Synthese (2007) 158:233–250
Knowledge, Rationality & Action 323–340
DOI 10.1007/s11229-006-9079-z

ORIGINAL PAPER

Axiomatizing collective judgment sets in a minimal logical language

Marc Pauly

Received: 21 September 2005 / Accepted: 4 July 2006 / Published online: 7 June 2007

Abstract We investigate under what conditions a given set of collective judgments can arise from a specific voting procedure. In order to answer this question, we introduce a language similar to modal logic for reasoning about judgment aggregation procedures. In this language, the formula $\Box\varphi$ expresses that φ is collectively accepted, or that φ is a group judgment based on voting. Different judgment aggregation procedures may be underlying the group decision making. Here we investigate majority voting, where $\Box\varphi$ holds if a majority of individuals accepts φ, consensus voting, where $\Box\varphi$ holds if all individuals accept φ, and dictatorship. We provide complete axiomatizations for judgment sets arising from all three aggregation procedures.

Keywords Judgment aggregation · Modal logic

1 Introduction

Social choice theory has traditionally been concerned with the aggregation of individual preferences. More recently some researchers have shifted the focus from preference aggregation to judgment aggregation. The difficulties arising in judgment aggregation are illustrated by the *discursive dilemma* (also known as *doctrinal paradox*): Consider a situation with three individuals each of which has an associated judgment set, the proposition he considers to be true. If the judgment sets for the three individuals are given by $\{p, q, p \wedge q\}$, $\{p, \neg q, \neg(p \wedge q)\}$ and $\{\neg p, q, \neg(p \wedge q)\}$, then each individual is logically consistent, but judgment aggregation based on proposition-wise majority voting will produce an inconsistent set of group judgments, namely $\{p, q, \neg(p \wedge q)\}$. This simple observation has been generalized in

M. Pauly (✉)
Department of Philosophy, Stanford University,
Stanford, USA
e-mail: Marc.Pauly@stanford.edu

List & Pettit (2002) and subsequent articles to impossibility results which go beyond majority voting.

From a different perspective, the discursive dilemma shows the importance of procedural effects in social decision making. If a group needs to decide whether or not to accept a conjunction $p \wedge q$, one can imagine at least two different procedures: The standard conclusion-driven procedure simply asks for a majority vote on $p \wedge q$. An alternative premise-driven procedure would accept $p \wedge q$ precisely when both p and q separately are accepted by majority vote. The discursive dilemma shows that these two procedures may produce different results, essentially because a majority for p and a majority for q do not imply a majority for $p \wedge q$.

In this paper, a logical model for judgment aggregation procedures is developed. The premise- and conclusion-driven procedures mentioned are only two examples in a wide range of possible procedures for deciding complex issues. What is the logical relationship between these different procedures? For majority voting, for instance, accepting $p \wedge q$ on the conclusion-driven procedure implies acceptance on the premise-driven procedure, but not vice versa. If consensus voting is applied instead of majority voting, however, both procedures are equivalent. Similar logical questions we might ask are: What are the views that can be consistently held by different majorities? What can we conclude from the fact that there is a majority for φ or a majority for ψ? Can a particular set of group judgments have been obtained by majority voting?

Our aim in this paper is to develop a logical framework that allows us to characterize precisely the logical relationships that exist between different procedures involving not only majority voting, but also consensus voting and dictatorship. On the one hand, this will yield an axiomatization of these voting mechanisms in the context of judgment aggregation. On the other hand, it will allow us to compare what logical properties distinguish, e.g., consensus voting from majority voting.

The axiomatizations obtained in this paper differ from those normally presented in social choice theory, like May's characterization of majority voting (May, 1952), in that they use a formal logical language for expressing axioms. We aim for finding an axiomatization which is syntactically minimal, i.e., in a logical language which extends propositional logic in a minimal way, where we can only talk about the propositions accepted by the group or society and nothing else. Thus, we may not, for instance, refer to propositions accepted by particular individuals, the fact that at least two individuals accept a certain proposition, and so on. The advantage of such a minimalist approach is that we can state, e.g., our characterization of collective judgment sets X arising from majority voting purely in terms of the formulas that need to be in X.

The paper is structured as follows: In Sect. 2, we begin by introducing the syntax and semantics of a minimal language for describing the results of aggregation procedures. Sect. 3 presents the axiomatization results for consensus voting and dictatorship; an axiomatization for majority voting is given in Sect. 4. Finally, Sect. 5 discusses some implications of the results obtained as well as future research.

2 A formal model of collective judgments

2.1 Individual and collective formulas

Given a finite nonempty set of propositional atoms Φ_0, we define the set of *individual formulas* Φ_I as the set of all formulas α generated by the following grammar, where $p \in \Phi_0$:

$$\alpha := p \mid \neg\alpha \mid \alpha_1 \wedge \alpha_2$$

In words, the set of formulas is the smallest set of expressions that contains all the propositional atoms and is closed under negation and conjunction.

An *individual valuation* is a function v: $\Phi_0 \to \{0,1\}$, and we let V_{I} be the set of all individual valuations. In the standard way, we extend an individual valuation v by induction to a function $\hat{v}$: $\Phi_{\text{I}} \to \{0,1\}$ which assigns truth values also to complex propositions:

$$\begin{aligned}\hat{v}(p) &= v(p) \text{ for } p \in \Phi_0\\ \hat{v}(\neg\alpha) &= 1 - \hat{v}(\alpha)\\ \hat{v}(\alpha \wedge \beta) &= min(\hat{v}(\alpha), \hat{v}(\beta))\end{aligned}$$

In general, we shall usually identify v and $\hat{v}$, simply writing $v(\alpha)$ instead of $\hat{v}(\alpha)$. Note that since Φ_0 is finite, valuations can be characterized completely by individual formulas. Given an individual valuation $v \in V_{\text{I}}$, we define $[v] = \bigwedge_{\{p\in\Phi_0|v(p)=1\}} p \wedge \bigwedge_{\{p\in\Phi_0|v(p)=0\}} \neg p$. For a finite set of valuations $V \subseteq V_{\text{I}}$, we define $[V] = \bigvee_{\{v\in V\}}[v]$. Finally, we shall let $V_{\text{I}}(\alpha) = \{v \in V_{\text{I}}|v(\alpha) = 1\}$ denote all the individual valuations satisfying $\alpha \in \Phi_{\text{I}}$. Analogously for a set of formulas $\Sigma \subseteq \Phi_{\text{I}}$, we let $V_{\text{I}}(\Sigma) = \{v \in V_{\text{I}}|\forall\sigma \in \Sigma : v(\sigma) = 1\}$.

In order to talk about collective judgments, we shall use the modal $\Box$-operator which will refer to the collective or aggregate judgment on a formula: the formula $\Box\alpha$ will be considered true whenever the group accepts α. Since in principle we want to allow for arbitrary collective judgments to be made, boxed formulas will be treated like atoms which can be assigned arbitrary truth values. Formally, we define the set of *collective formulas* Φ_{C} as the set of all formulas φ generated by the following grammar, where $\alpha \in \Phi_{\text{I}}$:

$$\varphi := \Box\alpha \mid \neg\varphi \mid \varphi_1 \wedge \varphi_2$$

In words, the set of collective formulas is the smallest set of expressions that is closed under negation and conjunction and that contains all $\Box\alpha$ expressions, where α is an individual formula. For both individual and collective formulas, we use the standard abbreviations for the remaining connectives: $\top := p \vee \neg p$, $\bot := \neg\top$, $\varphi \vee \psi := \neg(\neg\varphi \wedge \neg\psi)$, $\varphi \to \psi := \neg\varphi \vee \psi$ and $\varphi \leftrightarrow \psi := (\varphi \to \psi) \wedge (\psi \to \varphi)$.

