Franz Dietrich

Aggregation theory and the relevance of some issues to others

RM/07/024

(RM/07/002 -revised-)

Maastricht research school of **E**conomics
of **TE**chnology and **OR**ganizations

Universiteit Maastricht
Faculty of Economics and Business Administration
P.O. Box 616
NL - 6200 MD Maastricht

phone : ++31 43 388 3830
fax : ++31 43 388 4873

Aggregation theory and the relevance of some issues to others

Franz Dietrich
December 2006

Abstract. I propose a general collective decision problem consisting in many issues that are interconnected in two ways: by mutual constraints and by connections of relevance. Aggregate decisions should respect the mutual constraints, and be based on relevant information only. This general informational constraint has many special cases, including premise-basedness and Arrow's independence condition; they result from special notions of relevance. The existence and nature of (non-degenerate) aggregation rules depends on both types of connections. One result, if applied to the preference aggregation problem and adopting Arrow's notion of (ir)relevance, becomes Arrow's Theorem, without excluding indifferences unlike in earlier generalisations.

Keywords: aggregation, informational constraints, (ir)relevance, premise-based procedure

JEL Classification Numbers: D70, D71

1 Introduction

Most complex decision problems can be formalised as consisting of many binary decisions: decisions of accepting or rejecting certain propositions. For instance, establishing a preference relation R over a given set of alternatives Q consists in deciding, for each pair of alternatives $x, y \in Q$, whether or not xRy. Judging the values of different variables consists of judging, for each variable V and each of its potential values v, whether or not $V = v$. Producing a report that contains qualitative economic forecasts might involve deciding for or against many propositions: atomic ones like "inflation will increase" and compound ones like "*if* consumption will increase *and* foreign demand does *not* decrease, *then* inflation will increase" (where logical operators are italicised).

Although this division into binary issues is usually possible, there are arguably two distinct types of interconnections – to be called *logical* connections and *relevance* connections – that can prevent us from treating the issues independently. First, the decisions on the issues may logically constrain each other; in the above examples, the preference judgments must respect conditions like transitivity, the variables might constrain each other, and the propositions stated in the economic report must be logically consistent with each other, respectively. Second – and this is the topic of the paper – some issues may be *relevant* to (the decision on) other issues. The nature and interpretation of relevance connections is context-specific. A proposition r may be relevant to another one p on the grounds that r is an (argumentative) premise of p, or that r is a causal factor bringing about p, or that r and p share some other (semantic) relation. Relevance connections are not reducible to logical connections. Two issues – say, whether traffic lights are necessary and whether the diplomatic relations to a country should be interrupted – may be considered irrelevant to each other and yet be indirectly logically related via other issues under consideration. Conversely, an issue – say that of whether country X has weapons of mass destruction – may be considered relevant to another issue – say that of whether measure Y against country X is appropriate – without a (direct or indirect) logical connection in the complex decision problem considered.

Now suppose that the complex decision problem is faced by a group of individuals and should be settled by aggregating the individual judgments on each proposition (issue). Many concrete aggregation models and procedures in the literature in effect account, in different ways, both for logical connections and relevance connections. Logical connections are represented by delimiting the set of admissible decisions, for instance in the form of rationality conditions like transitivity in preference aggregation, or in the form of an overall budget constraint if different budget items are decided simultaneously. By contrast, relevance connections are accounted for through "informational" constraints on the way in which the decision (output of the aggregation rule) may depend on the individuals' input: only *relevant* information may be used. For instance, Arrow's condition of *independence of irrelevant alternatives* ("IIA") excludes the use of (arguably) irrelevant information. The *premise-based procedure* in judgment aggregation makes the decision on certain conclusion-type propositions dependent on people's judgments on other premise-type propositions considered relevant. In general, the question of "what is relevant to what?" may be controversial: some researchers reject Arrow's IIA condition, and in judgment aggregation it may be unclear which propositions to consider as premises and which as conclusions, and moreover the same conclusion-type proposition could be explained in more than one way in terms of premises.

