



## **Impressum**

Karlsruher Institut für Technologie (KIT)  
Fakultät für Wirtschaftswissenschaften  
Institut für Volkswirtschaftslehre (ECON)

Schlossbezirk 12  
76131 Karlsruhe

KIT – Universität des Landes Baden-Württemberg und  
nationales Forschungszentrum in der Helmholtz-Gemeinschaft

Working Paper Series in Economics  
**No. 50**, December 2013

ISSN 2190-9806

---

[econpapers.wiwi.kit.edu](http://econpapers.wiwi.kit.edu)

# Unanimity Overruled: Majority Voting and the Burden of History

Klaus Nehring\*, Marcus Pivato<sup>†</sup>, and Clemens Puppe<sup>‡</sup>

August 2013

## Abstract

Sequential majority voting over interconnected binary propositions can lead to the overruling of unanimous consensus. We characterize, within the general framework of judgement aggregation, under what circumstances this happens for some sequence of the voting process. It turns out that the class of aggregation spaces for which this difficulty arises is very large, including the aggregation of preference orderings over at least four alternatives, the aggregation of equivalence relations over at least four objects, resource allocation problems, and most committee selection problems.

We also ask whether it is possible to *design* respect for unanimity by choosing appropriate decision sequences. Remarkably, while this is not possible in general, it can be accomplished in interesting special cases. Adapting and generalizing a classic result by Shepsle and Weingast, we show that respect for unanimity can indeed be thus guaranteed in case of the aggregation of weak orderings, strict orderings and equivalence relations.

A well-known problem from the theory of preference aggregation is that sequential pairwise majority voting can yield Pareto-dominated alternatives, provided that there are at least four alternatives (Moulin, 1988, p.242). Within the general framework of judgement aggregation this is an instance of the fact that sequential majority voting over interconnected propositions can lead to the situation that a previous majority decision forces the rejection of a proposition that is unanimously accepted (see Section 1.5 below for examples).

In this paper, we characterize the aggregation spaces on which sequential majority voting always respects unanimous consent, no matter in which sequence the majority decisions are taken and no matter what the individual judgements are. We shall call such spaces *robustly unanimity consistent*. It turns out that the robustly unanimity consistent spaces represent a simple and natural generalization of the median spaces which play a distinguished role in the theory of aggregation (see Barthélémy and Monjardet (1981), Nehring

---

\*Department of Economics, University of California at Davis, U.S.A. [kdnehring@ucdavis.edu](mailto:kdnehring@ucdavis.edu).

<sup>†</sup>Department of Mathematics, Trent University, Canada. [marcuspivato@trentu.ca](mailto:marcuspivato@trentu.ca)

<sup>‡</sup>Department of Economics and Business Engineering, Karlsruhe Institute of Technology, Germany. [clemens.puppe@kit.edu](mailto:clemens.puppe@kit.edu)

and Puppe (2002, 2007)). Median spaces are characterized by the condition that the underlying feasibility restrictions among propositions can be described in terms of ‘simple’ implications each having exactly one proposition as its antecedent. By contrast, the robustly unanimity consistent spaces are characterized by the condition that the underlying feasibility restrictions among propositions are confined to implications in which the antecedent consists of the conjunction of at most *two* propositions. Evidently, this is still a rather special case; the result thus shows that potential unanimity violations are quite common in judgement aggregation.

While in most spaces violations of unanimity can occur for *some* sequence of majority decisions (given some profile of individual judgements), a weaker requirement is that they do not occur for *all* sequences. The underlying idea is that one can perhaps prevent unanimity violations to occur by an appropriate design of the decision path, i.e. by a suitable restriction on the set of admissible (sequential) agendas. Accordingly, we shall say that an aggregation space is *designably unanimity consistent* if there *exists* a sequence such that unanimous decisions are always respected provided that the majority decisions are taken in the prespecified order. It turns out that many spaces are not even designably unanimity consistent. However, there are interesting classes of aggregation problems that do have this property. Here, we prove that the spaces of all strict orderings, all weak orderings, and all equivalence relations on any finite set of alternatives are designably consistent, respectively. Our analysis generalizes classic results of Shepsle and Weingast (1984); in particular, we define an appropriate generalization of the ‘covering relation’ corresponding to a majority tournament, and show that it always respects unanimous decisions. We then construct a decision sequence such that the resulting transitive binary relation extends the covering relation, which yields the desired conclusion. While a full characterization of the class of all aggregation problems that are designably unanimity consistent appears to be difficult, we are able to derive a relatively simple necessary condition, which requires that certain ‘exposed’ feasibility restrictions have limited reach.

The agenda setting problem is a classic topic in political science, see e.g. Miller (1995). For previous considerations of sequential majority voting in judgement aggregation, see List (2004) and Dietrich and List (2007). The present paper complements the analysis of Nehring et al. (2013) where we analyze sequential majority voting in terms of the entailed extent of indeterminacy, following the classical approach of McKelvey (1979). Here, we concentrate on assessing the qualitative ‘severity’ of the path-dependence of sequential majority voting, as embodied by the violation of unanimous judgements. The widespread tension between unanimity on the one hand and ‘premise-based’ majoritarian and supermajoritarian voting rules on the other has been studied in the special case of truth-functional judgement aggregation problems in Nehring (2005).

This paper is organized as follows. In Section 1, we introduce the judgement aggregation model along with our terminology and notation. We also present some basic examples of aggregation spaces corresponding to preference aggregation, the aggregation of equivalence relations, the aggregation of general binary relations, committee selection and resource allocation, respectively. In Section 2, we characterize when sequential majority voting can override a unanimous consensus of voters on some issues for some sequence of the voting process, i.e. the class of robustly unanimity consistent spaces. In Section 3, we consider the

weaker condition that unanimous consensus be respected for *some* appropriate sequence of the voting process, i.e. designable unanimity consistency, and give positive as well as negative examples. For ease of reading, some technically more involved proofs are relegated to an appendix at the end of the paper.

# 1 Sequential majority voting and the Condorcet set

## 1.1 The judgement aggregation model

Let  $K \in \mathbb{N}$ , and let  $[1\dots K]$  index a set of propositions or issues. An element  $\mathbf{x} = (x_1, x_2, \dots, x_K) \in \{0, 1\}^K$  is called a *view*, and interpreted as an assignment of a truth value of ‘true’ or ‘false’ to each proposition.<sup>1</sup> Not all views are feasible, because there will be logical constraints between the propositions (determined by the structure of the underlying decision problem faced by society). Let  $X \subseteq \{0, 1\}^K$  be the set of ‘admissible’ or *consistent* views; a given set  $X \subseteq \{0, 1\}^K$  is also referred to as an *aggregation space*. An *anonymous profile* is a probability measure on  $X$  —that is, a function  $\mu : X \rightarrow [0, 1]$  such that  $\sum_{\mathbf{x} \in X} \mu(\mathbf{x}) = 1$  with the interpretation that, for all  $\mathbf{x} \in X$ ,  $\mu(\mathbf{x})$  is the proportion of the voters who hold the view  $\mathbf{x}$ . For any  $Y \subseteq X$ , define  $\mu(Y) := \sum_{\mathbf{y} \in Y} \mu(\mathbf{y})$ . Let  $\Delta(X)$  be the set of all anonymous profiles. For any  $\mu \in \Delta(X)$ , any  $k \in [1\dots K]$ , let

$$\mu_k(1) := \mu\{\mathbf{x} \in X ; x_k = 1\} \quad (1)$$

be the total ‘popular support’ for the position “ $x_k = 1$ ,” and let  $\mu_k(0) := 1 - \mu_k(1)$ . Let  $\Delta^*(X) := \{\mu \in \Delta(X); \mu_k(1) \neq \frac{1}{2}, \forall k \in [1\dots K]\}$  be the set of anonymous profiles where there is a strict majority supporting either 0 or 1 in each coordinate. For expository convenience, we will confine our analysis to profiles in  $\Delta^*(X)$ .<sup>2</sup> For every profile  $\mu$ , we denote by  $\text{Maj}(\mu) \in \{0, 1\}^K$  the *issuewise majority* view, i.e. for any  $\mu \in \Delta^*(X)$ ,  $\text{Maj}_k(\mu) := 1$  if  $\mu_k(1) > \frac{1}{2}$ , and  $\text{Maj}_k(\mu) := 0$  if  $\mu_k(1) < \frac{1}{2}$ .

## 1.2 The Condorcet set

In general, one has  $\text{Maj}(\mu) \notin X$  —the ‘majority ideal’ can be inconsistent with the underlying logical constraints faced by society.<sup>3</sup> However, a basic principle of majoritarianism is that we should try to satisfy the majority’s will in as many issues as possible. In Nehring et al. (2013), we have thus proposed the notion of the *Condorcet set*  $\text{Cond}(X, \mu) \subseteq X$  as the set of all views  $\mathbf{x} \in X$  such that no  $\mathbf{y} \in X$  agrees with  $\text{Maj}(\mu)$  on a strictly larger set of issues than  $\mathbf{x}$ . The elements of  $\text{Cond}(X, \mu)$  are also referred to as the *Condorcet admissible* views. A profile  $\mu$  is called *majority determinate* if  $\text{Cond}(X, \mu)$  is single-valued;

<sup>1</sup>The literature on judgement aggregation uses various models and terms, see List and Puppe (2009) for an overview. Here, we adopt the terminology of Nehring et al. (2013).

<sup>2</sup>Note that if the set of voters is large (resp. odd), then a profile in  $\Delta(X) \setminus \Delta^*(X)$  is highly unlikely (resp. impossible).

<sup>3</sup>Condorcet’s ‘voting paradox’ is a well-known instance; another example that has received considerable interest recently is the so-called ‘discursive dilemma,’ see Kornhauser and Sager (1986), List and Pettit (2002).

in this case,  $\text{Cond}(X, \mu) = \{\text{Maj}(\mu)\}$  (see Nehring et al. (2013)). A space  $X$  is called *majority determinate* if all profiles  $\mu \in \Delta^*(X)$  are majority determinate.

### 1.3 Sequential majority voting

For any  $\mathbf{y} = (y_k)_{k=1}^K \in \{0, 1\}^K$  and  $J \subset [1 \dots K]$ , define  $\mathbf{y}_J := (y_j)_{j \in J} \in \{0, 1\}^J$ . Furthermore, for any  $i \in [1 \dots K] \setminus J$ , say that  $y_i$  is  *$X$ -consistent* with  $\mathbf{y}_J$  if there exists some  $\mathbf{x} \in X$  with  $\mathbf{x}_J = \mathbf{y}_J$  and  $x_i = y_i$  —otherwise  $y_i$  is  *$X$ -inconsistent* with  $\mathbf{y}_J$ . A *path* through  $[1 \dots K]$  is a bijection  $\gamma : [1 \dots K] \rightarrow [1 \dots K]$ . We now define the  *$\gamma$ -sequential majority rule*  $F^\gamma : \Delta^*(X) \rightarrow X$ . Let  $\mu \in \Delta^*(X)$  and let  $\mathbf{z} := \text{Maj}(\mu)$ . Define  $\mathbf{y} := F^\gamma(\mu) \in X$  inductively as follows:

- Define  $y_{\gamma(1)} := z_{\gamma(1)}$ .
- Inductively, let  $J := \{\gamma(1), \gamma(2), \dots, \gamma(n)\}$ , and suppose we have already decided  $\mathbf{y}_J$ . Let  $i := \gamma(n+1)$ . If  $z_i$  is  $X$ -consistent with  $\mathbf{y}_J$ , then set  $y_i := z_i$ . Otherwise, set  $y_i = \neg z_i$ .<sup>4</sup>

The following result, proven in Nehring et al. (2013), establishes a close connection between sequential majority voting and the Condorcet set:

**Proposition 1.1 (Nehring, Pivato and Puppe, 2012)** *Let  $X \subseteq \{0, 1\}^K$ .*

- (a) *For any path  $\gamma$  through  $[1 \dots K]$ , the rule  $F^\gamma$  is Condorcet admissible.*
- (b) *Conversely, for any  $\mu \in \Delta^*(X)$  and  $\mathbf{x} \in \text{Cond}(X, \mu)$ , there exists a path  $\gamma$  such that  $F^\gamma(\mu) = \mathbf{x}$ .*

### 1.4 Majority determinacy and path-independence

A profile  $\mu \in \Delta^*(X)$  is *path-independent* if  $F^\gamma(\mu) = F^\xi(\mu)$  for any two paths  $\gamma$  and  $\xi$  through  $[1 \dots K]$ . The space  $X$  itself is called *path-independent* if every  $\mu \in \Delta^*(X)$  is path-independent.

Let  $J \subseteq [1 \dots K]$  and consider  $\mathbf{w} \in \{0, 1\}^J$  which corresponds to a subset of judgements on the issues in  $J$ . The set  $J$  is the *support* of  $\mathbf{w}$ , denoted  $\text{supp}(\mathbf{w})$ . We define  $|\mathbf{w}| := |J|$ . If  $I \subseteq J$  and  $\mathbf{v} \in \{0, 1\}^I$ , then we say  $\mathbf{v}$  is a *fragment* of  $\mathbf{w}$  (and write  $\mathbf{v} \sqsubseteq \mathbf{w}$ ) if  $\mathbf{v} = \mathbf{w}_I$ . Furthermore,  $\mathbf{w}$  is a *forbidden fragment* for  $X$  if, for all  $\mathbf{x} \in X$ , we have  $\mathbf{x}_J \neq \mathbf{w}$  where  $J = \text{supp}(\mathbf{w})$ . Finally,  $\mathbf{w}$  is a *critical fragment* if it is a minimal forbidden fragment —that is,  $\mathbf{w}$  is forbidden, and there exists no proper subfragment  $\mathbf{v} \sqsubset \mathbf{w}$  such that  $\mathbf{v}$  is forbidden.<sup>5</sup> Let  $C(X)$  be the set of critical fragments for  $X$ , and let  $\kappa(X) := \max\{|\mathbf{c}|; \mathbf{c} \in C(X)\}$ .

A particular role is played by spaces  $X \subseteq \{0, 1\}^K$  for which  $\kappa(X) = 2$ ; these are known as *median spaces* in combinatorial mathematics. Their importance in the theory of aggregation has been emphasized by (Barthélemy and Monjardet, 1981; Nehring and

<sup>4</sup>“ $\neg$ ” represents logical negation. That is:  $\neg 1 := 0$  and  $\neg 0 := 1$ .