Let $\Phi_{\Box\text{I}} = \{\Box\alpha|\alpha \in \Phi_{\text{I}}\}$. Then we can also define a *collective valuation* or *model* as a function v: $\Phi_{\Box\text{I}} \to \{0,1\}$, and we can analogously extend a collective valuation to a function $\hat{v}$: $\Phi_{\text{C}} \to \{0,1\}$. We shall let V_{C} denote the set of collective valuations. A formula γ is a *collective (individual) tautology* iff $v(\gamma) = 1$ for all collective (individual) valuations v. A collective (individual) formula γ is *satisfiable* if there is a collective (individual) valuation v such that $v(\gamma) = 1$. Similarly, a set of formulas Γ is satisfiable if there is a collective (individual) valuation v such that $v(\gamma) = 1$ for all $\gamma \in \Gamma$. Analogous to individual formulas, we let $V_{\text{C}}(\varphi) = \{v \in V_{\text{C}}|v(\varphi) = 1\}$ for $\varphi \in \Phi_{\text{C}}$ and $V_{\text{C}}(\Sigma) = \{v \in V_{\text{C}}|\forall\sigma \in \Sigma : v(\sigma) = 1\}$ for $\Sigma \subseteq \Phi_{\text{C}}$.

The notion of logical consequence is defined in the standard way: φ is a *logical consequence* of Γ, denoted as $\Gamma \models \varphi$, provided $V_{\text{C}}(\Gamma) \subseteq V_{\text{C}}(\varphi)$. We say that a set of collective formulas $\Delta \subseteq \Phi_{\text{C}}$ is *sound* for a class of models $\mathcal{C}$ iff $\mathcal{C} \subseteq V_{\text{C}}(\Delta)$, and Δ is *complete* for $\mathcal{C}$ iff $V_{\text{C}}(\Delta) \subseteq \mathcal{C}$. An *axiomatization* of a class of models is a set of formulas which is both sound and complete for that class.

2.2 Examples

To give some examples of aggregation procedures expressible in this language, note that the premise-driven procedure mentioned in the introduction can be written as $\Box p \wedge \Box q$ whereas the conclusion-driven procedure can be written as $\Box(p \wedge q)$. The formula $\Box p \wedge \Box q$ will be true whenever the collective accepts p and the collective accepts q. Hence, collective decision making is applied to the premises, and the final outcome will be the conjunction of the voting outcomes on the premises. On the other hand, the formula $\Box(p \wedge q)$ applies collective decision making to the conclusion $p \wedge q$ directly. The fact that a positive result in the conclusion-driven procedure implies a positive result in the premise-driven procedure is accordingly expressed by the formula $\Box(p \wedge q) \rightarrow (\Box p \wedge \Box q)$, and hence we would expect this formula to be valid in those systems that are based on majority voting (in fact, it will be an axiom). In fact, this formula can itself be viewed as an aggregation procedure, a meta-procedure that responds to the difference between premise- and conclusion-based procedure. This aggregation procedure will return “no” only in the case where the premise-base procedure returns “no” but the conclusion-driven procedure returns “yes”.

Collective valuations (models) represent outcomes of aggregation rules applied to particular profiles of individual valuations. Consider the case described in the introduction, where we have three individuals with individual valuations v_1, v_2 and v_3 such that $v_1(p) = v_1(q) = v_2(p) = v_3(q) = 1$ and $v_2(q) = v_3(p) = 0$. Suppose that the collective valuation v is the result of applying majority voting to these individual valuations, i.e., we define $v(\Box\alpha) = 1$ iff $v_i(\alpha) = 1$ for a majority of the three individuals. Hence, we have $v(\Box p) = v(\Box q) = 1$ while at the same time $v(\Box(p \wedge q)) = 0$ since only a minority of the individuals accepts both p and q. Thus, the conclusion-based procedure yields a negative outcome while the premise-based procedure yields a positive outcome: $v(\Box p \wedge \Box q) = 1$ since $v(\Box p) = v(\Box q) = 1$. Hence, this collective valuation captures precisely the situation of the discursive dilemma.

Some more remarks concerning logical consistency: Models require logical consistency only on the non-atomic level. On the atomic level, no logical consistency is required for the collective judgment, since the atoms of a collective valuation can be assigned arbitrary values. This also points to the raison d’être of the $\Box$-operator, for it is needed to separate the logically consistent part of a collective formula from the possibly inconsistent part. Within the scope of the $\Box$-operator, no logical consistency is required. This means that group judgments are not required to be logically consistent, since $\Box(p \wedge \neg p)$ may be true. Similarly, it is not necessarily the case that $v(\Box(p \wedge q)) = 1$ if and only if $v(\Box p) = v(\Box q) = 1$. Outside of the $\Box$-operator, however, we do require logical consistency: $\Box p \wedge \neg\Box p$ can never be true.

Finally, note that not every collective valuation can occur under every method of aggregation. The example above showed that the discursive dilemma model can occur under majority voting. Under consensus voting, however, this collective valuation can never arise: the group will accept $p \wedge q$ unanimously iff it accepts both p and q unanimously (given that we assume individuals to be logically consistent). Thus, we would expect $\Box(p \wedge q) \leftrightarrow (\Box p \wedge \Box q)$ to be a valid principle for consensus voting but not for majority voting. It is these logical differences we want to get at in our axiomatization results for the different voting methods.

2.3 Model classes: consensus, majority and dictatorship

We are interested in collective valuations which arise by means of certain decision methods. For a finite set of individuals $N = \{1, \ldots, n\}$, a *decision method* $D\colon \{0,1\}^N \rightarrow \{0,1\}$ maps n individual yes/no-decisions into a collective yes/no-decision. We shall be interested in axiomatizing models based on a few specific decision methods. An n-ary decision method D is a *dictatorship* if there is some individual $d \in N$ such that $D(x_1, \ldots, x_n) = x_d$ for all $x_1, \ldots, x_n \in \{0,1\}$. *Majority voting* is the decision method D where $D(x_1, \ldots, x_n) = 1$ iff $|\{x_i | x_i = 1\}| > \frac{1}{2}n$. Note that we have chosen here for strict majority, but the axiomatization results to be presented later have their analogues for weak majority, where for an even number of voters exactly half of the voters will also suffice. Finally, we call D *consensus voting* provided that $D(x_1, \ldots, x_n) = 1$ iff all $x_i = 1$.

The terminology we applied to decision methods can also be lifted to collective valuations. A model $v \in V_C$ is *n-systematic* iff there is some n-ary decision method D and there are n individual valuations $v_1, \ldots, v_n \in V_{\mathrm{I}}$ such that for all $\alpha \in \Phi_{\mathrm{I}}$, $v(\Box\alpha) = D(v_1(\alpha), \ldots, v_n(\alpha))$. We denote the class of n-systematic models as $\mathcal{SYS}_n$. In systematic models, the group judgment on some formula will only depend on the individual judgments concerning that formula. Furthermore, the way in which the group judgment depends on the individual judgments is uniform for all formulas. The class of all majority models, i.e., n-systematic models where the decision method is majority voting, will be denoted as $\mathcal{MAJ}_n$, and the class of all consensus models based on n individuals will be denoted as $\mathcal{CON}_n$. Finally, we let $\mathcal{DIC}$ denote the class of all dictatorial models. Note that the arity of the decision method does not matter in the case of dictatorships, so there is no need to specify the number of individuals for dictatorships on the level of collective valuations.

Figure 1 summarizes the definitions of the different model classes. We occasionally want to abstract over the number of individuals, using $\mathcal{CON}$ as an abbreviation for $\bigcup_n \mathcal{CON}_n$, and similarly for the other model classes. The relationships between the different model classes are given in Theorem 1.

