

The impossibility of unbiased judgment aggregation

Franz Dietrich and Christian List¹

10/2005

All existing impossibility theorems on judgment aggregation over logically connected propositions have one of two restrictions: they either use a controversial systematicity condition or apply only to special agendas of propositions with rich logical connections. An important open question is whether judgment aggregation faces any serious impossibilities without these restrictions. Here we prove the first impossibility theorem without systematicity that applies to all standard agendas: there exists no judgment aggregation rule satisfying universal domain, collective rationality, anonymity and a new condition called unbiasedness. For many agendas, anonymity can be weakened. Applied illustratively to (strict) preference aggregation represented in the judgment aggregation model, our result implies that every unbiased social welfare function with universal domain depends only on a single individual.

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

1 Introduction

In this paper, we prove a new impossibility theorem on the aggregation of individual judgments (acceptance or rejection) on logically connected propositions into collective judgments on these propositions. Judgment aggregation can represent many realistic collective decision problems due to the flexible notion of a proposition. 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. In our example, if individual judgments are as in Table 1, then a majority accepts a , a majority accepts $a \rightarrow b$, and yet a majority rejects b . This problem has become known as the "discursive paradox" (Pettit 2001).

Although there are parallels between judgment aggregation and the more familiar problem of preference aggregation in Condorcet's and Arrow's tradition, judgment aggregation can be shown to generalize preference aggregation²

¹F. Dietrich, Department of Quantitative Economics, University of Maastricht, P.O. Box 616, 6200 MD Maastricht, The Netherlands; C. List, Department of Government, London School of Economics, London WC2A 2AE, U.K.

²If we express preference relations as binary ranking propositions in predicate logic (of the form xPy), then preference aggregation becomes a special case of judgment aggregation (List and Pettit 2004; Dietrich and List 2005), as illustrated in section 6 below. For a related discussion, see Dokow and Holzman (2005).

	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

and faces some important additional complexities. A basic fact about Arrowian preference aggregation is that whenever the agenda – the set of alternatives under consideration – is such that majority voting generates irrational collective preferences for some profiles of individual preferences (i.e. when the agenda contains three or more alternatives), then so does *any* preference aggregation rule satisfying certain conditions:³ Arrow’s impossibility theorem. No such fact holds for judgment aggregation. Even if the agenda – the set of propositions under consideration – is such that majority voting generates irrational collective judgments for some profiles of individual judgments (i.e. if the agenda has a minimal inconsistent subset of size three or more), there may still exist other judgment aggregation rules that satisfy the exact counterparts of Arrow’s conditions and yet guarantee collective rationality; an impossibility result analogous to Arrow’s theorem need not apply. The agenda in our example above and many other agendas are of this kind. Neither the size of the agenda nor the size of its largest minimal inconsistent subset (which determines whether majority judgments can be irrational) determines whether or not a more general impossibility result applies. 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.

List and Pettit (2002) proved a first impossibility theorem on judgment aggregation, strengthened by Pauly and van Hees (2005), that makes a relatively unrestrictive 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 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 on an aggregation rule.

In response to this problem, several authors have proved impossibility theorems in which systematicity is weakened to independence (Pauly and van Hees 2005; Dietrich 2005; van Hees 2004; Gärdenfors 2005; Nehring and Puppe 2005; Dietrich and List 2005; Dokow and Holzman 2005; Mongin 2005). But these results make more restrictive assumptions on the agenda, notably ones that are violated in many standard examples of judgment aggregation, such as the example above. (They also require an additional responsiveness, unanimity or

³Universal domain, the weak Pareto principle, independence of irrelevant alternatives and non-dictatorship.

monotonicity condition.)

In an important recent contribution, Dokow and Holzman (2005) 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) and is also violated in the example above.

Thus an important question is still open: Is judgment aggregation free from any compelling impossibility results unless we consider special agendas with rich connections between propositions or impose the strong condition of systematicity? Here we prove the first impossibility result that applies to all standard agendas in the literature and yet relaxes systematicity to a weaker condition called unbiasedness. Our unbiasedness condition is inspired by May's (1952) condition of neutrality in a single binary choice and requires an equal treatment of each proposition and its negation, but not of different propositions.⁴ Specifically, under our weak agenda assumption, there exists no judgment aggregation rule satisfying universal domain, collective rationality, anonymity and unbiasedness. For many standard agendas, anonymity can be further weakened to the condition that the collective judgment set depends not only on a single individual. We also identify the weakest (i.e. necessary and sufficient) agenda assumption for which our impossibility result holds.