<sup>5</sup>Critical fragments are called ‘critical families’ in Nehring and Puppe (2007, 2010) and ‘minimal infeasible partial evaluations (MIPES)’ in Dokow and Holzman (2010).

Puppe, 2007, 2010). Note that a set of feasible views  $X$  is a median space if and only if all logical interrelations are confined to *simple* implications: for some  $j, k$  and all  $\mathbf{x} \in X$ ,  $x_j = 0$  implies that  $x_k = 0$ , or  $x_j = 0$  implies that  $x_k = 1$ . We have the following fundamental result (cf. (Nehring et al., 2013, Prop. 1.4), (Nehring and Puppe, 2007, Fact 3.4)).

**Proposition 1.2** (Nehring and Puppe, 2007; Nehring, Pivato and Puppe, 2012)  
*Let  $X \subseteq \{0, 1\}^K$ . The following are equivalent: [i]  $X$  is path-independent; [ii]  $X$  is majority determinate; [iii]  $X$  is a median space, i.e.  $\kappa(X) = 2$ .*

## 1.5 Examples

### 1.5.1 Aggregation of strict preference orderings

Let  $N \in \mathbb{N}$ , let  $A := [1\dots N]$  be some set of  $N$  social alternatives, let  $K := N(N - 1)/2$ , and bijectively identify  $[1\dots K]$  with a subset of  $A \times A$  which contains exactly *one* element of the set  $\{(a, b), (b, a)\}$  for each distinct  $a, b \in A$ . Then  $\{0, 1\}^K$  represents the space of all *tournaments* (i.e. complete, antisymmetric binary relations, or equivalently, complete directed graphs) on  $A$ . Let  $X_N^{\text{pr}} \subset \{0, 1\}^K$  be the set of all tournaments representing total orderings (i.e. permutations) of  $A$  (sometimes called the *permutahedron*). Classical Arrovian aggregation of strict preference orderings is simply judgement aggregation on  $X_N^{\text{pr}}$ . For any profile  $\mu \in \Delta^*(X_N^{\text{pr}})$ ,  $\text{Maj}(\mu)$  is the majority tournament corresponding to  $\mu$ . We will write  $a \succ^\mu b$  if the majority strictly prefers  $a$  to  $b$ , i.e. if  $\text{Maj}(\mu)$  entails a ‘yes’ on issue  $(a, b)$ . By definition, the Condorcet set  $\text{Cond}(X_N^{\text{pr}}, \mu)$  is the set of preference orderings on  $A$  such that no other ordering agrees with  $\text{Maj}(\mu)$  (resp.  $\succ^\mu$ ) on a larger set of pairwise comparisons.

As an illustrative example, consider the 4-permutahedron with alternatives  $a, b, c, d$ . Suppose that one third of the population endorses each of the preference orderings  $a \succ b \succ c \succ d$ ,  $b \succ c \succ d \succ a$  and  $c \succ d \succ a \succ b$ . For the corresponding majority tournament we have  $c \succ^\mu a$ ,  $d \succ^\mu a$ ,  $a \succ^\mu b$ ,  $b \succ^\mu c$ ,  $b \succ^\mu d$ , and  $c \succ^\mu d$  (see Figure 1). The Condorcet set is easily shown to consist of the following five orderings:  $a \succ b \succ c \succ d$ ,  $b \succ c \succ d \succ a$ ,  $c \succ d \succ a \succ b$ ,  $d \succ a \succ b \succ c$ ,  $c \succ a \succ b \succ d$ .

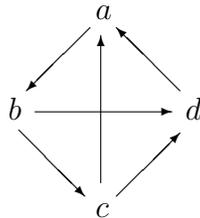


Figure 1: A majority tournament on four alternatives

### 1.5.2 Aggregation of equivalence relations

Let  $N \in \mathbb{N}$ ,  $A := [1\dots N]$ ,  $K := N(N-1)/2$ , and identify  $[1\dots K]$  with a subset of  $A \times A$  containing exactly one of the pairs  $(a, b)$  or  $(b, a)$  for each  $a \neq b \in A$ . Thus, an element of  $\{0, 1\}^K$  represents a symmetric, reflexive binary relation (i.e. undirected graph) on  $A$ . Denote by  $X_N^{\text{eq}} \subset \{0, 1\}^K$  be the set of all equivalence relations on  $A = [1\dots N]$ . Also in this case, one typically finds that the issue-wise majority view does not respect the underlying feasibility constraints, i.e. typically  $\text{Maj}(\mu) \notin X_N^{\text{eq}}$ . Consider, for instance, on the set  $A = \{a, b, c, d\}$  the three equivalence relations corresponding to the partitions  $\{\{a, b, c\}, \{d\}\}$ ,  $\{\{a\}, \{b, c, d\}\}$ , and  $\{\{a, b\}, \{c, d\}\}$ , respectively. If each of these three equivalence relations receives the support of  $1/3$  of the population, the corresponding issue-wise majority relation is evidently not transitive; e.g. either pair  $(a, b)$  and  $(b, c)$  is considered equivalent by a majority, while  $a$  and  $c$  are considered not equivalent by a majority.

### 1.5.3 Committee selection

For any  $\mathbf{x} \in \{0, 1\}^K$ , let  $\|\mathbf{x}\| := \#\{k \in [1\dots K]; x_k = 1\}$ . Let  $0 \leq I \leq J \leq K$ , and define

$$X_{I,J;K}^{\text{com}} := \{\mathbf{x} \in \{0, 1\}^K; I \leq \|\mathbf{x}\| \leq J\}. \quad (2)$$

Heuristically,  $[1\dots K]$  is a set of  $K$  ‘candidates,’ and  $X_{I,J;K}^{\text{com}}$  is the set of all ‘committees’ comprised of at least  $I$  and at most  $J$  of these candidates. In the context of the committee selection problem, the Condorcet set can be described as follows (see Nehring et al. (2013)).

For  $\mu \in \Delta^*(X_{I,J;K}^{\text{com}})$ , let  $M := \{k \in [1\dots K]; \mu_k(1) > \frac{1}{2}\}$  denote the set of all candidates receiving majority support. If  $I \leq |M| \leq J$ , then the Condorcet set corresponding to  $\mu$  consists of the single committee containing exactly the candidates in  $M$ ; if  $|M| < I$ , then the Condorcet set consists of all committees of size  $I$  that contain all members of  $M$ ; and if  $|M| > J$ , then the Condorcet set consists of all committees that contain exactly  $J$  members of  $M$ .

### 1.5.4 Resource allocation

Fix  $M, D \in \mathbb{N}$ , and consider the  $D$ -dimensional ‘discrete cube’  $[0\dots M]^D$ . Then each element  $\mathbf{x} \in [0\dots M]^D$  can be represented by a point  $\Phi(\mathbf{x}) := \tilde{\mathbf{x}} \in \{0, 1\}^{D \times M}$  defined as follows:

$$\text{for all } (d, m) \in [1\dots D] \times [1\dots M], \quad \tilde{x}_{(d,m)} := \begin{cases} 1 & \text{if } x_d \geq m; \\ 0 & \text{if } x_d < m. \end{cases} \quad (3)$$

This defines an injection  $\Phi : [0\dots M]^D \rightarrow \{0, 1\}^{D \times M}$ , and any subset of  $P \subset [0\dots M]^D$  can thereby be represented as a subset  $X := \Phi(P) \subset \{0, 1\}^{D \times M}$ . Judgement aggregation over  $X$  thus represents social choice over a  $D$ -dimensional ‘policy space’, where each voter’s position represents her ideal point in  $P$ , the set of feasible policies. This framework is especially useful for resource allocation problems, as we now illustrate. Let

$$\begin{aligned} \Delta_M^D &:= \left\{ \mathbf{x} \in [0\dots M]^D; \sum_{d=1}^D x_d = M \right\}, \\ \text{and } X_{M,D}^\Delta &:= \Phi[\Delta_M^D] \subset \{0, 1\}^{D \times M}. \end{aligned} \quad (4)$$

Geometrically,  $\Delta_M^D$  is a ‘discrete simplex’; points in  $\Delta_M^D$  represent all ways of allocating  $M$  indivisible dollars amongst exactly  $D$  claimants. Thus, judgement aggregation over  $X_{M,D}^\Delta$  describes a group which decides how to allocate a budget of  $M$  dollars towards  $D$  claimants by voting ‘yea’ or ‘nay’ to propositions of the form ‘ $x_d$  should be at least  $m$  dollars’ for each  $d \in [1 \dots D]$  and  $m \in [1 \dots M]$ ; see Lindner et al. (2010).

On the space  $X_{M,D}^\Delta$ , the Condorcet set can be described as follows. For each  $d \in [1 \dots D]$  and  $\mu$ , let  $m_d^* := \text{med}_d(\mu)$  denote the *median* in coordinate  $d$  (that is:  $m_d^*$  is the unique  $m \in [0 \dots M]$  such that  $\mu(x_{d,m}) > \frac{1}{2} > \mu(x_{d,m+1})$ }; this value exists because  $\mu \in \Delta^*(X_{M,D}^\Delta)$ ). It is easily shown that  $\text{Maj}(\mu) = \Phi(m_1^*, \dots, m_D^*)$ . Let  $D(\mu) := \left(\sum_{d=1}^D m_d^*\right) - M$  be the ‘majority deficit’ corresponding to the profile  $\mu$ . Note that the majority deficit can be positive or negative.

If  $D(\mu) = 0$ , then (and only then) the Condorcet consists of the single allocation  $\text{Maj}(\mu)$ . Moreover, if there is positive (negative) deficit, the Condorcet set arises by allocating at most (at least) the median amount to each claimant while distributing the slack in any feasible way. In the space  $X_{M,D}^\Delta$ , the size of the Condorcet set is thus directly determined by the absolute value of the majority deficit  $|D(\mu)|$ , which in effect thus measures the ‘degree of path-dependence.’

### 1.5.5 Aggregation of general binary relations

As above, consider  $N \in \mathbb{N}$  and  $A := [1 \dots N]$ . Let  $K := N(N - 1)$  and identify  $[1 \dots K]$  with the subset of  $A \times A$  that contains all pairs  $(a, b)$  with  $a \neq b$ . An element of  $\{0, 1\}^K$  can be identified with a reflexive binary relation on  $A$ .<sup>6</sup> Note that a number of common restrictions such as symmetry, asymmetry, or completeness, give rise to a median space structure, since they can all be described in terms of logical implications relating exactly *two* issues. In other words, each instance of one of these restrictions corresponds to critical fragment of length 2. For instance, symmetry corresponds to the implications  $x_{(a,b)} = 1 \Rightarrow x_{(b,a)} = 1$  for all  $a, b \in A$ , i.e. to the critical fragments of the form  $(1_{(a,b)}, 0_{(b,a)})$ . Similarly, asymmetry corresponds to the critical fragments of the form  $(1_{(a,b)}, 1_{(b,a)})$  for all  $a, b \in A$ , and completeness to the critical fragments of the form  $(0_{(a,b)}, 0_{(b,a)})$  for all  $a, b \in A$ .

The space  $X_N^{\text{pr}}$  defined in Section 1.5.1 above can thus be identified with the set of all transitive binary relations that satisfy the completeness and asymmetry conditions for all pairs  $a, b \in A$ ; similarly, the space  $X_N^{\text{eq}}$  corresponds to the set of all transitive binary relations that satisfy all symmetry conditions. Let  $X_N^{\text{w.o.}}$  denote the set of all weak orderings on  $[1 \dots N]$ , i.e. the set of all transitive binary relations that satisfy the completeness conditions for all  $a, b \in A$  (but not necessarily asymmetry). Let  $X_N^{\text{p.o.}}$  denote the set of all partial orderings on  $[1 \dots N]$ , i.e. the set of all transitive binary relations that satisfy the asymmetry conditions for all  $a, b \in A$  (but not necessarily completeness).

For notational convenience, we will henceforth denote the statement  $x_{(a,b)} = 1$  (‘ $a$  stands in relation to  $b$ ’) also by  $aRb$ , and the statement  $x_{(a,b)} = 0$  (‘ $a$  does not stand in relation

---

<sup>6</sup>For simplicity, we omit the potential issues  $(a, a)$  for all  $a \in A$ , and assume reflexivity of the underlying binary relations. All our results hold analogously for irreflexive binary relations, and they generalize in a straightforward way to binary relations that are neither reflexive nor irreflexive by incorporating in addition all issues of the form  $(a, a)$ .

to  $b'$ ) by  $aR^c b$ ; moreover, let  $aPb :\Leftrightarrow (aRb \ \& \ bR^c a)$ .

## 2 Robust unanimity consistency

Sequential majority voting, or equivalently: choosing from the Condorcet set in an unrestricted manner, can lead to violations of unanimous consent in some issues. A simple (and well-known) example is the following situation in the 4-permutahedron, cf. (Moulin, 1988, p.242).

**Example 2.1** As in Section 1.5.1 above, suppose that  $\mu(\succ_1) = \mu(\succ_2) = \mu(\succ_3) = \frac{1}{3}$  for the three preference orderings  $a \succ_1 b \succ_1 c \succ_1 d$ ,  $b \succ_2 c \succ_2 d \succ_2 a$ ,  $c \succ_3 d \succ_3 a \succ_3 b$ . Let  $\gamma$  be any path that first decides the three binary comparisons  $(d, a)$ ,  $(a, b)$ ,  $(b, c)$ . Then, we obtain  $d \stackrel{\mu}{\succ} a$ ,  $a \stackrel{\mu}{\succ} b$  and  $b \stackrel{\mu}{\succ} c$ , each with a (2/3) majority; thus, by transitivity, the ordering  $d \succ a \succ b \succ c$  is selected by  $F^\gamma(\mu)$ . In particular,  $d \succ c$  although there is unanimous agreement on the opposite ranking of  $d$  and  $c$ .  $\diamond$

The following analysis will show that the unanimity violation displayed in Example 2.1 is due to the existence of a critical fragment of length  $\geq 4$ . This explains, among other things, the well-known fact that unanimity violations cannot occur in the context of preference aggregation if there are only three alternatives.