Theorem 1

1. *$\mathcal{DIC} \subseteq \mathcal{MAJ}_n, \mathcal{CON}_n \subseteq \mathcal{SYS}_n$, for all n.*
2. *$\mathcal{DIC} = \mathcal{MAJ}_1 = \mathcal{CON}_1$ and $\mathcal{MAJ}_2 = \mathcal{CON}_2$.*
3. *$\mathcal{MAJ}_n \nsubseteq \mathcal{CON}$, for $|\Phi_0| \geq 2$ and $n \geq 3$.*
4. *$\mathcal{CON}_n \nsubseteq \mathcal{MAJ}$, for $|\Phi_0| \geq 2$ and $n \geq 3$.*

Proof Properties 1 and 2 are easy to verify, we shall only demonstrate properties 3 and 4. For property 3, consider the situation of the discursive dilemma generalized to n individuals, where there are individual valuations $v_1, \ldots, v_n$ defined as follows: $v_i(p) = 1$ iff $i \leq \frac{1}{2}n + 1$, and $v_i(q) = 1$ iff $i \geq \frac{1}{2}n$. If $v \in V_C$ is the collective valuation associated to majority voting, we have $v(\Box p) = v(\Box q) = 1$ while $v(\Box(p \wedge q)) = 0$

$\mathcal{SYS}_n$	systematic	$v(\Box\alpha) = D(v_1(\alpha), \ldots, v_n(\alpha))$
$\mathcal{DIC}$	dictatorial	$v(\Box\alpha) = v_0(\alpha)$, for some $v_0 \in V_I$
$\mathcal{MAJ}_n$	majority	$v(\Box\alpha) = 1$ iff $|\{i \leq n | v_i(\alpha) = 1\}| > \frac{1}{2}n$
$\mathcal{CON}_n$	consensus	$v(\Box\alpha) = 1$ iff $\forall i \leq n : v_i(\alpha) = 1$

Fig. 1 Model classes

which shows that $v \notin \mathcal{CON}$, since $\Box$ distributes over conjunction for consensus voting (see axiom C in Sect. 3.2).

For property 4, we will show in the proof of Theorem 8 that every collective valuation in $\mathcal{MAJ}$ will satisfy axiom T introduced in Sect. 4. Hence, it suffices to show that there are collective valuations based on consensus voting which fail to satisfy the axiom. Consider $n \geq 3$ individual valuations $v_1, \ldots, v_n$ such that we have $v_1 \neq v_2$, $v_1 \neq v_3$ and $v_2 \neq v_3$. Let $v \in V_C$ be the collective valuation associated with these individual valuations under consensus voting. Then the following formula

$$\neg\Box\bot \wedge \neg\Box[\{v_1, v_2\}] \wedge \bigwedge_{i \leq n} \neg\Box[v_i] \wedge \neg\Box\neg[v_i]$$

is true in v. If v satisfied axiom T, this would imply that $\neg\Box\top$ is also true in v, a contradiction. Hence, v cannot satisfy axiom T, and so $v \notin \mathcal{MAJ}$. □

3 Axiomatization results

We will now axiomatize a number of classes of models. Due to the link with non-normal modal logic, our terminology shall be close to the one used in Chellas (1980). We leave the most complex case of majority voting for later. Also, it will be instructive to compare the axiomatization results obtained here with the characterization results normally obtained in social choice theory. In order to facilitate this comparison, we will also state the results obtained in a way that gets rid of logic terminology as much as possible. The question we are asking then turns out to be this: what is a set of necessary and sufficient conditions that guarantee that a set of collective judgments can be the outcome of a particular voting procedure?

3.1 Systematicity

Systematicity is the minimal requirement which all our model classes share, so it is instructive to see what axioms can enforce systematicity. Let $\mathtt{E} = \{\Box\alpha \leftrightarrow \Box\beta \in \Phi_C \mid \alpha \leftrightarrow \beta \in \Phi_I \text{ is a tautology}\}$.

Theorem 2 *$V_C(\mathtt{E}) = \mathcal{SYS}_n$, provided $n \geq 2^{|\Phi_0|}$. Hence, E also axiomatizes $\mathcal{SYS}$.*

Proof For soundness, all axioms of E are easily seen to be true in any systematic model. For completeness, suppose that $v_c(\mathtt{E}) = 1$ for some collective valuation $v_c \in V_C$ and that $n \geq 2^{|\Phi_0|}$. We need to find individual valuations $v_1, \ldots, v_n \in V_I$ and a decision method D such that $v_c(\Box\alpha) = D(v_1(\alpha), \ldots, v_n(\alpha))$ for all $\alpha \in \Phi_I$.

As for the individual valuations, we will assign every possible valuation $v \in V_I$ to some individual, i.e., we will simply order all the valuations in V_I in some way, assigning the first valuation to individual 1, the second to individual 2, etc. and the last possible valuation to all the remaining players. Note that since $n \geq 2^{|\Phi_0|}$, there are sufficiently many players so that every world view (i.e., valuation) will be represented. As for D, we define $D(x_1, \ldots, x_n) = v_c(\Box\alpha)$, in case there is some $\alpha \in \Phi_I$ such that for all $i \leq n$ we have $v_i(\alpha) = x_i$, and we let $D(x_1, \ldots, x_n) = 0$ otherwise. Given that $v_c(\mathtt{E}) = 1$, D is well-defined: if there is also a second formula $\beta \in \Phi_I$ such that for all $i \leq n$, $v_i(\beta) = x_i$, we know (since all valuations are present among the individuals) that α and β are logically equivalent, so $v_c(\Box\alpha) = v_c(\Box\beta)$. By definition, we have $v_c(\Box\alpha) = D(v_1(\alpha), \ldots, v_n(\alpha))$ for all $\alpha \in \Phi_I$. □

We can identify a collective valuation v with the set of formulas collectively accepted, i.e., with $\{\varphi | v(\Box\varphi) = 1\}$. Then with less logic terminology, we can restate the content of Theorem 2 as follows:

> A set of formulas X can be the outcome of systematic voting iff for every two logically equivalent formulas α and β, $\alpha \in X$ iff $\beta \in X$.

In other words, the result implies that if we are interested in systematic judgment aggregation, group decision making is invariant under logical equivalence, i.e., within a decision procedure φ which contains $\Box\alpha$ as a subformula, we can substitute $\Box\beta$ for $\Box\alpha$ provided that α is logically equivalent to β.

Two further remarks concerning the result. First, the restriction that there are sufficiently many individuals to express all the possible valuations ($n \geq 2^{|\Phi_0|}$) is only needed to show completeness; soundness holds for an arbitrary number of individuals. Second, completeness does fail if there are not sufficiently many individuals. For $n = 1$, the formula $\Box p \wedge \Box\neg p \rightarrow \Box q$ holds in all systematic models, while there are systematic models with $n > 1$ where the formula fails to hold. More generally, the formula $\bigvee_{v \in V_I}(\Box[v] \leftrightarrow \Box\bot)$ holds in all systematic models if $n < 2^{|\Phi_0|}$, whereas it is easy to falsify the formula in a systematic model with $n \geq 2^{|\Phi_0|}$. Hence the bound on the number of individuals given in Theorem 2 is tight.

3.2 Consensus voting

Let EMCN or K denote the set of collective formulas containing E and all instances of the following three axiom schemes:

$$\text{M.}\quad \Box(\alpha \wedge \beta) \rightarrow (\Box\alpha \wedge \Box\beta)$$
$$\text{C.}\quad (\Box\alpha \wedge \Box\beta) \rightarrow \Box(\alpha \wedge \beta)$$
$$\text{N.}\quad \Box\top$$

If, in addition, the axiom D $\neg\Box\bot$ is added, we obtain the set of formulas KD.

Lemma 3 (Monotonicity) EM $\models \Box\alpha \rightarrow \Box\beta$, *provided* $\alpha \rightarrow \beta$ *is a tautology.*

Proof If $\alpha \rightarrow \beta$ is a tautology, then so is $\alpha \leftrightarrow (\alpha \wedge \beta)$. Consequently, any model v satisfying E will satisfy $\Box\alpha \leftrightarrow \Box(\alpha \wedge \beta)$, and applying the M axiom v must also satisfy $\Box\alpha \rightarrow \Box\beta$. □

Lemma 4 *For any set* $\Gamma \subseteq \Phi_I$, *if every finite subset of* Γ *is satisfiable, then so is* Γ.

Proof While this result is simply the compactness theorem for propositional logic treated in most logic textbooks, in our case, the result can be obtained in a simpler manner. Suppose that Γ is not satisfiable. Due to the fact that Φ_0 is finite, there must be a finite $\Gamma_0 \subseteq \Gamma$ such that for every $\alpha \in \Gamma$ there is some $\beta \in \Gamma_0$ such that α and β are logically equivalent. Hence, Γ_0 cannot be satisfiable. □

Theorem 5 $V_C(\text{KD}) = \mathcal{CON}_n$, *provided* $n \geq 2^{|\Phi_0|}$. *Hence,* KD *also axiomatizes* $\mathcal{CON}$.