While accounted for in concrete aggregation problems and procedures, the notions of relevance and of (ir)relevant information have not been treated in general terms. As relevance connections are not reducible to logical connections, both connections should be separate ingredients of a general aggregation model. More precisely, I propose to consider, in addition to logical connections, a (binary) *relevance* relation $\mathcal{R}$ between propositions (issues), and to aggregate in accordance with *independence of irrelevant information* ("III"). To allow broad applications, I leave general the type of complex decision problem and the interpretation and relation-theoretic properties of the relevance relation $\mathcal{R}$: it might be highly partial (few inter-relevances) or close to complete (many inter-relevances), and it need not be symmetric, or transitive, or reflexive (i.e. self-irrelevance is allowed).

In the special case that every proposition is considered relevant just to itself (i.e. $p\mathcal{R}q \Leftrightarrow p = q$ for any propositions p, q), III reduces to the restrictive condition of *proposition-wise independence* (often simply called *independence*): here, each proposition is decided via an isolated vote, using an arbitrary voting rule but ignoring people's judgments on other propositions. A number of general results have been obtained on proposition-wise independent aggregation, in abstract aggregation models (starting with Wilson 1975) or models of logic-based judgment aggregation (starting with List and Pettit 2002). Essentially, these results establish limits to the possibility of (non-degenerate) proposition-wise independent aggregation in the presence of logical connections between propositions. Impossibility results with necessary conditions on logical connections are derived, for instance, by Wilson (1975), List and Pettit (2002), Pauly and van Hees (2006), Dietrich (2006), Gärdenfors (2006), Mongin (2005-a) and van Hees (forthcoming). Nehring and Puppe (2002, 2005, 2006) derive the first results with minimal conditions on logical connections, and Dokow and Holzman (2005) introduce minimal conditions of an algebraic kind. Other (im)possibility results are given, for instance, in Dietrich (forthcoming), Dietrich and List (forthcoming-a, forthcoming-b) and Nehring (2005). Possibilities of proposition-wise independent aggregation arise if the individual judgments fall into particular domains (List 2003, Dietrich and List 2006) or if logical connections are modelled using subjunctive implications (Dietrich 2005).

The proposition-wise independence condition is often criticised (e.g., Chapman 2002, Mongin 2005-a), but has rarely been weakened in the general aggregation literature. The normative appeal of the condition is easily challenged by concrete examples: why, for instance, should the collective judgment on whether to introduce taxes on kerosene be independent of people's judgments on whether global warming should be prevented? All weaker independence conditions proposed in the literature are special cases of III: each implicitly uses some notion of relevance $\mathcal{R}$. Let me mention the literature's two most notable independence weakenings.[1]

One departure from proposition-wise independence aims to represent non-binary variables.[2] Suppose again the decision problem consists in estimating the values of different typically non-binary (interconnected) variables V like GDP growth. Then propositions take the form $V = v$, where V is a variable and v belongs to a set $Rge(V)$ of possible values of V. Suppose the collective estimate of each variable V must be a function of people's estimates of V (e.g. a weighted average). Then the collective judgment on whether $V = v$ depends on people's attitudes towards the propositions $V = v', v' \in Rge(V)$ (each individual accepts exactly one of them).[3] So aggregation is *variable-wise* independent – not *proposition-wise* as the decision on whether $V = v$ depends not just on people's views on whether $V = v$. Variable-wise independence is an example of III, where any $V = v$ and $V = v'$ are now inter-relevant. Variable-wise independence is often imposed: for instance in probability aggregation theory, where a variable is an event's probability and variable-wise independence leads (under other constraints) to *linear* aggregation rules (e.g. Genest and Zidek 1986); or in abstract aggregation theory, where Rubinstein and Fishburn (1986) derive more general linearity results on variable-wise independent aggregation; or in judgment aggregation, where Claussen and Roisland (2005) introduce a variable-wise version of the discursive paradox and show results on when it occurs. Also Pauly and van Hees' (2006) multi-valued logic approach can be viewed as using variable-wise independence.

A second weakening of proposition-wise independence aims to represent the different status of different propositions. Here the independence condition is applied

[1] A more radical move consists in imposing *no* independence condition (informational constraint) on aggregation. This route is taken in the literature on *belief merging* in artificial intelligence (e.g. Konieczni and Pino-Perez 2002, Eckert and Pigozzi 2005). Apart from the absence of informational constraints, belief merging is closely related to judgment aggregation: it also aims to merge sets of logical propositions.

