The impossibility of unbiased judgment aggregation

Franz Dietrich and Christian List[1]

9/2005, revised 4/2009

Standard impossibility theorems on judgment aggregation over logically connected propositions either use a controversial systematicity condition or apply only to agendas of propositions with rich logical connections. Are there any serious impossibilities without these restrictions? We prove an impossibility theorem without requiring systematicity that applies to most standard agendas: Every judgment aggregation function (with rational inputs and outputs) satisfying a condition called unbiasedness is dictatorial (or effectively dictatorial if we remove one of the agenda conditions). Our agenda conditions are tight. Applied illustratively to (strict) preference aggregation represented in our model, the result implies that every unbiased social welfare function with universal domain is effectively dictatorial.

Keywords: judgment aggregation, logic, impossibility, May's neutrality

1 Introduction

We prove a new impossibility theorem on the aggregation of individual judgments (acceptance or rejection) on logically connected propositions into corresponding collective judgments. Due to the flexible notion of a proposition, the model of judgment aggregation can represent many realistic decision problems. For example, the propositions could be the following:

a : "We can afford a budget deficit."

$a \rightarrow b$: "If we can afford a budget deficit, then we should increase spending on education."

b: "We should increase spending on education."

The interest in judgment aggregation was sparked by the observation that majority voting on logically connected propositions does not guarantee rational (i.e., complete and consistent) collective judgments: the "discursive paradox" (Pettit 2001).

	a	$a \rightarrow b$	b
Individual 1	True	True	True
Individual 2	True	False	False
Individual 3	False	True	False
Majority	True	True	False

Table 1: A discursive paradox

[1]F. Dietrich, Dept. of Quant. Economics, University of Maastricht, P.O. Box 616, 6200 MD Maastricht, The Netherlands; C. List, Dept. of Government, London School of Economics, London WC2A 2AE, U.K. We thank Klaus Nehring and the anonymous referees for very detailed and helpful comments on this paper.

In our example, if individual judgments are as shown in Table 1, a majority accepts a, a majority accepts $a \to b$, and yet a majority rejects b. Although judgment aggregation has many similarities to preference aggregation in Condorcet's and Arrow's tradition, judgment aggregation generalizes preference aggregation[2] and faces additional complexities.

In particular, a basic fact about Arrowian preference aggregation is the following:

> If, for a given decision problem (specified by the set of alternatives to be ranked), majority voting generates irrational collective preferences for some profiles of rational individual preferences (which happens just in case there are at least three alternatives), then so does *any* preference aggregation function satisfying Arrow's conditions (universal domain, weak Pareto, independence of irrelevant alternatives, non-dictatorship).

Thus any decision problem susceptible to *Condorcet's paradox* is, more generally, susceptible to *Arrow's theorem*. No such fact holds for judgment aggregation. Instead we find the following:

> Even if, for a given decision problem (here specified by the agenda of propositions under consideration), majority voting generates irrational collective judgments for some profiles of rational individual judgments (which happens just in case the agenda has a minimal inconsistent subset of three or more propositions), other judgment aggregation functions may still guarantee collective rationality while satisfying the counterparts of Arrow's conditions.

Thus a decision problem susceptible to a *discursive paradox* need not be susceptible to the counterpart of Arrow's theorem. Like most example agendas in the literature, the agenda above is susceptible to a discursive paradox but not to an Arrow-style impossibility. Neither the size of the agenda nor the sizes of its minimal inconsistent subsets alone determine whether or not an Arrow-style impossibility applies. The logical interconnections between the propositions in the agenda matter in more complex ways. The recent literature on judgment aggregation has explored this complexity, which also constitutes the motivation for this paper.

List and Pettit (2002, 2004) formalized judgment aggregation and proved a first impossibility theorem, strengthened by Pauly and van Hees (2006), Dietrich (2007) and Dietrich and List (2007a), which holds for most standard agendas, but imposes a strong condition of systematicity on the aggregation function. Systematicity is the conjunction of an Arrow-inspired independence condition

[2] As illustrated below, preference aggregation can be formally represented as a special case of judgment aggregation by expressing preference relations as binary ranking propositions in predicate logic of the form xPy (List and Pettit 2004, Dietrich and List 2007).

(requiring propositionwise aggregation) and a global neutrality condition (requiring equal treatment of all propositions). Thus the price for the theorem's applicability to many agendas is the strength of its systematicity condition.

Given this problem, several authors have proved impossibility theorems in which systematicity is weakened to independence (Pauly and van Hees 2006, Dietrich 2006, Gärdenfors 2006, Nehring and Puppe 2008, van Hees 2007, Dietrich and List 2007a, Dokow and Holzman forthcoming, Mongin forthcoming). But these results exclude many agendas, notably the one in the example above and many other standard agendas. This is not because the theorems have not been *proved* for these agendas, but because they do not *hold* for them. Indeed, Dokow and Holzman (forthcoming), extending an earlier result by Nehring and Puppe (2002) in the related "property space" model, have identified the *weakest* agenda conditions for an Arrow-style impossibility with independence (i.e., a result where Arrow-inspired conditions imply dictatorship); and these agenda conditions are still rather strong, although several oligarchy results are known for weaker agenda conditions (Nehring and Puppe 2008, Dokow and Holzman forthcoming).[3]

Once we give up systematicity, are all serious impossibilities restricted to special agendas? Unfortunately not. We introduce a condition of *unbiased* aggregation, inspired by May's (1952) neutrality condition and much weaker by itself than systematicity. Unbiasedness requires an equal treatment of each proposition and its negation, but not of any two propositions. For most agendas, every unbiased judgment aggregation function (with rational inputs and outputs) is dictatorial or at least effectively dictatorial. The lesson is that unbiasedness, although in isolation considerably weaker than systematicity, turns out to have very similar implications as systematicity. A mathematically related earlier result is a theorem by Nehring and Puppe (2005) on strategy-proof social choice functions that are *neutral-within-issues*, on which we comment later.

Our result is of interest in light of May's classic characterization of majority voting (1952), in so far as it shows that, as soon as May's binary agenda $\{p, \neg p\}$

[3] Crucially, for an Arrow-style impossibility result with independence (where the relevant conditions imply dictatorship), the agenda must be *totally blocked*, i.e., it must be possible to link any two propositions by a path of 'conditional entailments'. Under the different, yet not fully general condition that the agenda is *truth-functional*, impossibility results without assuming systematicity have been proved (Nehring and Puppe 2008, Dokow and Holzman forthcoming), although these are oligarchy results rather than dictatorship results. These oligarchy results, in turn, may lead us to expect that a condition such as unbiasedness will again generate a dictatorship result. For a review, see also List and Puppe (2009). Agendas describing standard discursive-dilemma situations, say agendas containing premise propositions and a conclusion proposition (and their negations), are typically not totally blocked, and only some of them are truth-functional, depending on the relationship between the premises and the conclusion. They are truth-functional, in particular, in two cases: the case in which the conclusion proposition is fully determined (in its truth value) by the premise propositions, and the case in which some premise proposition is a material conditional. (We have left open whether the conditional $a \to b$ in our leading example is a material one, i.e., logically equivalent to $\neg a \vee b$, but arguably it is not, in which case the agenda is not truth-functional.)

is just slightly enriched by additional propositions, May's possibility turns into an impossibility even if May's monotonicity condition is dropped and anonymity is significantly weakened. Our result also applies to (strict) preference aggregation problems with three or more alternatives: every unbiased social welfare function with universal domain is effectively dictatorial. Further, our results can be applied to two related aggregation problems: the *belief merging* problem discussed in computer science (e.g., Konieczny and Pino-Perez 2002), and the aggregation of binary evaluations (e.g., Wilson 1975 and Dokow and Holzman forthcoming). Throughout this paper we adopt Dietrich's (2007) general logics framework, which extends the model in List and Pettit (2002, 2004).

2 Definitions

Let $N = \{1, 2, \ldots, n\}$ be a group of individuals ($n \geq 2$) required to make collective judgments on logically connected propositions.

