The impossibility of unbiased judgment aggregation

Franz Dietrich and Christian List[1]

9/2005, revised 8/2006

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 compelling impossibilities without these restrictions? We prove an impossibility theorem without systematicity that applies to all standard agendas: Every regular judgment aggregation rule satisfying a condition called unbiasedness is dictatorial or inversely dictatorial on at least one proposition (for non-separable agendas, on all propositions). No unanimity, monotonicity or other responsiveness condition is needed. 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, judgment aggregation can represent many realistic decision problems. For example, the propositions could be the following:

a: "We can afford a budget deficit."

$a \to b$: "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). 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.

	a	$a \to b$	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 for comments on this paper.

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. A basic fact about Arrowian preference aggregation is the following. If, on a given agenda (set of alternatives under consideration), majority voting generates irrational collective preferences for some profiles of rational individual preferences (Condorcet's paradox),[3] then so does *any* preference aggregation rule satisfying Arrow's conditions (universal domain, weak Pareto, independence of irrelevant alternatives, non-dictatorship). Thus any agenda susceptible to Condorcet's paradox is also susceptible to Arrow's more general impossibility theorem. No such fact holds for judgment aggregation. Even if, on a given agenda (set of propositions under consideration), majority voting generates irrational collective judgments for some profiles of rational individual judgments,[4] other judgment aggregation rules may still satisfy the exact counterparts of Arrow's conditions and yet guarantee collective rationality. Thus an agenda susceptible to a "discursive paradox" is not necessarily susceptible to an exact counterpart of Arrow's theorem. The agenda in our example above and many others in the literature are of this kind. Neither the size of the agenda nor that of its largest minimal inconsistent subset, which determines whether irrational majority judgments can arise, determines whether or not an Arrow-style impossibility result holds. The logical interconnections between the propositions in the agenda matter in a surprisingly complex way. 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), that makes a relatively weak assumption on the agenda, but imposes a strong condition of systematicity on the aggregation rule. Systematicity is the conjunction of an Arrow-inspired independence condition (requiring propositionwise aggregation) and a global neutrality condition (requiring equal treatment of all propositions). Thus the price for the theorem's weak agenda assumption is the strength of its systematicity condition.

In response to this problem, several authors have proved impossibility theorems in which systematicity is weakened to independence (Pauly and van Hees 2006, van Hees 2004, Dietrich 2006, Gärdenfors 2005, Nehring and Puppe 2005a, Dietrich and List 2005, Dokow and Holzman 2005, Mongin 2005). But these results make stronger assumptions on the agenda, notably ones violated in many standard examples of judgment aggregation, such as the one above.[5]

[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 2005).

[3] An agenda has this property if it contains three or more alternatives.

[4] An agenda has this property if it has a minimal inconsistent subset of three or more propositions.

[5] They also require an additional responsiveness, unanimity or monotonicity condition.

Dokow and Holzman (2005), extending an earlier characterization result by Nehring and Puppe (2002) in the related model of "property spaces", have identified an agenda assumption that (in standard logics) is necessary and sufficient for an impossibility result with independence, together with a unanimity condition. This agenda assumption is much stronger than a necessary and sufficient assumption for an impossibility result with systematicity (Dietrich and List 2005). It is also violated in the example above.

Is judgment aggregation free from any compelling impossibilities unless we consider fairly rich agendas or impose systematicity? We prove an impossibility result without systematicity that applies to all standard agendas in the literature.[6] Instead of systematicity, we impose a weaker condition called unbiasedness. It is inspired by May's (1952) neutrality condition on a single binary choice and requires an equal treatment of each proposition and its negation, but not of different propositions. Under our weak agenda assumption, every regular unbiased judgment aggregation rule is dictatorial or inversely dictatorial on at least one proposition – and, if the agenda assumption is slightly strengthened, on all propositions. We also identify the necessary and sufficient agenda assumption for our result. A mathematically related earlier result is a theorem by Nehring and Puppe (2005b) on strategy-proof social choice functions that are neutral-within-issues, on which we comment below.