[2] It was originally not intended as a weakening of proposition-wise independence, but is one under our division of the decision problem into (binary) decisions on propositions of the form $V = v, v \in Rge(V)$.

[3] E.g. the collective accepts $V = v$ if and only if v is a certain weighted average of the individual estimates v' of V, i.e. of the values $v' \in Rge(V)$ for which $V = v'$ is accepted.

only to *some* propositions, e.g. to "premises" (Dietrich 2006) or to *atomic* propositions (Mongin 2005-a). Mongin (2005-a) argues that the collective judgment on a compound proposition like $p \wedge q$ should not ignore how the individuals judge p and judge q; our relevance relation $\mathcal{R}$ would then have to satisfy $p\mathcal{R}(p \wedge q)$ and $q\mathcal{R}(p \wedge q)$.

This paper has an expository and a technical focus. On the expository dimension, I introduce the relevance-based aggregation model; I discuss different types of relevance relations, including transitive relevance, asymmetric relevance, and relevance as premisehood; I introduce relevance-based conditions of III, agreement preservation and dictatorship (generalising for instance Arrow's conditions of IIA, weak Pareto and weak dictatorship); and I introduce significantly generalised forms of premise-based and prioritarian aggregation rules. On the technical dimension, I prove two possibility and four impossibility results on III aggregation. One result, if applied to the preference aggregation problem, becomes Arrow's Theorem. While Arrow's Theorem has been generalised earlier under the simplifying assumption that not only individuals but also the collective are never indifferent between distinct options,[4] one might view as an embarrassment of the growing literature that, despite its intended generality, its theorems do not generalise Arrow's (unrestricted) theorem; and its aggregation conditions do not have as special cases Arrow's conditions of IIA, weak Pereto and weak dictatorship.

2 Basic definitions

We consider a set $N = \{1, ..., n\}$ of individuals, where $n \geq 2$, faced with a collective decision problem of a general kind.

Agenda, judgment sets. The *agenda* is an arbitrary non-empty (possibly infinite) set X of *propositions* on which a decision (acceptance or rejection) is needed. The agenda includes negated propositions: $X = \{p, \neg p : p \in X^+\}$, where X^+ is some set of non-negated propositions and "$\neg p$" stands for "not p". Notationally, double-negations cancel each other out.[5] A *judgment set* is a set $A \subseteq X$ of (accepted) propositions; it is *complete* if it contains a member of each pair $p, \neg p \in X$ ("no abstentions").

Logical interconnections. Not all judgment sets are consistent. For the agenda $X = \{a, \neg a, b, \neg b, a \wedge b, \neg(a \wedge b)\}$, the (complete) judgment set $\{a, b, \neg(a \wedge b)\}$ is inconsistent. Let $\mathcal{J}$ be a non-empty set of judgment sets, each containing exactly one member of each pair $p, \neg p \in X$, and suppose the *consistent* judgment sets are precisely the sets in $\mathcal{J}$ and their subsets; all other judgment sets are *inconsistent*.[6] A judgment set $A \subseteq X$ *entails* a proposition $p \in X$ (written $A \vdash p$) if $A \cup \{\neg p\}$ is inconsistent. I write $q \vdash p$ for $\{q\} \vdash p$.

It is natural (though for the present results not necessary) to take the propositions in X to be *statements* of a formal language, and to take consistency/entailment to

[4] See Wilson (1975), Dietrich and List (forthcoming-b), and Dokow and Holzman (2005). Nehring (2003) shows an Arrow-like result.

[5] That is: whenever I write "$\neg q$" (where $q \in X$), I mean the other member of the pair $p, \neg p \in X$ to which q belongs; hence "$\neg\neg q$" stands for q.