Proof The proof is analogous to the proof of Theorem 2. Soundness is easy to check, and for completeness, let $n \geq 2^{|\Phi_0|}$ and suppose that $v_c(\text{KD}) = 1$ for some collective valuation $v_c \in V_C$. We will construct individual valuations $v_1, \ldots, v_n \in V_I$ such that $v_c(\Box\alpha) = 1$ iff for all $i \leq n$ we have $v_i(\alpha) = 1$.

Consider the sets $\Gamma = \{\alpha \in \Phi_I | v_c(\Box\alpha) = 1\}$ and $W = \{v \in V_I | \forall\alpha : v_c(\Box\alpha) = 1 \Rightarrow v(\alpha) = 1\}$. Since $v_c(\Box\top) = 1$ by axiom N, we know that Γ is nonempty. Also W must be nonempty. For suppose to the contrary that W were empty. Then there is no individual valuation satisfying all of Γ, and so by Lemma 4, there is some finite $\Gamma_0 = \{\gamma_1, \ldots, \gamma_m\} \subseteq \Gamma$ such that $\bigwedge \Gamma_0 \to \bot$ is an individual tautology. By Lemma 3, this implies that $v_c(\Box \bigwedge \Gamma_0 \to \Box\bot) = 1$, and by axiom C, $v_c((\Box\gamma_1 \wedge \ldots \wedge \Box\gamma_m) \to \Box\bot) = 1$. Consequently, $v_c(\Box\bot) = 1$, a contradiction with axiom D. Hence, W is indeed nonempty.

Now we will assign every valuation $v \in W$ to some individual, and since $n \geq 2^{|\Phi_0|}$, there are sufficiently many players to cover all valuations in W.

It remains to show that $v_c(\Box\alpha) = 1$ iff $\forall i \leq n$ we have $v_i(\alpha) = 1$. First, if $v_c(\Box\alpha) = 1$, then we have $v_i(\alpha) = 1$ by definition of W for all $i \leq n$. Conversely, suppose that $v_c(\Box\alpha) = 0$. In this case it suffices to find a single individual valuation $v_i \in W$ such that $v_i(\alpha) = 0$. For this, it suffices to show that $\Gamma \cup \{\neg\alpha\}$ is satisfiable. Suppose again to the contrary that it is not, then by Lemma 4 there must be a finite $\Gamma_0 = \{\gamma_1, \ldots, \gamma_m\} \subseteq \Gamma$ such that $\bigwedge \Gamma_0 \to \alpha$ is an individual tautology. Then by Lemma 3 and axiom C, $v_c((\Box\gamma_1 \wedge \ldots \wedge \Box\gamma_m) \to \Box\alpha) = 1$, and hence $v_c(\Box\alpha) = 1$, a contradiction. Hence, $\Gamma \cup \{\neg\alpha\}$ is indeed satisfiable, and so there must be some individual valuation $v_i \in W$ such that $v_i(\alpha) = 0$. □

As in the case of systematicity, completeness may fail if there are not enough individuals present. The formula $\bigvee_{v \in V_I} \Box[V_I - \{v\}]$ holds in all consensus models if $n < 2^{|\Phi_0|}$, but not necessarily if $n \geq 2^{|\Phi_0|}$. Hence, also the bound in Theorem 5 is tight. In terms of the formulas collectively accepted, Theorem 5 can be restated as follows:

> A set of formulas X can be the outcome of consensus voting iff (1) $\alpha \in X$ iff $\beta \in X$, whenever α and β are logically equivalent, (2) $\top \in X$, (3) $\bot \notin X$, and (4) $\alpha \wedge \beta \in X$ iff $\alpha \in X$ and $\beta \in X$.

Note that this result does not say that a set of formulas satisfying the four conditions mentioned *must* be the outcome of consensus voting. The set may equally well have been obtained, e.g., by a dictatorship (recall Theorem 1).

Note that collective judgments that arise from consensus voting must be logically consistent, but they may be incomplete, i.e., there may be a formula φ such that neither $\varphi \in X$ nor $\neg\varphi \in X$. This will be the case whenever there is neither a consensus for φ nor for $\neg\varphi$. Our next voting method, by contrast, will guarantee complete collective judgment sets.

3.3 Dictatorship

Let MCY denote the set of collective formulas containing M, C and all instances of the following axiom

$$\texttt{Y}. \quad \neg\Box\alpha \leftrightarrow \Box\neg\alpha.$$

Note first that MCY $\models$ E: Suppose that $v(\texttt{MCY}) = 1$. If $\alpha \leftrightarrow \beta \in \Phi_I$ is an individual tautology, then the collective formula φ, obtained by replacing every propositional atom $p \in \Phi_0$ in $\alpha \leftrightarrow \beta$ by $\Box p$, must be a collective tautology, and hence $v(\varphi) = 1$. But given the axioms v satisfies, $v(\varphi) = v(\Box\alpha \leftrightarrow \Box\beta)$, and hence axiom E is indeed a logical consequence of MCY.

Furthermore, $\mathtt{MCY} \models \mathtt{N}$: In the presence of axiom $\mathtt{Y}$, $\neg(\Box p \wedge \Box\neg p)$ must be valid and using axiom $\mathtt{C}$, $\neg\Box\bot$ must be valid which yields $\mathtt{N}$ using again axiom $\mathtt{Y}$. Note finally that we also have $\mathtt{MCY} \models \mathtt{D}$.

Theorem 6 *$V_C(\mathtt{MCY}) = \mathcal{DIC}$, i.e., $\mathtt{MCY}$ axiomatizes $\mathcal{DIC}$.*

Proof Soundness is easy to check. For completeness, given a model $v_c \in V_C$ satisfying $\mathtt{MCY}$, we simply need to find an individual valuation $v_1 \in V_I$ such that $v_c(\Box\alpha) = v_1(\alpha)$ for all $\alpha \in \Phi_I$. Simply define $v_1(p) = v_c(\Box p)$ for all $p \in \Phi_0$. We then verify by induction on α that $v_1(\alpha) = v_c(\Box\alpha)$ for all $\alpha \in \Phi_I$. The base case holds by definition, the inductive step for conjunction follows from axioms $\mathtt{M}$ and $\mathtt{C}$ and the inductive step for negation from axiom $\mathtt{Y}$. □

Intuitively, axioms $\mathtt{M}$, $\mathtt{C}$ and $\mathtt{Y}$ together require the group decision to be logically consistent. A dictatorship, being based on a logically consistent individual decision maker, clearly satisfies these requirements. Conversely, any logically consistent group decision can be represented by a dictatorship based on an individual who holds precisely the group's views. This reasoning is formalized in Theorem 6. Since in a distributive logic the $\Box$ operator distributes over all connectives, there are no procedural effects for dictatorship, i.e., no matter how a decision problem for a complex formula is proceduralized, the outcome will always be the same.

Like for the other voting methods, we can also restate Theorem 6 in a more intuitive manner:

> A set of formulas X can be the outcome of a dictatorial voting process iff (1) $\neg\alpha \in X$ iff $\alpha \notin X$, and (2) $\alpha \wedge \beta \in X$ iff $\alpha \in X$ and $\beta \in X$.

There is an immediate consequences we can draw from this characterization result: it follows that sets of formulas that arise from dictatorial voting must not only be logically consistent, but also complete, i.e., for any formula φ, the collective must accept either φ or $\neg\varphi$.

3.4 A simple example

Due to the axiomatization results obtained, we can now also easily derive some intuitive consequences from the axioms considered so far. We will only give one simple example taken from Snyder ("An axiomatic approach to judgment aggregation." Unpublished manuscript, 2006).

Consider the set of axioms $\Delta = \mathtt{EMCDV}$, where $\mathtt{V}$ is the axiom scheme $\Box\alpha \vee \Box\neg\alpha$. These axioms express that decision making is systematic ($\mathtt{E}$) and yields deductively closed ($\mathtt{MC}$) logically consistent ($\mathtt{D}$) and complete ($\mathtt{V}$) judgments. These axioms have a high degree of normative appeal. However, it turns out that $\Delta \models \mathtt{Y}$: If $v(\Box\neg\alpha) = v(\Box\alpha) = 1$, then by axiom $\mathtt{C}$, $v(\Box(\neg\alpha \wedge \alpha)) = 1$, so using axiom $\mathtt{E}$, this means that $v(\Box\bot) = 1$, contradicting axiom $\mathtt{D}$. Conversely, if $v(\neg\Box\alpha) = 1$, by axiom $\mathtt{V}$, we must have $v(\Box\neg\alpha) = 1$ as desired.

