



**HAL**  
open science

# On the subgame perfect implementability of voting rules

Matias Nunez, M. Remzi Sanver

► **To cite this version:**

Matias Nunez, M. Remzi Sanver. On the subgame perfect implementability of voting rules. *Social Choice and Welfare*, 2020, 56, pp.421-441. 10.1007/s00355-020-01293-9 . hal-03092402

**HAL Id: hal-03092402**

**<https://enpc.hal.science/hal-03092402>**

Submitted on 17 Sep 2021

**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.

# ON THE SUBGAME PERFECT IMPLEMENTABILITY OF VOTING RULES

MATÍAS NÚÑEZ<sup>a</sup> AND M. REMZI SANVER<sup>b</sup>

ABSTRACT. Abreu and Sen [1990] provide a necessary condition, called Condition  $\alpha$ , which is almost sufficient for a social choice rule to be implementable via subgame perfect equilibria. Yet, it is not straightforward to check the satisfaction of Condition  $\alpha$ . We contribute in this direction by establishing a nuanced picture over the subgame perfect implementability of compromise rules, as a function of the compromise threshold. This contrasts with scoring rules that all fail to be subgame perfect implementable and with several Condorcet rules which are subgame perfect implementable.

## 1. INTRODUCTION

The analysis of social choice rules which can be implemented via subgame perfect equilibria of extensive form mechanisms goes back to the characterization of Moore and Repullo [1988].<sup>1</sup> Their sufficient conditions involve domain restrictions and, for example, rule out classical voting environments, where every agent has strict preferences over a finite set of alternatives. The results of Moore and Repullo [1988] are extended by Abreu and Sen [1990] who provide a necessary condition, called Condition  $\alpha$ , which is also sufficient when combined with a weak no veto power condition. Condition  $\alpha$  involves no domain restriction and is analogous to the monotonicity condition of Maskin [1999] for Nash implementability. More recently, Vartiainen [2007a] gives a full characterization by identifying Condition  $\alpha^*$ , a strengthening of Condition  $\alpha$ , which is necessary and sufficient for the subgame perfect implementability of unanimous social choice rules.

Although Conditions  $\alpha$  and  $\alpha^*$  cover voting environments, it is not always straightforward to check whether a given voting rule satisfies these conditions. In fact, the literature is not very rich in analyzing subgame perfect implementability of voting rules by checking the satisfaction of Conditions  $\alpha$  and  $\alpha^*$ . Abreu and Sen [1990] shows that the plurality rule is not subgame perfect implementable whereas no scoring rule is subgame perfect

---

*Date:* September 21, 2020.

<sup>a</sup> CNRS & CREST, ECOLE POLYTECHNIQUE, UMR [9194], 91120 PALAISEAU, FRANCE.

<sup>b</sup> UNIVERSITÉ PARIS-DAUPHINE, PSL RESEARCH UNIVERSITY, CNRS, UMR [7243], LAMSADE, 75016 PARIS, FRANCE.

\* We thank Arunava Sen and Hannu Vartiainen for very fruitful exchanges and two anonymous reviewers for very useful comments. This project has been supported by the ANR-14-CE24-0007-01 (CoCoRIC-oCoDEC) and the project IDEX ANR-10-IDEX-0001-02 PSL MIFID. .

<sup>1</sup>A social choice rule  $f$  is subgame perfect implementable if there exists an extensive form mechanism  $\Gamma$  such that the set of subgame perfect equilibrium outcomes of  $\Gamma$  coincides with the outcome of  $f$  for any possible preference profile. As we do not use the definition throughout the paper, we do not give a formal expression, which can be found in Abreu and Sen [1990].

implementable as proven by Sen [1987]. Another work in this direction is by Sertel and Yılmaz [1999], which shows that the majoritarian compromise -a voting rule which we discuss in the sequel- satisfies condition  $\alpha$  and weak no veto power, hence being subgame perfect implementable. In Vartiainen [2007a], an example shows that plurality with a runoff is subgame perfect implementable over a given restricted domain (p.122) which does not enable us to derive a conclusion about what happens over the full domain.<sup>2</sup>

An alternative (and fruitful road) to check subgame-perfect implementability is to focus on implementation via backward induction, a more stringent requirement. An early work in this direction is Moulin [1986] who proposes a sophisticated agenda algorithm for tournament solutions. This path is more precisely followed by Herrero and Srivastava [1992] and Dutta and Sen [1993] which are able to show that a large class of Condorcet rules (such as selections from the uncovered set) are implementable via backward induction, from which we can deduce their subgame perfect implementability. These conditions are coined in terms of preference reversals (Herrero and Srivastava [1992]) or in terms of outcomes of binary voting procedures (Dutta and Sen [1993]) and therefore are not designed to be easy to check (see Horan [2013] for a review).

The appeal of all these analyses is self-justified, given the scarcity of voting rules which can be implemented via Nash equilibria. In fact, we know by Maskin [1999] the necessity of a certain monotonicity condition for social choice rules to be Nash implementable. However, Maskin monotonicity is demanding and violated by several well-known voting rules.<sup>3</sup> As Condition  $\alpha$  is weaker than Maskin monotonicity, determining voting rules which are subgame perfect implementable appears as a question of particular interest.

As a matter of fact, the subgame perfect implementability of several well-known voting rules, such as point runoff systems or compromise rules seems to remain an open question.<sup>4</sup> This situation is in contrast to Nash implementability where we know, arguably, for every voting rule whether it is Nash implementable or not. This contrast can be explained by the relative simplicity of Maskin monotonicity compared to Condition  $\alpha$ .

We aim to draw a broad picture on the subgame perfect implementability of voting rules. Our first set of results are negative. We introduce a weak and simple necessary condition, which we call Condition  $\beta$ , for subgame perfect implementability. Condition  $\beta$  applies just to profiles in which some alternative, say  $x$ , is ranked either first or last

<sup>2</sup>We show, nevertheless, in Section 3 that plurality with a runoff is not subgame perfect implementable over the full domain of linear orders.

<sup>3</sup>When singleton-valuedness is imposed, only dictatorial social choice rules are Maskin monotonic (Muller and Satterthwaite [1977]). When singleton-valuedness is relaxed and indifferences in individual preferences are ruled out, scoring rules (Erdem and Sanver [2005]), Condorcet consistent social choice rules (Jackson [2001], Özkal-Sanver and Sanver [2010]) and the majoritarian compromise (Sertel and Yılmaz [1999]) fail Maskin monotonicity. When indifferences in individual preferences are allowed, no Pareto optimal social choice rule is Maskin monotonic (Aşan and Sanver [2006]).

<sup>4</sup>We know more about subgame perfect implementability when randomization is allowed. Vartiainen [2007b] concludes about the subgame perfect implementability of several voting rules by allowing randomized mechanisms. In Section 6, we compare our results to those in this randomized environment.

by each of the voters. Take any such pair of profiles  $P$  and  $P'$  in which the set of voters who rank  $x$  first at  $P$  coincides with the set of voters who rank  $x$  first at  $P'$  (hence this also applies to the set of voters who rank  $x$  last). Condition  $\beta$  implies that  $x$  is selected by the rule at  $P$  if and only if it is selected by the rule at  $P'$ . Condition  $\beta$  is implied by Condition  $\alpha$ . However, in spite of its simplicity and weakness, it has a bite: all scoring rules violate Condition  $\beta$ , hence fail to be subgame perfect implementable. We also show that although plurality with a runoff is subgame perfect implementable when there are three alternatives only, this positive result does not extend to an arbitrary number of alternatives. In fact, we are able to show that no point runoff procedure is subgame perfect implementable when there are four or more alternatives. Our second set of results are positive: We consider Condorcet consistent rules and show that the top-cycle and many of its interesting refinements are subgame perfect implementable (in line with the previous results on backward induction). We find interesting to observe that the historical Borda - Condorcet dichotomy reappears when subgame perfect implementability of voting rules is considered. This observation is supported by the fact that the violation of Condition  $\beta$  reflects a strong violation of pairwise majoritarianism.

As a final set of results, we present a mixed picture regarding the subgame perfect implementability of a class of voting rules based on the trade-off between the number of voters behind an alternative (i.e., the quantity of support) and the rank of voter preferences at which this support is obtained (i.e., the quality of support). Merlin et al. [2019] propose a comprehensive classification of these rules as compromise rules and show that several voting rules, such as plurality, majoritarian compromise (Sertel and Yılmaz [1999]), median voting rule (Bassett and Persky [1999]), union of tops, fallback bargaining (Brams and Kilgour [2001]), Condorcet's practical method (Nurmi [1999]), can be expressed within this class. As a prominent subclass of compromise rules there is the family of  $q$ -approval fallback rules where  $q$  is an integer which can vary between 1 and the total number  $n$  of voters in the society. At any preference profile, among the alternatives that receive the support of  $q$  voters at the smallest rank  $r$ , those which receive the highest support at rank  $r$  are the  $q$ -approval fallback winners. Within this class of rules, we have the plurality rule when  $q = 1$ ; the majoritarian compromise when  $q = \lceil \frac{n}{2} \rceil$ ; and fallback bargaining when  $q = n$ . We again exploit Condition  $\beta$  to show that  $q$ -approval fallback rules fail to be subgame perfect implementable when  $q$  is less than a majority or  $q$  is between two thirds of the society and unanimity.<sup>5</sup> On the other hand, we show that when  $q$  is between a half and two thirds,  $q$ -approval fallback rules are subgame perfect implementable.<sup>6</sup> We find interesting to observe that  $q$ -approval fallback rules do

<sup>5</sup>This negative result covers the plurality rule which we already know from our Proposition 2 on scoring rules.

<sup>6</sup>This result includes the majoritarian compromise which we know by Theorem 1 of Sertel and Yılmaz [1999].

not seem to have a clear position within the Borda-Condorcet dichotomy, which is in turn reflected to their mixed behavior regarding their subgame perfect implementability.

Section 2 presents the basic notions and notation. Section 3 contains the negative results on the subgame perfect implementability of scoring rules and point runoff procedures. Section 4 contains the positive results on the subgame perfect implementability of Condorcet consistent rules. Section 5 presents the analysis on the subgame perfect implementability of compromise rules. Section 6 makes some final remarks.

## 2. BASIC NOTIONS AND NOTATION

Picking any two integers  $m \geq 3$ ,  $n \geq 2$ , we focus on a collective choice model where  $X$  with  $|X| = m$  is a set of alternatives and  $N = \{1, 2, \dots, n\}$  is a set of agents. Each agent  $i \in N$  is endowed with a preference relation  $P_i \in \mathcal{P}$  where  $\mathcal{P}$  is the set of all possible complete, transitive and antisymmetric binary relations on  $X$ . For each  $x \in X$  and each  $i \in N$ , we write  $L(x, P_i) = \{y \in X : x P_i y\}$  for the lower contour set of  $x$  at preference  $P_i$ . Note that  $x \in L(x, P_i)$  as  $P_i$  is reflexive. The rank of  $x$  at  $P_i$  equals  $r(x, P_i) = m + 1 - |L(x, P_i)|$ . Thus, the rank of the first ranked alternative equals 1, the rank of a second ranked alternative equals 2 and so on.