[6] So $\mathcal{J}$ contains the consistent *and complete* judgment sets. Any set $\{p, \neg p\} \subseteq X$ is inconsistent. Any subset of a consistent set is consistent. Finally, $\emptyset$ is consistent, and any consistent set has a superset that is consistent and complete (hence in $\mathcal{J}$).

be standard logical consistency/entailment, as is usually assumed in the judgment aggregation literature. The formal language, if sufficiently expressive, can mimic the natural language in which the real decision problem arises.[7]

A proposition $p \in X$ is a *contradiction* if $\{p\}$ is inconsistent, and a *tautology* if $\{\neg p\}$ is inconsistent. I call $A \subseteq X$ *consistent with* $B \subseteq X$ if $A \cup B$ is consistent; and I call $A \subseteq X$ *consistent with* $p \in X$ (and p *consistent with* A) if $A \cup \{p\}$ is consistent.

Aggregation. The *(judgment) aggregation rule* is a function F that assigns to every profile $(A_1, \ldots, A_n)$ of (individual) judgment sets in some domain of "admissible" profiles a (collective) judgment set $F(A_1, \ldots, A_n) = A \subseteq X$. It is often required that F has *universal* domain, i.e. allows as an input precisely all profiles $(A_1, ..., A_n)$ of consistent and complete (individual) judgment sets. An important question is how rational the (collective) judgment sets generated by F are: are they consistent? Complete? If F has universal domain and consistent and complete outputs, it is a function $F : \mathcal{J}^n \to \mathcal{J}$. Majority rule on $\mathcal{J}^n$, given by

$$F(A_1, ..., A_n) = \{p \in X : |\{i : p \in A_i\}| > n/2\} \text{ for all } (A_1, ..., A_n) \in \mathcal{J}^n,$$

may for most agendas generate inconsistent outputs: it is not a function $F : \mathcal{J}^n \to \mathcal{J}$.

Abstract aggregation. One may (re)interpret the elements of X as arbitrary *attributes*, which may but need not be *propositions/judgments*, and may but need not be expressed in formal logic. Then judgment sets become *attribute sets*, and the aggregation rule maps profiles of individual attribute sets to a collective attribute set. Of course, the attribute holders $i \in N$ need not be humans.

I give two examples here; more examples follow in the next section.

Example 1: preference aggregation. For a given set of (exclusive) alternatives Q ($|Q| \geq 3$), consider the agenda

$$X := \{xRy, \neg xRy : x, y \in Q\} \text{ (the \textit{preference agenda})},$$

where xRy is the proposition "x is at least as good as y". Throughout the paper, I often write xPy for $\neg yRx$. Let $\mathcal{J}$ be the set of all judgment sets $A \subseteq X$ that represent fully rational preferences, i.e. for which there exists a weak ordering[8] $\succeq$ on Q such that

$$A = \{xRy \in X : x \succeq y\} \cup \{\neg xRy \in X : x \not\succeq y\}.$$

Note that there is a bijective correspondence between weak orderings on Q and judgment sets in $\mathcal{J}$; and between judgment aggregation rules $F : \mathcal{J}^n \to \mathcal{J}$ and Arrowian social welfare functions (with universal domain). The agenda X and its consistency

[7] The formal language can be one of classical (propositional or predicate) logic or one of a non-classical logic such as a modal logic, as long as the logic satisfies certain regularity conditions. This follows Dietrich's (forthcoming) model of judgment aggregation in general logics, which generalises List and Pettit's (2002) original model in classical propositional logic.

[8] A weak ordering on Q is a binary relation $\succeq$ on Q that is reflexive, transitive, and connected (but not necessarily anti-symmetric, so that non-trivial indifferences are allowed).

notion belong to a predicate logic, as defined in Dietrich (forthcoming), drawing on List and Pettit (2004).[9]

Example 2: judging values of and constraints between variables. Suppose a group (e.g. a central bank's board or research panel) debates the values of different variables (e.g. macroeconomic variables measuring GDP, prices or consumption). Let $\mathbf{V}$ be a non-empty set of "variables". For each $V \in \mathbf{V}$ let $Rge(V)$ be a non-empty set of possible "values" of V (numbers or other objects), called the *range* of V. For any variable $V \in \mathbf{V}$ and any value $v \in Rge(V)$, the group has to judge the proposition $V = v$ stating that V takes the value v.[10] These judgments should respects the (causal) constraints between variables; but, not surprisingly, the nature of these constraints is itself disputed, for instance because the group members believe in different (e.g. econometric) estimation techniques. If the variables are real-valued, some *linear* constraints like $V + 3W - U = 5$, or non-linear ones like $V^2 = W$, might be debated. Let $\mathbf{C}$ be any non-empty set of "constraints" under consideration.[11] The agenda is given by

$$X = \{V = v, \neg(V = v) : V \in \mathbf{V}, v \in Rge(V)\} \cup \{c, \neg c : c \in \mathbf{C}\}.$$