Our result is of interest in light of May's classic characterization of majority voting (1952) as it shows that, as soon as May's agenda $\{p, \neg p\}$ is just slightly logically enriched, May's possibility result turns into an impossibility result even if May's monotonicity condition is dropped and anonymity significantly weakened.

To illustrate the result's generality, we show that it has a corollary for (strict) preference aggregation, represented in the judgment aggregation model: Every unbiased social welfare function with universal domain is effectively dictatorial. Unlike standard impossibility results on preference aggregation, this result holds not only for the aggregation of fully rational preferences, but also for that of merely acyclical ones. Throughout this paper we adopt Dietrich's (2004) general logics framework.[7]

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.

A *logic (with negation symbol $\neg$)* is given by a non-empty set $\mathbf{L}$ of formal expressions (*propositions*) closed under negation (i.e. $p \in \mathbf{L}$ implies $\neg p \in \mathbf{L}$) and an *entailment relation* $\vDash$, where, for each $A \subseteq \mathbf{L}$ and $p \in \mathbf{L}$, $A \vDash p$ is read as

[6] For a notorious exception, see a footnote below.

[7] It is easy to see that our result also applies to Wilson's (1975) model of the aggregation of binary evaluations, as recently revisited by Dokow and Holzman (2005).

"A entails p".[8] A set $A \subseteq \mathbf{L}$ is *inconsistent* if $A \vDash p$ and $A \vDash \neg p$ for some $p \in \mathbf{L}$, and *consistent* otherwise; $A \subseteq \mathbf{L}$ is *minimal inconsistent* if it is inconsistent and every proper subset $B \subsetneq A$ is consistent.

Our results hold for all logics with the following three minimal properties, including standard propositional, predicate, modal and conditional logics:

(L1) For all $p \in \mathbf{L}$, $\{p\} \vDash p$ (self-entailment).

(L2) For all $p \in \mathbf{L}$ and $A \subseteq B \subseteq \mathbf{L}$, if $A \vDash p$ then $B \vDash p$ (monotonicity).

(L3) $\emptyset$ is consistent, and each consistent set $A \subseteq \mathbf{L}$ has a consistent superset $B \subseteq \mathbf{L}$ containing a member of each pair $p, \neg p \in \mathbf{L}$ (completability).

For example, in propositional logic, $\mathbf{L}$ contains propositions such as a, b, $a \wedge b$, $a \vee b$, $\neg(a \rightarrow b)$, and $\vDash$ satisfies $\{a, a \rightarrow b\} \vDash b$, $\{a\} \vDash a \vee b$, but not $a \vDash a \wedge b$. Various realistic decision problems can be represented in our model, including preference aggregation problems as illustrated below.

The *agenda* is a non-empty subset $X \subseteq \mathbf{L}$, interpreted as the set of propositions on which judgments are to be made, where X is a union of proposition-negation pairs $\{p, \neg p\}$ (with p not itself a negated proposition). We assume that double negations cancel each other out, i.e. $\neg\neg p$ stands for p.[9] In the example above, the agenda is $X = \{a, \neg a, b, \neg b, a \rightarrow b, \neg(a \rightarrow b)\}$ in a standard propositional (or conditional) logic.

Call agenda X *weakly connected* if X has a minimal inconsistent subset Y of size at least three such that $(Y \backslash Z) \cup \{\neg z : z \in Z\}$ is consistent for some subset $Z \subseteq Y$ of even size. All standard agendas in the judgment aggregation literature are weakly connected, including agendas representing preference aggregation problems with three or more alternatives, as discussed below.[10] In our example above, take $Y = \{a, a \rightarrow b, \neg b\}$ and $Z = \{a, \neg b\}$. Below we consider an even weaker agenda assumption.[11]

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* as defined above, and (ii) *complete* in the sense that, for every proposition $p \in X$, $p \in A_i$ or $\neg p \in A_i$. A *profile* is an n-tuple $(A_1, \ldots, A_n)$ of individual judgment sets.