This little argument showing that $\mathtt{EMCDV} \models \mathtt{Y}$, and hence $\mathtt{EMCDV} \models \mathtt{MCY}$, in fact establishes that every collective valuation satisfying $\mathtt{EMCDV}$ must be in $\mathcal{DIC}$, by Theorem 6. This means that any systematic, consistent, complete and deductively closed collective judgment set is obtainable from a dictatorship. This should not come as much of a surprise, but it still serves to illustrate the simple deductive reasoning that can be carried out with our language.

This result also points out the difference between axiomatizing aggregation functions and axiomatizing collective valuations, the output of aggregation functions. An aggregation function analogue to our result would claim that an aggregation function that is systematic and always produces consistent, complete and deductively closed judgment sets must be a dictatorship. Such a result has been obtained in Pauly and van Hees (2006), but is very different from the result we have obtained here for collective valuations and much more surprising. Further comments on the relationship between axiomatizing aggregation functions vs. collective valuations can be found in Pauly ("On the role of language in social choice theory." Unpublished manuscript, 2006).

4 Axiomatizing majority voting

4.1 Simple games with ties

In order to obtain a complete axiomatization of $\mathcal{MAJ}$ we will make use of a characterization result for simple games with ties. A *simple game with ties* (SGT, also known as a prehypergraph with ties in Taylor and Zwicker (1992)) $G = (N, W, T, L)$ consists of a finite nonempty set of individuals N and sets of winning (W), tied (T) and losing (L) coalitions, where $W, T, L \subseteq \mathcal{P}(N)$ and these sets are pairwise disjoint. We call an SGT (N, W, T, L) *weighted* iff there exists a weight function $w : N \to \mathbb{R}$ and a threshold or quota $q \in \mathbb{R}$ such that for all $X \subseteq N$,

1. if $X \in W$ then $\sum_{x \in X} w(x) > q$,
2. if $X \in T$ then $\sum_{x \in X} w(x) = q$, and
3. if $X \in L$ then $\sum_{x \in X} w(x) < q$.

The UN security council and the original European Economic Community (EEC) provide examples of weighted voting games. In the EEC, for instance, France, West Germany, and Italy each had a weight of four votes, Belgium and the Netherlands had two votes and Luxembourg one vote. Every motion which got at least 12 of the 17 possible votes passed.

We call an SGT $G = (N, W, T, L)$ *k-trade robust* iff the following condition holds: For all sequences of coalitions $\langle X_1, \ldots, X_k \rangle$ and $\langle Y_1, \ldots, Y_k \rangle$ such that (1) for every $p \in N$, $|\{i : p \in X_i\}| = |\{i : p \in Y_i\}|$, (2) for every $i \leq k$ we have $X_i \in W \cup T$, and (3) for every $i \leq k$ we have $Y_i \in L \cup T$, we have $X_i \in T$ and $Y_i \in T$ for every $i \leq k$. G is *trade robust* iff it is k-trade robust for all k. Intuitively, suppose we have k non-losing coalitions and k non-winning coalitions. The non-losing coalitions can be obtained from the non-winning coalitions by trading, i.e., if an individual occurs x times in the non-losing coalitions, he will appear also x times in the non-winning coalitions. In this situation, all the coalitions involved must be tied.

Roughly speaking, trade robustness demands that you cannot move from a winning situation (where all sets are winning or tied) to a losing situation (where all sets are tied or losing) by simply trading members among the sets involved, unless all these sets are ties. To illustrate the notion of trade robustness, consider the following example: $N = \{1, 2, 3\}$, W consists of all sets containing two elements, $T = \{\{1, 2, 3\}\}$ and L contains all remaining sets. Then this game is not trade robust: Let $X_1 = \{2, 3\}$, $X_2 = \{1, 3\}$ and $X_3 = \{1, 2, 3\}$, and so all these sets are tied or winning. By trading members among the Xs, we can then obtain $Y_1 = Y_2 = \{1, 2, 3\}$ and $Y_3 = \{3\}$, all of which are tied or losing. Thus, this contradicts trade robustness since $\{2,3\}$ is in fact winning rather than tied. We can re-establish trade robustness by slightly changing the game: simply take $\{1,2,3\}$ to be winning rather than tied. In that case, the Ys given fail to be a counterexample given that Y_1 and Y_2 are now winning rather than tied. The fact that this modified game is indeed trade robust will follow from the following theorem and the observation that the modified game (in contrast to the original game) is weighted: assign weight 1 to each element of N and take the threshold q to be $3/2$.

The following combinatorial result due to Zwicker (pers. commun) generalizes a result of Taylor and Zwicker (1992) for simple games (without ties). Since the completeness theorem for majority voting relies essentially on this highly non-trivial result, we present its proof in the appendix in order to make the exposition self-contained, and also because the proof has not been published before.

Theorem 7 (Zwicker, pers. commun) *For any simple game with ties* $G = (N, W, T, L)$*, the following are equivalent:*

(i) *G is weighted.*
(ii) *G is trade robust.*
(iii) *G is 2^k-trade robust, with $k = 2^{|N|}$.*

4.2 Axiomatization

For the purposes of formulating our axiom system, it will be useful to define a number of abbreviations referring to non-strict majority, ties, etc. We will use $[>]\varphi$ for $\Box\varphi$ simply to remind the reader that our basic modality refers to strict majority. Furthermore, we define $[=]\varphi$ as $\neg\Box\varphi \wedge \neg\Box\neg\varphi$, $[\geq]\varphi$ as $[>]\varphi \vee [=]\varphi$, $[\leq]\varphi$ as $\neg[>]\varphi$ and $[<]\varphi$ as $\neg[\geq]\varphi$.

Let `STEM` denote the set of collective formulas containing `M`, `E` and all instances of the following two axiom schemes

`S.` $[>]\alpha \rightarrow \neg[>]\neg\alpha$

`T.` $([\geq]\alpha_1 \wedge \ldots \wedge [\geq]\alpha_k \wedge [\leq]\beta_1 \wedge \ldots \wedge [\leq]\beta_k) \rightarrow \bigwedge_{1\leq i\leq k} ([=]\alpha_i \wedge [=]\beta_i)$

where $\forall v \in V_{\mathrm{I}} : |\{i : v(\alpha_i) = 1\}| = |\{i : v(\beta_i) = 1\}|$

Axiom `T` is easily seen to express trade-robustness, and axiom `S` states that there can be no strict majority for φ and its negation at the same time.

Theorem 8 `STEM` *axiomatizes* $\mathcal{MAJ}$*, i.e.,* $V_{\mathrm{C}}(\texttt{STEM}) = \mathcal{MAJ}$.

Proof As for soundness, all axioms except `T` are straight forward to verify. For trade-robustness, consider a model v_c based on individual valuations $v_1, \ldots, v_n$ and majority voting. In order to obtain a contradiction, suppose that `T` is false in v_c, i.e., without loss of generality, one of the majorities for some α_j is strict. Then

$$\sum_{i\leq k}\sum_{p\leq n} v_p(\alpha_i) > k \cdot \frac{1}{2}n \geq \sum_{i\leq k}\sum_{p\leq n} v_p(\beta_i).$$

But since the trading property $\forall v \in V_{\mathrm{I}} : |\{i : v(\alpha_i) = 1\}| = |\{i : v(\beta_i) = 1\}|$ implies that $\sum_{p\leq n}\sum_{i\leq k} v_p(\alpha_i) = \sum_{p\leq n}\sum_{i\leq k} v_p(\beta_i)$, we have a contradiction.