A judgment set $A \subseteq X$ thus states that certain variables do (not) take certain values, and that the variables do (not) constrain each other in certain ways. To define logical connections, note first that some constraints may conflict with others (e.g., $V > W$ conflicts with $W > V$), and that some constraints may conflict with other *negated* constraints (e.g., $V \log(W) > 2$ conflicts with $\neg(V \log(W) > 0)$). Let $\mathcal{J}^*$ be some non-empty set containing for each constraint $c \in \mathbf{C}$ either c or $\neg c$ (not both); the sets in $\mathcal{J}^*$ represent consistent judgments on the constraints. Now let $\mathcal{J}$ be the set of all judgment sets $A \subseteq X$ containing exactly one member of each pair $p, \neg p \in X$ such that:[12]

(i) each variable $V \in \mathbf{V}$ has a single value $v \in Rge(V)$ with $V = v \in A$;
(ii) the family of values in (i) obeys all accepted constraints $c \in A \cap \mathbf{C}$;
(iii) the judgments on constraints are consistent: $A \cap \{c, \neg c : c \in \mathbf{C}\} \in \mathcal{J}^*$.

Note that it may be consistent to hold a negated constraint $\neg c$ and yet to assign values to variables in accordance with c. Indeed, variables can stand in certain relations by pure coincidence, i.e. without a constraint to this effect.[13]

---

[9] See Dietrich and List (forthcoming-b) for a logic representing *strict* preference aggregation.

[10] More generally, the group might consider propositions stating that V's value belongs to certain sets $S \subseteq val(V)$.

[11] A constraint might be formalised by a subset of the "joint range" $\Pi_{V \in \mathbf{V}} Rge(V)$ of the family of variables $(V)_{V \in \mathbf{V}}$ (e.g. a subset of $\mathbf{R}^3$ if $\mathbf{V}$ consists of three real-valued variables), or by an expression in a logical language (see below).

[12] It is easily possible to add *exogenous* constraints (which cannot be rejected, unlike those in $\mathbf{C}$), by further restricting in (iii) the allowed value assignments.

[13] More precisely, a constraint states not just an *actual* relation r between variables but a *necessary* relation "necessarily r", that is (in modal logical terms) "in all possible worlds r" ($\Box r$). The negation of this constraint ($\neg\Box r$) is equivalent to "possibly $\neg r$" ($\Diamond\neg r$), which is indeed consistent with "r", i.e. with the relation holding.

3 Independence of irrelevant information

The conditions I will impose on the aggregation rule are based on a *relevance relation*, whose nature and interpretation is context-specific, as indicated earlier. Such a relevance relation is not simply reducible to logical interconnections (of inconsistency or entailment). Suppose the proposition a : "country X has weapons of mass destruction" (and $\neg a$) is considered relevant to the proposition b: "country X should be attacked" (and to $\neg b$), but not vice versa. This asymmetry of relevance between the two issues need not be reflected in logical connections: $\mathcal{J}$ can be perfectly symmetric in the two issues. This is clear if X contains no issues other than these two (logically independent) ones, i.e. if $X = \{a, \neg a, b, \neg b\}$. But even additional propositions in X that create (indirect) logical links between the two issues need not reveal a direction of relevance, as is seen from examples.[14] Hence any relevance relation derived from logical interconnections would have to declare, against our intuition, the two issues as either mutually relevant or mutually irrelevant.

So relevance must be taken on board as an additional structure. I do this in the form of a *relevance relation.* Not any binary relation on X can reasonably count as a relevance relation. I call a binary relation $\mathcal{R}$ on the agenda X a *relevance relation* (where "$r\mathcal{R}p$" means "r is relevant to p") if the following condition holds.