To illustrate the generality of our result, we show that, if we represent (strict) preference aggregation in the judgment aggregation model, our result implies that every unbiased social welfare function with universal domain depends only on a single individual. Moreover, unlike standard impossibility results on preference aggregation, our result continues to apply even if we consider the aggregation of merely acyclical, as opposed to fully rational, preferences. Finally, 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). Throughout this paper we adopt Dietrich's (2004) general logics framework.

2 Definitions

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

A *logic* (with negation symbol \neg) consists of 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* \models , where, for each $A \subseteq \mathbf{L}$ and $p \in \mathbf{L}$, $A \models p$ is read as " A entails p ".⁵ A set $A \subseteq \mathbf{L}$ is *inconsistent* if $A \models p$ and $A \models \neg p$ for some $p \in \mathbf{L}$, and *consistent* otherwise; $A \subseteq \mathbf{L}$ is *minimal inconsistent* if it is inconsistent

⁴Unbiasedness is by itself logically independent from independence, though jointly with universal domain and collective rationality it implies independence.

⁵Formally, $\models \subseteq \mathcal{P}(\mathbf{L}) \times \mathbf{L}$.

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 $p \in \mathbf{L}$, $\{p\} \models p$ (self-entailment).
- (L2) For all $p \in \mathbf{L}$ and $A \subseteq B \subseteq \mathbf{L}$, if $A \models p$ then $B \models p$ (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 \models satisfies $\{a, a \rightarrow b\} \models b, \{a\} \models a \vee b$, but not $a \models a \wedge b$. 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 .⁶ 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 a conditional) logic.

We call an agenda *weakly connected* if X has a minimal inconsistent subset Y of size at least three such that $(Y \setminus Z) \cup \{\neg z : z \in Z\}$ is consistent for some subset $Z \subseteq Y$ of even size. All standard agendas in the judgment aggregation literature are weakly connected,⁷ including agendas representing preference aggregation problems with three or more alternatives, as discussed below. 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.⁸

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, \dots, A_n) of individual judgment sets.

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

⁶More 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 .

⁷The 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.

⁸List 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.

3 The theorem

For our result, we impose four conditions on an aggregation rule (below we relax anonymity).

Universal domain. $Domain(F)$ is the set of all possible profiles of rational individual judgment sets.

Collective rationality. $F(A_1, \dots, A_n)$ is rational for every profile $(A_1, \dots, A_n) \in Domain(F)$.

Anonymity. For any profile $(A_1, \dots, A_n) \in Domain(F)$ and any permutation $\sigma : N \rightarrow N$, $F(A_1, \dots, A_n) = F(A_{\sigma(1)}, \dots, A_{\sigma(n)})$.

Unbiasedness. For any proposition $p \in X$ and profiles $(A_1, \dots, A_n), (A_1^*, \dots, A_n^*) \in Domain(F)$, if [for all individuals i , $p \in A_i$ if and only if $\neg p \in A_i^*$] then [$p \in F(A_1, \dots, A_n)$ if and only if $\neg p \in F(A_1^*, \dots, A_n^*)$].⁹

While the first three conditions are standard conditions, unbiasedness is a new condition inspired by May's (1952) condition of neutrality. Unbiasedness requires an equal treatment of each proposition $p \in X$ and its negation $\neg p$, regardless of other judgments. If we interpret a judgment aggregation problem as consisting of multiple binary decisions between proposition-negation pairs, then unbiasedness can be seen as the application of May's neutrality condition to each such pair. Unbiasedness replaces List and Pettit's (2002) stronger 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.¹⁰

Unlike systematicity, 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 $\neg p$, such as majority voting on some pairs $p, \neg p \in X$, different dictatorships or inverse dictatorships on other pairs, symmetrical committee rules on even other pairs, erratic rules (such as ones making each minority decisive) on some pairs etc. Unbiasedness also differs from a global neutrality condition based on a permutation $\pi : X \rightarrow X$ of the agenda (e.g. van Hees 2004). It is by itself logically independent from independence; but, jointly with universal domain and collective rationality, it implies independence, as shown below. By contrast, systematicity implies both independence and global neutrality.

We can now state our main result.

⁹ 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 " $p \notin F(A_1, \dots, A_n)$ " for " $\neg p \in F(A_1, \dots, A_n)$ ". The two definitions are equivalent under universal domain and collective rationality.

¹⁰ Formally, systematicity requires that, for any propositions $p, q \in X$ and profiles $(A_1, \dots, A_n), (A_1^*, \dots, A_n^*) \in Domain(F)$, if [for all individuals i , $p \in A_i$ if and only if $q \in A_i^*$] then [$p \in F(A_1, \dots, A_n)$ if and only if $q \in F(A_1^*, \dots, A_n^*)$].