For any  $\mu \in \Delta^*(X)$  and  $k \in [1\dots K]$ , we say that  $\mu$  is *unanimous* in coordinate  $k$  if  $\mu_k(0) = 1$  or  $\mu_k(1) = 1$ . Let  $\mathbf{x} \in X$ ; we say that  $\mathbf{x}$  *violates unanimity* in issue  $k$  if  $\mu_k(x_k) = 0$ . We say that  $X$  is *robustly unanimity consistent* if, for all  $\mu \in \Delta^*(X)$  and all paths  $\gamma$ ,  $F^\gamma(\mu)$  does not violate unanimity in any issue. Note that the interest in a concept such as robust unanimity consistency goes well beyond the sequential interpretation of majority voting, as it rules out any potential conflict between majoritarianism and proposition-wise unanimity.

The following is this section's main result; it is reminiscent of Proposition 1.2 (recall that  $\kappa(X)$  is the size of the largest critical fragment in  $X$ ).

**Theorem 2.2** *An aggregation space  $X \subseteq \{0, 1\}^K$  is robustly unanimity consistent if and only if  $\kappa(X) \leq 3$ .*

*Proof.* To show the necessity of  $\kappa(X) \leq 3$ , by contraposition, suppose that  $X$  admits a critical fragment  $\mathbf{c}$  of length  $\geq 4$ . By reordering  $[1\dots K]$  if necessary, we can write  $\mathbf{c} = (c_1, c_2, \dots, c_m) \in \{0, 1\}^m$  with  $4 \leq m \leq K$ . By the criticality of  $\mathbf{c}$ , we can find three elements  $\mathbf{x}^1, \mathbf{x}^2, \mathbf{x}^3 \in X$  such that<sup>7</sup>

$$\begin{aligned} \mathbf{x}^1 &= (\neg c_1, c_2, c_3, c_4, \dots, c_m, \dots), \\ \mathbf{x}^2 &= (c_1, \neg c_2, c_3, c_4, \dots, c_m, \dots), \\ \mathbf{x}^3 &= (c_1, c_2, \neg c_3, c_4, \dots, c_m, \dots). \end{aligned}$$

Define  $\mu \in \Delta^*(X)$  by  $\mu(\mathbf{x}^j) = 1/3$  for  $j \in \{1, 2, 3\}$ . Then  $\mu_n(c_n) = 2/3$  for  $n \in \{1, 2, 3\}$ , while  $\mu_n(c_n) = 1$  for all  $n \in [4, \dots, m]$ , i.e.  $\mu$  is unanimous in issues  $4, \dots, m$ . Now, let  $\gamma$

<sup>7</sup>Recall from Footnote 4 that ' $\neg$ ' stands for the logical negation, i.e.  $\neg 0 = 1$  and  $\neg 1 = 0$ .

be the path  $(1, 2, 3, \dots, K)$ , and consider  $\mathbf{x} := F^\gamma(\mu) \in X$ . By the criticality of  $\mathbf{c}$ , one has  $x_n = c_n$  for all  $n < m$  and  $x_m \neq c_m$ . Thus,  $\mathbf{x} = F^\gamma(\mu)$  violates unanimity in issue  $m$ .

For the proof of sufficiency of  $\kappa(X) \leq 3$ , note first that, if  $\kappa(X) \leq \ell$ , then a majority on a particular issue can be overridden only as the entailment of  $\ell - 1$  earlier majority decisions. In particular, this can never happen if  $\kappa(X) = 2$ . If  $\kappa(X) = 3$ , it can happen only as the entailment of two earlier majority decisions; but in this case, at least one voter must affirm the entailed decision, so only supermajorities, not unanimities, can be overridden. The formal details of this argument are provided in the appendix.  $\square$

Note that the possible violations of respect for unanimity when  $\kappa(X) > 3$  constructed in the proof of Theorem 2.2 require only three agents and can thus occur quite easily.

**Example 2.3 (a)** (Aggregation of strict preference orderings) Let  $N \in \mathbb{N}$ , and let  $X_N^{\text{pr}}$  be as in Section 1.5.1 above. In the Appendix we show that

$$\kappa(X_N^{\text{pr}}) = N.$$

Thus,  $X_N^{\text{pr}}$  is robustly unanimity consistent if and only if  $N \leq 3$ .

**(b)** (Restricted preference domains) Let  $\triangleleft$  be a partial order on  $[1\dots N]$ , and let  $X_{\triangleleft}^{\text{pr}}$  be the set of all linear orders on  $[1\dots N]$  which extend  $\triangleleft$ . Judgement aggregation over  $X_{\triangleleft}^{\text{pr}}$  describes a variant of Arrovian preference aggregation where the voters unanimously agree with the preferences encoded in  $\triangleleft$ , and where the social preference order is also required (e.g. by the constitution) to agree with the preferences encoded in  $\triangleleft$ . For instance, suppose that all preferences are defined over the Cartesian product  $M \times B$  where  $M \subseteq \mathbb{R}$  represents a set of feasible amounts of public expenditure and  $B$  is a (finite) set of social states. More concretely, say that  $b_0 \in B$  stands for the ‘status quo’ (no public project) and each element in  $B \setminus \{b_0\}$  corresponds to a different version of a public project. Moreover, assume that the partial order  $\triangleleft$  expresses monotonicity with respect to the first component in each social state (‘more money is better’) while refraining from any judgement across states, i.e.  $(m, b) \triangleleft (m', b') :\Leftrightarrow [b = b' \text{ and } m < m']$ .

A  $\triangleleft$ -*antichain* is a subset  $Z \subseteq [1\dots N]$  such that every pair of elements in  $Z$  are  $\triangleleft$ -incomparable. Let  $\text{width}(\triangleleft)$  be the largest cardinality of any  $\triangleleft$ -antichain. For instance, in the above example  $\text{width}(\triangleleft) = |B|$ . In the Appendix we prove that

$$\kappa(X_{\triangleleft}^{\text{pr}}) = \text{width}(\triangleleft). \tag{5}$$

Thus,  $X_{\triangleleft}^{\text{pr}}$  is robustly unanimity consistent if and only if  $\text{width}(\triangleleft) \leq 3$ . Moreover, by equation (5), one obtains a median space in the concrete example if and only if  $|B| = 2$ , i.e. if there is only one version of the public project. In that case majority voting with a fixed budget amounts to social choice based on the median willingness to pay for the public project: if total budget is  $m_0$  and the public project costs  $m_1 \leq m_0$ , then the public project will be realized under majority voting if and only if at least half of the voters prefer

$(m_0, b_0)$  to  $(m_0 - m_1, b_1)$ . Moreover, if  $|B| = 2$ , the majority view is always feasible, i.e. is a linear order (that evidently agrees with the preferences encoded in  $\triangleleft$ ). Robust unanimity consistency, however, also prevails if  $|B| = 3$ , i.e. if there are two different versions of the public project.

**(c)** (Equivalence Relations) Let  $N \in \mathbb{N}$  and consider the spaces  $X_N^{\text{eq}} \subset \{0, 1\}^K$  as in Section 1.5.2. In the Appendix, we show that

$$\kappa(X_N^{\text{eq}}) = N.$$

Thus, the spaces  $X_N^{\text{eq}}$  are robustly unanimity consistent if and only if  $N \leq 3$ . An example of unanimity violation for  $N \geq 4$  is the profile considered in Section 1.5.2 above. Suppose again that each of the equivalence relations corresponding to the three partitions  $\{\{a, b, c\}, \{d\}\}$ ,  $\{\{a\}, \{b, c, d\}\}$ , and  $\{\{a, b\}, \{c, d\}\}$  receives popular support of  $1/3$ , respectively. Then, if issues are decided by sequential majority voting according to any path that starts with the issues  $(a, b)$ ,  $(b, c)$ ,  $(c, d)$ , we obtain that all elements of  $A = \{a, b, c, d\}$  are equivalent. But there is unanimous consent that  $a$  and  $d$  are not equivalent.

**(d)** (Committee Selection) Let  $0 \leq I \leq J \leq K$ , and let  $X_{I,J,K}^{\text{com}}$  be as defined in Section 1.5.3. A fragment  $\mathbf{c}$  is  $X_{I,J,K}^{\text{com}}$ -critical if and only if either (i)  $|\mathbf{c}| = J + 1$  and  $\mathbf{c} = (1, 1, \dots, 1)$  or (ii)  $|\mathbf{c}| = K - I + 1$  and  $\mathbf{c} = (0, 0, \dots, 0)$ . Thus,

$$\kappa(X_{I,J,K}^{\text{com}}) = 1 + \max\{J, K - I\}.$$

For example, suppose  $K = 4$  and  $J = I = 2$ ; then  $\kappa(X_{2,2,4}^{\text{com}}) = 3$ , so  $X_{2,2,4}^{\text{com}}$  is robustly unanimity consistent. However, if  $K \geq 5$ , then either  $J \geq 3$  or  $K - I \geq 3$ , so that  $\kappa(X_{I,J,K}^{\text{com}}) \geq 4$ ; then  $X_{I,J,K}^{\text{com}}$  is not robustly unanimity consistent.

**(e)** (Resource Allocation) Let  $D, M \in \mathbb{N}$ , and let  $X_{M,D}^{\Delta}$  be as in Section 1.5.4. Suppose  $M \geq D^2$  (which is always true if we denominate money in small enough units). In the Appendix we show that

$$\kappa(X_{M,D}^{\Delta}) = D.$$

Thus, the spaces  $X_{M,D}^{\Delta}$  are robustly unanimity consistent if and only if  $D \leq 3$ .  $\diamond$

### 3 Designable unanimity consistency

Theorem 2.2 above characterizes the aggregation spaces for which respect for a unanimous vote is guaranteed no matter *which* sequence of majority decisions one chooses. As we have seen, most aggregation spaces cannot guarantee respect for unanimity in this strong sense. In this section, we thus ask whether it is possible to overcome the problem *by design*. In other words, for a given aggregation space, do there *exist* paths of majority decisions that guarantee respect for unanimity for all profiles? One might expect a negative answer since *prima facie* it is not clear how, by design of the sequence, one could ‘detect’ unanimities in contrast to mere majorities (in a profile-independent way). However, it turns out that sometimes this is indeed possible.

Say that an aggregation space  $X \subseteq \{0, 1\}^K$  is *designably unanimity consistent* if there exists a path  $\gamma$  such that, for all  $\mu \in \Delta^*(X)$ ,  $F^\gamma(\mu)$  does not violate unanimity on any

issue. Our first result shows that this weakening of the notion of respect for unanimity does neither help in the case of committee selection nor in the case of resource allocation.

**Theorem 3.1** (a) *The spaces  $X_{I,J;K}^{\text{com}}$  as defined in Section 1.5.3 are designably unanimity consistent if and only if they are robustly unanimity consistent.*

(b) *The spaces  $X_{M,D}^{\Delta}$  as defined in Section 1.5.4 are designably unanimity consistent if and only if  $D \leq 3$ . Thus, again  $X_{M,D}^{\Delta}$  is designably unanimity consistent if and only if it is robustly unanimity consistent.*

*Proof of part (a).* Evidently, it suffices to show that if there is a unanimity violation for some profile  $\mu$  along some path  $\gamma$ , then there is a unanimity violation along every path  $\eta$ . Thus, let  $\eta$  be given and consider the permutation  $\eta \circ \gamma^{-1} : [1\dots K] \rightarrow [1\dots K]$ . For each feasible view  $\mathbf{x} \in X_{I,J;K}^{\text{com}}$ , consider the permuted view  $\eta \circ \gamma^{-1}(\mathbf{x})$ , and note that by the symmetry of the spaces  $X_{I,J;K}^{\text{com}}$ ,  $\eta \circ \gamma^{-1}(\mathbf{x})$  is also feasible. Denote by  $\eta \circ \gamma^{-1}(\mu)$  the profile that results from  $\mu$  by applying the permutation  $\eta \circ \gamma^{-1}$ . Evidently, since  $\mu$  entails a unanimity violation along  $\gamma$ ,  $\eta \circ \gamma^{-1}(\mu)$  entails a unanimity violation along the path  $\eta$ . This completes the proof of part (a) of Theorem 3.1.

The proof of part (b) is more involved and is therefore deferred to the Appendix.  $\square$

Given these negative results for  $X_{I,J;K}^{\text{com}}$  and  $X_{M,D}^{\Delta}$  it is quite remarkable that we have the following positive result for the aggregation of linear preference orders, weak preference orderings, and equivalence relations.

**Theorem 3.2** *The spaces  $X_N^{\text{pr}}$ ,  $X_N^{\text{w.o.}}$  and  $X_N^{\text{eq}}$  are designably unanimity consistent.*

The proof of Theorem 3.2 is instructive, so we will sketch the main ideas here. Let  $X$  be any of the three spaces in the statement of the theorem, and let  $[1\dots N]$  be the set of alternatives. A path  $\zeta$  on the set of all pairs  $(a, b) \in [1\dots N] \times [1\dots N]$  with  $a \neq b$  such that  $(a, b)$  occurs before  $(a', b')$  in  $\zeta$  whenever  $\min\{a, b\} < \min\{a', b'\}$  will be called *simply ordered*.<sup>8</sup> We will show that, for all  $\mu$  and all simply ordered paths  $\zeta$ , the transitive relation  $R^\zeta(\mu)$  generated by sequential majority voting along  $\zeta$  respects unanimity in all issues.

For any reflexive binary relation  $R$  with complement  $R^c$ , define two relations  $R^*$  and  $R^{**}$  as follows. For all  $a, b \in [1\dots N]$ ,

$$\begin{aligned} aR^*b & :\Leftrightarrow [aRb \text{ and, for all } d, (bRd \Rightarrow aRd) \ \& \ (dRa \Rightarrow dRb)] \\ aR^{**}b & :\Leftrightarrow [aR^cb \text{ and, for all } d, (aRd \Rightarrow dR^cb)] \end{aligned}$$

The relation  $R^*$  is the so-called ‘covering relation’, well-known in preference aggregation theory (Miller, 1980).<sup>9</sup> Note that  $R^*$  is always transitive, and  $R^*$  is symmetric whenever

<sup>8</sup>Observe that there is a multitude of different simply ordered paths, not only corresponding to different orderings of the alternatives but also with respect to a fixed ordering of the alternatives.