No underdetermination. Each $p \in X$ is settled by the judgments on the relevant propositions: for every consistent set $E \subseteq \{r, \neg r : r\mathcal{R}p\}$ containing a member of each pair $r, \neg r$ in $\{r, \neg r : r\mathcal{R}p\}$, either $E \vdash p$ or $E \vdash \neg p$. (I call such an E an *($\mathcal{R}$-)explanation* of p or *($\mathcal{R}$-)refutation* of p, respectively.)

This definition of a relevance relation has two main characteristics.

First, it requires no relation-theoretic properties like reflexivity or symmetry. This generality is essential to represent different notions of relevance (see below); and it is appropriate since no relation-theoretic property is uncontroversially adequate for all decision problems. Below I suggest relation-theoretic conditions on relevance for special decision problems, but different ones across decision problems.

Second, it requires "no underdetermination": a proposition's truth value must be fully determined by the relevant propositions' truth values. To illustrate this condition (which I justify in the next section), note first that it holds trivially for *self-relevant* propositions $p \in X$, as p's truth value settles p's truth value; here all explanations of p contain p, and all refutations of p contain $\neg p$.[15] In particular, all reflexive relations $\mathcal{R}$ satisfy "no underdetermination", i.e. are relevance relations. This said, "no underdetermination" is a weak condition. It only has a bite for propositions that are non-self-relevant, hence "externally" explained. For instance, suppose to a

[14] Suppose $X = \{a, \neg a, a \to b, \neg(a \to b), b, \neg b\}$. The symmetry argument is simple. A truth-value assignment $(t_1, t_2, t_3) \in \{T, F\}^3$ to the propositions $a, a \to b, \neg b$ is consistent if and only if (t_3, t_2, t_1) (in which the truth-values of a and $\neg b$ are interchanged) is consistent. This is so whether $a \to b$ represents a *subjunctive* or a *material* implication. In the first case, the only inconsistent truth-value assignment is (T, T, T). In the second case, there are other inconsistent truth-value assignments (as $a \to b$ is equivalent to $\neg a \vee b$), yet without breaking the symmetry between a and $\neg b$.

[15] If p is the *only* proposition relevant to p, p's only explanation is $\{p\}$ (except if p is a contradiction: then p has no explanation), and p's only refutation is $\{\neg p\}$ (except if $\neg p$ is a contradiction: then p has no refutation).

conjunction $a \wedge b$ only the conjuncts a and b, not $a \wedge b$ itself, are deemed relevant. (Such an idea underlies the premise-based procedure for the agenda given by $X^+ = \{a, b, a \wedge b\}$; see Example 4 below.) Here, $a \wedge b$'s truth value is indeed determined by a's and b's truth values; $a \wedge b$ has a single explanation ($\{a, b\}$) and three possible refutations ($\{\neg a, b\}$, $\{a, \neg b\}$, $\{\neg a, \neg b\}$). Dropping a's or b's relevance to $a \wedge b$ would lead to underdetermination.

Hereafter, let $\mathcal{R}$ be a given relevance relation. I denote the set of propositions relevant to $p \in X$ by $\mathcal{R}(p) := \{r \in X : r\mathcal{R}p\}$. The following condition requires the collective judgment on any proposition $p \in X$ to be formed on the basis of how the individuals judge the propositions relevant to p.

Independence of Irrelevant Information (III). For all propositions $p \in X$ and all profiles $(A_1, ..., A_n)$ and $(A'_1, ..., A'_n)$ in the domain, if $A_i \cap \mathcal{R}(p) = A'_i \cap \mathcal{R}(p)$ for every individual i then $p \in F(A_1, ..., A_n) \Leftrightarrow p \in F(A'_1, ..., A'_n)$.