Theorem 1. For a weakly connected agenda, there exists no aggregation rule satisfying universal domain, collective rationality, anonymity and unbiasedness.

Note that no responsiveness, unanimity or monotonicity condition is needed. Theorem 1 continues to hold if unbiasedness is restricted to the propositions in Y , as defined in weak connectedness. Below we identify the weakest agenda assumption for which the result holds.

Our proof also establishes two refinements. Call individual $i \in N$ *dictatorial for p* (respectively, *inversely dictatorial for p*) if, for any $(A_1, \dots, A_n) \in \text{Domain}(F)$, $p \in F(A_1, \dots, A_n)$ if and only if $p \in A_i$ (respectively, if and only if $p \notin A_i$).

Refinement 1. For a weakly connected agenda, if an aggregation rule satisfies universal domain, collective rationality and unbiasedness, then there exists a proposition $p \in X$ for which some individual $i \in N$ is (possibly inversely) dictatorial.

Call an 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 other agendas are non-separable.

Refinement 2. For a weakly connected and non-separable agenda (in a compact logic), if an aggregation rule satisfies universal domain, collective rationality and unbiasedness, then there exists some individual $i \in N$ such that, for every proposition $p \in X$, i is (possibly inversely) dictatorial for p (in particular, the collective judgment set depends only on individual i).

A disanalogy between our new result and earlier impossibility results on systematicity is that in those earlier results anonymity can be weakened to non-dictatorship and/or non-inverse-dictatorship (Pauly and van Hees 2005; Dietrich 2004; Dietrich and List 2005), while in the present result it cannot be weakened as much. But Refinement 2 shows that, for the large class of weakly connected and non-separable agendas, it can be weakened to the compelling condition that the collective judgment set depends not only on a single individual.

4 The proof

The proof of Theorem 1 is based on three lemmas. The first lemma requires the definition of independence.

Independence. For any proposition $p \in X$ and profiles $(A_1, \dots, A_n), (A_1^*, \dots, A_n^*) \in \text{Domain}(F)$, if [for all individuals i , $p \in A_i$ if and only if $p \in A_i^*$] then $[p \in F(A_1, \dots, A_n)$ if and only if $p \in F(A_1^*, \dots, A_n^*)]$.

Lemma 1. If an aggregation rule satisfies universal domain, collective rationality and unbiasedness, then it also satisfies independence.

Proof. Let F be as specified. Consider any $p \in X$ and any profiles (A_1, \dots, A_n) , $(A_1^*, \dots, A_n^*) \in \text{Domain}(F)$ in which the same set of individuals $C \subseteq N$ accepts p . We show that $p \in F(A_1, \dots, A_n)$ if and only if $p \in F(A_1^*, \dots, A_n^*)$, as required by independence. By collective rationality, if p is a tautology (i.e. $\{\neg p\}$ is inconsistent), p is contained in both $F(A_1, \dots, A_n)$ and $F(A_1^*, \dots, A_n^*)$; if p is a contradiction (i.e. $\{p\}$ is inconsistent), p is contained in neither of $F(A_1, \dots, A_n)$ and $F(A_1^*, \dots, A_n^*)$. Now suppose p is contingent. Then $\neg p$ is also contingent. By universal domain, there exists a profile $(A'_1, \dots, A'_n) \in \text{Domain}(F)$ such that exactly the individuals in C accept $\neg p$. By unbiasedness, $p \in F(A_1, \dots, A_n)$ is equivalent to $\neg p \in F(A'_1, \dots, A'_n)$, which, again by unbiasedness, is equivalent to $p \in F(A_1^*, \dots, A_n^*)$. ■

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

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

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

The second lemma requires the definition of the unanimity principle.

Unanimity principle. N is a winning coalition for each $p \in X$.

Lemma 2. Suppose F satisfies universal domain, collective rationality and unbiasedness, and 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} with universal domain defined by

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

satisfies collective rationality, unbiasedness and the unanimity principle.