A *(judgment) aggregation rule* is a function F that assigns to each admissible profile $(A_1, \ldots, A_n)$ a collective judgment set $F(A_1, \ldots, A_n) = A \subseteq X$, where $p \in A$ means that the group accepts proposition p. The set of admissible profiles is denoted $Domain(F)$. An example is *majority voting*, where, for each $(A_1, ..., A_n)$, $F(A_1, ..., A_n) = \{p \in X : |\{i \in N : p \in A_i\}| > |\{i \in N : p \notin A_i\}|\}$.

[8] Formally, $\vDash \subseteq \mathcal{P}(\mathbf{L}) \times \mathbf{L}$.

[9] More precisely, when we use the negation symbol $\neg$ hereafter, we mean a modified negation symbol $\sim$, where $\sim p := \neg p$ if p is unnegated and $\sim p := q$ if $p = \neg q$ for some q.

[10] The notorious exception is $X = \{a, \neg a, b, \neg b, a \leftrightarrow b, \neg(a \leftrightarrow b)\}$, where $\leftrightarrow$ is the material biconditional. However, for a strict or subjunctive biconditional in standard modal or conditional logics, X is weakly connected.

[11] List and Pettit's (2002) agenda assumption is also a special case of weak connectedness: here the agenda contains two or more atomic propositions, their conjunction (or disjunction or material implication) and the negations of these propositions.

Call aggregation rule F *regular* if (i) $Domain(F)$ is the set of all possible profiles of rational individual judgment sets, and (ii) $F(A_1, \ldots, A_n)$ is rational for every profile $(A_1, \ldots, A_n) \in Domain(F)$.

3 The result

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

Unbiasedness. For any proposition $p \in X$ and profiles $(A_1, \ldots, A_n)$, $(A_1^*, \ldots, A_n^*) \in Domain(F)$, 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 $\neg p \in F(A_1^*, \ldots, A_n^*)$].[12]

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 to each such pair. In the recent literature, unbiasedness is related to Nehring and Puppe's (2005b) neutrality-within-issues. (Specifically, a neutral-within-issues and strategy-proof social choice function induces a regular, unbiased and monotonic judgment aggregation rule.)

Unbiasedness is considerably weaker than List and Pettit's (2002) condition of systematicity, which requires an aggregation rule to be neutral between *any* two propositions $p, q \in X$, as in the case of majority voting, symmetrical supermajority rules, dictatorships or inverse dictatorships.[13] By contrast, unbiasedness permits aggregation rules 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, majority voting may be applied on some pairs $p, \neg p \in X$, dictatorships, inverse dictatorships or symmetrical committee rules on others, and erratic rules such minority rules on yet others etc. Unbiasedness also differs from a global neutrality condition based on a permutation $\pi : X \to X$ of the agenda (e.g. van Hees 2004). It is by itself logically independent from independence, but implies independence if regularity is also assumed, as shown below. Systematicity, by contrast, implies both independence and global neutrality.

To state our result, call individual $i \in N$ a *dictator for* p (respectively, an *inverse dictator for* p) if, for any $(A_1, ..., A_n) \in Domain(F)$, $p \in F(A_1, ..., A_n)$ if and only if $p \in A_i$ (respectively, if and only if $p \notin A_i$); call someone who is a dictator or inverse dictator for p an *effective dictator for* p; and call aggregation

[12] All our results also hold for a modified definition of unbiasedness, obtained by substituting "$p \notin A_i$" for "$\neg p \in A_i$" and "$p \notin F(A_1, ..., A_n)$" for "$\neg p \in F(A_1, ..., A_n)$". The two definitions are equivalent under universal domain and collective rationality.