Many informational constraints on aggregation used in social choice theory can be viewed as being the III condition relative to some notion of relevance. Roughly, the more propositions are relevant to each other, the weaker the informational constraint III is. III is empty if all propositions are relevant to all propositions, i.e. if $\mathcal{R} = X \times X$. III is the standard proposition-wise independence condition if each proposition is just self-relevant, i.e. $\mathcal{R}(p) = \{p\}$ for all $p \in X$. III is Gärdenfors' "weak" (yet still quite strong) independence if $\mathcal{R}(p) = \{p, \neg p\}$ for all $p \in X$. III is Dietrich's (2006) independence restricted to a subset $Y \subseteq X$ if $\mathcal{R}(p) = \{p\}$ for $p \in Y$ and $\mathcal{R}(p) = X$ for $p \in X \backslash Y$. III is Mongin's (2005-a) independence restricted to the atomic propositions (of an agenda X in a propositional language) if $\mathcal{R}(p) = \{p\}$ for atomic p and $\mathcal{R}(p) = X$ for compound p (e.g. $p = a \wedge \neg b$).

I now discuss further examples of relevance relations. These examples make the convenient assumption that relevance is *negation-invariant*:[16]

$$p\mathcal{R}q \Leftrightarrow \tilde{p}\mathcal{R}\tilde{q} \text{ for all } p, q \in X \text{ and all } \tilde{p} \in \{p, \neg p\},\ \tilde{q} \in \{q, \neg q\} \text{ (negation invariance).}$$

So $\mathcal{R}$ is determined by its restriction to the set $X^+ \subseteq X$ of non-negated propositions. Let $\mathcal{R}^+$ be this restriction, and for all $p \in X^+$ let $\mathcal{R}^+(p) := \mathcal{R}(p) \cap X^+$ ($= \{r \in X^+ : r\mathcal{R}p\} = \{r \in X^+ : r\mathcal{R}^+p\}$).

Example 1 (continued). For the preference agenda, III is equivalent to Arrow's *independence of irrelevant alternatives* ("IIA") in virtue of defining relevance by

$$\mathcal{R}^+(xRy) := \{xRy, yRx\} \text{ for all } xRy \in X. \tag{1}$$

I call this the *Arrowian* relevance relation. Indeed, to socially decide on xRy, Arrow considers as relevant whether people weakly prefer x to y *and also* whether they weakly prefer y to x. By contrast, the standard proposition-wise independence condition is stronger than IIA, as it denies the relevance of yRx to xRy.

[16] The relevance relation underlying proposition-wise independence (given by $\mathcal{R}(p) = \{p\}$ for all $p \in X$) is *not* negation-invariant.

Example 2 (continued). For the agenda of Example 2, one might put

$$\begin{array}{ll} \mathcal{R}^+(V=v)=\{V=v':v'\in Rge(V)\} & \text{for all } V=v\in X \\ \mathcal{R}^+(c)=\{c\} & \text{for all constraints } c\in\mathbf{C}. \end{array} \tag{2}$$

On a modified assumption, some distinct constraints $c, c' \in \mathbf{C}$ might be declared inter-relevant, for instance if they involve the same variables.

Example 3: relevance as an equivalence relation, and topic-wise independence. Examples 1 and 2 are instances of the general case where relevance is an equivalence relation: $\mathcal{R}$ is reflexive (which requires self-relevance), symmetric, and transitive. Each of these three conditions is a substantial assumption on the notion of relevance. The agenda X is then partitioned into equivalence classes (of inter-relevant propositions), each one interpretable as a *topic*; so III is a *topic-wise* (rather than proposition-wise) independence condition. A topic can be binary (of the form $\{p, \neg p\}$) or non-binary. For the preference agenda (Example 1), the Arrowian relevance relation creates topics of the form $\{xRy, \neg xRy, yRx, \neg yRx\}$ (for options $x, y \in Q$): the topic of x's and y's relative ranking.