Language. We consider a *language,* given by a non-empty set $\mathbf{L}$ of sentences (*propositions*) closed under negation: $p \in \mathbf{L}$ implies $\neg p \in \mathbf{L}$, where $\neg$ denotes "not". In addition to negation, the language may contain any other logical operators needed to express the decision problem, such as the classical operators $\wedge$ ("and"), $\vee$ ("or") and $\rightarrow$ ("if-then"), modal operators, subjunctive conditionals, etc. As usual in logic, every set $S \subseteq \mathbf{L}$ is either *consistent* or *inconsistent* (not both). Our results require some regularity axioms on the consistency notion (valid for many logics, classical or non-classical, propositional or predicate):[4]

C1 Pairs $\{p, \neg p\} \subseteq \mathbf{L}$ are inconsistent (*self-entailment*)
C2 Subsets of consistent sets $S \subseteq \mathbf{L}$ are consistent (*monotonicity*).
C3 $\emptyset$ is consistent, and each consistent set $S \subseteq \mathbf{L}$ has a consistent superset $T \subseteq \mathbf{L}$ containing a member of each pair $p, \neg p \in \mathbf{L}$ (*completability*).

For example, in a propositional logic, $\mathbf{L}$ contains propositions such as a, b, $a \rightarrow b$, $a \wedge b$, where $\{a, a \rightarrow b, \neg b\}$ is inconsistent, $\{a, \neg(a \wedge b)\}$ is consistent, etc. Various realistic decision problems can be represented in our model, including preference aggregation problems as illustrated below. Call a set $S \subseteq \mathbf{L}$ *minimal inconsistent* if it is inconsistent and every proper subset $T \subsetneq S$ is consistent. Call a proposition $p \in \mathbf{L}$ *contingent* if $\{p\}$ and $\{\neg p\}$ are consistent, a *contradiction* if $\{p\}$ is inconsistent, and a *tautology* if $\{\neg p\}$ is inconsistent.

Agenda. The *agenda* is a non-empty set $X \subseteq \mathbf{L}$ of propositions on which judgments are to be made, where X is a union of pairs $\{p, \neg p\}$ (with p not itself

[4]The conditions C1-C3 can be re-expressed in terms of (equivalent) conditions on the *entailment* relation in $\mathbf{L}$ (the conditions L1-L3 in Dietrich 2007, who uses the labels "I1-I3" for "C1-C3"). The reason is that a set $S \subseteq \mathbf{L}$ is consistent if and only if, for no $p \in \mathbf{L}$, S entails p and entails $\neg p$.

a negated proposition). If X is infinite, we require the logic to be *compact*: every inconsistent set has a finite inconsistent subset (this holds for many logics). In the example above, the agenda is $X = \{a, \neg a, b, \neg b, a \to b, \neg(a \to b)\}$ in a propositional logic. Notationally, double negations cancel each other out (i.e., $\neg\neg p$ stands for p).[5] A *subagenda* of the agenda X is a subset $Y \subseteq X$ that is itself an agenda, i.e., non-empty and a union of pairs $\{p, \neg p\}$.

Judgment sets. Each individual i's *judgment set* is a subset $A_i \subseteq X$, where $p \in A_i$ means that individual i accepts proposition p. A judgment set A_i is *rational* if it is (i) *consistent* and (ii) *complete* in the sense that, for every proposition $p \in X$, $p \in A_i$ or $\neg p \in A_i$. Let $\mathbf{U}$ be the set of all rational judgment sets.[6] A *profile* is an n-tuple $(A_1, \ldots, A_n)$ of individual judgment sets.

Aggregation functions. A *(judgment) aggregation function* is a function F that assigns to each profile $(A_1, \ldots, A_n)$ in some domain a collective judgment set $F(A_1, \ldots, A_n) = A \subseteq X$, where $p \in A$ means that the group accepts proposition p. An example is *majority voting*: here $F(A_1, ..., A_n) = \{p \in X : |\{i \in N : p \in A_i\}| > |\{i \in N : p \notin A_i\}|\}$ for all profiles $(A_1, ..., A_n)$ in the domain. All our results assume that the aggregation function is a function $F : \mathbf{U}^n \to \mathbf{U}$, that is:

- F accepts as inputs all possible profiles $(A_1, ..., A_n)$ of rational individual judgment sets ("universal domain"), and
- F generates as outputs rational collective judgment sets ("collective rationality").

This is a demanding rationality requirement on individuals and on the collective; but it is a standard requirement.[7] Call aggregation function F *unanimity-preserving* if $F(A, ..., A) = A$ for every unanimous profile $(A, ..., A)$ in the domain of F. Call an individual i a *dictator* (and F *dictatorial*) if $F(A_1, ..., A_n) = A_i$ for all profiles $(A_1, ..., A_n)$ in the domain of F. For any subset $Y \subseteq X$, call individual i a *dictator on* Y (and F *dictatorial on* Y) if $F(A_1, ..., A_n) \cap Y = A_i \cap Y$ for all profiles $(A_1, ..., A_n)$ in the domain of F.

3 Unbiasedness

Our central condition on an aggregation function is inspired by May's (1952) condition of neutrality:

[5] Hereafter, when we write $\neg p$ and p is already of the form $\neg q$, we mean q (rather than $\neg\neg q$).

[6] Some readers may prefer the notation $\mathbf{U}_X$ for the set of rational judgment sets on the agenda X, but since X is assumed to be fixed, there is no risk of ambiguity even without the subscript.

[7] Relaxations of full rationality within the logic-based model of judgment aggregation are considered, for example, in List and Pettit (2002), Gärdenfors (2006), Dietrich and List (2007b), and Dietrich and List (2008).

Unbiasedness. For any proposition $p \in X$ and profiles $(A_1, \ldots, A_n)$, $(A_1^*, \ldots, A_n^*)$ in the domain of F, if [for all individuals i, $p \in A_i$ if and only if $\neg p \in A_i^*$] then [$p \in F(A_1, \ldots, A_n)$ if and only if $\neg p \in F(A_1^*, \ldots, A_n^*)$].

Unbiasedness requires an equal treatment of each proposition $p \in X$ and its negation $\neg p$, regardless of other judgments. If we decompose an aggregation problem into multiple decisions between proposition-negation pairs, it can be interpreted as the application of May's neutrality condition (1952) to each such pair. Unbiasedness is also related to Nehring and Puppe's (2005) neutrality-within-issues.[8] Unbiasedness is considerably weaker than List and Pettit's (2002) condition of systematicity, which requires an aggregation function to be neutral between *any* two propositions $p, q \in X$:

Systematicity. For any propositions $p, q \in X$ and profiles $(A_1, \ldots, A_n)$, $(A_1^*, \ldots, A_n^*)$ in the domain of F, if [for all individuals i, $p \in A_i$ if and only if $q \in A_i^*$] then [$p \in F(A_1, \ldots, A_n)$ if and only if $q \in F(A_1^*, \ldots, A_n^*)$].

While systematicity permits all uniform propositionwise decision methods, such as majority voting, symmetrical supermajority voting, dictatorships or inverse dictatorships, unbiasedness also permits aggregation functions that apply different decision criteria to different propositions but the same criterion to each proposition $p \in X$ and its negation $\neg p$. For example, on some pairs $p, \neg p \in X$ one can apply majority voting, on others weighted majority voting, on yet others majority voting within some subgroup or dictatorships or erratic decision methods like inverse dictatorships, minority voting, or accepting a proposition if and only if it is supported by an odd number of individuals. Unbiasedness also differs from a global neutrality condition based on a permutation $\pi : X \to X$ of the agenda (e.g., van Hees 2007). It is by itself logically independent from independence, but implies independence under universal domain and collective rationality (see Lemma 1 below). Systematicity, by contrast, implies both independence and global neutrality.

4 A first impossibility of unbiased aggregation

Our theorems use two weak agenda conditions. Their precise form is justified by the tighness (necessity) of the conditions in our theorems, as shown below.

First, call agenda X *non-separable* if it cannot be partitioned into two logically independent subagendas X_1 and X_2, each containing at least one contingent proposition (where X_1 and X_2 are *logically independent* if $B_1 \cup B_2$ is consistent for any consistent subsets $B_1 \subseteq X_1$ and $B_2 \subseteq X_2$). Informally, non-separability requires that the decision problem cannot be split into two logically independent decision problems – a plausible condition in practice. The agenda in our

[8] A neutral-within-issues and strategy-proof social choice function induces an unbiased and monotonic judgment aggregation function.

example above and many others are non-separable. But if we extend our example agenda by adding a new pair $c, \neg c$, where c is an atomic proposition, the new agenda is separable, namely into the old agenda and the binary agenda $\{c, \neg c\}$.