[13] An aggregation rule F is *systematic* if, for any propositions $p, q \in X$ and profiles $(A_1, \ldots, A_n)$, $(A_1^*, \ldots, A_n^*) \in Domain(F)$, 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 $q \in F(A_1^*, \ldots, A_n^*)$].

rule F *effectively locally dictatorial* if some individual $i \in N$ is an effective dictator for some proposition $p \in X$.

Theorem 1. For a weakly connected agenda, every regular unbiased aggregation rule is effectively locally dictatorial.

In other words, under a regular unbiased aggregation rule (on a weakly connected agenda), some individual is a dictator or inverse dictator on at least one proposition. Theorem 1 continues to hold if unbiasedness is restricted to the propositions in Y, as defined in weak connectedness. Also, the proof shows that some individual is an effective dictator on every proposition in Y (and its negation). Below we identify the weakest agenda assumption for which the result holds.

If the agenda assumption is slightly strenghtened, we obtain a global dicatorship result. Call agenda X *non-separable* if it cannot be partitioned into two logically independent (sub)agendas 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 $B_2 \subseteq X_2$, and proposition $p \in \mathbf{L}$ is *contingent* if $\{p\}$ and $\{\neg p\}$ are consistent). The agenda in our example above and many others are non-separable.

Call aggregation rule F *effectively dictatorial* if some individual $i \in N$ is an effective dictator for every proposition $p \in X$.

Theorem 2. For a weakly connected and non-separable agenda (in a compact logic), every regular unbiased aggregation rule is effectively dictatorial.

Theorems 1 and 2 are related to Nehring and Puppe's (2005b) results on strategy-proof social choice functions that are neutral-within-issues (in the model of "property spaces"). Translated into our framework, their results imply that, for slightly weakened agenda assumptions, every regular unbiased aggregation rule which in addition satisfies monotonicity (and thus by implication also a unanimity condition) is locally dictatorial (respectively, globally dictatorial). In our theorems, no monotonicity, unanimity or other responsiveness condition is needed.

Below we show that the restriction to a compact logic can be dropped if the agenda assumption of non-separability is replaced by that of indirect connectedness (also defined below). Note the following corollary of Theorem 1. Call aggregation rule F *anonymous* if, for any profile $(A_1, \ldots, A_n) \in Domain(F)$ and any permutation $\sigma : N \to N$, $F(A_1, \ldots, A_n) = F(A_{\sigma(1)}, \ldots, A_{\sigma(n)})$.

Corollary 1. For a weakly connected agenda, there exists no regular, unbiased and anonymous aggregation rule.

Corollary 1 significantly strengthens List and Pettit's (2002) theorem by weakening systematicity to unbiasedness and weakening the agenda assumption.

Corollary 1 as well as Theorems 1 and 2 are particularly interesting in light of May's classic characterization of majority voting (1952). Translated into our terminology, May's theorem states that, for an agenda containing a single proposition-negation pair, an aggregation rule is majority voting if and only if it is regular, anonymous, unbiased and monotonic.[14] Our result shows that, if the agenda is just slightly enriched (in the sense of weak connectedness), May's theorem collapses into an impossibility result *even if* monotonicity is dropped altogether and anonymity is weakened to the requirement that there be no effective local dictator.

4 The proof

The proof of Theorem 1 is based on three lemmas.

Call aggregation rule F *independent* if, for any proposition $p \in X$ and profiles $(A_1, \ldots, A_n)$, $(A_1^*, \ldots, A_n^*) \in Domain(F)$, 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. A regular and unbiased aggregation rule is also independent.