In the sequel,  $P \in \mathcal{P}^n$  denotes a (preference) profile. For each integer  $\bar{r} = 1, \dots, m$ , each  $x \in A$  and each  $P \in \mathcal{P}^n$ , we write  $n(x, \bar{r}, P) = \{i \in N \mid r(x, P_i) \leq \bar{r}\}$  for the number of voters for whom the rank of alternative  $x$  is lower than or equal to  $\bar{r}$  in the profile  $P$ . We call  $n(x, \bar{r}, P)$  the support of  $x$  at rank  $\bar{r}$  for profile  $P$ . Note that  $n(x, \bar{r}, P) \in \{1, \dots, n\}$  is non-decreasing on  $\bar{r}$  and  $n(x, m, P) = n$ .

For each  $q = 1, \dots, n$ , each  $x \in A$  and each  $P \in \mathcal{P}^n$ , we define

$$r^*(x, q, P) = \min\{r \in \{1, \dots, m\} \mid n(x, \bar{r}, P) \geq q\}.$$

as the minimal rank at which  $x$  obtains a support of at least  $q$  voters.

A social choice rule (SCR) is a correspondence  $f : \mathcal{P}^n \rightarrow A$  such that  $f(P)$  is a non-empty subset of  $X$ , for each  $P \in \mathcal{P}$ . We know from Abreu and Sen [1990] that the following condition is necessary for a SCR to be subgame perfect implementable<sup>7</sup>:

**Definition 1.** *An SCR  $f : \mathcal{P}^n \rightarrow X$  satisfies Condition  $\alpha$  iff for any  $P, P' \in \mathcal{P}^n$  and any  $x \in f(P) \setminus f(P')$ , there exist an integer  $l \geq 0$ , a sequence  $(i_k)_{k=0}^l$  in  $N$  and a sequence  $(a_k)_{k=0}^{l+1}$  in  $X$  with  $a_0 = x$  such that*

$$(\alpha_1) a_k P_{i_k} a_{k+1} \text{ for every } k \in \{0, 1, \dots, l\},$$

$$(\alpha_2) a_{l+1} P'_{i_l} a_l,$$

$$(\alpha_3) r(a_k, P'_{i_k}) > 1 \text{ for each } k \in \{0, 1, \dots, l-1\},$$

$$(\alpha_4) \text{ if } r(a_{l+1}, P'_{i_k}) = 1 \text{ for every } k \in \{0, 1, \dots, l-1\}, \text{ then either } l = 0 \text{ or } i_{l-1} \neq i_l.$$

<sup>7</sup>As we consider SCRs with full range, the definition we adopt is simpler than the original one in Abreu and Sen [1990] which involves the range of the SCR.

Moreover, Abreu and Sen [1990] show that when there are at least three agents, the conjunction of Condition  $\alpha$  with the following weak no veto power condition suffices for Nash implementability: An SCR  $f : \mathcal{P}^n \rightarrow X$  satisfies *weak no veto power* (WNVP) iff given any  $i \in N$ ,  $r(x, P_j) = 1$  for all  $j \in N \setminus \{i\}$  implies  $x \in f(P)$ .

### 3. SCORING RULES AND POINT RUNOFF PROCEDURES

The satisfaction or violation of Condition  $\alpha$  is not always straightforward to check. We start by introducing a simple condition which we show to be weaker than condition  $\alpha$ .

**Definition 2.** An SCR  $f : \mathcal{P}^n \rightarrow A$  satisfies Condition  $\beta$  iff given any  $x \in A$ , any partition  $\{K, N \setminus K\}$  of  $N$  and any  $P, P' \in \mathcal{P}^n$  with  $r(x, P_i) = r(x, P'_i) = 1$  for all  $i \in K$  and  $r(x, P_i) = r(x, P'_i) = m$  for all  $i \in N \setminus K$ , we have  $x \in f(P)$  if and only if  $x \in f(P')$ .

**Proposition 1.** If a SCR  $f : \mathcal{P}^n \rightarrow A$  satisfies Condition  $\alpha$ , then  $f$  satisfies Condition  $\beta$ .

*Proof.* Consider a SCR  $f$  which fails Condition  $\beta$ . So there exist  $x \in A$ , a partition  $\{K, N \setminus K\}$  of  $N$  and some  $P, P' \in \mathcal{P}^n$  with  $r(x, P_i) = r(x, P'_i) = 1$  for all  $i \in K$  and  $r(x, P_i) = r(x, P'_i) = m$  for all  $i \in N \setminus K$ , while  $x \in f(P) \setminus f(P')$ . Suppose, for a contradiction, Condition  $\alpha$  holds. So there exist a sequence  $(i_k)_{k=0}^l$  of agents and a sequence  $(a_k)_{k=0}^{l+1}$  of alternatives with  $a_0 = x$ , such that

- ( $\alpha_1$ )  $a_k P_{i_k} a_{k+1}$  for every  $k \in \{0, 1, \dots, l\}$ ,
- ( $\alpha_2$ )  $a_{l+1} P'_{i_l} a_l$ ,
- ( $\alpha_3$ )  $r(a_k, P'_{i_k}) > 1$  for each  $k \in \{0, 1, \dots, l-1\}$ .

For  $a_0 P_{i_0} a_1$  to hold, we must have  $i_0 \in K$ , which implies  $r(a_k, P'_{i_k}) = 1$ , contradicting ( $\alpha_3$ ).  $\square$

Hence, Condition  $\beta$  is necessary for an SCR to be subgame perfect implementable. Interestingly, all scoring rules<sup>8</sup> violate Condition  $\beta$ , hence fail to be subgame perfect implementable, as we state and show below.

**Proposition 2.** There exists no scoring rule which is subgame perfect implementable.

*Proof.* We prove the statement by establishing the existence of a pair  $n$  and  $m$  where no subgame perfect implementable scoring rule can be defined. We let  $m = 3$  and write  $X = \{x, y, z\}$ . We will show that any scoring rule associated to the score vector  $(1, \lambda, 0)$  fails Condition  $\beta$  for any  $\lambda \in [0, 1]$ . We consider four cases regarding the value that  $\lambda$  can take.

First, let  $\lambda = 0$ . Consider a partition  $\{N_1, N_2, N_3\}$  of  $N$  with  $|N_1| = |N_2| = \lceil \frac{n}{3} \rceil$ . Note that  $|N_3| \leq \lceil \frac{n}{3} \rceil$ . Consider  $P \in \mathcal{P}^n$  with  $x P_i y P_i z$  for all  $i \in N_1$ ,  $y P_i z P_i x$  for all  $i \in N_2$

<sup>8</sup>With just three alternatives, the case in which we will focus, the class of all scoring rules can be expressed by the family of score vectors  $(1, \lambda, 0)$  with  $\lambda \in [0, 1]$ . Note that  $(1, 0, 0)$  is the plurality rule,  $(1, \frac{1}{2}, 0)$  is the Borda rule and  $(1, 1, 0)$  is the antiplurality rule. We write  $s(x, P)$  for the score of alternative  $x$  at profile  $P$ , that is the sum of the scores that  $x$  gets from each of the voters. The scoring rule declares as winners the alternatives with the highest score.

and  $z P_i y P_i x$  for all  $i \in N_3$ . Note that  $s(x, P) = s(y, P) = \lceil \frac{n}{3} \rceil$  and  $s(z, P) \leq \lceil \frac{n}{3} \rceil$ , thus  $f(P) \supseteq \{x, y\}$ . Now consider  $P' \in \mathcal{P}^n$  with  $P'_i = P_i$  for all  $i \in N_1 \cup N_3$  and  $z P'_i y P'_i x$  for all  $i \in N_2$ . Note that  $s(x, P') = \lceil \frac{n}{3} \rceil$  and  $s(z, P') > \lceil \frac{n}{3} \rceil$ , thus  $x \in f(P) \setminus f(P')$ . On the other hand, we have  $r(x, P_i) = r(x, P'_i) = 1$  for all  $i \in N_1$  and  $r(x, P_i) = r(x, P'_i) = m$  for all  $i \in N_2 \cup N_3$ , establishing the failure of Condition  $\beta$ .

Second, let  $\lambda \in (0, \frac{1}{2}]$ . Pick any  $n$  which is either even or odd but sufficiently large to satisfy  $\frac{2}{n+1} < \lambda$ . Check that this choice of  $n$  ensures  $\lceil \frac{n}{2} \rceil < \frac{n}{2-\lambda}$ . Now consider a partition  $\{N_1, N_2\}$  of  $N$  with  $|N_1| = \lceil \frac{n}{2} \rceil$ . Note that  $|N_2| \leq \lceil \frac{n}{2} \rceil$ . Consider  $P \in \mathcal{P}^n$  with  $x P_i y P_i z$  for all  $i \in N_1$  and  $z P_i y P_i x$  for all  $i \in N_2$ . Note that  $s(x, P) = \lceil \frac{n}{2} \rceil \geq \frac{n}{2}$ ,  $s(y, P) = n\lambda \leq \frac{n}{2}$  and  $s(z, P) = n - \lceil \frac{n}{2} \rceil \leq \frac{n}{2}$ , thus  $f(P) \supseteq \{x\}$ . Now consider  $P' \in \mathcal{P}^n$  with  $P'_i = P_i$  for all  $i \in N_1$  and  $y P'_i z P'_i x$  for all  $i \in N_2$ . Note that  $s(x, P') = \lceil \frac{n}{2} \rceil$  and  $s(y, P') = \lceil \frac{n}{2} \rceil \lambda + (n - \lceil \frac{n}{2} \rceil)$ . Recall that we have  $\lceil \frac{n}{2} \rceil < \frac{n}{2-\lambda}$ , which ensures  $s(y, P') > s(x, P')$ , thus  $x \in f(P) \setminus f(P')$ . On the other hand, we have  $r(x, P_i) = r(x, P'_i) = 1$  for all  $i \in N_1$  and  $r(x, P_i) = r(x, P'_i) = m$  for all  $i \in N_2$ , establishing the failure of Condition  $\beta$ .

Third, let  $\lambda \in (\frac{1}{2}, 1)$ . Pick any  $n$  which satisfies  $\lceil \lambda n \rceil < \frac{n}{2-\lambda}$ .<sup>9</sup> Now consider a partition  $\{N_1, N_2\}$  of  $N$  with  $|N_1| = \lceil \lambda n \rceil$ . Note that  $|N_2| < \lceil \lambda n \rceil$ . Consider  $P \in \mathcal{P}^n$  with  $x P_i y P_i z$  for all  $i \in N_1$  and  $z P_i y P_i x$  for all  $i \in N_2$ . Note that  $s(x, P) = \lceil \lambda n \rceil \geq \lambda n$ ,  $s(y, P) = \lambda n \leq \lceil \lambda n \rceil$  and  $s(z, P) < \lceil \lambda n \rceil$ , thus  $f(P) \supseteq \{x, y\}$ . Now consider  $P' \in \mathcal{P}^n$  with  $P'_i = P_i$  for all  $i \in N_2$  and  $x P'_i z P'_i y$  for all  $i \in N_1$ . Note that  $s(x, P') = \lceil \lambda n \rceil$  and  $s(z, P') = \lceil \lambda n \rceil \lambda + (n - \lceil \lambda n \rceil)$ . Recall that we have  $\lceil \lambda n \rceil < \frac{n}{2-\lambda}$ , which ensures  $s(z, P') > s(x, P')$ , thus  $x \in f(P) \setminus f(P')$ . On the other hand, we have  $r(x, P_i) = r(x, P'_i) = 1$  for all  $i \in N_1$  and  $r(x, P_i) = r(x, P'_i) = m$  for all  $i \in N_2$ , establishing the failure of Condition  $\beta$ .

