_1$  and  $a_2$  are respectively the best and last alternatives for player 1 while  $a_2$  and  $a_1$  are, respectively, the best and last alternatives for player 2. So the equilibrium outcomes of  $\mu$  belong to  $\{a_1, a_2\}$ . Note that  $\mu$  is **DE**, so no lottery with support  $\{a_1, a_2\}$  is an equilibrium outcome. Let, without loss of generality,  $a_1$  be an equilibrium outcome. This is the worst element for player 2 and also the worst lottery (due to **WEB**), hence player 1 must have an  $\{a_1\}$ -row.

Now, take a preference profile where  $a_2$  and  $a_3$  are, respectively, the best and last alternatives for player 1 while  $a_3$  and  $a_2$  are the respective top and bottom outcomes for player 2. So the equilibrium outcomes of  $\mu$  belong to  $\{a_2, a_3\}$ . We first show that  $a_3$  cannot be an equilibrium outcome. Suppose it is. As  $a_3$  is the worst element

and lottery for player 1, player 2 must have an  $\{a_3\}$ -column, due to **WEB**, which contradicts player 1 has an  $\{a_1\}$ -row. As a result,  $a_2$  is an equilibrium outcome and we argue, *mutatis mutandis*, player 1 has an  $a_2$ -row.

Iterate by making the arguments for  $a_3, a_4, \dots, a_{m-1}, a_m$ , proves that for each  $a \in A$ , player 1 has an  $\{a\}$ -row, showing that player 1 is the dictator. Repeating the argument assuming that  $a_2$  is an equilibrium outcome shows that player 2 is the dictator.  $\square$

## 7 Review of the literature

This section provides a short review of the two-player implementation problem (see Dutta [2019] for a recent and complete survey). As argued in the introduction, the pioneering works (Hurwicz and Schmeidler [1978] and Maskin [1999]) provide a provocative result: dictatorships are the only Pareto efficient rules that can be Nash implemented. Their proof builds on three key assumptions: (i) the preference domain is universal (any preference profile is allowed) while implementing mechanisms are (ii) simultaneous and (iii) deterministic.

The literature has explored the consequences of weakening each of these assumption.<sup>13</sup> The first strand relaxes condition (i), Dutta and Sen [1991] and Moore and Repullo [1990] are the key papers in this direction. They identify the domain restrictions under which one can design Pareto efficient and non-dictatorial Nash-implementable rules. While the full characterization is rather complex, the sufficient domain conditions for implementation often rely in the Euclidean space (see Section 5 in Dutta and Sen [1991] for instance); in the current work, we work in a setting where we do not impose any structure on the alternatives or on the preferences over them, beyond the fact that that preferences over alternatives are strict.

A second strand is concerned with (ii), that is, replacing simultaneous with dynamic mechanisms. This literature, in which Moore and Repullo [1988], Abreu and Sen [1991] and Herrero and Srivastava [1992] play a key role, shows that introducing an order of play expands the set of implementable rules with more than two players.

---

<sup>13</sup>Other approaches have modified the rationality notion, using “partial honesty”; see Dutta and Sen [2012] among others.

No characterization of implementable rules via subgame-perfect or via backward induction is available. By altering the notion of implementation (role-robust implementation), De Clippel et al. [2014]<sup>14</sup> show that a possibility arises with dynamic vetoes and randomized order of play (see also Barberà and Coelho [2019] who consider the implementation of the fallback-bargaining solution). However, while ex-ante fairness is achieved by randomizing the order of play, ex-post fairness fails. The order of play matters for determining the outcome, creating first, or second, mover advantages. As Moulin [1981] puts it, "voting by veto procedures introduce a strong asymmetry among agents: ... the ordering of the agents has a strong influence on the outcome" (see also Barberà and Coelho [2018] on first/second mover advantages).

The third and final strand of the literature deals with assumption (*iii*), as does the current work: it explores the consequences of modifying the type of mechanisms jointly with the notion of implementation.<sup>15</sup> Indeed, virtual implementation is a reformulation of the original problem: a social choice rule is virtually implementable if there exists a game form  $G$ , such that for all preference profiles  $G$  admits a unique equilibrium outcome (a lottery) which is  $\varepsilon$ -close to the outcome prescribed by the rule at this preference profile and this holds for every  $\varepsilon > 0$ . Following this approach, Matsushima [1988] and Abreu and Sen [1991] provide a strong possibility result: with at least three players, any rule is implementable. With two players, the result is more nuanced but some SCRs are virtually implementable (among which the Pareto-and-veto rule described in the current work). However, under the virtual implementation approach, "any alternative can be the outcome of the game as it receives positive probability in the equilibrium lottery" (Bochet and Maniquet [2010]). In other words, in order to virtually implement a social choice rule, one constructs game forms whose equilibrium outcome at every preference profile is a full-support

---

<sup>14</sup>A classic literature considers sequential voting by veto with many players (see Mueller [1978], Moulin [1981], Bloom and Cavanagh [1986a], Bloom and Cavanagh [1986b], Felsenthal and Machover [1992] and Anbarci [2006]) where each player is assigned a certain number of vetoes to be distributed freely among the alternatives. See also the rules of  $k$ -names (Barberà and Coelho [2010], Barberà and Coelho [2017]).