<sup>9</sup>See also Shepsle and Weingast (1984), Banks (1985), McKelvey (1986), Moulin (1986), Cox (1987), and p.254 of Moulin (1988).

$R$  is symmetric. Henceforth, we will refer to the fragment defined by the pair  $[R^*, R^{**}]$  as the *covering fragment* corresponding to the relation  $R$ .<sup>10</sup>

For any profile  $\mu \in \Delta^*(X)$ , let  $[R(\mu), R^c(\mu)]$  represent the issue-wise majority relation, i.e.  $aR(\mu)b$  if a majority affirms  $aRb$ , and  $aR^c(\mu)b$  if a majority rejects  $aRb$ . Moreover, denote by  $U(\mu)$  the ‘unanimity relation’ corresponding to  $\mu$ , i.e.  $aU(\mu)b$  if under  $\mu$  there is unanimous agreement that  $aRb$  or unanimous agreement that  $aR^c b$ .

**Lemma 1** *For all  $\mu$ ,  $R^*(\mu) \supseteq [U(\mu) \cap R(\mu)]$  and  $R^{**}(\mu) \supseteq [U(\mu) \cap R^c(\mu)]$ .*

*Proof of Lemma 1.* This follows from transitivity of  $R$ . Suppose that all agents affirm  $aRb$ . Then, if a majority affirms  $bRd$ , the same majority must affirm  $aRd$  by transitivity; similarly, if a majority affirms  $dRa$ , the same majority also affirms  $dRb$ . This shows the first part of Lemma 1. The second part follows from noting that, if a majority affirms  $aRd$  and another majority affirms  $dRb$ , then at least one agent must affirm  $aRd$  and  $dRb$ , hence by transitivity also  $aRb$ ; thus, we could not have unanimous support for  $aR^c b$ .  $\square$

The following lemma states that the transitive relation that results from sequential majority voting along the path  $\zeta$  extends the covering fragment corresponding to the issue-wise majority relation.

**Lemma 2** *For all  $\mu$ ,  $R^\zeta(\mu) \supseteq R^*(\mu)$  and  $(R^\zeta)^c(\mu) \supseteq R^{**}(\mu)$ .*

The proof of Lemma 2 is more involved and is therefore deferred to the Appendix.

*Proof of Theorem 3.2.* Lemmas 1 and 2 jointly imply that  $R^\zeta(\mu)$  extends the ‘positive’ unanimity relation corresponding to  $\mu$ , and  $(R^\zeta)^c(\mu)$  extends the ‘negative’ unanimity relation corresponding to  $\mu$ . The result follows.  $\square$

To illustrate the proof of Theorem 3.2, consider again Example 2.1 and Fig. 1 above. As is easily verified, the covering fragment of the majority tournament corresponding to the profile  $\mu$  is given by  $\{c \succ^\mu d\}$ .<sup>11</sup> In this example, the covering fragment thus corresponds to the single issue in which there is unanimous consent. If the majority decisions are taken according to the simply ordered path  $(a, b), (a, c), (a, d), (b, c), (b, d), (c, d)$ , we obtain the ordering  $c \succ d \succ a \succ b$ ; on the other hand, the path  $(b, c), (b, d), (b, a), (c, d), (c, a), (d, a)$  is also simply ordered (given a different ordering of the alternatives) and yields  $a \succ b \succ c \succ d$ . Thus, while different admissible paths may give different final outcomes, all resulting orderings must respect unanimous decisions as shown in the proof of Theorem 3.2; indeed, they all extend the covering fragment corresponding to the majority tournament, and the latter does not depend on the decision path.

For the profile  $\mu$  in Example 2.3(c) above, the covering fragment is given by  $(\succ^\mu)^* = \emptyset$  and  $(\succ^\mu)^{**} = U(\mu) = \{a \not\succeq d\}$ . Consider a bijection  $\pi : \{a, b, c, d\} \rightarrow [1..4]$  and any simply

<sup>10</sup>Formally, the covering fragment defined by  $[R^*, R^{**}]$  is given by the unique fragment  $\mathbf{w}$  such that  $\text{supp}(\mathbf{w}) = \text{supp}(R^*) \cup \text{supp}(R^{**})$ ,  $w_{(a,b)} = 1$  if  $aR^*b$ , and  $w_{(a,b)} = 0$  if  $aR^{**}b$  (no contradiction here, since  $R^*$  and  $R^{**}$  are disjoint).

<sup>11</sup>More precisely, we have  $R^*(\mu) = \{(c, d)\}$  and  $R^{**}(\mu) = \{(d, c)\}$ .

ordered path  $\zeta$  such that the issue  $(n, m)$  occurs before  $(n', m')$  in  $\zeta$  if  $\min\{\pi(n), \pi(m)\} < \min\{\pi(n'), \pi(m')\}$ , as in the proof of Theorem 3.2. It is easily verified that for the given profile, one obtains the following sets of equivalence classes as the result of sequential majority voting along the path  $\zeta$ : if  $\pi^{-1}(1) = b$ , then  $F^\zeta(\mu) = \{\{a, b, c\}, \{d\}\}$ ; if  $\pi^{-1}(1) = c$ , then  $F^\zeta(\mu) = \{\{a\}, \{b, c, d\}\}$ ; and if  $\pi^{-1}(1) \in \{a, d\}$ , then  $F^\zeta(\mu) = \{\{a, b\}, \{c, d\}\}$ . In any of these cases, the unanimous agreement that  $a \not\sim d$  is respected.

**Remark** In the case of the aggregation of strict preference orderings, the choice rule among alternatives induced by any path  $\zeta$  considered in Theorem 3.2 admits the following simple recursive description. First step: let  $A_1 := [1\dots N]$  be the initial pool of alternatives and set element  $a_1 := 1$  as the initial top alternative. Step  $\ell$ : Compare the current top alternative  $a_{\ell-1}$  to all alternatives  $m > a_{\ell-1}$  in the current pool  $A_{\ell-1}$ . If  $a_{\ell-1}$  wins against all such elements, then  $a_{\ell-1}$  is the final choice; otherwise, let the new top alternative  $a_\ell$  be the smallest majority winner against  $a_{\ell-1}$  and take  $A_\ell$  as the old pool minus the majority losers against  $a_{\ell-1}$ . Evidently, this procedure yields a final choice after a finite number of steps, which by Theorem 3.2 cannot be unanimously defeated by another alternative.

For any simply ordered path  $\zeta$ , the choice rule induced by  $F^\zeta(\mu)$  is the same as the one induced by the corresponding multi-stage elimination tree procedure of Shepsle and Weingast (1984) (cf. Section 9.4 in Moulin (1988)). In contrast to the multi-stage elimination tree which only produces a winning alternative, however,  $F^\zeta$  yields an entire ordering over all alternatives. Moreover, to accomplish this,  $F^\zeta$  only needs  $\frac{N(N-1)}{2}$  binary comparisons, whereas the multi-stage elimination tree involves  $(2^{N-1}) - 1$  binary comparisons. In particular, in the multi-stage elimination procedure, an alternative may face the same opponent several times; this does not happen along any path  $\zeta$ .

### 3.1 A necessary condition for designable unanimity consistency

As demonstrated in the proof of Theorem 2.2, any critical fragment of size greater than or equal to 4 can trigger violations of unanimity for some path of issues. Yet, as Theorem 3.2 shows, in certain spaces, for certain specially designed paths only critical fragments of size two or three will be ‘activated.’ But such spaces appear to be rather exceptional; typically, it will not be possible to ‘sidestep’ all critical fragments of size greater than three if such fragments exist. In particular, no path can sidestep any *exposed* critical fragment, as follows.

Recall that  $C(X)$  is the set of all critical fragments for  $X$ . Let  $\mathbf{c} \in C(X)$ , and let  $J := \text{supp}(\mathbf{c})$ . For all  $j \in J$ , let  $\mathbf{c}^j$  be the fragment obtained by negating the  $j$ th coordinate of  $\mathbf{c}$ . Since  $\mathbf{c}$  is critical, we know that each  $\mathbf{c}^j$  is contained in some element of  $X$ . We will say that  $\mathbf{c}$  is an *exposed critical fragment* if there exists a single ‘complementary’ fragment  $\mathbf{d} \in \{0, 1\}^{[1\dots K] \setminus J}$  such that, for all  $j \in \text{supp}(\mathbf{c})$ , the element  $(\mathbf{c}^j, \mathbf{d})$  is in  $X$ . Let  $C^*(X)$  be the set of all exposed critical fragments for  $X$ . Let  $\kappa^*(X)$  be the size of the largest element of  $C^*(X)$  (with  $\kappa^*(X) := 0$  if  $C^*(X)$  is empty).

**Example 3.3** (a) (*Preference relations*) Let  $N \in \mathbb{N}$ , and let  $X_N^{\text{pr}}$  be as in Section 1.5.1 above. For any  $m \in [3\dots N]$ , every critical fragment of size  $m$  for  $X_N^{\text{pr}}$  corresponds to a

preference cycle of length  $m$ . In the appendix, we show that such a critical fragment is exposed if and only if  $m = 3$ . Thus, all elements of  $C^*(X_N^{\text{pr}})$  have size 3, so  $\kappa^*(X_N^{\text{pr}}) = 3$ .

Likewise, let  $X_N^{\text{p.o.}}$  be the space of all strict partial orders on a set of  $N$  elements, as defined in Section 1.5.5. Then again  $\kappa^*(X_N^{\text{p.o.}}) = 3$ .

(b) (*Acyclic relations*) Let  $X$  be the space of all acyclic, antisymmetric, possibly incomplete binary relations on a set of  $N$  elements, defined similarly to the examples in Section 1.5.5. For every  $m \in [3 \dots N]$ , every cyclical relation of length  $m$  defines an exposed critical fragment for  $X$  of length  $m$ . Thus,  $C^*(X)$  contains elements of all lengths up to  $N$ , so  $\kappa^*(X) \geq N$ .

(c) (*Truth-functional aggregation*) Observe that any critical fragment of size  $K$  is automatically exposed (with its complementary fragment being the empty word). This observation yields exposed critical fragments for judgement aggregation problems associated with several common truth functions. For example, let  $X_{\&} := \{\mathbf{x} \in \{0,1\}^K; x_K = x_1 \& x_2 \& \dots \& x_{K-1}\}$ , where “&” denotes logical conjunction. Then the sequence  $(1, 1, \dots, 1, 0)$  is an exposed critical fragment. Next, let  $X_{\vee} := \{\mathbf{x} \in \{0,1\}^K; x_K = x_1 \vee x_2 \vee \dots \vee x_{K-1}\}$ , where “ $\vee$ ” denotes logical disjunction. Then the sequence  $(0, 0, \dots, 0, 1)$  is an exposed critical fragment. Finally, let  $\oplus$  denote the “exclusive OR” operation, and let  $X_{\oplus} := \{\mathbf{x} \in \{0,1\}^K; x_K = x_1 \oplus x_2 \oplus \dots \oplus x_{K-1}\}$ . If  $K$  is even, then the sequence  $(1, 1, \dots, 1, 1)$  is an exposed critical fragment. If  $K$  is odd, then the sequence  $(1, 1, \dots, 1, 0)$  is an exposed critical fragment.  $\diamond$

**Theorem 3.4** *If  $\kappa^*(X) \geq 4$ , then the aggregation space  $X$  is not designably unanimity consistent.*

When combined with Example 3.3(b), Theorem 3.4 implies that, for any  $N \geq 4$ , the space of acyclic relations on  $N$  is *not* designably unanimity consistent. Likewise, Example 3.3(c) shows that several common truth-function aggregation problems are not designably unanimity consistent.

Clearly, the positive examples in Theorem 3.2 above verify the necessary condition  $\kappa^*(X) \leq 3$  of Theorem 3.4. However, this does not mean that all spaces with  $\kappa^*(X) \leq 3$  share the simple structure of the spaces displayed in Theorem 3.2. The following example of all *separable* preference orderings shows that there are spaces of transitive relations with  $\kappa^*(X) \leq 3$  in which unanimity violations *do* occur along simply ordered paths.

**Example 3.5** Consider the space  $X^{\text{sep}}$  of all separable preference orderings on the set  $A = \{0,1\}^3$ . A preference ordering  $\succ$  on  $\{0,1\}^3$  is called *separable* if, for all  $a, b \in \{0,1\}^3$  and all  $j \in \{1, 2, 3\}$ ,  $[(a_j, a_{-j}) \succ (b_j, a_{-j})] \Leftrightarrow [(a_j, b_{-j}) \succ (b_j, b_{-j})]$ . In the Appendix, we verify that  $\kappa^*(X^{\text{sep}}) = 3$ .<sup>12</sup> Nevertheless, unanimity violations *do* occur along simply ordered paths. For instance, suppose that one third of the population has the alternative  $m_1 := (0, 1, 1)$  on top of their preference ordering, another third of the population the alternative  $m_2 = (1, 0, 1)$ , and the final third of the population the alternative  $m_3 :=$

<sup>12</sup>A stronger notion of separability frequently used in the literature requires  $[(a_j, a_{-j}) \succ (b_j, a_{-j})] \Leftrightarrow [(a_j, b_{-j}) \succ (b_j, b_{-j})]$  for all non-empty *subsets* of coordinates. One can show that under this stronger condition, no critical fragment is exposed; thus, in this case,  $\kappa^*(X) = 0$ .

$(1, 1, 0)$ , while all individuals agree that  $(0, 0, 0) \succ (1, 1, 1)$ . Such profile is compatible with separability.<sup>13</sup> Consider any simply ordered path that first compares the alternative  $(1, 1, 1)$  with its three neighbors  $m_1, m_2, m_3$ . In each case,  $(1, 1, 1)$  receives majority support of  $2/3$ . For instance, if  $(1, 0, 1)$  is the top alternative, then  $(1, 0, 1)$  is preferred to  $(0, 0, 1)$ , hence by separability,  $(1, 1, 1)$  is preferred to  $m_1$ ; similarly, if  $(1, 1, 0)$  is the top alternative, then  $(1, 1, 0)$  is preferred to  $(0, 1, 0)$ , hence again by separability,  $(1, 1, 1)$  is preferred to  $m_1$ . This shows that  $(1, 1, 1)$  gets majority support over  $m_1$ . By an analogous argument,  $(1, 1, 1)$  also receives majority support over  $m_2$  and over  $m_3$ . But if  $(1, 1, 1) \succ m_j$  for all  $j \in \{1, 2, 3\}$ , then separability and transitivity imply  $(1, 1, 1) \succ (0, 0, 0)$ , in violation of unanimity.  $\diamond$

Does the fact that Lemma 2 fails on the subspace of all separable preference orderings mean that this space is *not* designably unanimity consistent? After all, there could be other, non simply-ordered paths along which unanimity violations never occur. Can the argument given in Theorem 3.2 be generalized to demonstrate designable unanimity consistency of a larger class of spaces? We conclude this section with two conjectures that – without answering these questions – are meant to shed light on the issue. To formulate them, we need some additional concepts.