Fourth, let  $\lambda = 1$ . Consider a partition  $\{N_1, N_2, N_3\}$  of  $N$  with  $|N_1| = |N_2| = \lfloor \frac{n}{3} \rfloor$ . Note that  $|N_3| \geq \lfloor \frac{n}{3} \rfloor$ . Consider  $P \in \mathcal{P}^n$  with  $y P_i z P_i x$  for all  $i \in N_1$ ,  $x P_i z P_i y$  for all  $i \in N_2$  and  $x P_i y P_i z$  for all  $i \in N_3$ . Note that  $s(x, P) = s(y, P) \geq s(z, P)$ , thus  $f(P) \supseteq \{x, y\}$ . Now consider  $P' \in \mathcal{P}^n$  with  $P'_i = P_i$  for all  $i \in N_1 \cup N_2$  and  $x P'_i z P'_i y$  for all  $i \in N_3$ . Note that  $s(z, P') > s(x, P')$ , thus  $x \in f(P) \setminus f(P')$ . On the other hand, we have  $r(x, P_i) = r(x, P'_i) = 1$  for all  $i \in N_2 \cup N_3$  and  $r(x, P_i) = r(x, P'_i) = m$  for all  $i \in N_1$ , establishing the failure of Condition  $\beta$ .  $\square$

Note that although the negative result of Proposition 2 is established for  $m = 3$ , it can be extended to any  $m > 3$  by constructing profiles with a triplet of alternatives which are ranked above the rest by every voter. We also wish to remark that Proposition 2 holds for an arbitrary value of  $n$  when  $\lambda \in \{0, 1\}$  but for a restricted though infinite class of values that  $n$  can take when  $\lambda \in (0, 1)$ . As mentioned in the introduction, Proposition 2 was already found by Sen [1987]. Yet, the analysis here remains of interest since it shows the usefulness of condition  $\beta$ .

<sup>9</sup>The existence of such an  $n$  can be seen by observing that the inequality  $\lambda n + 1 < \frac{n}{2-\lambda}$  is equivalent to  $\frac{2-\lambda}{(1-\lambda)^2} < n$ .

We now turn to plurality with a runoff<sup>10</sup> and first show that this SCR is subgame perfect implementable when there are only three alternatives.

**Proposition 3.** *When  $n \geq 3$  and  $m = 3$ , plurality with a runoff is subgame perfect implementable.*

*Proof.* Let  $X = \{x, y, z\}$  and  $f : \mathcal{P}^n \rightarrow A$  be plurality with a runoff where ties are broken according to the alphabetical order of alternatives. When  $n \geq 3$ ,  $f$  satisfies WNVP. So we prove the proposition by showing that  $f$  satisfies Condition  $\alpha$  as well. Take any  $P, P' \in \mathcal{P}^n$  and any  $x \in f(P) \setminus f(P')$ . If there exists some  $j \in N$  and some  $y \in L(x, P_j) \setminus L(x, P'_j)$ , then the sequence  $(x, y)$  in  $X$  and  $j \in N$  establish Condition  $\alpha$ . Now we consider the case where  $L(x, P_i) \subseteq L(x, P'_i) \forall i \in N$ . Assume, without loss of generality, that at  $P$ ,  $x$  and  $y$  go for a runoff. Note that it is not possible that at  $P'$ , again  $x$  and  $y$  go for a runoff, because  $L(x, P_i) \subseteq L(x, P'_i) \forall i \in N$  would imply that  $x$  would be the runoff winner at  $P'$  which contradicts  $x \notin f(P')$ . So the two cases below are exhaustive:

Case 1: At  $P'$ ,  $y$  and  $z$  go for a runoff.

Let  $n_x$  be the number of voters who rank  $x$  first at  $P$ ;  $n_y = \#\{i \in N : y P_i z P_i x\}$ ,  $n'_y = \#\{i \in N : y P_i x P_i z\}$ ,  $n_z = \#\{i \in N : z P_i x P_i y\}$  and  $n'_z = \#\{i \in N : z P_i y P_i x\}$ . We first make the following observations:

Observation 1:  $n_x \leq \frac{n}{3}$ , as otherwise  $x$  would go for a runoff at  $P'$ .

Observation 2:  $n_z + n'_z \leq n_x$ , as  $z$  does not go for a runoff at  $P$ .

Observation 3:  $n_x + n_z$  is a majority, as  $x$  beats  $y$  at the runoff.

Observation 4:  $n_y + n'_y \geq \frac{n}{3}$  by simple arithmetic.

Observation 5:  $n_z > 0$ , as otherwise  $x$  would lose against  $y$  at the runoff.

Claim 1: Going from  $P$  to  $P'$ , some voters within  $\{i \in N : y P_i z P_i x\}$  swapped  $y$  and  $z$ .

*Proof of the claim:* Suppose that this is not the case. Since  $L(x, P_i) \subseteq L(x, P'_i)$  for every  $i \in N$ , the voters rank  $x$  first at  $P$  keep  $x$  at the top at  $P'$ . For the same reason, the voters in  $\{i \in N : y P_i x P_i z\}$  cannot lift  $z$  at the top. Hence it would not be possible that  $z$  goes to a runoff at  $P'$ , giving a contradiction.

Claim 2: Going from  $P$  to  $P'$ , not all voters in  $\{i \in N : z P_i x P_i y\}$  lifted  $x$  at the top.

By the two claims, the sequences  $(x, y, z) \in X$  and  $(i^*, j^*) \in N$  with  $i^* \in \{i \in N : z P_i x P_i y\}$  and  $j^* \in \{i \in N : y P_i z P_i x\}$  establish Condition  $\alpha$ .

Case 2: At  $P'$ ,  $x$  and  $z$  go for a runoff and  $z$  wins.

<sup>10</sup>Under plurality with a runoff, each voter announces a ranking of alternatives. The two alternatives with a highest plurality score go to the runoff, the rest of alternatives being removed. In the runoff, the majority winner among the two remaining alternatives is selected, where the majority winner is computed using the initial preference profile. Ties are broken lexicographically according to some linear order.



We proceed by similar arguments. It can only be the case that going from  $P$  to  $P'$ , some voters in  $\{i \in N : y P_i z P_i x\}$  swapped  $y$  and  $z$ ; also not all voters in  $\{i \in N : z P_i x P_i y\}$  lifted  $x$  at the top. In other words, the two claims above are valid by the same arguments and establish Condition  $\alpha$ .  $\square$

However, this positive result does not extend to an arbitrary number of alternatives. In fact, as we state and show below, when  $m \geq 4$ , no point runoff procedure as defined by Smith [1973] is subgame perfect implementable. The proof is based on building a pair of preference profiles with  $4k - 1$  voters and any number of alternatives.

**Theorem 1.** *When  $m \geq 4$ , no point runoff procedure is subgame perfect implementable.*

*Proof.* Let  $m \geq 4$ . Pick some integer  $k \geq 2$  and consider the preference profiles  $P$  and  $P'$  as follows:

$k - 1$  voters:  $x_3 P_i x_2 P_i \cdots P_i x_1$  (Block 1)

$k$  voters:  $x_4 P_i \cdots P_i x_1 P_i x_2$  (Block 2)

$k$  voters:  $x_1 P_i \cdots P_i x_2$  (Block 3)

$k$  voters:  $x_2 P_i \cdots P_i x_1$  (Block 4)

$k - 1$  voters:  $x_2 P'_i \cdots P'_i x_1$  (Block 1)

$k$  voters:  $x_3 P'_i \cdots x_1 P'_i x_2$  (Block 2)

$k$  voters:  $x_1 P'_i \cdots x_2$  (Block 3)

$k$  voters:  $x_2 P'_i \cdots x_1$  (Block 4)

We consider a point runoff procedure  $f$  with the tie-breaking linear order  $x_3 T_i x_1 T_i x_2 T_i x_4$ . Note that at  $P$  and  $P'$ ,  $f$  has to agree either with plurality with a runoff or with single transferable vote (STV<sup>11</sup>). Under plurality with a runoff, at the profile  $P$ ,  $x_1$ ,  $x_2$  and  $x_4$  tie for a runoff and  $x_1$  and  $x_2$  go for it according to  $T$ , which leads to  $f(P) = \{x_1\}$ . At the profile  $P'$ ,  $x_2$  goes for a runoff while  $x_1$  and  $x_3$  tie for it, hence by  $T$ ,  $x_2$  and  $x_3$  go for a runoff which leads to  $f(P') = \{x_3\}$ . One can also observe that we have  $f(P) = \{x_1\}$  and  $f(P') = \{x_3\}$  under STV as well.

In order to show that condition  $\alpha$  is violated, we need to show that there is no sequence  $(i_k)_{k=0}^l$  in  $N$  and a sequence  $(a_k)_{k=0}^{l+1}$  that satisfy Definition 1.

Notice first that in any such sequence  $a_0 = x_1$  by definition. Then  $x_1 P_{i_0} y$  for some  $y \in A$ . However,  $x_1$  is ranked last by the first and fourth block of voters in  $P$ . Thus,  $i_0$  is either in the second or in the third block of voters. Yet,  $x_1$  is ranked first in  $P'$  by all

<sup>11</sup>Under STV, each voter submits a preference ordering over the alternatives. The score of each alternative equals the number of voters who rank that alternative first. STV removes the alternatives with the lowest score and computes the new scores in the preference profile without the removed alternatives and continues until one alternative gets a majority.

voters in the third block, in contradiction with condition  $(\alpha_3)$ . Thus,  $i_0$  is a voter in the second block and hence  $y$  has to be  $x_2$ .

Now, there must be some  $i_1 \in N$  and some  $z \in A$  such that  $x_2 P_{i_1} z$ . Thus, by the same argument as before,  $i_1$  belongs to the first or the fourth block. However,  $x_2$  is ranked first in  $P'$  by any voter in these two blocks, again contradicting condition  $(\alpha_3)$ , completing the proof.  $\square$

#### 4. CONDORCET CONSISTENT RULES

At any  $P \in \mathcal{P}^n$ , we write  $\mu(P)$  for the majority relation. So for any  $x, y \in X$ , we have  $x \mu(P) y$  iff  $|\{i \in N : x P_i y\}| \geq |\{i \in N : y P_i x\}|$ . Throughout the section, we assume  $\mu(P)$  to be antisymmetric - for example let  $n$  be odd. Thus  $\mu(P)$  is a tournament. A comprehensive analysis of how to choose from a tournament is given by Laslier [1997]. However, we wish to note here that the unique maximal element of a tournament, if it exists, is the pairwise majority winner (or Condorcet winner). Even when a Condorcet winner fails to exist, there is always a unique minimal set (called the top-cycle) such that every alternative in the top-cycle is a majority winner against every alternative outside the top-cycle. We define this formally as follows.