Proof. Let F be as specified. Consider any $p \in X$ and any profiles $(A_1, ..., A_n)$, $(A_1^*, ..., A_n^*) \in Domain(F)$ 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 regularity, if p is a tautology (i.e. $\{\neg p\}$ is inconsistent), p is contained in both $F(A_1, ..., A_n)$ and $F(A_1^*, ..., A_n^*)$; if p is a contradiction (i.e. $\{p\}$ is inconsistent), 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. As F is regular, there exists a profile $(A_1', ..., A_n') \in Domain(F)$ 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 $p \in F(A_1^*, ..., A_n^*)$. ■

Call coalition $C \subseteq N$ *winning for* $p \in X$ (under F) if $p \in F(A_1, ..., A_n)$ for every profile $(A_1, ..., A_n) \in Domain(F)$ with $\{i : p \in A_i\} = C$. If an aggregation rule F is independent, it is uniquely determined by its winning coalitions, because

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

where, for each $p \in X$, $\mathcal{C}_p$ denotes the set of winning coalitions for p.

Call aggregation rule F *unanimity-respecting* if N is a winning coalition for every $p \in X$.

[14] Where n is odd. An aggregation rule F is *monotonic* if, for every $p \in X$, individual i, and pair of i-variant profiles $(A_1, ..., A_n)$, $(A_1, ..., A_i', ..., A_n)$ in $Domain(F)$ with $p \notin A_i$ and $p \in A_i'$, if $p \in F(A_1, ..., A_n)$ then $p \in F(A_1, ..., A_i', ..., A_n)$. (Two profiles are *i-variants* if they coincide for all individuals except possibly i.)

Lemma 2. Suppose F is regular and 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 rule $\widehat{F}$ on $Domain(\widehat{F}) := Domain(F)$ defined by

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

is regular, unbiased and unanimity-respecting.

(d) For any $p \in X$, either $\mathcal{C}_p = \widehat{\mathcal{C}_p}$ or $\mathcal{C}_p = \{C \subseteq N : C \notin \widehat{\mathcal{C}_p}\}$, where $\mathcal{C}_p$ and $\widehat{\mathcal{C}_p}$ are the sets of winning coalitions for p under F and $\widehat{F}$, respectively.

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 3); 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 complete and consistent judgment set $B \subseteq X$ 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 $(A_1, ..., A_n) \in Domain(\widehat{F})$, $\widehat{F}(A_1, ..., A_n)$ 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-respecting: for any $p \in X$ and any $(A_1, ..., A_n) \in Domain(\widehat{F})$ 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 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 $(A_1, ..., A_n)$, $(A_1^*, ..., A_n^*)$ $\in Domain(\widehat{F})$ 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. We distinguish two cases.

Case 1: $\widehat{p} = p$. Then $\mathcal{C}_p = \widehat{\mathcal{C}_p}$, because, for any profile $(A_1, ..., A_n) \in Domain(F)$, $p \in F(A_1, ..., A_n)$ is equivalent to $\widehat{p} \in \widehat{F}(A_1, ..., A_n)$ (using that $\widehat{\widehat{p}} = p$), i.e. to $p \in \widehat{F}(A_1, ..., A_n)$.

Case 2: $\widehat{p} = \neg p$. 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 $(A_1, ..., A_n) \in Domain(F)$ 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. ■

Call propositions $p, q \in X$ *connected* (in X) if 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^*.

Lemma 3. Suppose F is regular and unbiased. For any $p \in X$, let $\mathcal{C}_p$ be the set of winning coalitions for p. Then:

(a) If $p \in X$ is contingent, then $\mathcal{C}_p = \mathcal{C}_{\neg p}$, and, for any $C \subseteq N$, $C \in \mathcal{C}_p$ if and only if $N \backslash C \notin \mathcal{C}_p$.

(b) If $p, q \in X$ are connected and F is unanimity-respecting, $\mathcal{C}_p = \mathcal{C}_q$.

(c) If $p, q \in X$ are connected, either $\mathcal{C}_p = \mathcal{C}_q$ or $\mathcal{C}_p = \{C \subseteq N : C \notin \mathcal{C}_q\}$.