For completeness, consider any model $v_c \in V_{\mathrm{C}}$ satisfying `STEM`, and consider the simple game with ties $G = (V_{\mathrm{I}}, W, T, L)$ where $W = \{V_{\mathrm{I}}(\alpha) | v_c(\Box\alpha) = 1\}$. For the tied and losing coalitions, we define $L = \{V_{\mathrm{I}}(\alpha)|v_c(\Box\neg\alpha) = 1\}$ and $T = \{V_{\mathrm{I}}(\alpha)|v_c(\Box\alpha) = v_c(\Box\neg\alpha) = 0\}$. Note that due to axiom S, sets W and L are disjoint, and hence G is a well defined SGT. Axiom T guarantees that G is trade robust, and hence by applying Theorem 7 we know that G is weighted, with some weight function $w\colon V_{\mathrm{I}} \to \mathbb{R}$ and some threshold $q \in \mathbb{R}$. We now proceed to show a few properties of the weight function w and the threshold q.

First, we show that we can assume that $q = \frac{1}{2}\sum_{v\in V_{\mathrm{I}}} w(v)$. If $v_c(\Box\alpha) = 1$, $V_{\mathrm{I}}(\alpha) \in W$ and $V_{\mathrm{I}}(\neg\alpha) \in L$, and hence $\sum_{v\in V_{\mathrm{I}}:v(\alpha)=1} w(v) > q > \sum_{v\in V_{\mathrm{I}}:v(\alpha)=0} w(v)$, showing that $\sum_{v\in V_{\mathrm{I}}:v(\alpha)=1} w(v) > \frac{1}{2}\sum_{v\in V_{\mathrm{I}}} w(v)$. The case where $v_c(\Box\neg\alpha) = 1$ is analogous. Finally, if $v_c(\Box\alpha) = v_c(\Box\neg\alpha) = 0$, $V_{\mathrm{I}}(\alpha)$ is a tie, and hence $\sum_{v\in V_{\mathrm{I}}:v(\alpha)=1} w(v) = q = \sum_{v\in V_{\mathrm{I}}:v(\alpha)=0} w(v)$, showing that $\sum_{v\in V_{\mathrm{I}}:v(\alpha)=1} w(v) = \frac{1}{2}\sum_{v\in V_{\mathrm{I}}} w(v)$.

Second, we show that we can assume the weight function to be nonnegative, i.e., for all $v \in V_{\mathrm{I}}$, $w(v) \geq 0$. We distinguish two cases. (i) Suppose that v is irrelevant, i.e., for all $\alpha \in \Phi_{\mathrm{I}}$, $v_c(\Box\alpha) = v_c(\Box(\alpha \vee [v]))$. In that case, the presence of v never matters, and we can take $w(v) = 0$. (ii) If v is relevant, there is some $\alpha \in \Phi_{\mathrm{I}}$ such that $v_c(\Box\alpha) \neq v_c(\Box(\alpha \vee [v]))$. Using the monotonicity axiom M we know that $v_c(\Box\alpha) < v_c(\Box(\alpha \vee [v]))$, hence $w(v) > 0$.

Third, note that we can assume all weights and the threshold q to be integers. Due to the discreteness of the domain of the weight function w, all weights and the threshold can be assumed to be rational numbers, and hence we only need to multiply these by a sufficiently high integer to obtain integer weights and threshold.

Hence, taking these three observations together, we know that there exists a weight function $w\colon V_I \to \mathbb{N}$ such that

$$v_c(\Box\alpha) = 1 \text{ iff } \sum_{v\in V_{\mathrm{I}}:v(\alpha)=1} w(v) > \frac{1}{2}\sum_{v\in V_{\mathrm{I}}} w(v).$$

Now we can consider as individual valuations $v_1,\ldots,v_n$ precisely all those valuations v for which $w(v) > 0$, and we let the number of individuals with world-view v equal $w(v)$. Hence, in total, we have $\sum_{v\in V_I} w(v)$ individual valuations. If D is the decision method of majority voting, then we have for all $\alpha \in \Phi_{\mathrm{I}}$

$$D(v_1(\alpha),\ldots,v_n(\alpha)) = 1 \text{ iff } \sum_{v\in V_{\mathrm{I}}:v(\alpha)=1} w(v) > \sum_{v\in V_{\mathrm{I}}:v(\alpha)=0} w(v) \text{ iff } v_c(\Box\alpha) = 1,$$

thereby showing that v_c is a majority model based on individual valuations $v_1,\ldots,v_n$. □

Note that in contrast to the previous axiomatization results, Theorem 8 does not provide a characterization of $\mathcal{MAJ}_n$. Hence, the result is weaker than the others in the sense that we do not know beforehand how many individuals there need to be. The result can be thought of as a result *in the limit*. The model constructed in the proof will usually have a number of individuals exponential in $2^{|\Phi_0|}$, since the weights and quota obtained after transformation into integers may be exponential in $2^{|\Phi_0|}$ (cf. the results in Anthony (2003) on linear threshold units).

Finally, we also restate Theorem 8 in an alternative manner. Due to the need to distinguish weak from strict majorities, we will let $A = \{\alpha|v(\Box\alpha) = 1\}$ stand for the formulas accepted by the group, $R = \{\alpha|\neg\alpha \in A\}$ for the formulas rejected, and $T = \{\alpha|\alpha \notin A \cup R\}$ for the formulas neither rejected nor accepted (ties). Then we can formulate Theorem 8 as follows:

Consider sets of formulas A, T and R representing the formulas collectively accepted, tied and rejected, respectively. These sets can be the outcome of majority voting iff (1) $\alpha \in A$ iff $\beta \in A$, provided α and β are logically equivalent, (2) if $\alpha \wedge \beta \in A$ then $\alpha \in A$ and $\beta \in A$, (3) if $\alpha \in A$ then $\neg\alpha \notin A$, and (4) for all sets of formulas $X_1, \ldots, X_k$ and $Y_1, \ldots, Y_k$, if for all $i \leq k$, $X_i \in A \cup T$ and $Y_i \in T \cup R$, and for every v, $|\{i : v \in X_i\}| = |\{i : v \in Y_i\}|$, then for all $i \leq k$ we must have $X_i, Y_i \in T$.

Note that for majority voting, collective judgment sets need to be neither complete (due to the possibility of ties in case there are an even number of voters) nor logically consistent (as the discursive dilemma illustrates).

5 Conclusion

The present paper has provided axiomatizations of the possible outcomes of majority voting, consensus voting and dictatorship in the context of judgment aggregation. In other words, given a set of formulas X, these results tell us when X can arise, for instance, from majority voting. In contrast to standard axiomatization results in social choice theory, all axioms are expressed in a formal logical language (similar to modal logic), a language which is syntactically minimal in the sense that the language only allows for expressing what propositions have been accepted or rejected by the collective.

With respect to majority voting, the result obtained differs both from the classic result of May (1952) and from more recent attempts to define a modal logic of majority (Pacuit & Salame, 2004). The logic of majority considered in Pacuit and Salame (2004) aims at adding a majority operator to graded modal logic. The models considered are standard Kripke models rather than the models considered here, and infinite models are allowed. Furthermore, the language considered is much richer than the minimal language considered here, and no comparison between different voting procedures is made. As a consequence, the results obtained are at present more relevant to modal logic than to the axiomatization of judgment aggregation considered here.

When comparing Theorem 8 to May's classic characterization of majority voting (May, 1952), a number of differences become apparent. First, we focus on judgment aggregation rather than preference aggregation. Secondly, we did not axiomatize aggregation rules but rather collective valuations, instances of aggregation rules, collective opinions that may arise based on majority voting. The third and most important difference, however, is the use of a formal logical language to express our axioms. The advantage of this approach is that different axiomatizations can be compared along a new dimension, namely the richness of the language used in the axioms, or more generally, the concepts expressible in the formal language. In fact, it even becomes possible to show that something is *not* axiomatizable using a certain set of concepts, i.e., using a particular formal language. This issue of informal vs. formal language as well as the question what semantic concept to axiomatize both concern the methodology of social choice theory, discussed in detail in Pauly ("On the role of language in social choice theory." Unpublished manuscript, 2006).

There are a number of open issues arising out of the results presented in this paper. First, Theorems 2 and 5 require a minimum number of individuals to be present. If there are too few individuals, additional properties hold, as we showed. In other words, the axiomatization given is complete only for a big group of individuals. While these two result provide a bound, Theorem 8 is really a result in the limit. A closer analysis of the proof can yield an upper bound (Perry, L. "Strengthening the axiomatization of majority voting in MJAL." Unpublished manuscript, 2006), but this upper bound is extremely high. It remains to find out whether a smaller bound can be found and whether any *interesting* additional axioms become valid for small groups of individuals.