Recall that a critical fragment of  $X$  is a minimal forbidden binary sequence. Equivalently, a critical fragment of  $X$  is a *maximal* fragment of  $\{0, 1\}^K \setminus X$ , that is:  $\mathbf{w} \in \{0, 1\}^J$  is a critical fragment of  $X$  if, for all  $\mathbf{x} \in \{0, 1\}^K$ ,  $[\mathbf{x}_J = \mathbf{w} \Rightarrow \mathbf{x} \notin X]$ , and no proper subfragment of  $\mathbf{w}$  has this property. A collection  $\mathcal{B} \subseteq C(X)$  of critical fragments is called a *critical basis* for  $X$  if, for all  $\mathbf{x} \in \{0, 1\}^K \setminus X$ , there exists  $\mathbf{w} \in \mathcal{B}$  such that  $\mathbf{w} \sqsubseteq \mathbf{x}$ , while this fails for any proper subfamily of  $\mathcal{B}$ .

To illustrate, consider again the space of all strict preference orderings  $X_N^{\text{pr}}$ . Each critical fragment of the form  $(1_{(a,b)}, 1_{(b,c)}, 0_{(a,c)})$  corresponds to the transitivity requirement  $(a \succ b \ \& \ b \succ c) \Rightarrow a \succ c$ , and in fact, all critical fragments of length 3 have this form. While there are also critical fragments of length greater than 3, the family of all critical fragments of the form  $(1_{(a,b)}, 1_{(b,c)}, 0_{(a,c)})$  form a critical basis of  $X_N^{\text{pr}}$ . To see this, consider, say, the critical fragment  $\mathbf{v} = (1_{(a,b)}, 1_{(b,c)}, 1_{(c,d)}, 0_{(a,d)})$ . Evidently,  $\mathbf{v}$  corresponds to a 4-cycle, hence to an *indirect* implication of transitivity. Take any  $\mathbf{x} \in \{0, 1\}^K$  with  $\mathbf{v} \sqsubseteq \mathbf{x}$ ; since  $\mathbf{v}$  is a critical fragment, we have  $\mathbf{x} \notin X$ . To verify that the critical fragments of length 3 form a critical basis, we have to show that there exists a critical fragment  $\mathbf{w}$  of length 3 such that  $\mathbf{w} \sqsubseteq \mathbf{x}$ . Indeed, if  $x_{(a,c)} = 0$ , then  $\mathbf{x}$  contains the critical fragment  $(1_{(a,b)}, 1_{(b,c)}, 0_{(a,c)})$ , and if  $x_{(a,c)} = 1$ , then  $\mathbf{x}$  contains the critical fragment  $(1_{(a,c)}, 1_{(c,d)}, 0_{(a,d)})$ . This argument can be straightforwardly extended to cycles of any length.<sup>14</sup>

Let  $X_N^{\text{rel}} \subseteq \{0, 1\}^K$  be any space of reflexive binary relations on  $A = [1 \dots N]$  with  $K = N(N - 1)$  as in §1.5.5 above, and note that we now include again two separate issues  $(a, b)$  and  $(b, a)$  for any two distinct elements  $a$  and  $b$ . Denote the collection of all critical

<sup>13</sup>For example, the  $m_1$ -favouring voters could have the separable preference order  $(0, 1, 1) \succ (0, 1, 0) \succ (0, 0, 1) \succ (0, 0, 0) \succ (1, 1, 1) \succ (1, 1, 0) \succ (1, 0, 1) \succ (1, 0, 0)$ . Permuting the three coordinate yields separable preference orders for the other two groups.

<sup>14</sup>The elements of  $X_N^{\text{pr}}$  can be represented as complete and acyclic directed graphs on  $[1 \dots N]$ . In the language of directed graphs, every transitivity requirement corresponds to the absence of a 3-cycles. These form a basis of  $X_N^{\text{pr}}$ , since a complete directed graph is acyclic if (and only if) it contains no 3-cycle.

fragments corresponding to transitivity by

$$\mathcal{C}_{trans} := \{(1_{(a,b)}, 1_{(b,c)}, 0_{(a,c)}) : a, b, c \in [1 \dots N], a \neq c\}.$$

Moreover, say that a critical fragment  $\mathbf{w}$  of  $X_N^{rel}$  is *paired* if its support  $\text{supp}(\mathbf{w})$  has the form  $\{(a, b), (b, a)\}$  for distinct elements  $a$  and  $b$ . Note that there are only three different types of paired critical fragments: for all  $a, b$ , criticality of the fragment  $(1_{(a,b)}, 0_{(b,a)})$  corresponds to the symmetry requirement  $aRb \Rightarrow bRa$ ; similarly, criticality of  $(1_{(a,b)}, 1_{(b,a)})$  corresponds to the asymmetry requirement  $aRb \Rightarrow bR^c a$ ; and finally, criticality of  $(0_{(a,b)}, 0_{(b,a)})$  corresponds to the completeness requirement  $aR^c b \Rightarrow bRa$ .

**Definition** Say that  $X_N^{rel}$  is a *simple space of transitive relations* if  $X_N^{rel}$  has a critical basis  $\mathcal{B}$  of the form  $\mathcal{B} = \mathcal{C}_{trans} \cup \mathcal{D}$  where  $\mathcal{D}$  consists only of critical fragments of length two. Moreover, say that  $X_N^{rel}$  is a *paired simple space of transitive relations* if  $X_N^{rel}$  has a critical basis  $\mathcal{B}$  of the form  $\mathcal{B} = \mathcal{C}_{trans} \cup \mathcal{D}$  where  $\mathcal{D}$  consists of paired fragments only.

Obviously, every paired simple space is a simple space. The converse does not hold as Example 3.5 shows. Examples of paired simple spaces of transitive relations are the set of all weak partial orderings (corresponding to  $\mathcal{D} = \emptyset$ ), the set of all strict partial orderings (corresponding to the case in which  $\mathcal{D}$  consists of all asymmetry requirements of the form  $aRb \Rightarrow bR^c a$ ), the set of all weak orderings (when  $\mathcal{D}$  consists of all completeness requirements of the form  $aR^c b \Rightarrow bRa$ ), the set of all strict preference orderings (when  $\mathcal{D}$  consists of all asymmetry and completeness requirements), and the set of all equivalence relations (when  $\mathcal{D}$  consists of all symmetry requirements of the form  $aRb \Rightarrow bRa$ ).

**Conjecture 1** *All paired simple spaces of transitive relations are designably unanimity consistent.*

**Conjecture 2** *All simple spaces of transitive relations are designably unanimity consistent.*

The envisaged proof strategy of Conjecture 1 is to adapt our proof of Theorem 3.2, showing via Lemma 2 that unanimity violations cannot occur along simply ordered paths. However, our present proof of Lemma 2 (provided in the appendix) is not ready-made for this purpose since in general paired simple spaces the structure of critical fragments is more complex than in the three exemplary spaces considered in Theorem 3.2.

As Example 3.5 shows, the stronger Conjecture 2 – if correct – requires a different argument; in particular, one would first have to find an appropriate class of paths along which unanimity violations cannot occur. The statement of Conjecture 2 seems to be considerably less likely to be true.

## Conclusion

In this paper, we have provided a simple characterization of the class of aggregation spaces in which sequential majority voting guarantees respect for unanimity, no matter in which sequence the (majority) decisions are taken (‘robust unanimity consistency’).

We have also investigated the weaker condition that there *exists* a decision path along which unanimous judgements are always respected (‘designable unanimity consistency’),

and provided positive as well as negative examples. The positive examples that we found are all placed in spaces of transitive binary relations with a particularly simple structure. Finally, we offered two conjectures about the scope of designable unanimity consistency; it seems to be a worthwhile problem for future research to settle these conjectures and to clarify whether there exist further interesting and relevant positive examples.

## Appendix: Proofs

*Notation A.1.* Let  $J \subset [1\dots K]$ , and let  $\mathbf{w} \in \{0, 1\}^J$ . We define  $[\mathbf{w}] := \{\mathbf{x} \in \{0, 1\}^K; \mathbf{w} \sqsubseteq \mathbf{x}\}$ . Thus,  $\mathbf{w}$  is  $X$ -forbidden if and only if  $X \cap [\mathbf{w}] = \emptyset$ . It is sometimes convenient to express  $\mathbf{w}$  as an element of  $\{0, 1, *\}^K$ , where we define  $w_k = *$  for all  $k \notin J$ . For example, suppose  $J = \{i, i+1, \dots, j\}$  for some  $i \leq j \leq K$ ; then we would write  $\mathbf{w} = (\underbrace{*, \dots, *}_{i-1}, w_i, w_{i+1}, \dots, w_j, \underbrace{*, \dots, *}_{K-j})$ .

*Proof of Theorem 2.2.* Let  $\mu \in \Delta^*(X)$ , and let  $\text{Unan}(\mu) := \{\mathbf{x} \in \{0, 1\}^K; \mathbf{x} \text{ does not violate } \mu\text{-unanimity}\}$ .

Necessity of  $\kappa(X) \leq 3$  for robust unanimity consistency has already been shown in the main text. In order to demonstrate sufficiency, suppose that all critical fragments have order 3 or less. If  $\text{Maj}(\mu) \in X$ , then  $\text{Cond}(X, \mu) = \{\text{Maj}(\mu)\}$  and we're done, because  $\text{Maj}(\mu) \in \text{Unan}(\mu)$ .

Suppose  $\text{Maj}(\mu) \notin X$ , and let  $\mathbf{y} \in \text{Cond}(X, \mu)$ .

**Claim 1:** *Let  $k \in [1\dots K]$ . If  $y_k \neq \text{Maj}_k(\mu)$ , then  $\mu_k(\text{Maj}_k(\mu)) < 1$ .*

*Proof:* If  $y_k \neq \text{Maj}_k(\mu)$ , then there exist  $i, j \in [1\dots K] \setminus \{k\}$  and some  $X$ -critical fragment  $\mathbf{c} = (c_i, c_j, c_k) \sqsubset \text{Maj}(\mu)$  such that  $y_i = c_i$  and  $y_j = c_j$ , thereby forcing  $y_k \neq c_k$ . (If there was no such critical fragment, then we could change  $y_k$  to  $\text{Maj}_k(\mu)$  without leaving  $X$ —hence  $\mathbf{y} \notin \text{Cond}(X, \mu)$ ).

By contradiction, suppose  $\mu_k[c_k] = 1$ . Note that  $\mu_i[c_i] > \frac{1}{2}$  and  $\mu_j[c_j] > \frac{1}{2}$  (because  $\mathbf{c} \sqsubset \text{Maj}(\mu)$ ). Thus,  $\mu_i[\neg c_i] < \frac{1}{2}$  and  $\mu_j[\neg c_j] < \frac{1}{2}$ , and of course,  $\mu_k[\neg c_k] = 0$ . But  $X \setminus [\mathbf{c}] \subset [\neg c_i] \cup [\neg c_j] \cup [\neg c_k]$ ; thus,  $\mu(X \setminus [\mathbf{c}]) \leq \mu[\neg c_i] + \mu[\neg c_j] + \mu[\neg c_k] < \frac{1}{2} + \frac{1}{2} + 0 = 1$ . But if  $\mu(X \setminus [\mathbf{c}]) < 1$ , then  $\mu(X \cap [\mathbf{c}]) > 0$ , which means  $[\mathbf{c}] \cap X \neq \emptyset$ , contradicting the fact that  $\mathbf{c}$  is an  $X$ -forbidden fragment.

By contradiction, we must have  $\mu_k[c_k] < 1$ . ◇ **Claim 1**

The contrapositive of Claim 1 says that  $y_k = \text{Maj}_k(\mu)$  whenever  $\mu_k(\text{Maj}_k(\mu)) = 1$ ; thus,  $\mathbf{y} \in \text{Unan}(\mu)$ . This holds for all  $\mathbf{y} \in \text{Cond}(X, \mu)$ ; hence  $\text{Cond}(X, \mu) \subseteq \text{Unan}(\mu)$ , as desired. □

*Proof of Example 2.3(a).* (This is a special case of the proof of Example 2.3(b) below).

“ $\kappa(X_N^{\text{pr}}) \leq N$ ”: Any  $X_N^{\text{pr}}$ -forbidden fragment corresponds to a ‘preference cycle’  $a_1 \prec a_2 \prec \dots \prec a_M \prec a_1$  (creating a transitivity violation), for some  $a_1, \dots, a_M \in [1\dots N]$ ;

such a cycle can have at most  $N$  elements. Furthermore, this  $X_N^{\text{pr}}$ -forbidden fragment is critical if it only specifies the ordering between nearest neighbours in the preference cycle (otherwise it has redundant coordinates which can be removed to obtain a forbidden subfragment). Thus a  $X_N^{\text{pr}}$ -critical fragment requires at most  $N$  coordinates.

“ $\kappa(X_N^{\text{pr}}) \geq N$ ”: Let  $\mathbf{c}$  be the fragment of order  $N$  representing the set of  $N$  assertions:  $\{1 \prec 2, 2 \prec 3, 3 \prec 4, \dots, (N-1) \prec N, N \prec 1\}$ . Then  $\mathbf{c}$  is forbidden (it creates a cycle), but no subfragment of  $\mathbf{c}$  is forbidden; hence  $\mathbf{c}$  is  $X_N^{\text{pr}}$ -critical.  $\square$

*Notation:* For any subset  $I \subset [1\dots K]$ , we define  $\mathbf{1}_I^*$  to be the fragment  $\mathbf{x}$  such that  $x_i := 1$  for all  $i \in I$  and  $x_k := *$  for all  $k \in [1\dots K] \setminus I$ .