Proof. Let F be as specified. 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 $(A_1, ..., A_n), (A_1^*, ..., A_n^*) \in Domain(F)$ 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 $(A_1, ..., A_n) \in Domain(F)$ 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-respecting. 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 (using L1-L3) 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) there exists a profile $(A_1, ..., A_n) \in Domain(F)$ 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 $v \in F(A_1, ..., A_n)$ by $C \in \mathcal{C}_v$, and $Y \subseteq F(A_1, ..., A_n)$ by $N \in \mathcal{C}_v$. By

$\{v\} \cup Y \subseteq F(A_1, ..., A_n)$ 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.

(c) Suppose $p, q \in X$ are connected. Let $\widehat{F}$ and $\widehat{\mathcal{C}}_r$, $r \in X$, be as defined in Lemma 2. By Lemma 2, $\widehat{F}$ is regular, unbiased and unanimity-respecting. So, by part (b), $\widehat{\mathcal{C}}_p = \widehat{\mathcal{C}}_q$. This together with part (d) of Lemma 2 implies the claim. ■

We can now prove Theorem 1. Let X be weakly connected, and let F be regular and unbiased. Let $Y \subseteq X$ be as defined in weak connectedness, and let $\widehat{F}$ and $\widehat{p}$ (for any $p \in X$) be defined as in Lemma 2. By Lemma 2, $\widehat{F}$ is regular, unbiased and unanimity-respecting; hence $\widehat{F}$ is also independent by Lemma 1.

As $\widehat{F}$ is independent, $\widehat{F}$ induces a unique aggregation rule F^* for the sub-agenda $X^* := \{p, \neg p : p \in Y\}$; specifically, F^* is the aggregation rule for X^* on the domain of all possible profiles of rational individual judgment sets on X^* defined by

$$F^*(A_1, ..., A_n) = \widehat{F}(B_1, ..., B_n) \cap X^* \text{ for any } (A_1, ..., A_n) \in Domain(F^*),$$

where $(B_1, ..., B_n) \in Domain(\widehat{F})$ satisfies $B_i \cap X^* = A_i$ for each i, and where by independence $F^*(A_1, ..., A_n)$ does not depend on the particular choice of $(B_1, ..., B_n)$.

Claim 1. F^* is regular, unanimity-respecting and systematic.

The rule F^* inherits its regularity, respect for unanimity and unbiasedness from $\widehat{F}$, which has these properties by Lemma 2. So, by part (b) of Lemma 3, as any two propositions in X^* are connected, each proposition in X^* has the same set of winning coalitions (under F^*). So F^* is systematic.

Claim 2. F^* is dictatorial, say with dictator i on every proposition $p \in X^*$.

This claim follows from claim 1 by applying Proposition 1 in Dietrich and List (2005).

Claim 3. Under F, for each $p \in X^*$, i is a (possibly inverse) dictator for p, and hence an effective dictator for p, which completes the proof of Theorem 1.

By Claim 2 and as F^* is the restriction of $\widehat{F}$ to X^*, i is under $\widehat{F}$ a dictator for each $p \in X^*$. So, by part (d) of Lemma 2, under F, for each $p \in X^*$, i is a (possibly inverse) dictator for p. ■

Theorem 2 can be proved by replacing in Claims 1 to 3 X^* by X and F^* by $\widehat{F}$, where the new Claim 1 still holds because, using Proposition 2 below, any two propositions in X are still *indirectly* connected as defined below.

5 A necessary and sufficient agenda assumption

How much further can our agenda assumption be weakened? We now identify an (essentially) necessary and sufficient agenda assumption for our impossibility

result. 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, for each $t \in \{1, ..., k-1\}$, p_t and p_{t+1} are connected (as defined above).