Second, we implicitly assumed that the agenda, i.e., the set of propositions considered by the decision makers, includes the whole language. This choice was mainly made for the sake of mathematical elegance. But while there are certainly important cases where this assumption is not unrealistic (e.g., merging large databases), in many practical situations, both the individuals and the group will only care about a relatively small number of propositions. Hence, we could introduce the agenda as an explicit parameter into the model and results.

Third, one might want to consider 3-valued rather than 2-valued logic. This would allow individuals to abstain on propositions, where the third truth value represents abstention. More generally, the framework may be generalized from propositional logic to a very abstract logical framework like the one used in Dietrich and List (forthcoming), which also covers predicate logic and modal logic. This would allow us to model cases of judgment aggregation where the propositions involved are more complex.

Acknowledgements I would like to thank Bill Zwicker for his permission to incorporate Theorem 7 together with its proof into the paper, and Christian List, Franz Dietrich, Josep Freixas and three anonymous reviewers for their helpful comments and suggestions.

Appendix

The following theorem generalizes the result in Taylor and Zwicker (1992) from simple games to simple games with ties. The proof closely follows Zwicker (pers. commun).

Theorem 7 *For any simple game with ties $G = (N, W, T, L)$, the following are equivalent:*

(i) *G is weighted.*
(ii) *G is trade robust.*
(iii) *G is 2^k-trade robust, with $k = 2^{|N|}$.*

Proof (*i*) $\Rightarrow$ (*ii*): Consider a weighted game $G = (N, W, T, L)$ with weight function w and quota q. Suppose that there are two sequences $\langle X_1, \ldots, X_k\rangle$ and $\langle Y_1, \ldots, Y_k\rangle$ which satisfy the antecedent of the trade-robustness condition. Since $w(X_i) \geq q$ and $w(Y_i) \leq q$ for every i, we have $q \leq \frac{1}{k}\sum_{i\leq k} w(X_i) = \frac{1}{k}\sum_{i\leq k} w(Y_i) \leq q$, where the equality is due to the fact $|\{i : p \in X_i\}| = |\{i : p \in Y_i\}|$ for every $p \in N$. Hence, the average weight of the Xs and the average weight of the Ys are both q. But then if some X_i were actually winning, the average weight of the Xs would have to be strictly bigger than q, analogously if some Y_i were actually losing.

The proof that (*ii*) $\Rightarrow$ (*iii*) is trivial, so it only remains to show that (*iii*) $\Rightarrow$ (*i*). We inductively construct the weighting in such a way that at each stage, the unweighted part $N - A$ acts like it is reasonably trade robust, and the weighted part A behaves as if it were part of a correct weighting. The following definition formalizes this. □

Definition Suppose that $G = (N, W, T, L)$ is a simple game with ties where $N = \{1, \ldots, n\}$. If $A \subseteq N$ and $f\colon A \to \mathbb{R}$, then we call *f trade robust for A* iff the following holds: Whenever $k \leq 2^s$, where $s = 2^{|N|-|A|} - 1$, and $\langle X_1 \cup Y_1, \ldots, X_k \cup Y_k \rangle$ and $\langle X'_1 \cup Y'_1, \ldots, X'_k \cup Y'_k \rangle$ are two sequences of (not necessarily distinct) coalitions satisfying

(1) $\forall i \leq k : \ X_i \cap Y_i = X'_i \cap Y'_i = \emptyset$
(2) $\forall i \leq k : \ Y_i, Y'_i \subseteq A$
(3) $\sum_{i=1}^{k} \sum_{p \in Y_i} f(p) \ \leq \ \sum_{i=1}^{k} \sum_{p \in Y'_i} f(p)$
(4) $\forall p \in N : \ |\{i : p \in X_i\}| = |\{i : p \in X'_i\}|$
(5) $\forall i \leq k : \ X_i \cup Y_i \in W \cup T$ and $X'_i \cup Y'_i \in L \cup T$,
then
(A) $\forall i \leq k : \ X_i \cup Y_i \in T$
(B) $\forall i \leq k : \ X'_i \cup Y'_i \in T$
(C) $\sum_{i=1}^{k} \sum_{p \in Y_i} f(p) \ = \ \sum_{i=1}^{k} \sum_{p \in Y'_i} f(p)$.

Note first that if G is $2^{2^{|N|}}$ trade robust, then the empty function is trade robust for $A = \emptyset$. Second, if G is trade robust for $A = N$, then G must be weighted. To see this, observe that G is weighted if winning coalitions weigh more than both tied and losing coalitions, tied coalitions weigh more than losing coalitions, and all tied coalitions have the same weight. In this case, one may choose the quota q as the weight of a tied coalition, thereby showing that G is weighted. Now if G is trade robust for $A = N$, we can apply the definition for $k = 1$ to the sequences $\langle \emptyset \cup Y_1 \rangle$ and $\langle \emptyset \cup Y'_1 \rangle$ in the cases just mentioned, e.g., with $Y_1 \in W$ and $Y'_1 \in T$. Given these observations, in order to prove the theorem, it suffices to show the following:

Main Claim: Suppose that $G = (N, W, T, L)$ is a simple game with ties, $A \subseteq N$, f is trade robust for A and $a \in N - A$. Then there exists a $c \in \mathbb{R}$ such that $f \cup \{(a, c)\}$ is trade robust for $A \cup \{a\}$.

The idea behind the following argument is that when a real number fails to be an appropriate weight for the new individual a to be added, it can be classified as being either too light or too heavy. After showing that every failure is indeed either a low or a high failure (but never both), a number of claims are proved about these failures. In the end, we will have shown the existence of a weight which is less than all the high failures and greater than all the low failures. We now proceed to make this argument precise.

We call a number $c \in \mathbb{R}$ a *low failure of type A* if there exist sequences $\langle X_1 \cup Y_1, \ldots, X_k \cup Y_k \rangle$ and $\langle X'_1 \cup Y'_1, \ldots, X'_k \cup Y'_k \rangle$ showing that $f \cup \{(a, c)\}$ is not trade robust for $A \cup \{a\}$ such that $|\{i : a \in Y_i\}| > |\{i : a \in Y'_i\}|$, clauses (1)–(5) of the definition are satisfied, and clause (A) of the definition is violated. Intuitively, c is chosen too low as a weight for a, and this is exploited in the witnessing sequences by using a excessively in the winning coalitions. Low failures of type B and C are defined similarly; a number can be a low failure of more than one type. Finally, we analogously define a *high failure* of type A, B and C, if violating sequences exist with the inequality reversed.

Claim 1: Every failure is either a high or a low failure.
Proof: Suppose c is a failure, but neither high nor low. Then if two sequences of coalitions witness the failure, there is an equal number of a among the Y_i and the Y'_i. But then the occurrences of a can be shifted from Y_i to X_i and from Y'_i to X'_i with

conditions (3) and (4) of the definition preserved. This would show that f is not trade robust for A, a contradiction.

Claim 2: $c \in \mathbb{R}$ is a low failure iff there are witnessing sequences for which a occurs in none of the Y_i'.
Proof: Given any two sequences showing that c is a low failure, shift each occurrence of a among the Y_i' to X_i', and an equal number of occurrences of a among Y_i to X_i. Again, this modification preserves (3) and (4) of the definition, so the resulting sequences are as desired and still witness the low failure of c.

Claim 3: $c \in \mathbb{R}$ is a high failure iff there are witnessing sequences for which a occurs in none of the Y_i.
Proof: Analogous to the proof of claim 2.

Claim 4: If $c \in \mathbb{R}$ is a low failure and $c' < c$, then c' is also a low failure.
Proof: Sequences witnessing a low failure c will also witness a low failure c', given that condition (3) is preserved by decreasing the failure.

Claim 5: If $c \in \mathbb{R}$ is a high failure and $c' > c$, then c' is also a high failure.
Proof: Analogous to the proof of claim 4.