A set  $Y \subseteq X$  is *dominant* at  $P \in \mathcal{P}^n$  iff  $y \mu(P) z \forall y \in Y, \forall z \in X \setminus Y$ . The *top-cycle* at  $P$ , denoted  $TC(P)$ , is the dominant set which is minimal with respect to set inclusion. A SCR  $f : \mathcal{P}^n \rightarrow A$  is *Condorcet consistent* iff  $f(P) \subseteq TC(P)$  at every  $P \in \mathcal{P}^n$ .

There is a literature on the implementation of Condorcet consistent SCRs which gave birth to a variety of monotonicity conditions. For example, set monotonicity, which is weaker than Maskin monotonicity, has been identified by Özkal-Sanver and Sanver [2006] as a member of a broader class of monotonicity conditions which they call “cover monotonicity”. A stronger version of the condition<sup>12</sup> has been later used by Brandt [2015] in his analysis of strategy-proof set-valued SCRs. A SCR  $f : \mathcal{P}^n \rightarrow A$  is *set monotonic* iff given any  $P, P' \in \mathcal{P}^n$  with  $L(x, P_i) \subseteq L(x, P'_i) \forall i \in N, \forall x \in f(P)$ , we have  $f(P) \subseteq f(P')$ .

We now show that every Condorcet consistent and set monotonic SCR is subgame perfect implementable.

**Theorem 2.** *Let  $n \geq 3$ . Every Condorcet consistent and set monotonic SCR  $f : \mathcal{P}^n \rightarrow A$  is subgame perfect implementable.*

*Proof.* Let  $f$  be Condorcet consistent and set monotonic. It is straightforward to check that Condorcet consistency implies WNVP. We will show that  $f$  satisfies Condition  $\alpha$  as well. Take any  $P, P' \in \mathcal{P}^n$  and any  $x \in f(P) \setminus f(P')$ . If there exists some  $j \in N$  and some  $y \in L(x, P_j) \setminus L(x, P'_j)$ , then the sequence  $(x, y)$  in  $X$  and  $j \in N$  establish Condition

<sup>12</sup>This stronger condition requires the same antecedent but a stronger consequence  $f(P) = f(P')$  instead of  $f(P) \subseteq f(P')$

$\alpha$ . Now we consider the case where  $L(x, P_j) \subseteq L(x, P'_j) \forall j \in N$ . Remark that  $|f(P)| > 1$ , as otherwise we would have  $f(P) = \{x\}$  which, conjoined with  $L(x, P_j) \subseteq L(x, P'_j) \forall j \in N$  and the set monotonicity of  $f$  would imply  $x \in f(P')$ , contradicting  $x \in f(P) \setminus f(P')$ . We treat the cases  $|f(P')| > 1$  and  $|f(P')| = 1$  separately. First, let  $|f(P')| > 1$ . As  $f$  is set monotonic,  $x \in f(P) \setminus f(P')$  implies the existence of some  $w \in X$  and some  $j^* \in N$  with  $z P_{j^*} w$  and  $w P'_{j^*} z$  for some  $z \in f(P)$ . As  $x, z \in f(P)$  and  $f(P) \subseteq TC(P)$ , we have  $x, z \in TC(P)$  and by definition of  $TC(P)$ , there is a sequence  $(x_0, \dots, x_{s-1})$  of  $s \geq 2$  distinct alternatives in  $TC(P)$  with  $x_0 = x, x_{s-1} = z$ , and  $x_k \mu(P) x_{k+1}$  for  $k \in \{0, \dots, s-2\}$ . For each  $k \in \{0, \dots, s-2\}$ , let  $N^k = \{i \in N : x_k P_i x_{k+1}\}$ . By definition of the majority relation, we have  $|N^k| > \frac{n}{2}$ . Let  $\widetilde{N}^k = \{i \in N^k : r(x_k, P'_i) > 1\}$ . Suppose  $|\widetilde{N}^k| = \emptyset$ . Thus,  $r(x_k, P'_i) = 1$  for all  $i \in N^k$ . As  $|N^k| > \frac{n}{2}$  and  $f$  is Condorcet consistent, we have  $f(P') = \{x_k\}$ , contradicting  $|f(P')| > 1$ . Hence,  $\widetilde{N}^k \neq \emptyset$ . For each  $k \in \{0, \dots, s-2\}$ , pick some  $i_k \in \widetilde{N}^k$ . The sequence  $(i_k)_{k=0}^{s-1}$  of agents with  $i_{s-1} = j^*$  and the sequence  $(x_k)_{k=0}^s$  of alternatives with  $x_s = w$  establishes Condition  $\alpha$ . We now consider the case  $|f(P')| = 1$ . Let  $f(P') = \{z\}$ . As  $|f(P)| > 1$ , by set monotonicity, there is some  $y \in f(P)$  and some  $j^* \in N$  with  $y P_{j^*} w$  and  $w P'_{j^*} y$  for some  $w \in X$ . Note that  $y \neq x$ , as  $L(x, P_j) \subseteq L(x, P'_j) \forall j \in N$ . As  $x, y \in f(P)$  and  $f(P) \subseteq TC(P)$ , we have  $x, y \in TC(P)$  and by definition of  $TC(P)$ , there is a sequence  $(x_0, \dots, x_{s-1})$  of  $s \geq 2$  distinct alternatives in  $TC(P)$  with  $x_0 = x, x_{s-1} = y$ , and  $x_k \mu(P) x_{k+1}$  for  $k \in \{0, \dots, s-2\}$ . We establish the proof by treating the cases  $z \notin \{x_0, \dots, x_{s-1}\}$  and  $z \in \{x_0, \dots, x_{s-1}\}$  separately. For the former case, by similar arguments,  $\widetilde{N}^k \neq \emptyset$  for any  $k \in \{0, \dots, s-2\}$ . The sequence  $(i_k)_{k=0}^{s-1}$  of agents with  $i_k \in \widetilde{N}^k$  and  $i_{s-1} = j^*$ , together with the sequence  $(x_k)_{k=0}^s$  of alternatives with  $x_s = z$  establish Condition  $\alpha$ . Finally, we consider the case  $z \in \{x_0, \dots, x_{s-1}\}$ . Let  $z = x_l$  for some  $l \in \{1, \dots, s-2\}$ . The sequence  $(i_k)_{k=0}^l$  of agents with  $i_l = j^*$  and the sequence  $(x_k)_{k=0}^{l+1}$  of alternatives with  $x_{l+1} = y$  establishes Condition  $\alpha$ .  $\square$

Theorem 2 covers several interesting tournament solutions: We know from Theorems 3.1, 3.7, 3.8, 3.9 of Özkal-Sanver and Sanver [2010] that the top-cycle, the iterated uncovered set, the minimal covering set (Dutta [1990]) and the bipartisan set (Laffond et al. [1993]) are set monotonic.<sup>13</sup> As to some rather negative results, we know that the uncovered set and the Banks set (Banks [1985]) are set monotonic if and only if  $|X| \in \{3, 4\}$  (Theorems 3.3 and 3.4 of Özkal-Sanver and Sanver [2010]); Copeland rule and Slater rule are set monotonic if and only if  $|X| = 3$  (Theorems 3.5 and 3.6 of Özkal-Sanver and Sanver [2010]).

<sup>13</sup>As an earlier result in this direction, Theorem 4.4 of Özkal-Sanver and Sanver [2006] shows the set monotonicity of the top-cycle. The same conclusion is later derived in Remark 1 of Brandt [2015] where the set monotonicity of the minimal covering set and the bipartisan set is also mentioned.

## 5. COMPROMISE RULES

We adopt the terminology from Merlin et al. [2019]. Given any  $q \in \{1, \dots, n\}$ , the  $q$ -threshold rule is the SCR  $g_q : \mathcal{P}^n \rightarrow X$  which is defined for each  $P \in \mathcal{P}^n$  as

$$g_q(P) = \{x \in A \mid r^*(x, q, P) \leq r^*(y, q, P) \text{ for any } y \in X\}.$$

So at each  $P$ ,  $g_q$  selects the alternatives that obtain a support of  $q$  voters at the smallest rank. Let  $r^*(q, P)$  denote this rank. We quote the following result from Brams and Kilgour [2001] (Theorem A2, p.308):

**Proposition 4.**  $r^*(q, P) \leq \lfloor \frac{mq-m+n}{n} \rfloor$

The SCR  $f_q : \mathcal{P}^n \rightarrow X$  which refines  $g_q(P)$  by selecting the alternatives that obtain the highest support at rank  $r^*(q, P)$  is called  $q$ -approval fallback bargaining. So we have

$$f_q(P) = \{x \in A \mid n(x, r^*(q, P), P) \geq n(y, r^*(q, P), P) \text{ for any } y \in X\}.$$

Sertel and Yilmaz [1999] show that the majoritarian compromise, a particular case of  $f_q$  where  $q = \frac{n}{2}$ , exemplifies an SCR which is subgame perfect implementable but not Nash implementable. As we will see throughout the paper, the failure of Nash implementability extends to all values of  $q$  while subgame perfect implementability exhibits a mixed picture. For example, when  $q$  is less than majority,  $f_q$  fails Condition  $\alpha$ , hence subgame perfect implementability, as we state and show below<sup>14</sup>:

**Theorem 3.**  $f_q : \mathcal{P}^n \rightarrow X$  fails to be subgame perfect implementable when  $q \in \{1, \dots, \lfloor \frac{n}{2} \rfloor - 1\}$ .

*Proof.* Consider first the case with  $q \in \{1, \dots, \lfloor \frac{n}{3} \rfloor\}$ . Take a partition  $\{N_1, N_2, N_3\}$  of  $N$  with  $\#N_1 = \#N_2 = \lfloor \frac{n}{3} \rfloor$ , hence  $\#N_3 \leq \#N_1$ . Take any triplet  $x, y, z \in X$  and construct a profile  $P \in \mathcal{P}^n$  with  $r(x, P_i) = 1 \forall i \in N_1$ ,  $r(y, P_i) = 1$  and  $r(x, P_i) = m \forall i \in N_2$ , and  $r(z, P_i) = 1$  and  $r(x, P_i) = m \forall i \in N_3$ . We have  $f_q(P) = \{x, y\}$  when  $\#N_3 < \#N_1$  and  $f_q(P) = \{x, y, z\}$  when  $\#N_3 = \#N_1$ . In any case,  $x \in f_q(P)$ . Now consider  $P' \in \mathcal{P}^n$  with  $P'_i = P_i$  for all  $i \in N_1 \cup N_2$  while  $r(y, P'_i) = 1$  and  $r(x, P'_i) = m \forall i \in N_3$ . We have  $f_q(P') = \{y\}$ , hence  $x \in f_q(P) \setminus f_q(P')$ . On the other hand, we have  $r(x, P_i) = r(x, P'_i) = 1$  for all  $i \in N_1$  and  $r(x, P_i) = r(x, P'_i) = m$  for all  $i \in N_2 \cup N_3$ , establishing the failure of Condition  $\beta$ .