Second, call agenda X *minimally connected* if it has these two properties:[9]

(i) It has a minimal inconsistent subset Y of size at least three (the *non-simplicity* condition).

(ii) It has a minimal inconsistent subset Y such that $(Y \backslash Z) \cup \{\neg p : p \in Z\}$ is consistent for some set $Z \subseteq Y$ of even size (the *even-number negation* condition in Dietrich 2007 and Dietrich and List 2007a; equivalent, for finite X, to Dokow and Holzman's forthcoming algebraic *non-affineness* condition).

All standard example agendas in the judgment aggregation literature are minimally connected, including our example agenda above (take $Y = \{a, a \rightarrow b, \neg b\}$ in (i) and (ii) and $Z = \{a, \neg b\}$) and agendas representing preference aggregation problems with three or more alternatives (as discussed below). The notorious exception is $X = \{a, \neg a, b, \neg b, a \leftrightarrow b, \neg(a \leftrightarrow b)\}$, where $\leftrightarrow$ is taken to be a material biconditional (i.e., $a \leftrightarrow b$ is logically equivalent to $(a \wedge b) \vee (\neg a \wedge \neg b)$): (ii) fails since Z does not exist. However, even this agenda becomes minimally connected once $\leftrightarrow$ is taken to be a subjunctive biconditional, which is arguably more realistic (Dietrich forthcoming): take $Y = \{a, \neg b, a \leftrightarrow b\}$ in (i) and (ii) and $Z = \{a, a \leftrightarrow b\}$. Also the agenda X assumed by List and Pettit (2002) – containing distinct atomic propositions a, b and their conjunction $a \wedge b$ – is minimally connected: take $Y = \{a, b, \neg(a \wedge b)\}$ in (i) and (ii), and $Z = \{a, b\}$.

Theorem 1 *For (and only for) a non-separable minimally connected agenda, every unbiased unanimity-preserving aggregation function $F : \mathbf{U}^n \rightarrow \mathbf{U}$ is dictatorial.*

In Theorem 1, and also in Theorems 2 and 3 below, the necessity of the agenda conditions (the "only for") assumes a group of size $n \geq 3$. Theorem 1 immediately implies the following corollary.

Corollary 1 *For any agenda X (without restriction), every unbiased unanimity-preserving aggregation function $F : \mathbf{U}^n \rightarrow \mathbf{U}$ is dictatorial on each non-separable minimally connected subagenda $Y \subseteq X$.*

By Theorem 1, for most agendas, all unbiased aggregation functions (with rational inputs and outputs) are degenerate: they are dictatorial or overrule unanimities. Thus non-degenerate aggregation functions must favour some propositions over their negations.

This theorem (and part of its proof) is related to Nehring and Puppe's (2005) results on strategy-proof social choice functions that are neutral-within-issues.

[9] The sets Y may or may not differ between parts (i) and (ii).

Translated into our framework, their results imply that, under further weakened agenda conditions (namely dropping property (ii) of minimal connectedness), every aggregation function $F : \mathbf{U}^n \to \mathbf{U}$ that is unbiased *and monotonic* (hence also unanimity-preserving) is dictatorial. So our result uses weaker aggregation conditions but stronger agenda conditions. (This situation is somewhat similar to the standard cases in which independence and systematicity are required: there, too, the additional agenda condition (ii) is needed to obtain a dictatorship result without monotonicity. For an overview, see List and Puppe 2009.) Similar remarks apply also to our later theorems, which do not even require a unanimity condition.

To show the necessity part of Theorem 1 (the "only for"), which holds for group size $n \geq 3$, we simply specify counterexamples ("possibilities").

First, suppose agenda X is separable, say into subagendas X_1 and X_2. Then an unbiased and unanimity-preserving but not dictatorial aggregation function $F : \mathbf{U}^n \to \mathbf{U}$ can be defined by

$$F(A_1, ..., A_n) = F_1(A_1 \cap X_1, ..., A_n \cap X_1) \cup F_2(A_1 \cap X_2, ..., A_n \cap X_2),$$

where F_1 and F_2 are dictatorships for the agendas X_1 and X_2, respectively, with a *different* dictator each time.

Next suppose X violates part (i) of minimal connectedness; so all minimal inconsistent sets contain at most two propositions. Then majority voting among any non-singleton odd subset of the individuals, for simplicity the first three, given on the universal domain by

$$F(A_1, ..., A_n) = \{p \in X : \text{ a majority of } A_1, A_2, A_3 \text{ contains } p\},$$

generates consistent (and complete) judgment sets, hence defines an unbiased unanimity-preserving aggregation function $F : \mathbf{U}^n \to \mathbf{U}$ that is not dictatorial.

Finally, suppose X violates part (ii) of minimal connectedness. Then Dokow and Holzman's (forthcoming) *parity function* among the first three individuals, defined on the universal domain by

$$F(A_1, ..., A_n) = \{p \in X : \text{an odd number of } A_1, A_2, A_3 \text{ contains } p\},$$

generates consistent (and complete) judgment sets (see Dokow and Holzman forthcoming; for a non-algebraic proof see Dietrich 2007). So F defines an unbiased unanimity-preserving aggregation function $F : \mathbf{U}^n \to \mathbf{U}$ that is not dictatorial. Crucially, this function is non-monotonic, and, by Nehring and Puppe's result (2005), there cannot exist a monotonic such function, assuming that all the other conditions of the theorem are met.

The sufficiency proof in Theorem 1 rests on some lemmas. We begin by establishing the standard property of "independence" or "propositionwise aggregation".

Independence. For any proposition $p \in X$ and profiles $(A_1, \ldots, A_n), (A_1^*, \ldots, A_n^*)$ in the domain of F, if [for all individuals i, $p \in A_i$ if and only if $p \in A_i^*$] then [$p \in F(A_1, \ldots, A_n)$ if and only if $p \in F(A_1^*, \ldots, A_n^*)$].

Lemma 1 *Every unbiased aggregation function $F : \mathbf{U}^n \to \mathbf{U}$ is independent.*[10]

Proof. Let F be as specified. Consider any $p \in X$ and any profiles $(A_1, ..., A_n), (A_1^*, ..., A_n^*) \in \mathbf{U}^n$ in which the same set of individuals $C \subseteq N$ accepts p. We show that $p \in F(A_1, ..., A_n)$ if and only if $p \in F(A_1^*, ..., A_n^*)$, as required by independence. By collective rationality, if p is a tautology, p is contained in both $F(A_1, ..., A_n)$ and $F(A_1^*, ..., A_n^*)$; if p is a contradiction, p is contained in neither of $F(A_1, ..., A_n)$ and $F(A_1^*, ..., A_n^*)$. Suppose p is contingent. Then $\neg p$ is also contingent. There exists a profile $(A_1', ..., A_n') \in \mathbf{U}^n$ such that exactly the individuals in C accept $\neg p$. By unbiasedness, $p \in F(A_1, ..., A_n)$ is equivalent to $\neg p \in F(A_1', ..., A_n')$, which, again by unbiasedness, is equivalent to $p \in F(A_1^*, ..., A_n^*)$. ■

Using Lemma 1, we can easily see why Corollary 1 follows from Theorem 1. Let $F : \mathbf{U}^n \to \mathbf{U}$ be unbiased and unanimity-preserving on some agenda X (without restriction). Consider a non-separable minimally connected subagenda $Y \subseteq X$. As F is independent (by Lemma 1), F induces a unique aggregation function F^* for this subagenda: F^* has universal domain relative to agenda Y – call this $\mathbf{U}_Y^n$ – and is for all $(A_1^*, ..., A_n^*) \in \mathbf{U}_Y^n$ given by

$$F^*(A_1^*, ..., A_n^*) = F(A_1, ..., A_n) \cap Y, \tag{1}$$

where $(A_1, ..., A_n)$ is a profile in the original domain $\mathbf{U}^n$ with $A_i \cap Y = A_i^*$ for all i; by independence it does not matter which such profile $(A_1, ..., A_n)$ is chosen. The function F^* inherits from F the properties of unbiasedness, unanimity-preservation and collective rationality. So, by Theorem 1, F^* is dictatorial. Hence, by (1) and independence, F is dictatorial on Y, as required. ■