Theorems 1 and 2 (and Corollary 1) continue to hold if *weak connectedness* is weakened to the assumption that (i) X has a minimal inconsistent subset Y of size at least three, and (ii) X has a minimal inconsistent subset Y^* such that $(Y^* \backslash Z) \cup \{\neg z : z \in Z\}$ is consistent for some subset $Z \subseteq Y^*$ of even size, where (iii) some $p \in Y$ is indirectly connected to some $q \in Y^*$.[15] To see this, adapt the proof by redefining the subagenda X^* as $\{p, \neg p : p \in Y\} \cup \{p, \neg p : p \in Y^*\} \cup \{p_t, \neg p_t : t = 1, ..., k\}$, where $p_1, ..., p_k$ is a path indirectly connecting some $p \in Y$ to some $q \in Y^*$.

The conjunction of (i), (ii) and (iii) is not only sufficient, but also essentially necessary for obtaining the impossibility result of Theorem 1.

Proposition 1. For a compact logic or a finite agenda, and n odd, if the agenda violates (i), (ii) or (iii), then there exists a regular, unbiased and anonymous aggregation rule.

Proof. Suppose not all of (i)-(iii) hold. Let X_1 be the set of all $p \in X$ that either belong to a set $Y^* \subseteq X$ of the type in (ii) or are indirectly connected to an element of such a set. Further, define $X_2 := X \backslash X_1$. (X_1 or X_2 can be empty.) By assumption, (*) X_1 has no subset Y of the type in (i), and (**) X_2 has no subset Y^* of the type in (ii).

Claim: 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 is 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 $s \in B \cap X_2$. 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 contradiction proves the claim.

Define the rule F on the domain of all profiles of rational individual judgment sets by $F(A_1, ..., A_n) := B_1 \cup B_2$, where

$$\begin{aligned} B_1 &: = \{p \in X_1 : |\{i \in N : p \in A_i\}| > |\{i \in N : p \notin A_i\}|\}, \\ B_2 &: = \{p \in X_2 : |\{i \in N : p \in A_i\}| \text{ is odd}\}. \end{aligned}$$

F is unbiased and anonymous, and the output $B_1 \cup B_2$ is complete as n is odd. To complete the proof, we show that $B_1 \cup B_2$ is consistent. B_1 is consistent by (*) and as X is finite or the logic compact. B_2 is is consistent by (**) (see

[15] If $Y = Y^*$, this reduces to the earlier definition of weak connectedness. The conjunction of (i) and (ii) is the condition of *minimal connectedness*, the weakest agenda assumption for which an impossibility result with List and Pettit's original conditions holds (Dietrich and List 2005).

Dokow and Holzman (2005), from where we had the insight for how to define F on X_2). So, by the above claim, $B_1 \cup B_2$ is consistent. ■

Condition (iii) holds in particular if the agenda is *indirectly connected* in the sense that *any* contingent propositions $p, q \in X$ are indirectly connected. To show that indirect connectedness of the agenda, although stronger than (iii), is still undemanding, we prove that, under standard assumptions, it is equivalent to non-separability, as defined above. Moreover, Theorem 2 continues to hold if non-separability is replaced with indirect connectedness and the compactness requirement is dropped.

Proposition 2. For any agenda, indirect connectedness implies non-separability, and the two are equivalent if the agenda is finite or the logic compact.

Proof. It is sufficient to prove the claim for $X_0 := \{p \in X : p \text{ contingent}\}$, because X is indirectly connected if and only if X_0 is, and X is non-separable if and only if X_0 is.

1. First assume X_0 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.

2. Now suppose X is finite or the logic is compact, and 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 $X_1 := \{r \in X : p \text{ and } r \text{ are indirectly connected}\}$ 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. Further, each of X_1 and X_2 is closed under negation. If follows that X_1 and X_2 are subagendas of X. Finally, it can be shown (see the "claim" in the proof of Proposition 1) that X_1 and X_2 are logically independent, as required. ■

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

(i) $\mathbf{L}$ 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).