Now consider the case with  $q \in \{\lfloor \frac{n}{3} \rfloor + 1, \dots, \lfloor \frac{n}{2} \rfloor - 1\}$ . Take a partition  $\{N_1, N_2, N_3\}$  of  $N$  with  $\#N_1 = \#N_2 = q$ , hence  $\#N_3 < \#N_1$ . Take any triplet  $x, y, z \in X$  and construct a profile  $P \in \mathcal{P}^n$  with  $r(x, P_i) = 1 \forall i \in N_1$ ,  $r(y, P_i) = 1$  and  $r(x, P_i) = m \forall i \in N_2$ , and  $r(z, P_i) = 1$  and  $r(x, P_i) = m \forall i \in N_3$ . Note that  $f_q(P) = \{x, y\}$ . Now consider  $P' \in \mathcal{P}^n$  with  $P'_i = P_i$  for all  $i \in N_1 \cup N_2$  while  $r(y, P'_i) = 1$  and  $r(x, P'_i) = m \forall i \in N_3$ . Now  $f_q(P') = \{y\}$  since  $\#N_1 < \#N_2 + \#N_3$ . Thus,  $x \in f_q(P) \setminus f_q(P')$ . On the other hand, we have  $r(x, P_i) = r(x, P'_i) = 1$

<sup>14</sup>This in turn implies that any such  $f_q$  fails to be Nash implementable, since Condition  $\alpha$  is weaker than Maskin monotonicity.

for all  $i \in N_1$  and  $r(x, P_i) = r(x, P'_i) = m$  for all  $i \in N_2 \cup N_3$ , establishing the failure of Condition  $\beta$ .  $\square$

When  $q \in \{\lceil \frac{n}{2} \rceil, \dots, n\}$ , subgame perfect implementability of  $f_q$  necessitates to obtain the support of  $q$  before reaching the last rank. We formally state this below:

**Proposition 5.** *For any  $q \in \{\lceil \frac{n}{2} \rceil, \dots, n\}$ ,  $f_q : \mathcal{P}^n \rightarrow X$  is subgame perfect implementable only if  $r^*(q, P) < m$  for every  $P \in \mathcal{P}^n$ .*

*Proof.* Suppose there exists some  $\tilde{P} \in \mathcal{P}^n$  with  $r^*(q, \tilde{P}) = m$ . We write  $X = \{x_1, \dots, x_m\}$  and let  $N_k = \{i \in N : r(x_k, \tilde{P}_i) = m\}$  for each  $k \in \{1, \dots, m\}$ . As  $r^*(q, \tilde{P}) = m$ , we have  $\#N_k \geq n - q + 1$  for every  $k \in \{1, \dots, m\}$ . Now, consider a profile  $P \in \mathcal{P}^n$  with  $r(x_1, P_i) = m \forall i \in N_1$ ,  $r(x_1, P_i) = 1$  and  $r(x_k, P_i) = m \forall i \in N_k, \forall k \in \{2, \dots, m\}$ . As  $\#N_k \geq n - q + 1$  for every  $k \in \{1, \dots, m\}$ , we have  $f_q(P) = X$ . Now consider  $P' \in \mathcal{P}^n$  with  $P'_i = P_i \forall i \in N \setminus N_m$ ,  $r(x_1, P'_i) = 1$  and  $r(x_{m-1}, P'_i) = m \forall i \in N_m$ . Note that  $f_q(P') = \{x_m\}$ . Thus,  $x_1 \in f_q(P) \setminus f_q(P')$ . On the other hand, we have  $r(x_1, P_i) = r(x_1, P'_i) = 1$  for all  $i \in N \setminus N_1$  and  $r(x_1, P_i) = r(x_1, P'_i) = m$  for all  $i \in N_1$ , establishing the failure of Condition  $\beta$ .  $\square$

Proposition 5, combined with Proposition 4, leads to the following proposition:

**Proposition 6.** *Take any  $r \in [\frac{1}{2}, 1]$  and let  $q = \lceil rn \rceil$ . If  $rn$  is an integer, then  $f_{rn} : \mathcal{P}^n \rightarrow X$  fails to be subgame perfect implementable when  $r > \frac{m-1}{m}$ . If  $rn$  is not an integer, then  $f_{rn} : \mathcal{P}^n \rightarrow X$  fails to be subgame perfect implementable when  $r \geq \frac{m-1}{m}$ .*

*Proof.* By Proposition 5,  $f_q : \mathcal{P}^n \rightarrow X$  fails to be subgame perfect implementable if  $r^*(q, P) = m$ . By Proposition 4, we have  $r^*(q, P) = m$  iff  $\lfloor \frac{mq-m+n}{n} \rfloor \geq m$  which implies  $\frac{mq-m+n}{n} \geq m$ . Therefore,  $q \geq n + 1 - \frac{n}{m}$ . So  $\lceil rn \rceil \geq n + 1 - \frac{n}{m}$ , as  $\lceil rn \rceil = q$ . If  $\lceil rn \rceil$  is an integer, then  $rn \geq n + 1 - \frac{n}{m}$ , which implies  $rn > n - \frac{n}{m}$ , which in turn implies  $r > \frac{m-1}{m}$ . If  $\lceil rn \rceil$  is not an integer, then  $rn \geq n - \frac{n}{m}$ , which implies  $r \geq \frac{m-1}{m}$ .  $\square$

Proposition 6 has the following corollary.

**Theorem 4.** *If  $f_q : \mathcal{P}^n \rightarrow X$  is subgame perfect implementable for any  $m \geq 3$  and  $n \geq 2$ , then  $q \notin \{\lceil \frac{2n}{3} \rceil + 1, \dots, n\}$ .*

Theorem 3 and Theorem 4 announce that the subgame perfect implementability of  $f_q$  cannot be ensured when  $q \in \{1, \dots, \lceil \frac{n}{2} \rceil - 1\} \cup \{\lceil \frac{2n}{3} \rceil + 1, \dots, n\}$ . However, we have a positive result for the values of  $q$  which remain in between, i.e., when  $q \in \{\lceil \frac{n}{2} \rceil, \dots, \lceil \frac{2n}{3} \rceil\}$ .<sup>15</sup>

**Theorem 5.** *Let  $n \geq 3$ .  $f_q : \mathcal{P}^n \rightarrow X$  is subgame perfect implementable when  $q \in \{\lceil \frac{n}{2} \rceil, \dots, \lceil \frac{2n}{3} \rceil\}$ .*

<sup>15</sup>This exemplifies a class of SCRs which are subgame perfect implementable but not Nash implementable, as  $f_q$  fails Maskin monotonicity when  $q \geq \frac{n}{2}$ . This can be seen through Figure ?? where any  $f_q$  with  $q \geq \frac{n}{2}$  picks all three alternatives at profile  $P$  and only  $c$  at profile  $P'$ .

*Proof.* Take any  $q \in \{\lceil \frac{n}{2} \rceil, \dots, \lceil \frac{2n}{3} \rceil\}$ . It is straightforward to see that  $f_q$  satisfies WNWP. Therefore, we will establish the subgame perfect implementability of  $f_q$  by showing that it satisfies Condition  $\alpha$ . In the sequel, we write  $f$  instead of  $f_q$  to simplify notation. Take any  $P, P' \in \mathcal{P}^n$  and any  $x \in f(P) \setminus f(P')$ . We will consider several mutually exclusive and exhaustive cases. For each case, we will identify, for some non-negative integer  $l$ , a sequence  $(i_k)_{k=0}^l$  in  $N$  and a sequence  $(a_k)_{k=0}^{l+1}$  in  $X$  with  $a_0 = x$  such that the conditions  $(\alpha_1)$ ,  $(\alpha_2)$ ,  $(\alpha_3)$  and  $(\alpha_4)$  of Condition  $\alpha$  are satisfied.  $r(a_{l+1}, P'_{i_k}) = 1$  for every  $k \in \{0, 1, \dots, l-1\}$ , then either  $l = 0$  or  $i_{l-1} \neq i_l$ .

**Case A.** There exists some  $i^* \in N$  and some  $z \in L(x, P_{i^*}) \setminus L(x, P'_{i^*})$ . The sequence  $(x, z)$  in  $X$  and  $i^* \in N$  establish Condition  $\alpha$ .

**Case B.**  $L(x, P_i) \subseteq L(x, P'_i)$  for every  $i \in N$ .

Let  $\tilde{N} := \{i \in N \mid r(x, P'_i) > 1\}$  denote the set of agents who do not rank  $x$  first in  $P'$ . Note that  $\tilde{N} \neq \emptyset$ , as otherwise  $x$  is ranked first by all agents in  $P'$  which implies  $x \in f(P')$ , contradicting  $x \in f(P) \setminus f(P')$ . In the sequel, we write  $r^*(P)$  instead of  $r^*(q, P)$  to simplify notation.

Note that  $r^*(P) \notin \{1, m\}$ .<sup>16</sup> As  $L(x, P_i) \subseteq L(x, P'_i) \forall i \in N$ , we have  $n(x, q, P') \geq n(x, q, P)$ . Therefore  $r^*(P') \leq r^*(P)$ . Since  $x \in f(P) \setminus f(P')$ , there is some  $c \in f(P')$ . If  $r^*(P') < r^*(P)$ , then  $n(c, k', P') \geq q$  at some  $k' < r^*(P)$ . Otherwise  $r^*(P') = r^*(P)$ , implying  $n(c, r^*(P), P') > n(c, r^*(P), P)$ . Thus, there is some  $b \in X$  with  $b P_{j^*} c$  and  $c P'_{j^*} b$  for some  $j^* \in N$  while  $b \neq x$ , as  $L(x, P_i) \subseteq L(x, P'_i) \forall i \in N$ . We consider two cases regarding a relationship between  $b$  and  $\tilde{N}$ .

**Case B.I** There is  $i^* \in \tilde{N}$  with  $x P_{i^*} b$ . First let  $i^* = j^*$ . The sequence  $(j^*, j^*)$  in  $N$  and the sequence  $(x, b, c)$  in  $X$  satisfy conditions  $(\alpha_1)$ ,  $(\alpha_2)$  and  $(\alpha_3)$ . As  $x P_{i^*} b P_{i^*} c$  and  $L(x, P_i) \subseteq L(x, P'_i)$  for every  $i \in N$ , we have  $r(c, P'_{j^*}) \neq 1$ , establishing condition  $(\alpha_4)$ , which in turn establishes Condition  $\alpha$ . Now let  $i^* \neq j^*$ . The sequence  $(i^*, j^*)$  in  $N$  and the sequence  $(x, b, c)$  in  $X$  establish Condition  $\alpha$ .

**Case B.II**  $b P_i x$  for every  $i \in \tilde{N}$ . Observe that there is  $i^* \in \tilde{N}$  with  $r(x, P_{i^*}) \in \{2, \dots, r^*(P)\}$ .<sup>17</sup> Thus,  $m - r^*(P) + 1 \leq |L(x, P_{i^*})| \leq m - 1$ . We again consider two cases: Either there is  $h \in N$  with  $r(b, P_h) > r^*(P)$  or  $r(b, P_i) \leq r^*(P)$  for all  $i \in N$ .

**Case B.II.1** There is  $h \in N$  with  $r(b, P_h) > r^*(P)$ . So  $|L(b, P_h)| < m - r^*(P) + 1$ , which implies  $|X \setminus L(b, P_h)| > m - (m - r^*(P) + 1) = r^*(P) - 1$ . Since  $|L(x, P_{i^*})| \geq m - r^*(P) + 1$ , it follows that  $|L(x, P_{i^*})| + |X \setminus L(b, P_h)| > m$ , hence  $L(x, P_{i^*}) \cap (X \setminus L(b, P_h))$  is non-empty. In fact, we now state a lemma (proven in the appendix) which shows that  $L(x, P_{i^*}) \cap (X \setminus L(b, P_h))$  contains at least two elements.