For the next lemma, call propositions $p, q \in X$ *connected* (in X) if they are conditionally dependent: there exist $p^* \in \{p, \neg p\}$ and $q^* \in \{q, \neg q\}$ such that $\{p^*, q^*\} \cup Y$ is inconsistent for some $Y \subseteq X$ consistent with p^* and with q^*. And call coalition $C \subseteq N$ *winning for* $p \in X$ (under F) if $p \in F(A_1, ..., A_n)$ for every profile $(A_1, ..., A_n)$ in the domain with $\{i : p \in A_i\} = C$. If an aggregation function F is independent, it is uniquely determined by the family $(\mathcal{C}_p)_{p \in X}$, where $\mathcal{C}_p$ is the set of coalitions winning for $p \in X$. Specifically, for all profiles $(A_1, ..., A_n)$ in the domain,

$$F(A_1, ..., A_n) = \{p \in X : \{i : p \in A_i\} \in \mathcal{C}_p\}.$$

Lemma 2 *Suppose $F : \mathbf{U}^n \to \mathbf{U}$ is unbiased. For any $p \in X$, let $\mathcal{C}_p$ be the set of coalitions winning for p. Then:*
(a) *If $p \in X$ is contingent, $\mathcal{C}_p = \mathcal{C}_{\neg p}$ and $C \in \mathcal{C}_p \Leftrightarrow N \backslash C \notin \mathcal{C}_p$.*
(b) *If $p, q \in X$ are connected and F is unanimity-preserving, $\mathcal{C}_p = \mathcal{C}_q$.*

[10] Unbiasedness does not imply independence without assuming universal domain and collective rationality.

Proof. Let $F : \mathbf{U}^n \rightarrow \mathbf{U}$ be unbiased. By Lemma 1, F is independent.

(a) Let $p \in X$ be contingent. To show $\mathcal{C}_p = \mathcal{C}_{\neg p}$, consider any $C \subseteq N$, and let us prove that $C \in \mathcal{C}_p$ if and only if $C \in \mathcal{C}_{\neg p}$. As p is contingent, there exist profiles $(A_1, ..., A_n), (A_1^*, ..., A_n^*) \in \mathbf{U}^n$ such that $C = \{i : p \in A_i\} = \{i : \neg p \in A_i^*\}$. By unbiasedness, $p \in F(A_1, ..., A_n)$ if and only if $\neg p \in F(A_1^*, ..., A_n^*)$, whence $\{i : p \in A_i\} \in \mathcal{C}_p$ if and only if $\{i : \neg p \in A_i^*\} \in \mathcal{C}_{\neg p}$, i.e., $C \in \mathcal{C}_p$ if and only if $C \in \mathcal{C}_{\neg p}$.

To prove the second part of (a), let $C \subseteq N$ again. As p is contingent, there exists a profile $(A_1, ..., A_n) \in \mathbf{U}^n$ such that $C = \{i : p \in A_i\}$, hence $N \backslash C = \{i : \neg p \in A_i\}$. Now $C \in \mathcal{C}_p$ is equivalent to $p \in F(A_1, ..., A_n)$, hence to $\neg p \notin F(A_1, ..., A_n)$, hence to $N \backslash C \notin \mathcal{C}_{\neg p}$, hence to $N \backslash C \notin \mathcal{C}_p$, as shown above.

(b) Suppose $p, q \in X$ are connected and F is unanimity-preserving. Then there exist $v \in \{p, \neg p\}$ and $w \in \{q, \neg q\}$ and $Y \subseteq X$ such that (i) each of $\{v\} \cup Y$ and $\{w\} \cup Y$ is consistent, and (ii) $\{v, w\} \cup Y$ is inconsistent. It follows that (iii) each of $\{v, \neg w\} \cup Y$ and $\{\neg v, w\} \cup Y$ is consistent. By (iii), v and w are contingent. So, by part (a), it is sufficient to show that $\mathcal{C}_v = \mathcal{C}_w$. We only show that $\mathcal{C}_v \subseteq \mathcal{C}_w$, as the converse inclusion is analogous. Suppose $C \in \mathcal{C}_v$. By (iii) and universal domain there exists a profile $(A_1, ..., A_n) \in \mathbf{U}^n$ such that $\{v, \neg w\} \cup Y \subseteq A_i$ for all $i \in C$ and $\{\neg v, w\} \cup Y \subseteq A_i$ for all $i \in N \backslash C$. We have $v \in F(A_1, ..., A_n)$ as $\mathcal{C}_v$ contains C, and $Y \subseteq F(A_1, ..., A_n)$ as each $\mathcal{C}_y$, $y \in Y$, contains N by unanimity-preservation. Because of $\{v\} \cup Y \subseteq F(A_1, ..., A_n)$ and (ii), $w \notin F(A_1, ..., A_n)$. So $N \backslash C \notin \mathcal{C}_w$, and hence $C \in \mathcal{C}_w$ by part (a), as required. ■

We now show that non-separability of the agenda X is equivalent to another structural property. Call propositions $p, q \in X$ *indirectly connected* if there exist $p_1, ..., p_k \in X$ with $p_1 = p$ and $p_k = q$ such that any two neighbours p_t, p_{t+1} are connected (as defined above). And call agenda X *indirectly connected* if any two contingent propositions $p, q \in X$ are indirectly connected.

Lemma 3 *An agenda is non-separable if and only if it is indirectly connected.*

Proof. We may assume without loss of generality that all $p \in X$ are contingent, because if X also has non-contingent members then X is indirectly connected if and only if the subagenda $\{p \in X : p \text{ is contingent}\}$ is indirectly connected, and X is non-separable if and only if $\{p \in X : p \text{ is contingent}\}$ is non-separable (if X has *only* non-contingent members, the claim is trivial).

First, assume X is separable. Then there is a partition of X into logically independent subagendas X_1, X_2. Consider any $p \in X_1$ and $q \in X_2$. We show that p and q are not indirectly connected. Suppose for a contradiction that $p_1, ..., p_m \in X$ $(m \geq 1)$ are such that $p = p_1$, $q = p_m$, and p_t and p_{t+1} are connected for any $t \in \{1, ..., m-1\}$. As $p_1 \in X_1$ and $p_m \in X_2$, there must be a $t \in \{1, ..., m-1\}$ such that $p_t \in X_1$ and $p_{t+1} \in X_2$. As p_t and p_{t+1} are connected, there are $p_t^* \in \{p_t, \neg p_t\}$, $p_{t+1}^* \in \{p_{t+1}, \neg p_{t+1}\}$ and $Y \subseteq X$ such that (i) $\{p_t^*, p_{t+1}^*\} \cup Y$ is inconsistent and (ii) each of $\{p_t^*\} \cup Y$ and $\{p_{t+1}^*\} \cup Y$

is consistent. By (ii), each of the sets $B_1 := (\{p_t^*\} \cup Y) \cap X_1$ and $B_1 := (\{p_{t+1}^*\} \cup Y) \cap X_2$ is consistent. So $B_1 \cup B_2$ is consistent, as X_1 and X_2 are logically independent. But

$$\begin{aligned} B_1 \cup B_2 &= [(\{p_t^*\} \cup Y) \cap X_1] \cup [(\{p_{t+1}^*\} \cup Y) \cap X_2] \\ &= [(\{p_t^*, p_{t+1}^*\} \cup Y) \cap X_1] \cup [(\{p_t^*, p_{t+1}^*\} \cup Y) \cap X_2] \\ &= (\{p_t^*, p_{t+1}^*\} \cup Y) \cap [X_1 \cup X_2] = \{p_t^*, p_{t+1}^*\} \cup Y, \end{aligned}$$

which is inconsistent by (ii), a contradiction.

Secondly, let X be not indirectly connected. We show that X is separable. By assumption, there exist $p, q \in X$ that are not indirectly connected. Define $X_1 := \{r \in X : p$ and r are indirectly connected$\}$ and $X_2 := X \backslash X_1$. Since p is indirectly connected to itself (as p is contingent), $p \in X_1$. Further, $q \in X_2$. So each of X_1 and X_2 is non-empty. Moreover, each of X_1 and X_2 is closed under negation (like X). If follows that X_1 and X_2 are subagendas of X.