Claim 6: No failure can be both a high and a low failure.
Proof: Suppose that c is a low failure of type α as witnessed by

$$\langle X_1 \cup Y_1, \ldots, X_k \cup Y_k \rangle \text{ and } \langle X_1' \cup Y_1', \ldots, X_k' \cup Y_k' \rangle,$$

chosen as in claim 2, and a high failure of type β as witnessed by

$$\langle U_1 \cup V_1, \ldots, U_l \cup V_l \rangle \text{ and } \langle U_1' \cup V_1', \ldots, U_l' \cup V_l' \rangle$$

chosen as in claim 3. Note also that $k, l \leq 2^z$ with $z = 2^{|N-(A\cup\{a\})|} - 1$. Let $|\{i : a \in Y_i\}| = s$ and $|\{i : a \in V_i'\}| = t$. We now construct two new sequences of coalitions, first the sequence of unprimed coalitions where we repeat each unprimed coalition $X_i \cup Y_i$ t times and $U_i \cup V_i$ s times, second the sequence of primed coalitions where we repeat each primed coalition $X_i' \cup Y_i'$ t times and $U_i' \cup V_i'$ s times. In these new sequences, a occurs $s \cdot t$ times among both Y_i and V_i', and not at all among Y_i' and V_i. Hence, we can shift a from the Y_is to the X_is and from the V_i's to the U_i's while preserving conditions (3) and (4) of the definition. Also, a violation of condition (A), (B), or (C) by the original sequences will carry over to the new combined sequences. Finally, the length of the new sequences is at most $2 \cdot (2^z)^2$, where $z = 2^{|N-(A\cup\{a\})|} - 1$ which is equivalent to 2^w with $w = 2^{|N-A|} - 1$. Consequently, f cannot be trade robust for A, a contradiction.

Claim 7: Let $c = 2 \cdot 2^{2^{|N|}} \cdot \sum_{p \in A} |f(p)|$. If $c' > c$, then c' is not a low failure. If $c' < -c$, then c' is not a high failure.
Proof: Assume that c' is a low failure with $c' > c$ and witnessing sequences chosen as in claim 2. Since a appears among the Y_i but not among the Y_i',

$$\sum_{i=1}^{k} \sum_{p \in Y_i} f(p) > c - 2^{2^{|N|}} \sum_{p \in A} |f(p)| = 2^{2^{|N|}} \sum_{p \in A} |f(p)| \geq \sum_{i=1}^{k} \sum_{p \in Y_i'} f(p),$$

thus contradicting condition (3). The argument for the case where $c' < -c$ is analogous.

Claim 8: The low failures are bounded above, and hence there is a supremum c_L of the low failures. Analogously, the high failures are bounded below, and hence there is an infimum c_H of the high failures.
Proof: A direct consequence of claim 7.

Claim 9: If c_L is the supremum of the low failures of type A, then c_L is itself a low failure of type A. The same holds for low failures of type B.
Proof: Since we have a bound on the length of coalition sequences, only finitely many can witness failures. Hence, there must be a sequence $c_1 < c_2 < \ldots$ converging on c_L for which there is a single pair of coalition sequences witnessing that each c_i is a low failure of type A. But this means that we have $\sum_{i=1}^{k}\sum_{p\in Y_i} f(p) \leq \sum_{i=1}^{k}\sum_{p\in Y'_i} f(p)$ for each $c_j = f(a)$, and so the inequality must still hold for $f(a) = c_L$. Hence, c_L must be a low failure of type A. The case for failures of type B is analogous.

Claim 10: If c_L is the supremum of the low failures of type C, then c_L is itself not a low failure of type C. Instead, there is a *lower margin witness* pair of coalition sequences for $A \cup \{a\}$ and $f \cup \{(a, c_L)\}$ satisfying conditions (1)–(5) of the definition such that $|\{i : a \in Y_i\}| > |\{i : a \in Y'_i\}|$.
Proof: If c_L were a low failure of type C, then by continuity, there is also some sufficiently small $\varepsilon > 0$ such that $c_L + \varepsilon$ is also a low failure of type C, contradicting our assumption that c_L is the supremum. Like in claim 9, there must be a sequence $c_1 < c_2 < \ldots$ converging on c_L for which there is a single pair of coalition sequences witnessing that each c_i is a low failure of type C. This now means that we have $\sum_{i=1}^{k}\sum_{p\in Y_i} f(p) < \sum_{i=1}^{k}\sum_{p\in Y'_i} f(p)$ for each $c_j = f(a)$, and so the weak inequality $\sum_{i=1}^{k}\sum_{p\in Y_i} f(p) \leq \sum_{i=1}^{k}\sum_{p\in Y'_i} f(p)$ must still hold for $f(a) = c_L$. Since c_L, is not a low failure of type C, this inequality cannot be strict for $f(a) = c_L$. Hence, the pair of coalition sequences is a lower margin witness.

Claim 11: If c_H is the infimum of the high failures of type A, then c_H is itself a high failure of type A. The same holds for high failures of type B.
Proof: Analogous to the proof of claim 9.

Claim 12: If c_H is the infimum of the high failures of type C, then c_H is itself not a high failure of type C. Instead, there is an *upper margin witness* pair of coalition sequences for $A \cup \{a\}$ and $f \cup \{(a, c_H)\}$ satisfying conditions (1)–(5) of the definition such that $|\{i : a \in Y_i\}| < |\{i : a \in Y'_i\}|$.
Proof: Analogous to the proof of claim 10.

To complete the proof of our main claim, we consider four different cases depending on whether c_L and c_H are themselves failures.

Case (i)Assume that c_L is a low failure and c_H is a high failure. Then the high failures constitute an interval closed on the left and the low failures constitute a disjoint interval closed on the right. Hence, there is an open interval between the two which contains real numbers that are not failures and that can be chosen as a suitable weight for a.

Case (ii) Assume that c_L is not a low failure and c_H is not a high failure. Then the low failures constitute an interval open on the right, and the high failures a disjoint interval open on the left. Hence, there is a nonempty closed interval between the two, and so there must be at least one point which is not a failure.

Case (iii) Assume that c_L is not a low failure but c_H is a high failure. Then c_H must be a high failure of type A or B, and c_L must be the supremum of low failures of type C. It suffices to show that $c_L < c_H$ for obtaining a nonempty half open interval of values which are not failures. In order to obtain a contradiction, suppose that $c_L = c_H$. Suppose that

$$\langle X_1 \cup Y_1, \ldots, X_k \cup Y_k \rangle \text{ and } \langle X_1' \cup Y_1', \ldots, X_k' \cup Y_k' \rangle,$$

chosen as in claim 2, serve as a lower margin witness for c_L, and

$$\langle U_1 \cup V_1, \ldots, U_l \cup V_l \rangle \text{ and } \langle U_1' \cup V_1', \ldots, U_l' \cup V_l' \rangle,$$

chosen as in claim 3, witness that c_H is a failure of type A or B. Now apply the construction of claim 6, multiplying, combining and shifting a from the Y_i to the X_i and from V_i' to the U_i'. The combined system satisfies conditions (1)–(5) of the definition, and fails to satisfy (A) or (B), depending on the type of failure of c_H. Consequently, f cannot be trade robust for A, a contradiction.

Case (iv) Assume that c_L is a low failure but c_H is not a high failure. Then the proof is analogous to case (iii). □

References

Anthony, M. (2003). Boolean functions and artificial neural networks. CDAM Research Report LSE-CDAM-2003-01, London School of Economics.

Chellas, B. (1980). *Modal logic—An introduction*. Cambridge: Cambridge University Press.

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

List, C., & Pettit, P. (2002). Aggregating sets of judgements: An impossibility result. *Economics and Philosophy, 18*, 89–110.

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

Pacuit, E., & Salame, S. (2004). Majority logic. In *Proceedings of the Ninth International Conference on Principles of Knowledge Representation and Reasoning (KR)* (pp. 598–605).

Pauly, M., & van Hees, M. (2006) Logical constraints on judgement aggregation. *Journal of Philosophical Logic, 35*(6), 569–585.

Taylor, A., & Zwicker, W. (1992). A characterization of weighted voting. *Proceedings of the American Mathematical Society, 115*(4), 1089–1094.

Taylor, A., & Zwicker, W. (1999). *Simple games: Desirability relations, trading, pseudoweightings*. Princeton, NJ: Princeton University Press.