**Lemma A:**  $|L(x, P_{i^*}) \cap (X \setminus L(b, P_h))| \geq 2$ .

<sup>16</sup>If  $r^*(P) = 1$ ,  $n(x, 1, P') \geq n(x, 1, P) \geq q$  so that  $x \in f(P)$ . This implies that  $n(x, 1, P') \geq q$  and hence  $x \in f(P')$ , a contradiction. Thus  $r^*(P) \neq 1$ . We deduce  $r^*(P) \neq 1$  from Brams and Kilgour [2001].

<sup>17</sup>Otherwise, for every  $i \in N$  with  $r(x, P_i) \in \{2, \dots, r^*(P)\}$ , we have  $i \notin \tilde{N}$ , thus  $r(x, P'_i) = 1$ . Moreover, being in Case B, we have  $r(x, P'_i) = 1$  for every  $i \in N$  with  $r(x, P_i) = 1$  as well. Since  $n(x, r^*(P), P) \geq q$ , it follows that  $n(x, 1, P') \geq q$  implying  $x \in f(P')$ , contradicting  $x \in f(P) \setminus f(P')$ .

As a consequence of Lemma A, there exists  $y \in L(x, P_{i^*}) \cap X \setminus L(b, P_h)$  with  $y \neq x$ .

**Case B.II.1.i**  $r(y, P'_h) > 1$ .

The sequence  $(i^*, h, j^*)$  in  $N$  and the sequence  $(x, y, b, c)$  in  $X$  satisfy conditions  $(\alpha_1)$ ,  $(\alpha_2)$  and  $(\alpha_3)$ . If  $h \neq j^*$ , then condition  $(\alpha_4)$  is also satisfied. Now let  $h = j^*$  and  $r(c, P'_{i^*}) = r(c, P'_{j^*}) = 1$ . As  $L(x, P_i) \subseteq L(x, P'_i)$  for every  $i \in N$ , we have  $c P_{j^*} x$ , thus  $b P_{j^*} c P_{j^*} x$ , implying the existence of  $b' \in X \setminus \{b, c, x\}$  with  $r(b', P_{j^*}) = 1$ . Thus,  $b' P_{j^*} c$  and  $c P'_{j^*} b'$ . Being in Case B.II.1, we admit the existence of  $h' \in N$  with  $r(b', P_{h'}) > r^*(P)$ .<sup>18</sup>Note that  $h' \neq j^* = h$ . Again by Lemma A, there exists  $y \in L(x, P_{i^*}) \cap X \setminus L(b', P_{h'})$  with  $y \neq x$ . Being in Case B.II.1.i, we admit  $r(y, P'_{h'}) > 1$ .<sup>19</sup>The sequence  $(i^*, h', h = j^*)$  in  $N$  and the sequence  $(x, y, b', c)$  in  $X$  satisfy Condition  $\alpha$ .

**Case B.II.1.ii**  $r(y, P'_h) = 1$ .

So  $r(x, P'_h) > 1$ , hence  $h \in \tilde{N}$  which implies  $b P_h x$  as we are in Case B.II. Thus  $x \notin L(x, P_{i^*}) \cap (X \setminus L(b, P_h))$ . By Lemma A, there exists  $z \in L(x, P_{i^*}) \cap (X \setminus L(b, P_h))$  with  $z \notin \{x, y\}$ . Thus,  $z \in L(x, P_{i^*})$  and  $b \in L(z, P_h)$ . Moreover,  $r(z, P'_h) > 1$ . The sequence  $(i^*, h, j^*)$  in  $N$  and the sequence  $(x, y, b, c)$  in  $X$  satisfy conditions  $(\alpha_1)$ ,  $(\alpha_2)$  and  $(\alpha_3)$ . Moreover, as  $r(y, P'_h) = 1$ , we have  $r(c, P'_h) > 1$ , establishing condition  $(\alpha_4)$ , thus Condition  $\alpha$ .

**Case B.II.2**  $r(b, P_i) \leq r^*(P)$  for all  $i \in N$ .

We handle Case B.II.2 in two subcases:

**Case B.II.2.a**  $\#\{z \in A \mid z P_i c \text{ and } c P'_i z \text{ for some } i \in N\} \geq 2$ .

Take some  $b' \in \{z \in A \mid z P_i c \text{ and } c P'_i z \text{ for some } i \in N\}$  with  $b' \neq b$ . Suppose, for a contradiction, that  $r(b', P_i) \leq r^*(P)$  for all  $i \in N$ . As  $x$  and  $b$  are chosen by  $f$  at  $P$  at rank  $r^*(P)$ , it must be the case that  $n(y, r^*(P) - 1, P) \leq q - 1$  for  $y \in \{x, b, b'\}$ . However,  $n(y, r^*(P), P) = n$  for  $y \in \{x, b, b'\}$ , which implies  $3(n - (q - 1)) \leq n$ , which in turn implies  $q \geq 2n/3 + 1$ , giving a contradiction. Hence,  $r(b', P_h) > r^*(P)$  for some  $h \in N$ , which brings us to case B.II.1.

**Case B.II.2.b**  $\#\{z \in A \mid z P_i c \text{ and } c P'_i z \text{ for some } i \in N\} = 1$ .

So  $b$  is the unique alternative which satisfies  $b P_i c$  and  $c P'_i b$  for some  $i \in N$ . Let  $K(b, c) = \{i \in N \mid b P_i c \text{ and } c P'_i b \text{ for some } i \in N\}$ . We have  $r(c, P'_i) = r(c, P_i) - 1$  for every  $i \in K(b, c)$ , as otherwise we would contradict  $\#\{z \in A \mid z P_i c \text{ and } c P'_i z \text{ for some } i \in N\} = 1$ . As  $n(y, r^*(P), P) = n$  for  $y \in \{x, b\}$ , we have  $r^*(P') < r^*(P)$ , and more precisely  $r^*(P') = r^*(P) - 1$ . As a result,  $r(c, P_i) = r^*(P)$ , hence  $r(c, P'_i) = r^*(P) - 1$  for some  $i \in K(b, c)$ .

Now, let  $S_1 = \{i \in N : r(c, P_i) = r^*(P) \text{ and } r(c, P'_i) = r^*(P) - 1\}$ ,  $S_2 = \{i \in N : r(b, P_i) = r^*(P)\}$ ,  $S_3 = \{i \in N : r(x, P_i) = r^*(P)\}$ ,  $S_4 = N \setminus (S_1 \cup S_2 \cup S_3)$ . Let  $s_i$  stand for the cardinality of  $S_i$ ,  $i \in \{1, 2, 3, 4\}$ . Note that  $s_2$  and  $s_3$  are each at least  $n - q + 1$ , by the definition of  $r^*(P)$ . So we write  $s_2 = n - q + 1 + t_2$  and  $s_3 = n - q + 1 + t_3$  for some integer  $t_2, t_3 \geq 0$ . Thus,  $s_4 = 2q - n - 2 - s_1 - t_2 - t_3$ .

Now, we make five observations regarding  $S_1, S_2$ , and  $S_3$ .

<sup>18</sup>If  $r(b', P_i) \leq r^*(P)$  for all  $i \in N$ , then we proceed to Case B.II.2.

<sup>19</sup>If  $r(y, P'_{h'}) = 1$ . then we proceed to Case B.II.1.ii.

**Observation 1:** As we have  $n(y, r^*(P), P) = n$  for  $y \in \{x, b\}$ , we observe  $x P_i c$  and  $b P_i c$  for every  $i \in S_1$ ;  $x P_i b$  for every  $i \in S_2$ ; and  $b P_i x$  for every  $i \in S_3$ .

**Observation 2:** Suppose  $r(c, P_i) \leq r^*(P)$  for every  $i \in S_2 \cup S_3$  which implies  $r(c, P_i) \leq r^*(P) - 1$  for every  $i \in S_2 \cup S_3$ . As  $s_2 + s_3 = 2n - 2q + 2 + t_2 + t_3 \geq q$ , this contradicts the definition of  $r^*(P)$ . Hence, we observe the existence of some  $i \in S_2 \cup S_3$  with  $r(c, P_i) > r^*(P)$ .

**Observation 3:** As  $c \in f(P')$  with  $r^*(P') = r^*(P) - 1$ , we have  $n(c, r^*(P) - 1, P') \geq q$ . Moreover,  $q - (\#S_1 + \#S_4) = n - q + 2 + t_2 + t_3 > \#S_2, \#S_3$ . Thus, we observe that  $\exists i \in S_2$  with  $c P_i b$  and  $\exists i \in S_3$  with  $c P_i x$ .

**Observation 4:** As we are in Case B,  $r(x, P'_i) \leq r(x, P_i) \forall i \in N$ . Suppose  $r(x, P'_i) < r(x, P_i) \forall i \in S_3$ , which implies  $n(x, r^*(P) - 1, P') \geq s_1 + s_2 + s_3 \geq q$ , which would contradict  $x \notin f(P')$ . Thus, we observe that  $r(x, P'_i) = r(x, P_i) = r^*(P)$  for some  $i \in S_3$ .

**Observation 5:** As  $r^*(P) < m$ , it follows that  $m \geq 4$ , so we observe the existence of some  $d \in X \setminus \{x, b, c\}$ .

Under these five observations, we establish Condition  $\alpha$  by considering the following four exhaustive and mutually exclusive cases (which are all subcases of Case B.II.2.b):

**Case i:**  $r(c, P_i) \leq r^*(P)$  for every  $i \in S_3$ .

Let  $S'_2 = \{i \in S_2 \mid c P_i b\}$  and  $S''_2 = \{i \in S_2 \mid b P_i c\}$ . Note that  $S'_2$  and  $S''_2$  partition  $S_2$ . Also, combining  $r(c, P_i) \leq r^*(P) \forall i \in S_3$  with Observation 2 ensures the non-emptiness of  $S''_2$ . Moreover, given any  $i \in S''_2$  and any  $j \in S_3$ , there exists  $d_{ij} \in X \setminus \{x, b, c\}$  with  $r(d_{ij}, P_i) < r^*(P)$  and  $r(d_{ij}, P_j) > r^*(P)$ .

**Case i.1:** There is some  $j_1 \in S''_2, j_2 \in S_3$  with  $r(d_{j_1 j_2}, P'_{j_1}) > 1$ . Condition  $\alpha$  is established by the sequence  $(x, d, b, c)$  in  $X$  and the sequence  $(i_0, i_1, i_2)$  in  $N$  with  $i_0 = j_2 \in S_3, i_1 = j_1 \in S''_2$  and  $i_2 \in S_1$ .

**Case i.2:**  $r(d_{ij}, P'_i) = 1$  for each  $i \in S''_2, j \in S_3$ .

Condition  $\alpha$  is established by the sequence  $(x, b, c)$  in  $X$  and the sequence  $(i_0, i_1)$  in  $N$  with  $i_0 \in S''_2$  and  $i_1 \in S_1$ .

**Case ii:**  $r(c, P_i) > r^*(P)$  for some  $i \in S_3$ .

Let  $S'_3 = \{i \in S_3 \mid x P_i c\}$ . Note that  $S'_3 \neq \emptyset$  since  $r(x, P_j) = r^*(P)$  for each  $j \in S_3$  and  $r(c, P_i) > r^*(P)$  for some  $i \in S_3$ .