We complete the proof by showing that X_1 and X_2 are logically independent. Suppose for a contradiction that $B_1 \subseteq X_1$ and $B_2 \subseteq X_2$ are each consistent but that $B_1 \cup B_2$ is inconsistent. As X is finite or the logic compact, there exists a minimal inconsistent subset $B \subseteq B_1 \cup B_2$. We have neither $B \subseteq B_1$ nor $B \subseteq B_2$, since otherwise B would be consistent. So there exist $r \in B \cap X_1$ and $s \in B \cap X_2$. Propositions r and s are connected, because, putting $Y := B \backslash \{r, s\}$, $\{r, s\} \cup Y = B$ is inconsistent, but each of $\{r\} \cup Y = B \backslash \{s\}$ and $\{s\} \cup Y = B \backslash \{r\}$ is consistent by B's minimal inconsistency. This is a contradiction. ■

Essentially by combining Lemma 3 with part (b) of Lemma 2, we now deduce *systematicity* (which brings us into the terrain of known results of the literature).

Lemma 4 *For a non-separable agenda X, every unbiased unanimity-preserving aggregation function $F : \mathbf{U}^n \to \mathbf{U}$ is systematic.*

Proof. Let X and F be as specified. Consider propositions $p, q \in X$ and profiles $(A_1, \ldots, A_n), (A_1^*, \ldots, A_n^*) \in \mathbf{U}^n$, such that (*) for all individuals i, $p \in A_i \Leftrightarrow q \in A_i^*$. We show that (**) $p \in F(A_1, ..., A_n) \Leftrightarrow q \in F(A_1^*, ..., A_n^*)$. If p and q are both contingent, then they are indirectly connected by Lemma 3, hence have the same set of winning coalitions by iterated applications of part (b) of Lemma 2; so, using (*), we obtain (**). If p or q is a tautology then, by individual rationality and (*), all A_i contain p and all A_i^* contain q; so by unanimity-preservation (and independence) $p \in F(A_1, ..., A_n)$ and $q \in F(A_1^*, ..., A_n^*)$, implying (**). Finally, if p or q is a contradiction then, by individual rationality and (*), all A_i contain $\neg p$ and all A_i^* contain $\neg q$; so by unanimity-preservation (and independence) $\neg p \in F(A_1, ..., A_n)$ and $\neg q \in F(A_1^*, ..., A_n^*)$, which again implies (**) (by collective rationality). ■

Lemma 4 together with the following known result implies the sufficiency part of Theorem 1, hence completes the proof.

Lemma 5 *(Dietrich and List 2007a) For a minimally connected agenda, every systematic unanimity-preserving aggregation function* $F : \mathbf{U}^n \to \mathbf{U}$ *is dictatorial.*

To make the present argument self-contained, we sketch the proof of this result. Let the agenda X and the function F be as specified. By systematicity, there is a set $\mathcal{C}$ of ("winning") coalitions $C \subseteq N$ such that

$$F(A_1, ..., A_n) = \{p \in X : \{i : p \in A_i\} \in \mathcal{C}\} \text{ for all } (A_1, ..., A_n) \in \mathbf{U}^n.$$

Steps (a)-(e) below show the existence of a dictator i, i.e., $\mathcal{C} = \{C \subseteq N : i \in C\}$.

(a) $N \in \mathcal{C}$, by unanimity-preservation.

(b) $C \in \mathcal{C} \Leftrightarrow N \backslash C \notin \mathcal{C}$, as exactly one member of any $p, \neg p \in X$ is accepted.

(c) $[C \in \mathcal{C}\ \&\ C \subseteq C^* \subseteq N] \Rightarrow C^* \in \mathcal{C}$. Suppose $C \in \mathcal{C}\ \&\ C \subseteq C^* \subseteq N$. Let $Y \subseteq X$ be as in part (ii) of "minimal connectedness". For illustrative purposes, let Y be binary, say $Y = \{p, q\}$ (see Dietrich and List 2007a for the general case). Then $\{p, q\}$ is inconsistent, but $\{p, \neg q\}, \{\neg p, q\}, \{\neg p, \neg q\}$ are each consistent. Consider a profile $(A_1, ..., A_n) \in \mathbf{U}^n$ such that $\{p, \neg q\} \subseteq A_i$ if $i \in C$, $\{\neg p, \neg q\} \subseteq A_i$ if $i \in C^* \backslash C$, and $\{\neg p, q\} \subseteq A_i$ otherwise. As $C \in \mathcal{C}$, $p \in F(A_1, ..., A_n)$. So, by $\{p, q\}$'s inconsistency, $\neg q \in F(A_1, ..., A_n)$. Hence $C^* \in \mathcal{C}$.

(d) $C, C^* \in \mathcal{C} \Rightarrow C \cap C^* \in \mathcal{C}$. Let $C, C^* \in \mathcal{C}$. Let $Y \subseteq X$ be as in part (i) of "minimal connectedness". Consider distinct $p, q, r \in Y$. Consider a profile in which each individual accepts all $s \in Y \backslash \{p, q, r\}$, and rejects exactly one of p, q, r: all $i \in C \cap C^*$ reject p, all $i \in C^* \backslash C$ reject q, and all others reject r. By (a), all $s \in Y \backslash \{p, q, r\}$ are collectively accepted. As exactly all $i \in C^*$ ($\in \mathcal{C}$) accept r, the collective accepts r. As at least all $i \in C$ ($\in \mathcal{C}$) accept q, the collective accepts q by (c). So, as Y is inconsistent, the collective accepts $\neg p$. Hence $C \cap C^* \in \mathcal{C}$.

(e) $\mathcal{C} = \{C \subseteq N : i \in C\}$ for some individual (dictator) i. By (d), $\tilde{C} := \cap_{C \in \mathcal{C}} C \in \mathcal{C}$. So $\tilde{C} \neq \emptyset$ (as $\emptyset \notin \mathcal{C}$ by (a) and (b)). $\tilde{C}$ must be singleton, say $\tilde{C} = \{i\}$: otherwise $\tilde{C}$ could be partitioned into two non-empty coalitions C, C', where $N \backslash C, N \backslash C' \in \mathcal{C}$ by (b), hence $N \backslash \tilde{C} = (N \backslash C) \cap (N \backslash C') \in \mathcal{C}$ by (d), and so $\tilde{C} \notin \mathcal{C}$ by (b), a contradiction. By $\{i\} \in \mathcal{C}$ and (c), $\{C \subseteq N : i \in C\} \subseteq \mathcal{C}$. This inclusion also implies the converse inclusion by (b). ■

5 Two impossibilities without the unanimity condition

To strengthen the impossibility of unbiased aggregation further, we now show that even without requiring unanimity-preservation we still run into impossibility. We prove two theorems, both of which merely assume the aggregation function $F : \mathbf{U}^n \to \mathbf{U}$ to be unbiased. In the first result, we strengthen Theorem 1's agenda condition in such a way that unanimity-preservation can be *derived*; so Theorem 1 applies and implies dictatorship. In the second result, we keep Theorem 1's mild agenda condition, and show that an equally undesirable variant of dictatorship follows: an "effective dictatorship".

Given that we now only require the aggregation function $F : \mathbf{U}^n \to \mathbf{U}$ to be unbiased, it may seem that at least some such F can overrule at least some unanimous judgments. Surprisingly, for many agendas this is not the case. First assume the agenda X is *asymmetric*: there exists a consistent set $A \subseteq X$ such that $\{\neg p : p \in A\}$ is inconsistent (Dietrich 2007, Dietrich and List 2007a). Then F cannot overrule unanimous judgments for *every* $p \in A$: otherwise, if all individuals accept all $p \in A$, the collective accepts all $\neg p$ (with $p \in A$), a collective inconsistency. But F may still overrule unanimous judgments on some $p \in X \backslash A$. To prevent any overruling of unanimity, the agenda must be *locally asymmetric*: for every subagenda $Y \subseteq X$, there exists a consistent set $A \subseteq X$ that becomes inconsistent by negating those propositions that are in Y (i.e., $(A \backslash Y) \cup \{\neg p : p \in A \cap Y\}$ is inconsistent). (Local asymmetry is a stronger condition than asymmetry: by taking $Y = X$, the former implies the latter.) Examples are discussed shortly.