<sup>15</sup>See also the papers on approval voting with two players as Núñez and Laslier [2015] and Laslier et al. [2017]. See also Jackson and Sonnenschein [2007] who show that linking decisions (that is, a common decision on several independent problems) can help overcoming incentive constraints in Bayesian collective decision problems.

lottery, arbitrarily close to the outcome prescribed by the social choice rule. This represents a threat to the relevance of these solutions since it involves that socially undesirable alternative, even with a small probability, can be selected.

## 8 Concluding remarks

Strike mechanisms arise as a solution to the two-person implementation problem. This solution is obtained by altering two key elements of the classic framework: (i) considering mechanisms that allow in equilibrium pure alternatives and off equilibrium lotteries and (ii) restricting efficiency to the ex-post Pareto notion.

Our class of **DE** mechanisms is a simultaneous version of the dynamic veto mechanisms (see Moulin [1981]) which, by allowing off-equilibrium set-valued outcomes, resolve the unfairness generated by dynamic mechanisms. To see the difference between our solution and the one based on dynamic veto mechanisms, consider a dynamic game that allows player 1 to veto  $n+1-k$  alternatives and player 2 to veto  $k-1$  of the remaining  $k$  alternatives, where  $k \in \{1, \dots, n+1\}$ . At each preference profile  $\succ$ , the subgame perfect equilibrium outcome of this game is the most preferred alternative of player 1 among  $\text{pv}_v(\succ)$  where  $v_1 = n+1-k$  and  $v_2 = k-1$ . In other words, this dynamic veto mechanism subgame perfect implements a subcorrespondence of  $\text{pv}_v$  by refining it with respect to the true preference of the first mover. One could argue that fairness here could be achieved by selecting randomly the first-mover. Yet, this needs qualification since this randomization prevents some alternatives to arise as the following example shows. When  $A = \{a, b, c, d, e\}$ , at the preference profile  $a \succ_1 b \succ_1 c \succ_1 d \succ_1 e$  and  $c \succ_2 b \succ_2 a \succ_2 d \succ_2 e$ , the dynamic veto mechanism which gives 2 vetoes to each voter implements, by alternating first movers, either  $a$  or  $c$  but excludes  $b$ . However,  $\text{pv}_v$  picks all three of  $a$ ,  $b$  and  $c$ . Thus, our simultaneous direct veto mechanisms allow for the implementation of the compromise alternative  $b$  whereas their dynamic counterparts fail to do so. This constitutes a strong argument in favor of using simultaneous mechanisms.

We close by noting three limitations of our analysis. First, it is restricted to Nash

implementation in pure strategies. Allowing for mixed strategies and exploring the existence of interesting **DE** mechanisms for settings with two or more players seems to be a promising research avenue (see Mezzetti and Renou [2012]). Second, the set of implementable SCRs expands if one considers implementation through non-**DE** mechanisms. Indeed, as long as **BEB** holds, the game-form associated to plurality rule Nash implements the union of tops<sup>16</sup> which selects at each preference profile all alternatives that are top-ranked by at least one player.<sup>17</sup> Third, we have considered implementation through ex-post Pareto efficient **DE** mechanisms. Other notions of efficiency are present in the literature such as stochastic dominance. Whether other SCRs can be Nash implemented through **DE** mechanisms by considering different notions of efficiency remains to be explored.

## References

- D. Abreu and A. Sen. Virtual implementation in Nash equilibrium. *Econometrica*, 59(4):997–1021, 1991.
- N. Anbarci. Finite alternating-move arbitration schemes and the equal area solution. *Theory and decision*, 61(1):21–50, 2006.
- S. Barberà and D. Coelho. On the rule of  $k$ -names. *Games and Economic Behaviour*, 70:44–61, 2010.
- S. Barberà and D. Coelho. Balancing the power to appoint officers. *Games and Economic Behavior*, 101:189–203, 2017.
- S. Barberà and D. Coelho. On the advantages and disadvantages of being the first mover under rules of  $k$  names. *International Journal of Economic Theory*, 14(1): 51–60, 2018.

---

<sup>16</sup>See Yeh [2008] for an axiomatization of this rule.

<sup>17</sup>In this game form, each player announces a single alternative and one of them is selected randomly. Since it is a dominant strategy to announce one's best alternative, this mechanism is not **DE** as we may have several alternatives selected with positive probability in equilibrium.