- (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 $\widehat{\mathcal{C}}_p$ denote the set of winning coalitions for p under F and under \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 p 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 $p \in A$, $F(B, \dots, B)$ contains \widehat{p} because:

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

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

(c) \widehat{F} satisfies collective rationality: for any $(A_1, \dots, A_n) \in \text{Domain}(\widehat{F})$, $\widehat{F}(A_1, \dots, A_n)$ is

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

\widehat{F} satisfies the unanimity principle: for any $p \in X$ and any $(A_1, \dots, A_n) \in \text{Domain}(\widehat{F})$ with $p \in A_i$ for all individuals i ,

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

To show that \widehat{F} satisfies unbiasedness, consider any $p \in X$ and (A_1, \dots, A_n) , $(A_1^*, \dots, A_n^*) \in \text{Domain}(\widehat{F})$ 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 $p \in \widehat{F}(A_1, \dots, A_n)$ is equivalent to $\widehat{p} \in F(A_1, \dots, A_n)$, by definition of \widehat{F} and as $p = \widehat{\widehat{p}}$. The latter is equivalent to $\neg \widehat{p} \in F(A_1^*, \dots, A_n^*)$, by (*) and as F satisfies unbiasedness. This, in turn, is equivalent to $\widehat{\neg \widehat{p}} \in \widehat{F}(A_1^*, \dots, A_n^*)$ by definition of \widehat{F} , i.e. to $\neg p \in \widehat{F}(A_1^*, \dots, A_n^*)$ 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, \dots, A_n) in the universal domain, $p \in F(A_1, \dots, A_n)$ is equivalent to $\widehat{p} \in \widehat{F}(A_1, \dots, A_n)$ (using that $\widehat{p} = p$), i.e. to $p \in \widehat{F}(A_1, \dots, A_n)$.

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, \dots, A_n) in the universal domain such that $\{i : p \in A_i\} = C$. Now $C \in \mathcal{C}_p$ is equivalent to $p \in F(A_1, \dots, A_n)$, which is equivalent to $\widehat{p} \in \widehat{F}(A_1, \dots, A_n)$ (as in case 1), i.e. to $\neg p \in \widehat{F}(A_1, \dots, A_n)$, hence to $p \notin \widehat{F}(A_1, \dots, A_n)$, and so to $C \notin \widehat{\mathcal{C}}_p$, as desired. ■

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 satisfies collective rationality, universal domain and unbiasedness, and, 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 $N \setminus C \notin \mathcal{C}_p$.
- (b) If $p, q \in X$ are connected and F satisfies the unanimity principle, $\mathcal{C}_p = \mathcal{C}_q$
- (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 satisfies independence.

(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 $(A_1, \dots, A_n), (A_1^*, \dots, A_n^*) \in \text{Domain}(F)$ such that $C = \{i : p \in A_i\} = \{i : \neg p \in A_i^*\}$. By unbiasedness, $p \in F(A_1, \dots, A_n)$ if and only if $\neg p \in F(A_1^*, \dots, A_n^*)$, 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 $(A_1, \dots, A_n) \in \text{Domain}(F)$ such that $C = \{i : p \in A_i\}$, hence $N \setminus C = \{i : \neg p \in A_i\}$. Now $C \in \mathcal{C}_p$ is equivalent to $p \in F(A_1, \dots, A_n)$, hence to $\neg p \notin F(A_1, \dots, A_n)$, hence to $N \setminus C \notin \mathcal{C}_{\neg p}$, hence to $N \setminus C \notin \mathcal{C}_p$, as shown above.

(b) Suppose $p, q \in X$ are connected and the unanimity principle holds. 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 $(A_1, \dots, A_n) \in \text{Domain}(F)$ 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 $i \in N \setminus C$. We have $v \in F(A_1, \dots, A_n)$ by $C \in \mathcal{C}_v$, and $Y \subseteq F(A_1, \dots, A_n)$ by $N \in \mathcal{C}_v$. By $\{v\} \cup Y \subseteq F(A_1, \dots, A_n)$ and (ii), $w \notin F(A_1, \dots, A_n)$. So $N \setminus C \notin \mathcal{C}_w$, and hence $C \in \mathcal{C}_w$ by part (a), as desired.

(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} satisfies collective rationality, universal domain, unbiasedness, and also the unanimity principle. 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 satisfy universal domain, collective rationality and unbiasedness. We show that F is not anonymous. 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} satisfies collective rationality, universal domain, unbiasedness and the unanimity principle; hence \widehat{F} also satisfies independence by Lemma 1.

As \widehat{F} satisfies independence, \widehat{F} induces a unique aggregation rule F^* for the subagenda $X^* := \{p, \neg p : p \in Y\}$; specifically, F^* is the aggregation rule for X^*

with universal domain given by

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

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

Claim 1. F^* satisfies collective rationality, the unanimity principle and systematicity.

The rule F^* inherits collective rationality, the unanimity principle 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 .

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 (possibly inversely) dictatorial for p . In particular, F violates anonymity, which completes the proof.