Lemma 6 *For (and only) for a locally asymmetric agenda, every unbiased aggregation function $F : \mathbf{U}^n \to \mathbf{U}$ is unanimity-preserving.*

Proof. First suppose $F : \mathbf{U}^n \to \mathbf{U}$ is unbiased and *not* unanimity-preserving. Then $Z := \{p \in X : N \text{ is not a winning coalition for } p\}$ is non-empty. Z is closed under negation: if $p \in Z$ then p is contingent, and hence $\neg p \in Z$ by part (a) of Lemma 2. So Z is a subagenda. To show that X is *not* locally asymmetric, we show that, for every consistent set $A \subseteq X$, also $(A \backslash Z) \cup \{\neg p : p \in A \cap Z\}$ is consistent. It obviously suffices to show this for the case that A is *complete*. So let $A \subseteq X$ be complete and consistent, i.e., $A \in \mathbf{U}$. By definition of Z, $F(A, ..., A)$ contains all $p \in A \backslash Z$ and all $\neg p$ with $p \in A \cap Z$, the latter because by part (a) of Lemma 2 the empty coalition is winning for $\neg p$. So $F(A, ..., A)$ contains all members of the set $(A \backslash Z) \cup \{\neg p : p \in A \cap Z\}$. Hence this set is consistent, as required.

Conversely, suppose X is not locally asymmetric. Then there is a subagenda $Z \subseteq X$ such that (*) for all consistent sets $A \subseteq X$, also $(A \backslash Z) \cup \{\neg p : p \in A \cap Z\}$ consistent. Let i be an individual and F the aggregation function with universal domain $\mathbf{U}^n$ that makes i a dictator on $X \backslash Z$ and an inverse dictator on Z; that is, for all $(A_1, ..., A_n) \in \mathbf{U}^n$,

$$F(A_1, ..., A_n) = (A_i \backslash Z) \cup \{\neg p : p \in A_i \cap Z\}.$$

By construction, F is unbiased, because i is a dictator on p if and only if i is a dictator on $\neg p$, and generates complete judgment sets. It also generates consistent ones by (*); hence $F : \mathbf{U}^n \to \mathbf{U}$. But, as required, F is not unanimity-preserving, because for all $(A, ..., A) \in \mathbf{U}^n$ we have $F(A, ..., A) = (A \backslash Z) \cup \{\neg p : p \in A \cap Z\} \neq A$. ■

This lemma together with Theorem 1 implies our next theorem.

Theorem 2 *For (and only for) a non-separable, minimally connected and locally asymmetric agenda, every unbiased aggregation function* $F : \mathbf{U}^n \to \mathbf{U}$ *is dictatorial.*

As in the case of Theorem 1, a corollary follows (using Lemma 1):

Corollary 2 *For any agenda* X *(without restriction), every unbiased aggregation function* $F : \mathbf{U}^n \to \mathbf{U}$ *is dictatorial on each non-separable, minimally connected and locally asymmetric subagenda* $Y \subseteq X$.

Theorem 2 applies to our example agenda $X = \{a, \neg a, b, \neg b, a \to b, \neg(a \to b)\}$, which is locally asymmetric, as we now check by going through the subagendas $Y \subseteq X$. For $Y = \{a, \neg a\}$, let $A = \{\neg a, a \to b, \neg b\}$ (as A is consistent but $\{a, a \to b, \neg b\}$ is not). For $Y = \{b, \neg b\}$, let $A = \{a, a \to b, b\}$. In general, for any subagenda Y, let $A \subseteq X$ be the unique set for which $(A \backslash Y) \cup \{\neg p : p \in A \cap Y\} = \{a, a \to b, \neg b\}$, i.e., let A be the set arising from $\{a, a \to b, \neg b\}$ by negating all $p \in A \cap Y$. The so-constructed set A is indeed consistent since $\{a, a \to b, \neg b\}$ becomes consistent after negating one, two or all of its member(s), assuming here that "$\to$" is a *subjunctive* implication (see Dietrich forthcoming). If, less realistically, "$\to$" is a *material* implication (i.e., $a \to b$ is equivalent to $\neg a \vee b$), then this way to construct A does not work for all subagendas Y,[11] but other constructions work.

But Theorem 2 does *not* apply to the strict preference aggregation problem discussed below: this agenda is not locally asymmetric, in fact not even asymmetric *simpliciter* because inverting (negating) rational strict preferences yields rational strict preferences. This illustrates that the strong impossibility of Theorem 2 ("unbiasedness implies dictatorship") applies to fewer agendas than the impossibilities of Theorem 1 above and Theorem 3 below.

Now we turn to our last impossibility result, which requires *neither* the unanimity condition of Theorem 1 *nor* the extra agenda condition of Theorem 2.

For any subset $Y \subseteq X$, call individual i an *inverse dictator on* Y if $F(A_1, ..., A_n) \cap Y = Y \backslash A_i$ for all profiles $(A_1, ..., A_n)$ in the domain. Call individual i an *effective dictator* (and F *effectively dictatorial*) if there is a partition of X into subsets X_+, X_- (each one possibly empty) such that individual i is a dictator on X_+ and an inverse dictator on X_-; that is, $F(A_1, ..., A_n) = (X_+ \cap A_i) \cup (X_- \backslash A_i)$ for all profiles $(A_1, ..., A_n)$ in the domain. For any subset $Y \subseteq X$, call individual i an *effective dictator on* Y (and F *effectively dictatorial on* Y) if there is a partition of Y into subsets Y_+, Y_- (each one possibly empty)

[11] Because $\{a, a \to b, \neg b\}$ stays inconsistent after negating $a \to b$ as well as one or both of a and $\neg b$ (e.g., $\{a, \neg(a \to b), b\}$ is inconsistent as the negated material implication $\neg(a \to b)$ is equivalent to $a \wedge \neg b$). Under the subjunctive interpretation of "$\to$", by contrast, the set $\{a, a \to b, \neg b\}$ becomes consistent as soon as at least one of its members is negated.

such that individual i is a dictator on Y_+ and an inverse dictator on Y_-; that is, $F(A_1, ..., A_n) \cap Y = (Y_+ \cap A_i) \cup (Y_- \backslash A_i)$ for all profiles $(A_1, ..., A_n)$ in the domain.

Theorem 3 *For (and only for) a non-separable minimally connected agenda, every unbiased aggregation function* $F : \mathbf{U}^n \to \mathbf{U}$ *is effectively dictatorial.*

Again, a corollary follows (using Lemma 1):

Corollary 3 *For any agenda X (without restriction), every unbiased aggregation function* $F : \mathbf{U}^n \to \mathbf{U}$ *is effectively dictatorial on each non-separable minimally connected subagenda* $Y \subseteq X$*.*

Note a further corollary. Call aggregation function F *anonymous* if, for any profiles $(A_1, \ldots, A_n)$ and $(A_{\sigma(1)}, ..., A_{\sigma(n)})$ in the domain, where $\sigma : N \to N$ is a permutation, $F(A_1, \ldots, A_n) = F(A_{\sigma(1)}, \ldots, A_{\sigma(n)})$.

Corollary 4 *If the agenda has a non-separable and minimally connected subagenda (possibly the entire agenda itself), there exists no anonymous unbiased aggregation function* $F : \mathbf{U}^n \to \mathbf{U}$*.*

Corollary 4 significantly strengthens List and Pettit's (2002) theorem by individually weakening the conditions generating the impossibility: namely by weakening systematicity to unbiasedness and weakening the agenda condition. Without weakening systematicity to unbiasedness, the agenda condition in List and Pettit's result can be weakened further than in Corollary 4, namely to minimal connectedness alone (Dietrich and List 2007a).

Theorem 3 and Corollary 4 are particularly interesting in light of May's classic characterization of majority voting (1952). Translated into our terminology, May's theorem (without indifference) states that, for any binary agenda $X = \{p, \neg p\}$, an aggregation function $F : \mathbf{U}^n \to \mathbf{U}$ is majority voting if and only if it is anonymous, unbiased and monotonic.[12] Our result shows that, if the agenda is just slightly enriched beyond binariness, May's theorem collapses into an impossibility result *even if* monotonicity is dropped and anonymity is weakened to the requirement that there be no effective dictator.

In Theorem 3, the necessity of the agenda conditions (the "only for") follows from the aggregation functions constructed earlier to show necessity in Theorem 1. These aggregation functions are not only *non*-dictatorial but also *not* effectively dictatorial.

Regarding sufficiency, we first show that the aggregation function F in Theorem 3 induces a unanimity-preserving aggregation function $\widehat{F}$:

[12] Where n is odd. An aggregation function F is *monotonic* if, for all individuals i and all profiles $(A_1, ..., A_n), (A_1, ..., A_i^*, ..., A_n)$ in the domain differing only in i's judgment set, if $F(A_1, ..., A_n) = A_i^*$ then $F(A_1, ..., A_i^*, ..., A_n) = A_i^*$.