- S. Barberà and D. Coelho. On the Selection of Compromise Arbitrators. mimeo, UAB, 2019.
- J.-P. Benoît and E.A. Ok. Nash implementation without no-veto power. *Games and Economic Behavior*, 64(1):51–67, 2008.
- D. Bloom and C. Cavanagh. An analysis of the selection of arbitrators. *American Economic Review*, 76(3):408–22, 1986a.
- D. Bloom and C. Cavanagh. *An analysis of alternative mechanisms for selecting arbitrators*. Number 1224. Harvard Institute of Economic Research, Harvard University, 1986b.
- O. Bochet. Nash implementation with lottery mechanisms. *Social Choice and Welfare*, 28(1):111–125, 2007.
- O. Bochet and F. Maniquet. Virtual Nash implementation with admissible support. *Journal of mathematical economics*, 46(1):99–108, 2010.
- A. Bogomolnaia and H. Moulin. A new solution to the random assignment problem. *Journal of Economic theory*, 100(2):295–328, 2001.
- T. Börgers and P. Postl. Efficient Compromising. *Journal of Economic Theory*, 144: 2057–2076, 2009.
- J. Chan, A. Lizzeri, W. Suen, and L. Yariv. Deliberating collective decisions. *The Review of Economic Studies*, 85(2):929–963, 2017.
- G. De Clippel, K. Eliaz, and B. Knight. On the selection of arbitrators. *American Economic Review*, 104:3434–58, 2014.
- B. Dutta. Recent results on implementation with complete information. In *Social Design*, pages 249–260. Springer, 2019.
- B. Dutta and A. Sen. A Necessary and sufficient condition for two-person Nash implementation. *Review of Economic Studies*, 58:121–128, 1991.



- B. Dutta and A. Sen. Nash Implementation with Partially Honest Individuals. *Games and Economic Behavior*, 74:154–169, 2012.
- S.D. Felsenthal and M. Machover. Sequential voting by veto: making the mueller-moulin algorithm more versatile. *Theory and Decision*, 33(3):223–240, 1992.
- D. Fudenberg and D.K. Levine. Whither game theory? towards a theory of learning in games. *Journal of Economic Perspectives*, 30(4):151–70, 2016.
- P. Gärdenfors. Manipulation of social choice functions. *Journal of Economic Theory*, 13(2):217–228, 1976.
- J.K. Goeree and L. Yariv. An experimental study of collective deliberation. *Econometrica*, 79(3):893–921, 2011.
- M.J. Herrero and S. Srivastava. Implementation via backward induction. *Journal of Economic Theory*, 56(1):70–88, 1992.
- L. Hurwicz and D. Schmeidler. Construction of outcome functions guaranteeing existence and Pareto optimality of Nash equilibria. *Econometrica*, 46:1447–1474, 1978.
- M.O. Jackson. A Crash Course in Implementation Theory. *Social Choice and Welfare*, 18:655–708, 2001.
- M.O. Jackson and H.F. Sonnenschein. Overcoming incentive constraints by linking decisions. *Econometrica*, 75(1):241–257, 2007.
- J.S Kelly. Strategy-proofness and social choice functions without singlevaluedness. *Econometrica*, 45(2):439–446, 1977.
- J.-F. Laslier, M. Núñez, and C. Pimienta. Reaching consensus through approval bargaining. *Games and Economic Behavior*, 104:241–251, 2017.
- E. Maskin. Nash equilibrium and welfare optimality. *Review of Economic Studies*, 66: 23–38, 1999.

- H. Matsushima. A new approach to the implementation problem. *Journal of Economic Theory*, 45(1):128–144, 1988.
- C. Mezzetti and L. Renou. Implementation in mixed Nash equilibrium. *Journal of Economic Theory*, 147(6):2357–2375, 2012.
- J. Moore and R. Repullo. Subgame perfect implementation. *Econometrica*, pages 1191–1220, 1988.
- J. Moore and R. Repullo. Nash implementation: A full characterization. *Econometrica*, 58:1083–1099, 1990.
- H. Moulin. Prudence versus sophistication in voting strategy. *Journal of Economic theory*, 24(3):398–412, 1981.
- D.C. Mueller. Voting by veto. *Journal of Public Economics*, 10(1):57–75, 1978.
- M. Núñez and J.-F. Laslier. Bargaining through approval. *Journal of Mathematical Economics*, 60:63–73, 2015.
- I. Özkal-Sanver and M.R. Sanver. Nash implementation via hyperfunctions. *Social Choice and Welfare*, 26(3):607–623, 2006.
- M. R. Sanver. Nash implementing non-monotonic social choice rules by awards. *Economic Theory*, 28(2):453–460, 2006.
- M. R. Sanver. Implementing Pareto optimal and individually rational outcomes by veto. *Group Decision and Negotiation*, 27(2):223–233, 2018.
- D. Schmeidler. Subjective probability and expected utility without additivity. *Econometrica*, 57(3):571–587, 1989.
- E. van Damme. *Stability and perfection of Nash equilibria*. Springer-Verlag, 2nd edition, 1991.
- J.B. Van Huyck, R.C. Battalio, and R.O. Beil. Tacit coordination games, strategic uncertainty, and coordination failure. *The American Economic Review*, 80(1):234–248, 1990.

C.-H. Yeh. An efficiency characterization of plurality rule in collective choice problems. *Economic Theory*, 34(3):575–583, 2008.