*Proof of Example 2.3(b).* Let  $W := \text{width}(\triangleleft)$ . We must show that  $\kappa(X_{\triangleleft}^{\text{pr}}) = W$ .

Recall that  $K = N(N-1)/2$  and  $X_N^{\text{pr}} \subset \{0, 1\}^K$  is defined by identifying  $[1\dots K]$  with some subset of  $[1\dots N] \times [1\dots N]$  containing, for all distinct  $n, m \in \mathbb{N}$ , exactly one element of each pair  $\{(n, m), (m, n)\}$ . By reordering  $[1\dots K]$  if necessary, we can suppose that the partial ordering  $\triangleleft$  is obtained by fixing the values of the coordinates  $[J+1 \dots K]$  for some  $J \in [1\dots K]$ . Thus, we can regard  $X_{\triangleleft}^{\text{pr}}$  as a subset of  $\{0, 1\}^J$ .

“ $\kappa(X_{\triangleleft}^{\text{pr}}) \geq W$ ” Let  $A := \{a_1, \dots, a_W\} \subseteq [1\dots N]$  be an antichain of cardinality  $W$ . Let  $I \subset [1\dots J]$  be the set of coordinates representing the  $W$  pairs  $\{(a_1, a_2), (a_2, a_3), \dots, (a_{W-1}, a_W), (a_W, a_1)\}$ . Thus, the fragment  $\mathbf{1}_I^*$  represents the (cyclical) binary relation “ $a_1 \triangleleft a_2 \triangleleft \dots \triangleleft a_N \triangleleft a_1$ ”. Thus,  $\mathbf{1}_I^*$  is a  $(X_{\triangleleft}^{\text{pr}})$ -forbidden fragment.

Now, for any linear ordering of  $A$ , there is some element of  $X_{\triangleleft}^{\text{pr}}$  which encodes an extension of this ordering (because  $A$  is an antichain of  $\triangleleft$ ). Thus, no subfragment of  $\mathbf{1}_I^*$  is  $(X_{\triangleleft}^{\text{pr}})$ -forbidden; hence  $\mathbf{1}_I^*$  is a  $(X_{\triangleleft}^{\text{pr}})$ -critical fragment. Clearly,  $|\mathbf{1}_I^*| = W$ . Thus,  $\kappa(X_{\triangleleft}^{\text{pr}}) \geq W$ .

“ $\kappa(X_{\triangleleft}^{\text{pr}}) \leq W$ ” (by contradiction) Let  $\mathbf{v}$  be a  $(X_{\triangleleft}^{\text{pr}})$ -critical fragment, and suppose  $|\mathbf{v}| > W$ . Then  $\mathbf{v}$  must encode a cyclical binary relation “ $b_1 \succ b_2 \succ \dots \succ b_V \succ b_1$ ”, for some subset  $B := \{b_1, \dots, b_V\} \subset [1\dots N]$ . Let  $I \subset [1\dots J]$  be the set of  $V$  coordinates representing the  $V$  pairs  $\{(b_1, b_2), (b_2, b_3), \dots, (b_{V-1}, b_V), (b_V, b_1)\}$ ;

**Claim 1:**  $\mathbf{v} = \mathbf{1}_I^*$ ; hence  $|\mathbf{v}| = V$ .

*Proof:* (by contradiction) If  $\text{supp}(\mathbf{v})$  contained any coordinates not linking two elements from the set  $\{b_1, \dots, b_V\}$ , then we could remove this coordinate without destroying the cycle. Likewise, if  $\text{supp}(\mathbf{v})$  contained  $(b_n, b_m)$  for any  $n, m \in [1\dots V]$  with  $m \not\equiv n \pm 1 \pmod{V}$ , then we could remove this coordinate without breaking the cycle. Either way we would obtain a shorter  $(X_{\triangleleft}^{\text{pr}})$ -forbidden fragment, contradicting the minimality of  $\mathbf{v}$ .  $\diamond$  **Claim 1**

Now,  $|B| = V$ , and Claim 1 implies that  $V = |\mathbf{v}|$ , while  $|\mathbf{v}| > W$  by hypothesis. Thus,  $|B| > W$ , so  $B$  cannot be a  $\triangleleft$ -antichain. Thus, there exist some  $n, m \in [1\dots V]$  such that  $b_n \triangleleft b_m$ . Clearly,  $n \not\equiv m \pm 1 \pmod{V}$ , because  $(n, n+1)$  and  $(n, n-1)$  are represented

by elements of  $[1\dots J]$  (because they form part of the support of  $\mathbf{v}$ ). If  $n < m$ , then we can remove the elements  $\{b_{n+1}, \dots, b_{m-1}\}$  and get a shorter cycle:

$$b_1 \prec b_2 \prec \dots \prec b_n \triangleleft b_m \prec b_{m+1} \prec \dots \prec b_{V-1} \prec b_V \prec b_1.$$

If  $n > m$ , then we can remove the elements  $\{b_1, \dots, b_{m-1}\}$  and  $\{b_{n+1}, \dots, b_V\}$  and get a shorter cycle:

$$b_m \prec b_{m+1} \prec \dots \prec b_{n-1} \prec b_n \triangleleft b_m.$$

Either way, we can construct a *smaller* ( $X_{\triangleleft}^{\text{pr}}$ )-forbidden fragment, which contradicts the minimality of  $\mathbf{v}$ . By contradiction, we must have  $V \leq W$ . This argument holds for any critical fragment  $\mathbf{v}$ . Thus,  $\kappa(X_{\triangleleft}^{\text{pr}}) \leq W$ .  $\square$

*Proof of Example 2.3(c).* “ $\kappa(X_N^{\text{eq}}) \leq N$ ”: Any  $X_N^{\text{eq}}$ -forbidden fragment corresponds to some ‘broken cycle’ of the form:  $a_1 \sim a_2 \sim \dots \sim a_M \not\sim a_1$  (creating a transitivity violation), for some  $a_1, \dots, a_M \in [1\dots N]$ ; such a broken cycle can have at most  $N$  elements. Furthermore, this  $X_N^{\text{eq}}$ -forbidden fragment is critical only if it specifies only the (non)equivalences between nearest neighbours in the broken cycle (otherwise it has redundant coordinates which can be removed to obtain a forbidden subfragment). Thus an  $X_N^{\text{eq}}$ -critical fragment requires at most  $N$  coordinates.

“ $\kappa(X_N^{\text{eq}}) \geq N$ ”: let  $\mathbf{c} \in X_N^{\text{eq}}$  be the fragment of order  $N$  representing the set of  $N$  assertions  $\{1 \sim 2, 2 \sim 3, 3 \sim 4, \dots, (N-1) \sim N, N \not\sim 1\}$ . Then  $\mathbf{c}$  is  $X_N^{\text{eq}}$ -forbidden (it violates transitivity), but no subfragment of  $\mathbf{c}$  is forbidden; hence  $\mathbf{c}$  is  $X_N^{\text{eq}}$ -critical.  $\square$

*Proof of Example 2.3(e).* “ $\kappa(X_{M,D}^{\triangleleft}) \leq D$ ”: Let  $\mathbf{c}$  be a  $X_{M,D}^{\triangleleft}$ -critical fragment. Suppose there exist  $d \in [1\dots D]$  and  $n < m \in [1\dots M]$  such that  $c_{(d,n)} = 0$  and  $c_{(d,m)} = 1$ . Then we could eliminate all other coordinates from  $\mathbf{c}$  to obtain a  $X_{M,D}^{\triangleleft}$ -forbidden fragment containing only these two coordinates. Thus, if  $\mathbf{c}$  was  $X_{M,D}^{\triangleleft}$ -critical, then we must have  $\mathbf{c} = (c_{(d,n)}, c_{(d,m)})$ ; hence  $|\mathbf{c}| = 2$ .

So, for all  $d \in [1\dots D]$  assume that there do *not* exist any  $n < m \in [1\dots M]$  such that  $c_{(d,n)} = 0$  and  $c_{(d,m)} = 1$ .

**Claim 1:** For each  $d \in [1\dots D]$ ,  $\text{supp}(\mathbf{c})$  contains at most one of the coordinates  $\{(d, 1), (d, 2), \dots, (d, M)\}$ .

*Proof:* (by contradiction) Let  $1 \leq n < m \leq M$ , and suppose  $(d, n) \in \text{supp}(\mathbf{c})$  and  $(d, m) \in \text{supp}(\mathbf{c})$ . If  $c_{(d,n)} = 1 = c_{(d,m)}$  then we can remove coordinate  $c_{(d,n)}$  and still have a  $X_{M,D}^{\triangleleft}$ -forbidden fragment. If  $c_{(d,n)} = 0 = c_{(d,m)}$ , then we can remove coordinate  $c_{(d,m)}$  and still have a  $X_{M,D}^{\triangleleft}$ -forbidden fragment. Now suppose  $c_{(d,n)} = 1$  and  $c_{(d,m)} = 0$ . If  $\|\mathbf{c}\| > M$  then we can remove  $c_{(d,m)}$  and still have a  $X_{M,D}^{\triangleleft}$ -forbidden fragment. If  $\|\mathbf{c}\| < M$ , then we can remove  $c_{(d,n)}$  and still have a  $X_{M,D}^{\triangleleft}$ -forbidden fragment.

In any case, we can remove a coordinate from  $\mathbf{c}$  to obtain a smaller  $X_{M,D}^{\triangleleft}$ -forbidden fragment; hence  $\mathbf{c}$  is not  $X_{M,D}^{\triangleleft}$ -critical.  $\diamond$  **Claim 1**

Claim 1 implies  $|\mathbf{c}| \leq D$ . This holds for any  $X_{M,D}^\Delta$ -critical fragment, so  $\kappa(X_{M,D}^\Delta) \leq D$ .

“ $\kappa(X_{M,D}^\Delta) \geq D$ ”: Let  $m \in [1 \dots M]$  be the smallest number such that  $mD > M$ , and consider the fragment  $\mathbf{c}$  of order  $D$  defined by the  $D$  assertions  $\{x_1 \geq m, x_2 \geq m, \dots, x_D \geq m\}$ . This fragment is  $X_{M,D}^\Delta$ -forbidden because it requires allocating a total of  $mD > M$  ‘dollars’. No subfragment of  $\mathbf{c}$  is  $X_{M,D}^\Delta$ -forbidden, because  $(D-1)m = mD - m \stackrel{(*)}{\leq} mD - D = (m-1)D \stackrel{(\dagger)}{\leq} M$ . (Here  $(*)$  is because  $D \leq m$  because  $D^2 \leq M$ , and  $(\dagger)$  is because  $m$  is the *smallest* number with  $mD > M$ ). Thus,  $\mathbf{c}$  is critical. Thus,  $\kappa(X_{M,D}^\Delta) \geq |\mathbf{c}| = D$ .  $\square$

*Proof of Theorem 3.1(b).* By Example 2.3(e) it suffices to show that the spaces  $X_{M,D}^\Delta$  are not designably unanimity consistent whenever  $D \geq 4$ . Recall from equation (3) in Section 1.5.4 the way in which  $X_{M,D}^\Delta$  is embedded in  $\{0, 1\}^{D \times M}$ . Consider any path  $\gamma : [1 \dots DM] \rightarrow [1 \dots D] \times [1 \dots M]$ . For all  $d \in [1 \dots D]$ , let  $t(d) := \min\{t \in [1 \dots DM] : \gamma(t) = (d, m) \text{ for some } m \in [1, M/3]\}$ . Find  $d_1, d_2, d_3 \in [1 \dots D]$  such that  $t(d_1) < t(d_2) < t(d_3) < t(d)$  for all other  $d \in [1 \dots D]$ . For all  $j \in \{1, 2, 3\}$ , let  $\mathbf{x}^j$  be the element of  $X_{M,D}^\Delta$  which allocates all  $M$  dollars towards claimant  $d_j$ . (Thus, for all  $m \in [1 \dots M]$ , we have  $x_{d_j, m}^j = 1$ , while  $x_{c, m}^j = 0$  for all  $c \in [1 \dots D] \setminus \{d_j\}$ .) Then let  $\mu \in \Delta(X_{M,D}^\Delta)$  be the profile with  $\mu[\mathbf{x}^j] = \frac{1}{3}$  for  $j \in \{1, 2, 3\}$ . Observe that  $\text{Maj}(\mu) = \mathbf{0}$ . Furthermore, observe that  $\mu_{d,m}(0) = 1$  (unanimity) for all  $m \in [1 \dots M]$  and  $d \in [1 \dots D] \setminus \{d_1, d_2, d_3\}$  (this set is nonempty because  $D \geq 4$ ).

By construction, we must have  $F^\gamma(\mu)_{d_1, m_1} = F^\gamma(\mu)_{d_2, m_2} = F^\gamma(\mu)_{d_3, m_3} = 0$  for some  $m_1, m_2, m_3 \in [1, M/3]$ . Thus,  $F^\gamma(\mu)$  allocates less than  $M/3$  dollars towards each of claimants  $d_1, d_2, d_3$ . Thus,  $F^\gamma(\mu)$  must allocate more than 0 dollars towards some other claimant. That is,  $F^\gamma(\mu)_{d,m} = 1$  for some  $m \in [1 \dots M]$  and  $d \in [1 \dots D] \setminus \{d_1, d_2, d_3\}$ . But this violates unanimity in coordinate  $(d, m)$ .  $\square$

*Proof of Lemma 2.* We prove this by induction on  $N$ , first for equivalence relations. Evidently, the statement holds if  $N = 2$ . By induction, suppose the statement holds for all sets of alternatives with less than  $N$  elements. For notational convenience, we write  $aI(\mu)b$  [resp.  $aI^*(\mu)b$ ,  $aI^\zeta(\mu)b$ ] if  $aR(\mu)b$  and  $bR(\mu)a$  [resp.  $(aR^*(\mu)b$  and  $bR^*(\mu)a$ ),  $(aR^\zeta(\mu)b$  and  $bR^\zeta(\mu)a$ ]. Along the path  $\zeta$ , the alternative 1 is first compared with every other alternative in  $[2 \dots N]$ . Evidently, the two sets  $E_1 := \{a \in [1 \dots N] : aI(\mu)1\}$  and  $N_1 := \{a \in [1 \dots N] : \text{not } aI(\mu)1\}$  partition the set  $[1 \dots N]$ . By construction, one has  $aI^\zeta(\mu)1 \Leftrightarrow aI(\mu)1$ , i.e.  $aI^\zeta(\mu)1 \Leftrightarrow a \in E_1$ . By transitivity, one therefore has  $aI^\zeta(\mu)b$  for all  $a, b \in E_1$  and not  $(cI^\zeta(\mu)d)$  for all  $c \in E_1$  and all  $d \in N_1$ , i.e.  $E_1$  is the  $R^\zeta(\mu)$ -equivalence class to which 1 belongs. Note that at this moment,  $R^\zeta(\mu)$  is determined exactly for those issues  $(a, b)$  for which  $a \in E_1$  or  $b \in E_1$ .