(ii) $A \vDash p$ if and only if $A \cup Z$ entails p in the standard sense of predicate logic, with Z defined as the set of rationality conditions on strict preferences.[16]

For details of this construction, see Dietrich and List (2005) (also List and Pettit 2004). The agenda X thus defined is weakly connected and non-separable. Each rational judgment set $A_i \subseteq X$ uniquely represents a strict (i.e. asymmetric, transitive and connected) preference relation $\succ_i \subseteq K \times K$, 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 $A_i = \{xPy, yPz, xPz, \neg yPx, \neg zPy, \neg zPx\}$.

Now a regular judgment aggregation rule uniquely represents a social welfare function (with universal domain) taking strict preferences as input and output. Unbiasedness, applied to a social welfare function, becomes the condition that, for any pair of alternatives $x, y \in K$ and any two preference profiles $(\succ_1, ..., \succ_n)$, $(\succ_1^*, ..., \succ_n^*)$, 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 $y \succ^* x$]. Theorem 2 immediately yields the following corollary.

Corollary 2. For the agenda $X = \{xPy, \neg xPy \in \mathbf{L} : x, y \in K \text{ with } x \neq y\}$, every regular unbiased aggregation rule is effectively dictatorial.

In the language of preference aggregation, every unbiased social welfare function with universal domain is effectively dictatorial. Again, 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 our 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 weakly connected and non-separable in a modified predicate logic obtained by weakening the conditions in the set Z above so as to capture acyclicity alone.

[16] 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 each pair of distinct contants $x, y \in K$, $\neg x{=}y$.

7 Concluding remarks

In judgment aggregation, we face not only a logical trade-off between different conditions on an aggregation rule (as in preference aggregation), but also a logical trade-off between these conditions and the generality of the agendas of propositions for which the aggregation rules in question are used. We have proved an impossibility theorem that applies to all standard agendas in the literature and yet does not impose systematicity, a condition often criticized as being too strong. The weaker condition of unbiasedness allows the rule to treat different propositions differently, while preserving neutrality between each proposition and its negation. Unbiasedness can be seen as the application of a May-type neutrality condition to each proposition-negation pair. Like May's condition, unbiasedness is a plausible requirement in many, but not all, aggregation problems.

Our result shows that, for all weakly connected agendas, unbiasedness leads to an effective local dictatorship (in a regular aggregation rule). If the agenda is also non-separable or indirectly connected, as in our initial example and in preference aggregation problems with three or more alternatives (and in many other standard aggregation problems), unbiasedness leads to an effective (global) dictatorship. Finally, we have identified the weakest agenda assumption for which our result holds. Unlike some other related impossibility results of social choice, our result requires no unanimity, monotonicity or other responsiveness condition.

The identified impossibility appears significant, as it implies that, in virtually all realistic judgment aggregation problems, any aggregation rule with commonly accepted properties must favour some propositions over their negations.

8 References

Dietrich, F. (2004) A generalised model of judgment aggregation. Social Choice and Welfare (forthcoming)

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

Dietrich, F., List, C. (2005) Arrow's theorem in judgment aggregation. Social Choice and Welfare (forthcoming)

Dokow, E., Holzman, R. (2005) Aggregation of binary evaluations, Working paper, Technion Israel Institute of Technology

Gärdenfors, P. (2005) An Arrow-like theorem for voting with logical consequences. Economics and Philosophy (forthcoming)

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

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

Mongin, P. (2005) Factoring out the impossibility of logical aggregation. Working paper, CNRS, Paris

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. (2005a) Consistent Judgment Aggregation: A Characterization. Working paper, University of Karlsruhe

Nehring, K., Puppe, C. (2005b) On the Possibility of Strategy-Proof Social Choice: Non-dictatorship, Anonymity and Neutrality. Working paper, University of Karlsruhe

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. (2004) The limits of epistemic democracy. Working paper, University of Groningen

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