**Case ii.1:**  $r(c, P'_{j_1}) > 1$  for some  $j_1 \in S_2$ .

We know, by Observation 4, that  $r(x, P'_{j_0}) > 1$  for some  $j_0 \in S_3$ . First, let  $j_0 \in S'_3$ . Condition  $\alpha$  is established by the sequence  $(x, c, b, c)$  in  $X$  and the sequence  $(i_0, i_1, i_2)$  in  $N$  with  $i_0 = j_0 \in S'_3, i_1 = j_1 \in S_2$  and  $i_2 \in S_1$ .

Now, let  $j_0 \notin S'_3$  hence  $j_0 \in S''_3$ . Note that there is some  $d \in X \setminus \{x, b, c\}$  with  $x P_{j_0} d$  and  $d P_{j_1} x$  for some  $j_1 \in S'_3$ . Condition  $\alpha$  is established by the sequence  $(x, d, c, b, c)$  in  $X$  and the sequence  $(i_0, i_1, i_2, i_3)$  in  $N$  with  $i_0 = j_0 \in S''_3, i_1 = j_1 \in S'_3, i_2 = j_1 \in S'_2$  and  $i_3 \in S_1$ .

**Case ii.2:**  $r(c, P'_i) = 1$  for all  $i \in S_2$ .



Condition  $\alpha$  is established by the sequence  $(x, b, c)$  in  $X$  and the sequence  $(i_0, i_1)$  in  $N$  with  $i_0 \in S_2, i_1 \in S_1$ .  $\square$

Remark that the positive result of Theorem 5 (i.e., that  $f_q$  is subgame perfect implementable for  $q \in \{\lceil \frac{n}{2} \rceil, \dots, \lceil \frac{2n}{3} \rceil\}$ ) is valid for any choice of  $m, n \geq 3$ . The negative result of Theorem 3 is also of the same spirit:  $f_q$  fails to be subgame perfect implementable for  $q \in \{1, \dots, \lceil \frac{n}{2} \rceil - 1\}$  for any  $m \geq 3$  and  $n \geq 2$ . On the other hand, the negative result of Theorem 4 is less general, as it states that when  $q \in \{\lceil \frac{2n}{3} \rceil + 1, \dots, n\}$ , one cannot ensure the subgame perfect implementability of  $f_q$  for any  $m \geq 3$  and  $n \geq 2$ . This raises the question whether there are choices of  $m \geq 3$  and  $n \geq 2$  where  $f_q$  is subgame perfect implementable even when  $q \in \{\lceil \frac{2n}{3} \rceil + 1, \dots, n\}$ . By Proposition 5, the only alternatives are those which ensure  $r^*(q, P) < m$  for every  $P \in \mathcal{P}^n$ . As a case in point, we denote in the sequel  $FB : \mathcal{P}^2 \rightarrow X$  for the SCR which is a particular case of  $q$ -approval fallback bargaining where  $q = n = 2$ .<sup>20</sup> By Proposition 4,  $FB$  ensures  $r^*(q, P) < m$  for every  $P \in \mathcal{P}^n$ . Moreover, and interestingly,  $FB$  satisfies Condition  $\alpha$ , as we state and show below.

**Proposition 7.**  $FB : \mathcal{P}^2 \rightarrow X$  satisfies Condition  $\alpha$ .

*Proof.* Let  $N = \{1, 2\}$ . Take any  $P, P' \in \mathcal{P}^2$  and any  $x \in FB(P) \setminus FB(P')$ . We write  $r^*(P)$  instead of  $r^*(q, P)$  to simplify notation. If there exists some  $i^* \in N$  and some  $z \in L(x, P_{i^*}) \setminus L(x, P'_{i^*})$ , then the sequence  $(x, z)$  in  $X$  and  $i^* \in N$  establish Condition  $\alpha$ . So from now on, we assume  $L(x, P_i) \subseteq L(x, P'_i)$  for every  $i \in N$ . There are four mutually exclusive and exhaustive cases to consider:

**Case 1:**  $|FB(P)| > 1$ .

By definition of  $FB$ , we can only have  $|FB(P)| = 2$ . Write  $FB(P) = \{x, y\}$ . Again by definition of  $FB$ ,  $r(x, P_i) = r^*(P)$  for some  $i \in N$  and  $r(y, P_j) = r^*(P)$  for some  $j \in N \setminus \{i\}$ . Let, without loss of generality,  $i = 1$  and  $j = 2$ . Note  $y P_1 x$  and  $x P_2 y$ .

**Case 1.1:**  $z \in FB(P')$  for some  $z \notin FB(P)$ .

As  $z \in FB(P') \setminus$  and  $z \notin FB(P)$ ,  $\exists i^* \in N$  with  $r(z, P_{i^*}) > r^*(P)$  and  $r(z, P_{i^*}) \leq r^*(P)$ . As  $L(x, P_i) \subseteq L(x, P'_i) \forall i \in N$  and  $r(x, P_1) = r^*(P)$ , we have  $i^* = 2$ . So  $y P_2 z$  and  $z P'_2 y$ . So,  $z P_1 x$ . Moreover,  $y P_1 x$ . Thus,  $\exists t \in X \setminus \{x, y, z\}$  with  $x P_1 t$  and  $t P_2 y$ . Condition  $\alpha$  is established by the sequence  $(x, t, y, z)$  in  $X$  and the sequence  $(1, 2, 2)$  in  $N$ .

**Case 1.2:**  $FB(P') = \{y\}$ .

As  $L(x, P_i) \subseteq L(x, P'_i) \forall i \in N$ , we have  $t P_2 y$  and  $y P'_2 t$  for some  $t \in X \setminus \{x, y\}$ . Moreover,  $x P_1 t$ , as otherwise we would have  $FB(P) = \{t\}$ . Condition  $\alpha$  is established by the sequence  $(x, t, y)$  in  $X$  and the sequence  $(1, 2)$  in  $N$ .

**Case 2:**  $FB(P) = \{x\}$ .

<sup>20</sup> $FB$  stands for ‘‘Fallback Bargaining’’ which Brams and Kilgour [2001] suggest as a bargaining solution in an environment where agents confront a finite set of alternatives.  $FB$  appears in the literature under different names, such as ‘‘Rawlsian arbitration rule’’ in Sprumont [1993], ‘‘Kant-Rawls social compromise’’ in Hurwicz and Sertel [1999] and ‘‘unanimity compromise’’ in Kibris and Sertel [2007].

**Case 2.1:**  $r(x, P_1) = r(x, P_2) = r^*(P)$ .

As  $L(x, P_i) \subseteq L(x, P'_i) \forall i \in N$ , we have  $FB(P') = \{x\}$ , contradicting  $x \in FB(P) \setminus FB(P')$ .

**Case 2.2:**  $r(x, P_i) = r^*(P)$  for some  $i \in N$  and  $r(x, P_j) < r^*(P)$  for some  $j \in N \setminus \{i\}$ .

Let, without loss of generality,  $i = 1$  and  $j = 2$ . So  $r(y, P_2) = r^*(P)$  for some  $y \in X \setminus \{x\}$ . As  $FB(P) = \{x\}$ , we have  $x P_1 y$ . As  $L(x, P_i) \subseteq L(x, P'_i) \forall i \in N$ , we have  $z \in FB(P')$  for some  $z \in X \setminus \{x, y\}$ . Note that  $z P_1 x$ , as  $L(x, P_i) \subseteq L(x, P'_i) \forall i \in N$ . Moreover,  $y P_2 z$ , as  $z \notin FB(P)$ . Also,  $z P'_2 y$ , as  $z \in FB(P')$ . Condition  $\alpha$  is established by the sequence  $(x, y, z)$  in  $X$  and the sequence  $(1, 2)$  in  $N$ .  $\square$

The sufficiency of Condition  $\alpha$  and WNVP for subgame perfect implementability requires at least three agents and the literature does not admit a general sufficient condition for subgame perfect implementation with two agents.<sup>21</sup> As a result, we can only deduce from Proposition 7 a possibility for the subgame perfect implementability of  $FB$ . However, Anbarci [2006] suggests two two-player mechanisms, namely “alternate-strikes” and “voting by alternating offers”, whose subgame perfect equilibria outcomes are subsets of (without necessarily coinciding with) the  $FB$  outcomes at every preference profile.

## 6. FINAL REMARKS

In a literature where relatively little is known about which SCRs are subgame perfect implementable, the picture we draw can be qualified as broad. We show, on one hand, a general incompatibility between scoring rules and subgame perfect implementability; and on the other hand, we establish the existence of several interesting Condorcet consistent subgame perfect implementable SCRs. It is interesting to observe the historical Borda - Condorcet tension being reflected to the problem of implementation by extensive form mechanisms.

Our results can be interpreted as positive, at least compared to the literature on Nash implementation where several interesting SCRs fail Nash implementability. It is known that scoring rules, point runoff systems, SCRs based on the majority relation (such as the top-cycle, the uncovered set, the minimal covering set, the bi-partisan set), compromise rules (such as majoritarian compromise, fallback bargaining) all fail to be Nash implementable. Our findings show that the top-cycle, the minimal covering set, the bi-partisan set are all examples of SCRs which are subgame perfect implementable while they fail to be Nash implementable.<sup>22</sup>

<sup>21</sup>The only analysis we know in this direction is Vartiainen [2006] which is an incomplete working paper. We thank Hannu Vartiainen for generously providing us with his work.

<sup>22</sup>Interestingly, the set monotonicity condition which ensures the subgame perfect implementability of Condorcet consistent rules has been shown to ensure the Nash implementability of SCRs via set-valued normal form mechanisms in the environments of Bochet and Maniquet [2010] and Özkal-Sanver and Sanver [2006]. Moreover, Brandt [2015] shows that a slightly stronger version of that condition, together with particular assumptions on how voters extend their preferences over sets, ensures strategy-proofness of set valued SCRs.

An early example in this direction is given by Sertel and Yılmaz [1999] who show that the majoritarian compromise, which fails Nash implementability, is subgame perfect implementable. We analyze the subgame perfect implementability of the whole class of  $q$ -approval fallback bargaining rules (of which the majoritarian compromise is the member with  $q = \frac{1}{2}$ ) and show that these rules are subgame perfect implementable iff  $q$  is between the half and two thirds of the total number of agents. This mixed picture is an observation of interest, as compromise rules stand apart of the Borda - Condorcet tension<sup>23</sup> which is also reflected to their implementability. As a particular instance of  $q$ -approval fallback bargaining rules, setting  $q = n$ , we have fallback bargaining which has first been proposed as bargaining solution by Brams and Kilgour [2001] and later been characterized by Congar and Merlin [2012] in the context of voting. Our results show that the subgame perfect implementability of fallback bargaining depends on the size of the social choice problem. In particular, with two agents, fallback bargaining satisfies Condition  $\alpha$  and, as Anbarci (2006) shows, there exists mechanisms that implement fallback bargaining via subgame perfect equilibria.<sup>24</sup>

An analysis close to ours is by Vartiainen [2007b] who considers randomized mechanisms which renders subgame perfect implementation easier to achieve. As such, he shows the subgame perfect implementability of several Condorcet consistent rules, which is implied by our positive results on Condorcet consistent rules. In a similar vein, his negative result on scoring rules (Proposition 12 in Vartiainen [2007b]) can be transferred to our framework. Nevertheless, our Proposition 2 on the non-existence of subgame perfect implementable scoring rules is more general with respect to the size of the society, as Proposition 12 of Vartiainen [2007b] is established for 12 voters only. It is also worth noting that plurality with a runoff stands as a nice example which shows the difference between the two environments: while Vartiainen [2007b] shows its subgame perfect implementability by randomized mechanisms, we show that it is not subgame perfect implementable in our deterministic environment.