**Claim 1:** *The relation  $R^\zeta(\mu)$  as determined so far respects the covering fragment, that is: for all issues  $(a, b)$  such that  $a \in E_1$  or  $b \in E_1$ ,  $aR^*(\mu)b \Rightarrow aR^\zeta(\mu)b$ , and  $aR^{**}(\mu)b \Rightarrow \text{not } (aR^\zeta(\mu)b)$ .*

*Proof:* Let  $aR^*(\mu)b$ . If  $a \in E_1$ , then  $1R(\mu)a$ , and hence  $1R(\mu)b$ , i.e.  $b \in E_1$ , which implies  $aR^\zeta(\mu)b$ . Similarly, if  $b \in E_1$  one has  $bR(\mu)1$ , and hence  $aR(\mu)1$ , which implies  $a \in E_1$  and thus again  $aR^\zeta(\mu)b$ .

Next, let  $aR^{**}(\mu)b$ . If  $a \in E_1$ , then  $b \notin E_1$ , and hence not  $(aR^\zeta(\mu)b)$ . Similarly, if  $b \in E_1$ , then  $a \notin E_1$ , and hence again not  $(aR^\zeta(\mu)b)$ .  $\diamond$  claim 1

Now consider the set  $N_1 \subset [1\dots N]$ , and note that the issues  $(a, b) \in N_1 \times N_1$  are exactly the issues which are still ‘free,’ i.e. the issues on which  $R^\zeta(\mu)$  is not yet determined. Evidently, the restriction  $\zeta'$  of  $\zeta$  to  $N_1 \times N_1$  is simply ordered, and  $R^\zeta(\mu)$  restricted to  $N_1 \times N_1$  coincides with  $R^{\zeta'}(\mu')$ , i.e. with the equivalence relation on  $N_1$  that results from sequential majority voting applied to the restriction  $\mu'$  of  $\mu$  to  $N_1$  along the path  $\zeta'$ .

For all  $a, b \in N_1$ , it is clear that  $aR^*(\mu)b \Rightarrow aR^*(\mu')b$  and  $aR^{**}(\mu)b \Rightarrow aR^{**}(\mu')b$ . By the induction hypothesis,  $R^{\zeta'}(\mu')$  extends the covering fragment  $[R^*(\mu'), R^{**}(\mu')]$  on  $N_1$ . Thus,  $R^\zeta(\mu)$  extends the restriction to  $N_1$  of the covering fragment  $[R^*(\mu), R^{**}(\mu)]$ . This, together with Claim 1, completes the proof of Lemma 2 for equivalence relations.

Next, consider the case of weak orderings; from this, the case of linear preference orderings follows as a special case. Again, we prove Lemma 2 by induction on  $N$ . For  $N = 2$  it holds trivially. By induction, assume that Lemma 2 holds for all sets of alternatives with less than  $N$  elements. Consider alternative 1, and define  $U_1 := \{a \in [1\dots N] : aP(\mu)1\}$ ,  $E_1 := \{a \in [1\dots N] : aI(\mu)1\}$ ,  $L_1 := \{a \in [1\dots N] : 1P(\mu)a\}$ ; recall that  $aP(\mu)b$  means  $[aR(\mu)b$  and not  $(bR(\mu)a)]$ . Evidently,  $1 \in E_1$  (in the case of linear orderings one has  $E_1 = \{1\}$ ); moreover, the collection of the non-empty sets from  $\{U_1, E_1, L_1\}$  forms a partition of  $[1\dots N]$ . Sequential majority voting along the path  $\zeta$  yields  $aP^\zeta(\mu)1$  for all  $a \in U_1$ ,  $aI^\zeta(\mu)1$  for all  $a \in E_1$ , and  $1P^\zeta(\mu)a$  for all  $a \in L_1$ . Hence, by transitivity,

- (i)  $aP^\zeta(\mu)b$  for all  $(a, b) \in U_1 \times (E_1 \cup L_1)$ ;
- (ii)  $aP^\zeta(\mu)b$  for all  $(a, b) \in E_1 \times L_1$ ; and
- (iii)  $aI^\zeta(\mu)b$  for all  $(a, b) \in E_1 \times E_1$ .

Define  $Z_1 := (U_1 \times U_1) \cup (L_1 \times L_1) \subset [1\dots N] \times [1\dots N]$ , and note that at this moment,  $R^\zeta(\mu)$  is determined exactly for all issues  $(a, b) \in ([1\dots N] \times [1\dots N]) \setminus Z_1$ .

**Claim 2:** *The relation  $R^\zeta(\mu)$  as determined so far respects the covering fragment, that is: for all issues  $(a, b) \notin Z_1$ ,  $aR^*(\mu)b \Rightarrow aR^\zeta(\mu)b$ , and  $aR^{**}(\mu)b \Rightarrow \text{not}(aR^\zeta(\mu)b)$ .*

*Proof:* Let  $(a, b) \notin Z_1$ . The proof consists of several cases and subcases.

- First suppose  $aR^*(\mu)b$ . There are two subcases; either  $a \in U_1$  or  $a \in E_1 \cup L_1$ .
  1. If  $a \in U_1$ , then  $b \in E_1 \cup L_1$  (because  $(a, b) \notin Z_1$ ), and hence  $aR^\zeta(\mu)b$  by observation (i) above.
  2. If  $a \in E_1 \cup L_1$ , then  $1R(\mu)a$ , and therefore, since  $aR^*(\mu)b$ , we also have  $1R(\mu)b$ , i.e.  $b \in E_1 \cup L_1$ . This yields two further subcases.

- (a) If  $b \in E_1$ , then  $bR(\mu)1$ , and therefore, since  $aR^*(\mu)b$ , we also have  $aR(\mu)1$ , i.e.  $a \in E_1$ , hence  $aR^\zeta(\mu)b$  by observation (iii) above.
- (b) If  $b \in L_1$  then  $a \in E_1$  since  $(a, b) \notin Z_1$ , hence  $aR^\zeta(\mu)b$  by observation (ii).

In any of the three cases considered, we have  $aR^\zeta(\mu)b$ , as desired.

- Now suppose  $aR^{**}(\mu)b$ . There are two subcases; either  $a \in E_1 \cup U_1$  or  $a \in L_1$ .
  1. If  $a \in E_1 \cup U_1$  then  $aR(\mu)1$  and therefore, since  $aR^{**}(\mu)b$ , we deduce that  $\text{not}(1R(\mu)b)$ , i.e.  $b \in U_1$ . Since  $(a, b) \notin Z_1$ , we must thus have  $a \in E_1$ , in which case  $bP^\zeta(\mu)a$  by observation (i). This means that  $\text{not}(aR^\zeta(\mu)b)$ .
  2. If  $a \in L_1$  then  $b \in E_1 \cup U_1$  (because  $(a, b) \notin Z_1$ ), in which case again  $bP^\zeta(\mu)a$  by observations (i) and (ii), and hence  $\text{not}(aR^\zeta(\mu)b)$ .

In both cases, we conclude that  $\text{not}(aR^\zeta(\mu)b)$ , as desired. ◇ claim 2

Now consider the set  $Z_1 = (U_1 \times U_1) \cup (L_1 \times L_1)$ , and recall that the issues  $(a, b) \in Z_1$  are exactly the issues which are still ‘free,’ i.e. the issues on which  $R^\zeta(\mu)$  is not yet determined. Evidently, the restrictions  $\zeta'$  of  $\zeta$  to  $U_1 \times U_1$  and  $\zeta''$  of  $\zeta$  to  $L_1 \times L_1$  are simply ordered. Moreover,  $R^\zeta(\mu)$  restricted to  $U_1 \times U_1$  coincides with  $R^{\zeta'}(\mu')$ , i.e. with the weak ordering on  $U_1$  that results from sequential majority voting applied to the restriction  $\mu'$  of  $\mu$  to  $U_1$  along the path  $\zeta'$ , and  $R^\zeta(\mu)$  restricted to  $L_1 \times L_1$  coincides with  $R^{\zeta''}(\mu'')$ , i.e. with the weak ordering on  $L_1$  that results from sequential majority voting applied to the restriction  $\mu''$  of  $\mu$  to  $L_1$  along the path  $\zeta''$ .

For all  $a, b \in U_1$ , it is easy to see that  $aR^*(\mu)b \Rightarrow aR^*(\mu')b$  and  $aR^{**}(\mu)b \Rightarrow aR^{**}(\mu')b$ . By the induction hypothesis,  $R^{\zeta'}(\mu')$  extends the covering fragment  $[R^*(\mu'), R^{**}(\mu')]$  on  $U_1$ . Thus,  $R^\zeta(\mu)$  extends the restriction to  $U_1$  of the covering fragment  $[R^*(\mu), R^{**}(\mu)]$ .

Likewise, for all  $a, b \in L_1$ , we have  $aR^*(\mu)b \Rightarrow aR^*(\mu'')b$  and  $aR^{**}(\mu)b \Rightarrow aR^{**}(\mu'')b$ . By the induction hypothesis,  $R^{\zeta''}(\mu'')$  extends the covering fragment  $[R^*(\mu''), R^{**}(\mu'')]$  on  $L_1$ . Thus,  $R^\zeta(\mu)$  extends the restriction to  $L_1$  of the covering fragment  $[R^*(\mu), R^{**}(\mu)]$ . This observation, together with Claim 2, completes the proof of Lemma 2 for weak orderings.

An inspection of the proof of the lemma for weak orderings shows at once that the result holds also for the case of linear preference orderings. In this case, the argument simplifies somewhat due to the fact that all indifference classes are singletons, as already noted. □

*Proof of Example 3.3(a)* First suppose that  $X = X_N^{\text{pr}}$ . Let  $\mathbf{c}$  be a critical fragment for  $X_N^{\text{pr}}$ . First suppose  $|\mathbf{c}| = 3$ . Then  $\mathbf{c}$  has the form  $(a \succ b, b \succ c, c \succ a)$ , for some three alternatives  $a, b, c \in A$ . The fragment  $\mathbf{c}$  specifies the (cyclic) preference relations amongst  $a, b$ , and  $c$ ; a complementary fragment  $\mathbf{d}$  will specify the preference relations between all elements of  $A \setminus \{a, b, c\}$ , and also specify the preferences between the elements of  $A \setminus \{a, b, c\}$  and those of  $\{a, b, c\}$ . Suppose we construct  $\mathbf{d}$  to specify the following preferences:

- $\mathbf{d}$  specifies a specific transitive ordering on  $A \setminus \{a, b, c\}$ . (For example, if  $A = \{a, b, c, d, e, f, \dots, z\}$ , then  $\mathbf{d}$  might specify  $d \succ e \succ f \succ \dots \succ z$ .)
- $\mathbf{d}$  stipulates that  $a \succ h$ ,  $b \succ h$  and  $c \succ h$  for all  $h \in A \setminus \{a, b, c\}$ .

Now, define  $\mathbf{c}^1 := (a \prec b, b \succ c, c \succ a)$ ,  $\mathbf{c}^2 := (a \succ b, b \prec c, c \succ a)$ , and  $\mathbf{c}^3 := (a \succ b, b \succ c, c \prec a)$ . Then each of  $(\mathbf{c}^1, \mathbf{d})$ ,  $(\mathbf{c}^2, \mathbf{d})$  and  $(\mathbf{c}^3, \mathbf{d})$  is an element of  $X_N^{\text{pr}}$ . To see this, note that,  $(\mathbf{c}^1, \mathbf{d})$  corresponds to the preference order  $b \succ c \succ a \succ d \succ e \succ f \succ \dots \succ z$ , while  $(\mathbf{c}^2, \mathbf{d})$  corresponds to the preference order  $c \succ a \succ b \succ d \succ e \succ f \succ \dots \succ z$ , and  $(\mathbf{c}^3, \mathbf{d})$  corresponds to the preference order  $a \succ b \succ c \succ d \succ e \succ f \succ \dots \succ z$ .

We conclude that all critical fragments of length 3 are exposed critical fragments for  $X_N^{\text{pr}}$ .

To see the converse, suppose  $|\mathbf{c}| = 4$ . Then  $\mathbf{c}$  has the form  $(a \succ b, b \succ c, c \succ d, d \succ a)$ , for some four alternatives  $a, b, c, d \in A$ . We claim there is no single complementary fragment  $\mathbf{d}$  which is compatible with every subfragment of three out of the four coordinates of  $\mathbf{c}$ . To see this, note that, in addition to specifying relations between elements of  $A \setminus \{a, b, c, d\}$ , the fragment  $\mathbf{d}$  must also specify the relationship between  $a$  and  $c$ , and the relationship between  $b$  and  $d$ . If we negate the last coordinate of  $\mathbf{c}$  to obtain  $\mathbf{c}' := (a \succ b, b \succ c, c \succ d, d \prec a)$ , then  $\mathbf{d}$  must stipulate that  $a \succ c$  and  $b \succ d$  (corresponding to the transitive ordering  $a \succ b \succ c \succ d$ ). But if instead, we negate second coordinate of  $\mathbf{c}$  to obtain  $\mathbf{c}'' := (a \succ b, b \prec c, c \succ d, d \succ a)$ , then  $\mathbf{d}$  must stipulate that  $a \prec c$  and  $b \prec d$  (corresponding to the transitive ordering  $c \succ d \succ a \succ b$ ). Thus, there is no single complementary fragment  $\mathbf{d}$  which is compatible with both  $\mathbf{c}'$  and  $\mathbf{c}''$ . Thus,  $\mathbf{c}$  is *not* exposed. A similar problem occurs if  $|\mathbf{c}| \geq 5$ . Thus, all exposed critical fragments of  $X_N^{\text{pr}}$  must have length 3.