Lemma 7 *Suppose $F : \mathbf{U}^n \to \mathbf{U}$ is unbiased. Define, for each $p \in X$,*

$$\widehat{p} := \begin{cases} p & \text{if } N \text{ is a winning coalition for } p, \\ \neg p & \text{if } N \text{ is not a winning coalition for } p. \end{cases}$$

Then:

(a) *For any $p \in X$, $\widehat{\neg p} = \neg\widehat{p}$ and $\widehat{\widehat{p}} = p$.*

(b) *For any $A \subseteq X$, A is consistent if and only if $\{\widehat{p} : p \in A\}$ is consistent.*

(c) *The aggregation function $\widehat{F}$ with universal domain $\mathbf{U}^n$ defined by*

$$\widehat{F}(A_1, ..., A_n) = \{\widehat{p} : p \in F(A_1, ..., A_n)\}$$

satisfies unbiasedness, unanimity-preservation and collective rationality.

(d) *For every $p \in X$,*

$$\mathcal{C}_p = \begin{cases} \widehat{\mathcal{C}}_p & \text{if } \widehat{p} = p \\ \{C \subseteq N : C \notin \widehat{\mathcal{C}}_p\} & \text{if } \widehat{p} = \neg p, \end{cases}$$

where $\mathcal{C}_p$ ($\widehat{\mathcal{C}}_p$) is the set of coalitions winning for p under F ($\widehat{F}$).

Proof. Let F be as specified.

(a) Suppose $p \in X$. N is winning for p if and only if N is winning for $\neg p$; if p is contingent this follows easily from unbiasedness (see also part (a) of Lemma 2); if p is not contingent it holds because N is winning for every tautology and (vacuously) for every contradiction. As N is winning for p if and only if N is winning for $\neg p$, we have $\widehat{p} = p$ if and only if $\widehat{\neg p} = \neg p$, whence $\widehat{\neg p} = \neg\widehat{p}$. Moreover, if $\widehat{p} = p$ then $\widehat{\widehat{p}} = \widehat{p} = p$, and if $\widehat{p} = \neg p$ then $\widehat{\widehat{p}} = \widehat{\neg p} = \neg\widehat{p} = \neg\neg p = p$.

(b) Let $A \subseteq X$. By (a) it is sufficient to show one direction of the implication. Let A be consistent. Then there exists a set $B \in \mathbf{U}$ such that $A \subseteq B$. For each $p \in A$, $F(B, ..., B)$ contains $\widehat{p}$ because:

- if $N \in \mathcal{C}_p$ then $\widehat{p} = p \in F(B, ..., B)$;
- if $N \notin \mathcal{C}_p$ then $p \notin F(B, ..., B)$, and so $\widehat{p} = \neg p \in F(B, ..., B)$.

By $\{\widehat{p} : p \in A\} \subseteq F(B, ..., B)$, $\{\widehat{p} : p \in A\}$ is consistent.

(c) For any $(A_1, ..., A_n) \in \mathbf{U}^n$, $\widehat{F}(A_1, ..., A_n)$ is

- consistent by (b) and the consistency of $F(A_1, ..., A_n)$;
- complete as, for any $p \in X$, if $p \notin \widehat{F}(A_1, ..., A_n)$ then $\widehat{p} \notin F(A_1, ..., A_n)$ by $p = \widehat{\widehat{p}}$, hence $\neg\widehat{p} \in F(A_1, ..., A_n)$, and so $\widehat{F}(A_1, ..., A_n)$ contains $\widehat{\neg\widehat{p}} = \widehat{\widehat{\neg p}} = \neg p$.

$\widehat{F}$ is unanimity-preserving: for any $p \in X$ and any $(A_1, ..., A_n) \in \mathbf{U}^n$ with $p \in A_i$ for all individuals i,

- if $p \in F(A_1, ..., A_n)$, then N is a winning coalition for p under F by independence (see Lemma 1), hence $\widehat{p} = p$, and so $p \in \widehat{F}(A_1, ..., A_n)$;
- if $p \notin F(A_1, ..., A_n)$, then $\neg p \in F(A_1, ..., A_n)$, hence $\widehat{\neg p} \in \widehat{F}(A_1, ..., A_n)$, where $\widehat{\neg p} = \neg\widehat{p} = \neg\neg p = p$ (since $\widehat{p} = \neg p$).

To show that $\widehat{F}$ is unbiased, consider any $p \in X$ and $(A_1, ..., A_n), (A_1^*, ..., A_n^*) \in \mathbf{U}^n$ such that $p \in A_i$ if and only if $\neg p \in A_i^*$. Then (*) $\widehat{p} \in A_i$ if and only if

$\neg\widehat{p} \in A_i^*$. Now $p \in \widehat{F}(A_1, ..., A_n)$ is equivalent to $\widehat{p} \in F(A_1, ..., A_n)$, by definition of $\widehat{F}$ and as $p = \widehat{\widehat{p}}$. The latter is equivalent to $\neg\widehat{p} \in F(A_1^*, ..., A_n^*)$, by (*) and as F is unbiased. This, in turn, is equivalent to $\widehat{\neg\widehat{p}} \in \widehat{F}(A_1^*, ..., A_n^*)$ by definition of $\widehat{F}$, i.e., to $\neg p \in \widehat{F}(A_1^*, ..., A_n^*)$ as $\widehat{\neg\widehat{p}} = \widehat{\widehat{\neg p}} = \neg p$ by part (a).

(d) Let p, $\mathcal{C}_p$ and $\widehat{\mathcal{C}}_p$ be as specified.

First let $\widehat{p} = p$. Then $\mathcal{C}_p = \widehat{\mathcal{C}}_p$, because, for any profile $(A_1, ..., A_n) \in \mathbf{U}^n$, $p \in F(A_1, ..., A_n)$ is equivalent to $\widehat{p} \in \widehat{F}(A_1, ..., A_n)$ (using that $\widehat{\widehat{p}} = p$), i.e., to $p \in \widehat{F}(A_1, ..., A_n)$.

Now let $\widehat{p} = \neg p$. To show that $\mathcal{C}_p = \{C \subseteq N : C \notin \widehat{\mathcal{C}}_p\}$, we consider any $C \subseteq N$, and prove that $C \in \mathcal{C}_p$ is equivalent to $C \notin \widehat{\mathcal{C}}_p$. By $\widehat{p} = \neg p$, p is contingent, and so there exists a profile $(A_1, ..., A_n) \in \mathbf{U}^n$ such that $\{i : p \in A_i\} = C$. Now $C \in \mathcal{C}_p$ is equivalent to $p \in F(A_1, ..., A_n)$, which is equivalent to $\widehat{p} \in \widehat{F}(A_1, ..., A_n)$ (as in case 1), i.e., to $\neg p \in \widehat{F}(A_1, ..., A_n)$, hence to $p \notin \widehat{F}(A_1, ..., A_n)$, and so to $C \notin \widehat{\mathcal{C}}_p$, as required. ■

To prove (the sufficiency part of) Theorem 3, let X be non-separable and minimally connected, and let $F : \mathbf{U}^n \to \mathbf{U}$ be unbiased. Let $\widehat{F}$ and $\widehat{p}$ (for any $p \in X$) be defined as in Lemma 7. By part (c) of Lemma 7, $\widehat{F}$ satisfies all conditions required in Theorem 1. So $\widehat{F}$ is dictatorial by Theorem 1, say with dictator i. Hence, by part (d) of Lemma 7, i is under F a dictator on $Z_+ := \{p \in X : \widehat{p} = p\}$ and an inverse dictator on $Z_- := \{p \in X : \widehat{p} = \neg p\}$ $(= X \backslash Z_+)$. So F is effectively dictatorial. ■

6 An illustration

To illustrate the generality of our result, we apply Theorem 1 to the aggregation of (strict) preferences, represented in the judgment aggregation model. We consider the agenda $X = \{xPy, \neg xPy \in \mathbf{L} : x, y \in K \text{ with } x \neq y\}$, where

- $\mathbf{L}$ is a predicate logic for representing preferences, with
  - a two-place predicate P (representing strict preference), and
  - a set of constants $K = \{x, y, z, ...\}$ with $|K| \geq 3$ (representing alternatives).
- A is consistent if and only if $A \cup Z$ is consistent in the standard sense of predicate logic, with Z defined as the set of rationality conditions on strict preferences.[13]

For details of this construction, see Dietrich and List (2007a) (also List and Pettit 2004). Each rational judgment set $A_i \subseteq X$ uniquely represents a strict (i.e., asymmetric, transitive and connected) preference relation $\succ_i \subseteq K \times K$,

[13] Formally, Z contains $(\forall v_1)(\forall v_2)(v_1Pv_2 \rightarrow \neg v_2Pv_1)$ (asymmetry), $(\forall v_1)(\forall v_2)(\forall v_3)((v_1Pv_2 \wedge v_2Pv_3) \rightarrow v_1Pv_3)$ (transitivity), $(\forall v_1)(\forall v_2)(\neg v_1 = v_2 \rightarrow (v_1Pv_2 \vee v_2Pv_1))$ (connectedness) and, for any distinct constants $x, y \in K$, $\neg x = y$ (exclusiveness of alternatives).

where, for any $x, y \in K$, $xPy \in A_i$ if and only if $x \succ_i y$. For example, if $K = \{x, y, z\}$, the preference relation $x \succ_i y \succ_i z$ is represented by the judgment set $A_i = \{xPy, yPz, xPz, \neg yPx, \neg zPy, \neg zPx\}$.