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

Claim 3 actually establishes Refinement 1 above. Refinement 2 can be shown 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 The weakest possible agenda assumption for our result

We now identify the weakest agenda assumption for which our impossibility result holds. Call propositions $p, q \in X$ *indirectly connected* if there exist $p_1, \dots, p_k \in X$ with $p_1 = p$ and $p_k = q$ such that, for each $t \in \{1, \dots, k-1\}$, p_t and p_{t+1} are connected (as defined above).

Theorem 1 and Refinements 1 and 2 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 $(Y^* \setminus Z) \cup \{\neg z : z \in Z\}$ is consistent for some subset $Z \subseteq Y^*$ of even size, where (iii) some $p \in Y$ is indirectly connected to some $q \in Y^*$. If $Y = Y^*$, this reduces to the earlier definition of weak connectedness.¹¹

¹¹The 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 four conditions holds (Dietrich and List 2005).

To see that our result continues to hold if weak connectedness is weakened in this way, one needs to adapt the above proof by redefining the subagenda X^* as $\{p, \neg p : p \in Y\} \cup \{p, \neg p : p \in Y^*\} \cup \{p_t, \neg p_t : t = 1, \dots, k\}$, where p_1, \dots, 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 an aggregation rule satisfying universal domain, collective rationality, anonymity and unbiasedness.

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 $X_2 := X \setminus X_1$. (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 $Y := B \setminus \{r, s\}$, $\{r, s\} \cup Y = B$ is inconsistent, but each of $\{r\} \cup Y = B \setminus \{s\}$ and $\{s\} \cup Y = B \setminus \{r\}$ is consistent by B 's minimal inconsistency. This contradiction proves the claim.

Define the rule F with universal domain by $F(A_1, \dots, A_n) := B_1 \cup B_2$, 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 consistent by (**) (see 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, Refinement 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 $p_1, \dots, p_m \in X$ ($m \geq 1$) are such that $p = p_1$, $q = p_m$, and p_t and p_{t+1} are connected for any $t \in \{1, \dots, m-1\}$. As $p_1 \in X_1$ and $p_m \in X_2$, there must be a $t \in \{1, \dots, m-1\}$ 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 $B_2 := (\{p_{t+1}^*\} \cup Y) \cap X_2$ 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 (i), 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 $X_2 := X \setminus X_1$. 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. It 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 desired. ■

6 An illustration

To illustrate the generality of our result, we apply Theorem 1 to the aggregation of (strict) preferences, embedded into 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 $K = \{x, y, z, \dots\}$ with $|K| \geq 3$ (representing alternatives), and (ii) $A \models p$ 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.¹² (Details of the construction are given in Dietrich and List 2005; see also List and Pettit 2004.)

¹²Formally, 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 constants $x, y \in K$, $\neg x=y$.

The agenda X thus defined is weakly connected and non-separable. Also, 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 judgment aggregation rule satisfying collective rationality uniquely represents a social welfare function with strict preferences as input and output. Universal domain and anonymity become equivalent to the equally named standard conditions on a social welfare function, and 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 $(\succ_1, \dots, \succ_n)$, $(\succ_1^*, \dots, \succ_n^*)$, 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 $y \succ^* x$].

We can now state the corollary of Theorem 1 for strict preference aggregation, using Refinement 2: Every unbiased social welfare function with universal domain depends only on a single individual. 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 theorem on judgment aggregation should illustrate the theorem's generality. Interestingly, unlike Wilson's and Arrow's theorems, our result continues to hold even if the rationality conditions on preferences 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.

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 the first impossibility theorem on judgment aggregation that applies to all standard agendas in the literature and yet does not impose systematicity, a condition often criticized as being too strong. Our 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 is

inconsistent with anonymity (under universal domain and collective rationality); no responsiveness, monotonicity or unanimity condition is needed for this result. 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 implies that the collective judgment set depends only on a single individual. Finally, we have identified the weakest agenda assumption for which our result holds.

Our impossibility finding 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) Judgment aggregation in general logics. Working paper, PPM Group, University of Konstanz

Dietrich, F. (2005) Judgment Aggregation: (Im)Possibility Theorems. *Journal of Economic Theory* (forthcoming)

Dietrich, F., List, C. (2005) Arrow's theorem in judgment aggregation. Working paper, LSE

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: 680-4

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

Nehring, K., Puppe, C. (2005) Consistent Judgment Aggregation: A Characterization. Working paper, University of Karlsruhe

Pauly, M., van Hees, M. (2005) Logical Constraints on Judgment Aggregation. *Journal of Philosophical Logic* (forthcoming)

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