The proof that  $\kappa^*(X_N^{\text{p.o.}}) = 3$  is similar. □

*Proof of Example 3.3(b).* Suppose  $A = \{1, 2, \dots, N\}$ . For any  $m \leq N$ , consider the fragment  $(1 \succ 2, 2 \succ 3, \dots, (m-1) \succ m, m \succ 1)$ . This fragment is obviously forbidden, since it describes a cyclic order. To see that it is exposed, let  $\mathbf{d}$  be a complementary fragment which stipulates that  $i \not\succeq j$  and  $j \not\succeq i$  for all  $i \in [m+1 \dots N]$  and all  $j \in [1 \dots N]$ , and also  $i \not\succeq j$  for all non-adjacent  $i, j \in [1 \dots m]$ . If  $\mathbf{c}'$  is obtained by negating any single coordinate of  $\mathbf{c}$ , then  $(\mathbf{c}', \mathbf{d})$  is an element of  $X_N^{\text{p.o.}}$ ; it describes an acyclic relation where the elements  $[1 \dots m]$  are linearly ordered, whereas all other elements of  $A$  remain completely unordered. For example, if we negate the last coordinate of  $\mathbf{c}$ , we get  $\mathbf{c}' = (1 \succ 2, 2 \succ 3, \dots, (m-1) \succ m, m \not\succeq 1)$ , which corresponds to the partial order  $1 \succ 2 \succ \dots \succ m$ .

In this way, any cycle of elements in  $A$  of any length  $m$  defines an exposed critical fragment of size  $m$ . Thus,  $\kappa^*(X_N^{\text{p.o.}}) \geq N$ . □

*Proof of Theorem 3.4.* Suppose  $\kappa^*(X) > 4$ ; we will show that  $X$  is *not* designably unanimity consistent. Let  $\mathbf{c} \in C^*(X)$  be an exposed critical fragment with  $|\mathbf{c}| \geq 4$ . Without loss of generality (by reordering  $[1 \dots K]$  if necessary), suppose  $\text{supp}(\mathbf{c}) = [1 \dots J+1]$ , where  $3 \leq J < K$ ; we will write  $\mathbf{c} = (c_1, c_2, \dots, c_{J+1})$ . By hypothesis, there exists some ‘‘complementary’’ fragment  $\mathbf{d} = (d_{J+2}, \dots, d_K) \in \{0, 1\}^{K-J-1}$  such that, for all  $j \in [1 \dots J+1]$ , the sequence  $\mathbf{x}^j := (c_0, c_1, \dots, c_{j-1}, \neg c_j, c_{j+1}, \dots, c_{J+1}, d_{J+2}, \dots, d_K)$  is an element of  $X$ .

Now, let  $\zeta$  be any path through  $K$ . The path  $\zeta$  encounters the elements of  $[1 \dots J+1]$  in some order. Let  $j^*$  be the *last* element of  $[1 \dots J+1]$  encountered by  $\zeta$ . Now let  $\mu \in \Delta^*(X)$  be the profile such that  $\mu(\mathbf{x}^j) = 1/J$  for all  $j \in [1 \dots J+1]$  except  $j^*$ .

**Claim 1:** For all  $k \in [J+2 \dots K]$ , we have  $F^\zeta(\mu)_k = d_k$ . For all  $j \in [1 \dots J+1] \setminus \{j^*\}$ , we have  $F^\zeta(\mu)_j = c_j$ .

*Proof:* For all  $k \in [J+2 \dots K]$ , we have  $\mu_k(d_k) = 1$  by construction; thus  $\text{Maj}(\mu)_k = d_k$ .

Likewise for all  $j \in [1 \dots J+1] \setminus \{j^*\}$ , we have  $\mu_j(c_j) = (J-1)/J > 1/2$  (because  $J \geq 3$ ); thus  $\text{Maj}(\mu)_j = c_j$ .

Finally, by hypothesis, the entire fragment  $(c_0, c_1, \dots, c_{j^*-1}, *, c_{j^*+1}, \dots, c_{J+1}, d_{J+2}, \dots, d_K)$  is not forbidden, because it is contained in  $\mathbf{x}^{j^*}$ , which is an element of  $X$ . Thus  $F^\zeta(\mu)$  agrees with  $\text{Maj}(\mu) = (\mathbf{c}, \mathbf{d})$  in all coordinates of  $[1 \dots K] \setminus \{j^*\}$ .  $\diamond$  **claim 1**

The word  $\mathbf{c}$  is forbidden. Thus, Claim 1 implies that  $F^\zeta(\mu)_{j^*} \neq c_{j^*}$ . But by construction,  $\mu[c_{j^*}] = 1$ . Thus,  $F^\zeta$  must violate unanimity in coordinate  $j^*$ .

We can construct such a profile  $\mu$  for *any* path  $\zeta$ . Thus,  $F$  is *not* designably unanimity consistent.  $\square$

**Remark.** The proof of Theorem 3.4 seems very similar to the proof of Theorem 2.2, but there is a subtle difference. In the proof of Theorem 2.2, we could specify a path which moved through *all* elements of  $[1 \dots J+1]$  before moving on to *any* elements of  $[J+2 \dots K]$ . Thus, it was easy to construct a profile which forced this path to violate unanimity. But in the proof of Theorem 3.4 we don’t have this freedom; the proof must work for *any* path. This is why  $\mathbf{c}$  must be an *exposed* critical fragment.

*Proof of Example 3.5.* We show that all exposed critical fragments of  $X^{\text{sep}}$  have length equal to 3. Let  $\mathbf{c}$  be a critical fragment of  $X^{\text{sep}}$ . First, suppose that  $|\mathbf{c}| = 2$ ; then,  $\mathbf{c} = (c_1, c_2)$  corresponds to a separability restriction of the form  $[(a_j, a_{-j}) \succ (b_j, a_{-j})] \Leftrightarrow [(a_j, b_{-j}) \succ (b_j, b_{-j})]$  for some  $a, b \in \{0, 1\}^3$  and some  $j \in \{1, 2, 3\}$ . Say,  $c_1$  corresponds to the affirmation of ‘ $(a_j, a_{-j}) \succ (b_j, a_{-j})$ ’ while  $c_2$  corresponds to the negation of ‘ $(a_j, b_{-j}) \succ (b_j, b_{-j})$ .’ Consider any  $f_{-j}$  different from both  $a_{-j}$  and  $b_{-j}$  (such an  $f$  exists because  $a, b \in \{0, 1\}^3$ ), and the issue  $(a_j, f_{-j}) \succ (b_j, f_{-j})$ . Any complementary fragment  $\mathbf{d} \in \{0, 1\}^{K-2}$  must either affirm or negate the proposition ‘ $(a_j, f_{-j}) \succ (b_j, f_{-j})$ ,’ hence one cannot have both  $(c_1, \neg c_2, \mathbf{d}) \in X^{\text{sep}}$  and  $(\neg c_1, c_2, \mathbf{d}) \in X^{\text{sep}}$ . Thus,  $\mathbf{c}$  is not exposed.

Next, suppose  $|\mathbf{c}| > 2$ , and suppose that  $\mathbf{c}$  is exposed.

**Claim 1:**  $\mathbf{c}$  does not contain any issue of the form  $(a_j, a_{-j}) \succ (b_j, a_{-j})$  for some  $j \in \{1, 2, 3\}$ .

*Proof:* By contraposition, suppose that, say,  $c_1$  corresponds to the issue  $(a_j, a_{-j}) \succ (b_j, a_{-j})$  for some alternatives  $a, b \in \{0, 1\}^3$  and some  $j \in \{1, 2, 3\}$ . Consider any  $f_{-j} \neq a_{-j}$  and the issue  $(a_j, f_{-j}) \succ (b_j, f_{-j})$ ; since  $|\mathbf{c}| > 2$  and since  $\mathbf{c}$  is *minimally* forbidden,  $\mathbf{c}$  cannot involve the issue  $(a_j, f_{-j}) \succ (b_j, f_{-j})$ . Thus, any ‘complementary’ fragment  $\mathbf{d}$  must either affirm or negate this issue; by separability,  $\mathbf{d}$  cannot be compatible with both  $c_1$  and  $\neg c_1$ . Therefore,  $\mathbf{c}$  is not exposed.  $\diamond$  Claim 1

By Claim 1, any exposed critical fragment  $\mathbf{c}$  must correspond to a direct violation of transitivity. But we already know from Example 3.3(a) that in this case one must have  $|\mathbf{c}| = 3$ .  $\square$

An example of an exposed critical fragment in  $X^{\text{sep}}$  is the preference cycle  $(1, 0, 0) \succ (0, 1, 0) \succ (0, 0, 1) \succ (1, 0, 0)$ . A ‘complementary’ fragment  $\mathbf{d}$  can be obtained by specifying a partial order with  $(1, 1, 1)$  as the best alternative,  $(0, 0, 0)$  as the worst alternative, all elements of the set  $A_1 := \{(1, 1, 0), (1, 0, 1), (0, 1, 1)\}$  strictly preferred to all elements of the set  $A_2 := \{(1, 0, 0), (0, 1, 0), (0, 0, 1)\}$ , and in addition some transitive ranking of the alternatives in  $A_1$ . As is easily seen, such  $\mathbf{d}$  is compatible with separability and any transitive ranking within  $A_2$ .

## References

- Banks, J., 1985. Sophisticated voting outcomes and agenda control. *Soc. Choice Welfare* 3, 295–306.
- Barthélémy, J.-P., Monjardet, B., 1981. The median procedure in cluster analysis and social choice theory. *Math. Social Sci.* 1 (3), 235–267.
- Cox, G., 1987. The uncovered set and the core. *Amer. J. Polit. Sci.* 31, 408–422.
- Dietrich, F., List, C., 2007. Judgement aggregation by quota rules: Majority voting generalized. *Journal of Theoretical Politics* 19 (4), 391–424.
- Dokow, E., Holzman, R., 2010. Aggregation of binary evaluations. *J.Econ.Theory* 145, 495–511.
- Kornhauser, L., Sager, L., 1986. Unpacking the court. *Yale Law Journal* 96, 82–117.
- Lindner, T., Nehring, K., Puppe, C., 2010. Allocating public goods via the midpoint rule. (preprint).
- List, C., 2004. A model of path-dependence in decisions over multiple propositions. *American Political Science Review* 98 (3), 495–513.
- List, C., Pettit, P., 2002. Aggregating sets of judgements: An impossibility result. *Economics and Philosophy* 18, 89–110.
- List, C., Puppe, C., 2009. Judgement aggregation: a survey. In: Anand, P., Pattanaik, P., Puppe, C. (Eds.), *Oxford handbook of rational and social choice*. Oxford University Press, Oxford, UK., Ch. 19, pp. 457–481.

- McKelvey, R., 1979. General conditions for global intransitivities in formal voting models. *Econometrica* 47 (5), 1085–1112.
- McKelvey, R., 1986. Covering, dominance, and institution-free properties of social choice. *Amer. J. Polit. Sci.* 30, 283–314.
- Miller, N. R., 1980. A new solution set for tournaments and majority voting: further graph-theoretical approaches to the theory of voting. *American Journal of Political Science* 24, 68–69.
- Miller, N. R., 1995. *Committees, Agendas, and Voting*. Routledge, London.
- Moulin, H., 1986. Choosing from a tournament. *Soc. Choice Welf.* 3 (4), 271–291.
- Moulin, H., 1988. *Axioms of cooperative decision making*. Cambridge University Press, Cambridge, UK.
- Nehring, K., 2005. The impossibility of a paretian rational. (preprint).
- Nehring, K., Pivato, M., Puppe, C., 2013. The Condorcet set: Majority voting over interconnected propositions. (preprint).
- Nehring, K., Puppe, C., 2002. Strategy-proof social choice on single-peaked domains: Possibility, impossibility and the space between. (preprint).
- Nehring, K., Puppe, C., 2007. The structure of strategy-proof social choice I: General characterization and possibility results on median spaces. *J.Econ.Theory* 135, 269–305.
- Nehring, K., Puppe, C., 2010. Abstract Arrowian aggregation. *J.Econ.Theory* 145, 467–494.
- Shepsle, K., Weingast, B., 1984. Uncovered sets and sophisticated voting outcomes, with implications for agenda institutions. *American Journal of Political Science* 28 (1), 49–75.

# Working Paper Series in Economics

---

recent issues

- No. 50** *Klaus Nehring, Marcus Pivato, Clemens Puppe: Unanimity overruled: majority voting and the burden of history, December 2013*
- No. 49** *Andranik S. Tangian: Decision making in politics and economics: 5. 2013 election to German Bundestag and direct democracy, December 2013*
- No. 48** *Marten Hillebrand, Tomoo Kikuchi, Masaya Sakuragawa: Bubbles and crowding-in of capital via a savings glut, November 2013*
- No. 47** *Dominik Rothenhäusler, Nikolaus Schweizer, Nora Szech: Institutions, shared guilt, and moral transgression, October 2013*
- No. 46** *Marten Hillebrand: Uniqueness of Markov equilibrium in stochastic OLG models with nonclassical production, November 2012*
- No. 45** *Philipp Schuster and Marliese Uhrig-Homburg: The term structure of bond market liquidity conditional on the economic environment: an analysis of government guaranteed bonds, November 2012*
- No. 44** *Young Shin Kim, Rosella Giacometti, Svetlozar T. Rachev, Frank J. Fabozzi, Domenico Mignacca: Measuring financial risk and portfolio optimization with a non-Gaussian multivariate model, August 2012*
- No. 43** *Zuodong Lin, Svetlozar T. Rachev, Young Shin Kim, Frank J. Fabozzi: Option pricing with regime switching tempered stable processes, August 2012*
- No. 42** *Siegfried K. Berninghaus, Werner Güth, Stephan Schosser: Backward induction or forward reasoning? An experiment of stochastic alternating offer bargaining, July 2012*
- No. 41** *Siegfried Berninghaus, Werner Güth, King King Li: Approximate truth of perfectness - an experimental test, June 2012*
- No. 40** *Marten Hillebrand and Tomoo Kikuchi: A mechanism for booms and busts in housing prices, May 2012*