The agenda X thus defined is non-separable. It is also minimally connected (take $Y = \{xPy, yPz, zPx\}$ in parts (i) and (ii) and $Z = \{xPy, yPz\}$, where $x, y, z \in K$ are distinct alternatives). So Theorems 1 and 3 apply (whereas Theorem 2 does not apply, as X is not locally asymmetric, in fact not even asymmetric). Let us state Theorem 3 for this agenda:

Corollary 5 *For the agenda* $X = \{xPy, \neg xPy \in \mathbf{L} : x, y \in K \text{ with } x \neq y\}$, *every unbiased aggregation function* $F : \mathbf{U}^n \to \mathbf{U}$ *is effectively dictatorial.*

What does this mean in the language of preference aggregation? A judgment aggregation function $F : \mathbf{U}^n \to \mathbf{U}$ uniquely represents a social welfare function with universal domain (taking strict preferences as input and output). Unbiasedness, applied to such a social welfare function, becomes the condition that, for any pair of alternatives $x, y \in K$ and any two preference profiles $(\succ_1, ..., \succ_n)$, $(\succ_1^*, ..., \succ_n^*)$ in the universal domain, if [for all individuals i, $x \succ_i y$ if and only if $y \succ_i^* x$] then [$x \succ y$ if and only if $y \succ^* x$]. Corollary 5 thus implies that every unbiased social welfare function with universal domain is effectively dictatorial. No unanimity (Pareto) condition is needed. Although this result could also be obtained in standard social choice theory (for example, via Wilson's 1972 result on social choice without the Pareto principle), the observation that it is a corollary of our new result on judgment aggregation should illustrate the result's generality. Interestingly, unlike Wilson's and Arrow's theorems, our result continues to hold even if the rationality conditions on preference relations are relaxed to acyclicity alone (giving up full transitivity and connectedness). The reason is that the agenda X, as specified above, remains non-separable and minimally connected in a modified predicate logic obtained by weakening the conditions in the set Z above so as to capture acyclicity alone.[14]

7 Concluding remarks

In judgment aggregation, we face not only a trade-off between different conditions on an aggregation function for any given agenda, as in preference aggregation, but also a trade-off between these conditions and the generality of the agendas for which they can be met by an aggregation function. We have proved three impossibility results. Two of them (Theorems 1 and 3) apply to all standard example agendas in the literature, including the agendas representing preference agggregation problems. This generality is surprising, as we do not

[14]Formally, Z then contains, for any sequence of distinct constants $x_1, x_2, ..., x_k \in K$, the proposition $(x_1Px_2 \wedge ... \wedge x_{k-1}Px_k) \to \neg x_kPx_1$ (no cycle over $x_1, x_2, ..., x_k$) and, for any distinct constants $x, y \in K$, $\neg x = y$ (exclusiveness of alternatives).

impose the requirement of systematicity, a condition often criticized as being too strong. We impose the weaker condition of unbiasedness: a May-type neutrality condition applied to each proposition-negation pair, without requiring neutrality across pairs.

Our results show that, for many agendas, unbiasedness leads to dictatorship, or at least to effective dictatorship. Thus the implications of unbiasedness are surprisingly similar to those of systematicity, despite the fact that unbiased is by itself considerably weaker than systematicity. Our agenda conditions (non-separability, minimal connectedness, and in Theorem 2 also local asymmetry) are tight: agendas that violate any of them avoid the impossibility. Theorems 2 and 3 require no unanimity, monotonicity or other responsiveness condition,[15] unlike many related impossibility results of social choice. But we retain full rationality of individuals and the group. (If we give up full rationality, unbiasedness does become satisfiable together with our other conditions, e.g., by symmetrical special majority voting, as discussed in List and Pettit 2002 and Dietrich and List 2007b, and by other oligarchic rules, e.g., Gärdenfors 2006, Dietrich and List 2008.)

Our results appear significant, as they imply that, in virtually all realistic judgment aggregation problems, any aggregation function with commonly accepted properties must favour some propositions over their negations.

8 References

Dietrich, F. (2006) Judgment Aggregation: (Im)Possibility Theorems. Journal of Economic Theory 126(1): 286-298

Dietrich, F. (2007) A generalised model of judgment aggregation. Social Choice and Welfare 28(4): 529-565

Dietrich, F. (forthcoming) The possibility of judgment aggregation on agendas with subjunctive implications. Journal of Economic Theory

Dietrich, F., List, C. (2007a) Arrow's theorem in judgment aggregation. Social Choice and Welfare 29(1): 19-33

Dietrich, F., List, C. (2007b) Judgment aggregation by quota rules: majority voting generalized. Journal of Theoretical Politics 19(4): 391-424

Dietrich, F., List, C. (2008) Judgment aggregation without full rationality. Social Choice and Welfare 31: 15-39

Dokow, E., Holzman, R. (forthcoming) Aggregation of binary evaluations. Journal of Economic Theory

Gärdenfors, P. (2006) An Arrow-like theorem for voting with logical consequences. Economics and Philosophy 22(2): 181-190

Konieczny, S., Pino-Perez, R. (2002) Merging information under constraints:

[15] By this, we mean that no such condition is assumed or imposed on the 'left-hand side' of our theorems. To be sure, if our conditions imply dictatorship, then of course they also imply such other conditions, in so far as dictatorships meet them.

a logical framework, Journal of Logic and Computation 12(5): 773-808

List, C., Pettit, P. (2002) Aggregating Sets of Judgments: An Impossibility Result. Economics and Philosophy 18: 89-110

List, C., Pettit, P. (2004) Aggregating Sets of Judgments: Two Impossibility Results Compared. Synthese 140(1-2): 207-235

List, C., Puppe, C. (2009) Judgment aggregation: a survey. In Anand, P., Puppe, C., Pattanaik, P. (eds.) Oxford Handbook of Rational and Social Choice. Oxford: Oxford University Press

May, K. O. (1952) A set of independent, necessary and sufficient conditions for simple majority decision. Econometrica 20: 680-684

Mongin, P. (forthcoming) Factoring out the impossibility of logical aggregation. Journal of Economic Theory

Nehring, K., Puppe, C. (2002) Strategy-Proof Social Choice on Single-Peaked Domains: Possibility, Impossibility and the Space Between. Working paper, University of California at Davies

Nehring, K., Puppe, C. (2005) The Sructure of Strategy-Proof Social Choice. Part II: Non-dictatorship, Anonymity and Neutrality. Working paper, University of Karlsruhe

Nehring, K., Puppe, C. (2008). Consistent judgement aggregation: the truth-functional case. Social Choice and Welfare 31: 41-57

Pauly, M., van Hees, M. (2006) Logical Constraints on Judgment Aggregation. Journal of Philosophical Logic 35: 569-585

Pettit, P. (2001) Deliberative Democracy and the Discursive Dilemma. Philosophical Issues 11: 268-299

van Hees, M. (2007) The limits of epistemic democracy. Social Choice and Welfare 28(4): 649-666

Wilson, R. (1972) Social Choice Theory without the Pareto Principle. Journal of Economic Theory 5: 478-486

Wilson, R. (1975) On the Theory of Aggregation. Journal of Economic Theory 10: 89-99