Dutta and Sen [1993] also establish positive results on the subgame perfect implementability of Condorcet social choice functions. They consider implementation with game trees which are completely sequential, i.e., all information sets are singletons. This

<sup>23</sup>As Merlin et al. [2019] discuss, compromise rules are not Condorcet consistent, nor they can be qualified as scoring rules, although they can be expressed in terms of elementary scoring rules which vary as a function of the preference profile.

<sup>24</sup>Anbarci (2006) does not derive general subgame perfect implementability results, but suggests simple extensive form mechanisms which implements fallback bargaining. His analysis inspires a question of interest within the voting context: Our positive results are based on the rather complicated canonical mechanism of Abreu and Sen [1990] which implements any SCR which satisfies the sufficient conditions of subgame perfect implementability. Is there simple mechanisms, specific to each of the SCRs we show to be subgame perfect implementable, which implement the SCR in question? As another analysis in this direction, we have Suh and Wen [2008] who consider matching problems and without deriving general subgame perfect implementability results, suggest a simple extensive form mechanism which implements a given matching solutions under (rather strong) domain restriction assumptions.

makes implementation harder compared to the Abreu and Sen [1990] environment where mechanisms use simultaneous games. In fact, when all moves are required to be sequential, one cannot establish a sufficiency result by checking Condition  $\alpha$ . As a result, the findings of Dutta and Sen [1993] on Condorcet social choice functions can be qualified as “stronger” than ours. On the other hand, as they consider single-valued social choice functions (as it must be the case with completely sequential game trees in voting environment with linear orders), our results are logically independent.<sup>25</sup>

An interesting question is to see whether our positive results prevail when singleton-valued SCRs are considered. We know by Muller and Satterthwaite [1977] that there is a severe tension between singleton-valuedness and Nash implementability of SCRs: when singleton-valuedness is imposed, only dictatorial or constant SCRs are Maskin monotonic, hence Nash implementable. It is worth noting that the set monotonicity condition which ensures the subgame perfect implementability of Condorcet consistent rules becomes equivalent to Maskin monotonicity when SCRs are singleton-valued. So whether the Muller-Satterthwaite impossibility has a counterpart for extensive form mechanisms seems as an interesting question to elaborate.

#### REFERENCES

- D. Abreu and A. Sen. Subgame perfect implementation: A necessary and almost sufficient condition. *Journal of Economic Theory*, 50(2):285–299, 1990.
- N. Anbarci. Finite Alternating-Move Arbitration Schemes and the Equal Area Solution. *Theory and Decision*, 61:21–50, 2006.
- G. Aşan and M.R. Sanver. Maskin monotonic aggregation rules. *Economics Letters*, 91(2):179–183, 2006.
- J.S. Banks. Sophisticated voting outcomes and agenda control. *Social Choice and Welfare*, 1(4):295–306, 1985.
- G.W. Bassett and J. Persky. Robust voting. *Public Choice*, 99(3):299–310, 1999.
- O. Bochet and F. Maniquet. Virtual Nash implementation with admissible support. *Journal of Mathematical Economics*, 46(1):99–108, 2010.
- S.J. Brams and D. M. Kilgour. Fallback bargaining. *Group Decision and Negotiation*, 10(4):287–316, 2001.
- F. Brandt. Set-monotonicity implies Kelly-strategyproofness. *Social Choice and Welfare*, 45(4):793–804, 2015.
- R. Congar and V. Merlin. A characterization of the maximin rule in the context of voting. *Theory and Decision*, 72(1):131–147, 2012.

---

<sup>25</sup>Dominance solvable voting schemes are also related to our analysis. As argued by Dutta and Sen [1993], “a social choice function that can be implemented via backward induction can also be implemented via sophisticated equilibrium”. However, implementability via sophisticated equilibrium and subgame perfect implementability are logically independent.

- B Dutta. On the tournament equilibrium set. *Social Choice and Welfare*, 7(4):381–383, 1990.
- B. Dutta and A. Sen. Implementing generalized Condorcet social choice functions via backward induction. *Social Choice and Welfare*, 10(2):149–160, 1993.
- O. Erdem and M. R. Sanver. Minimal monotonic extensions of scoring rules. *Social Choice and Welfare*, 25(1):31–42, 2005.
- M.J. Herrero and S. Srivastava. Implementation via backward induction. *Journal of Economic Theory*, 56(1):70–88, 1992.
- S.. Horan. Implementation of Majority Voting Rules. mimeo, Universite de Montréal, 2013.
- L. Hurwicz and M.R. Sertel. Designing mechanisms, in particular for electoral systems: the Majoritarian Compromise. In R. Sertel Murat, editor, *Economic Design and Behaviour*. London: MacMillan, 1999.
- M.O. Jackson. A crash course in implementation theory. *Social Choice and Welfare*, 18(4):655–708, 2001.
- O. Kibris and M.R. Sertel. Bargaining Over a Finite Set of Alternatives. *Social Choice and Welfare*, 28:421–437, 2007.
- G. Laffond, J.-F. Laslier, and M. Le Breton. The bipartisan set of a tournament game. *Games and Economic Behavior*, 5(1):182–201, 1993.
- J.F. Laslier. *Tournament Solutions and Majority Voting*. Springer Verlag, 1997.
- E. Maskin. Nash Equilibrium and Welfare Optimality. *Review of Economic Studies*, 66: 23–38, 1999.
- V. Merlin, I. Özkal-Sanver, and M. R. Sanver. Compromise rules revisited. *Group Decision and Negotiation*, 28:63–78, 2019.
- J. Moore and R. Repullo. Subgame perfect implementation. *Econometrica*, 56(5):1191–1220, 1988.
- H. Moulin. Choosing from a tournament. *Social Choice and Welfare*, 3(4):271–291, 1986.
- E. Muller and M.A Satterthwaite. The equivalence of strong positive association and strategy-proofness. *Journal of Economic Theory*, 14(2):412–418, 1977.
- H. Nurmi. *Voting Paradoxes and How to Deal with Them*. Springer, 1999.
- I. Özkal-Sanver and M. R. Sanver. Nash implementation via hyperfunctions. *Social Choice and Welfare*, 26(3):607–623, 2006.
- I. Özkal-Sanver and M.R. Sanver. A new monotonicity condition for tournament solutions. *Theory and Decision*, 69(3):439–452, 2010.
- A. Sen. *Two Essays in the theory of implementation*. PhD thesis, Princeton University, 1987.
- M.R. Sertel and B. Yılmaz. The majoritarian compromise is majoritarian-optimal and subgame-perfect implementable. *Social Choice and Welfare*, 16(4):615–627, 1999.

- J. Smith. Aggregation of preferences with a variable electorate. *Econometrica*, 41:1027–1041, 1973.
- Y. Sprumont. Intermediate Preferences and Rawlsian Arbitration Rules. *Social Choice and Welfare*, 10:1–15, 1993.
- S.C. Suh and Q. Wen. Subgame perfect implementation of stable matchings in marriage problems. *Social Choice and Welfare*, 31(1):163–174, 2008.
- H. Vartiainen. A necessary and sufficient condition for two person subgame perfect implementability. mimeo, 2006.
- H. Vartiainen. Subgame perfect implementation: A full characterization. *Journal of Economic Theory*, 133(1):111–126, 2007a.
- H. Vartiainen. Subgame perfect implementation of voting rules via randomized mechanisms. *Social Choice and Welfare*, 29(3):353–367, 2007b.

#### APPENDIX A. PROOF OF LEMMA A

By definition, there is some  $i \in \tilde{N}$  with  $r(x, P_i) \in \{2, \dots, r^*(P)\}$  and some  $h \in N$  with  $r(b, P_h) \in \{r^*(P) + 1, \dots\}$ . Thus,  $r(x, P_i) < r(b, P_h)$ .

In the sequel of the proof, we write

$$\begin{aligned} a_v &= \{y \in X \mid r(y, P_i) = r(x, P_i) + v\} \text{ for } v = 1, \dots, m - r(x, P_i) \\ \text{and } b_w &= \{y \in X \mid r(y, P_h) = w\} \text{ for } w = 1, \dots, r(b, P_h) - 1. \end{aligned}$$

Observe that  $L(x, P_i) = \{x\} \cup \{a_1, \dots, a_{m-r(x, P_i)}\}$  and  $X \setminus L(b, P_h) = \{b_1, \dots, b_{r(b, P_h)-1}\}$  with  $s < t \implies a_s P_i a_t$  and  $b_s P_h b_t$ . The proof is now divided in two cases: if  $x \in L(x, P_i) \cap (X \setminus L(b, P_h))$  and otherwise.

**Case 1:** Consider first the case where  $x \in L(x, P_i) \cap (X \setminus L(b, P_h))$ . We will show that  $L(x, P_i) \cap (X \setminus L(b, P_h)) \neq \{x\}$ .

By definition,  $|X \setminus L(b, P_h)| = r(b, P_h) - 1$ . Yet, since  $x \in L(x, P_i) \cap (X \setminus L(b, P_h))$  and  $b \notin X \setminus L(b, P_h)$ , there are  $r(b, P_h) - 2$  different alternatives in  $X \setminus L(b, P_h)$  different from  $b$  and  $x$ . Similarly  $|L(x, P_i)| = m - r(x, P_i) + 1$  jointly with  $x \in L(x, P_i) \cap (X \setminus L(b, P_h))$  and  $b \notin X \setminus L(b, P_h)$  implies that there are  $m - r(x, P_i)$  different alternatives in  $L(x, P_i)$  different from  $b$  and  $x$ . However  $m - r(x, P_i) + r(b, P_h) - 2 = m + (r(b, P_h) - r(x, P_i)) - 2 \geq m - 1$  since  $r(x, P_i) < r(b, P_h)$ . However, there cannot be  $m - 1$  alternatives different from  $b$  and  $x$  in  $X$ . Therefore, there is at least one alternative in common between  $L(x, P_i) \setminus \{b, x\}$  and  $(X \setminus L(b, P_h)) \setminus \{b, x\}$ , concluding case 1.

**Case 2:** Consider now the case that  $x \notin L(x, P_i) \cap (X \setminus L(b, P_h))$ .

There are again  $m - r(x, P_i)$  different alternatives in  $L(x, P_i)$  which are different from  $b$  and  $x$ . As  $x \notin X \setminus L(b, P_h)$ , there are now  $r(b, P_h)$  different alternatives in  $X \setminus L(b, P_h)$  different from  $b$  and  $x$ . As  $r(x, P_i) < r(b, P_h)$ , we have  $(m - r(x, P_i)) + (r(b, P_h)) > m$  but

there cannot be  $m$  alternatives different from  $b$  and  $x$ . Thus  $|L(x, P_i) \cap (X \setminus L(b, P_h))| \geq 2$ , as wanted, concluding the proof.