An example of topic-wise independence is the *variable-wise* independence condition mentioned in the introduction. Consider a variant of Example 2, in which the inter-variable constraints are exogenously imposed rather than under decision. So the agenda is given by $X^+ = \{V = v : V \in \mathbf{V} \text{ and } v \in Rge(V)\}$, and relevance by $\mathcal{R}^+(V = v) = \{V = v' : v' \in Rge(V)\}$. To each variable $V \in \mathbf{V}$ corresponds an equivalence class: $\{V = v, \neg(V = v) : v \in Rge(V)\}$, the topic of V's value. Judging this topic boils down to specifying a value $v \in Rge(V)$ of V (i.e. $V = v$ is accepted and all $V = v', v' \in Rge(V)\backslash\{v\}$ are rejected). So a judgment set $A \in \mathcal{J}$ can be identified with a function b assigning to each variable $V \in \mathbf{V}$ a value $v \in Rge(V)$. Then $\mathcal{J}$ becomes a set B of such functions, and an aggregation rule $F : \mathcal{J}^n \to \mathcal{J}$ becomes a function $f : B^n \to B$.[17]

Example 4: relevance as premisehood, and generalised premise-based rules. If we interpret "$r\mathcal{R}p$" as "r is a premise/reason/argument for (or against) p", III is the condition that the aggregation rule be *premise-based*: that the collective judgment on any proposition $p \in X$ be determined by people's reasons for their judgments on p.

In principle, $\mathcal{R}$ could define an arbitrarily complex premisehood structure over a possibly complex agenda, generalising the classical *premise-based procedure* (PBP) usually defined for simple agendas like agendas 1 and 2 in Figure 1. For agenda 1, the classical PBP decides each "premise" a and b by a majority vote, and decides the "conclusion" $a \wedge b$ by logical entailment from the decisions on a and b. This PBP is III for the relevance ("premisehood") relation indicated in Figure 1:

$$\mathcal{R}^+(a) = \{a\},\ \mathcal{R}^+(b) = \{b\},\ \mathcal{R}^+(a \wedge b) = \{a, b\}. \tag{3}$$

[17] Rubinstein and Fishburn (1986) analyse variable-wise independent aggregators $f : B^n \to B$, assuming that there are only finitely many variables V, all with the same range, an algebraic field $\mathcal{F}$ (e.g. $\mathcal{F} = \mathbf{R}$). So $B \subseteq \mathcal{F}^{\mathbf{V}}$. Their two main results establish correspondences between algebraic properties of B, like being a *hyperplane* of the $\mathcal{F}$-vector space $\mathcal{F}^{\mathbf{V}}$, and algebraic properties of "admissible" aggregators $f : B^n \to B$, like *linearity* or *additivity*. In practice, the hyperplane condition on B seems restrictive: variables are interconnected by exactly one equation, and this equation is linear. But note that linearity can sometimes be achieved by transforming the variables.

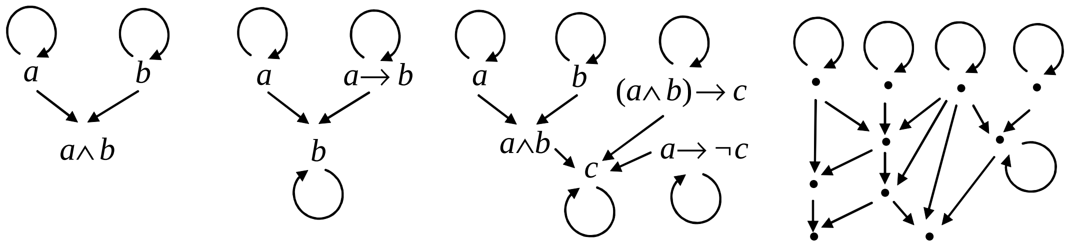

Figure 1: Relevance ("premisehood") relations over four agendas. Arrows indicate relevance. Agenda 1: $X^+ = \{a, b, a \wedge b\}$. Agenda 2: $X^+ = \{a, a \to b, b\}$. Agenda 3: $X^+ = \{a, b, c, a \wedge b, (a \wedge b) \to c, a \to \neg c\}$. Agenda 4: X^+ contains ten propositions indicated by "$\cdot$".

For agenda 2 in Figure 1, the classical PBP takes majority votes on each "premise" a and $a \to b$; if the resulting decisions logically constrain the "conclusion" b,[18] b is decided accordingly; otherwise b is (for instance) decided by a majority vote on b. This PBP is III for relevance as given in Figure 1. Unlike in (3), the conclusion is self-relevant: individual judgments on b may matter for deciding b.

In general, call $p \in X$ a *root* proposition if p has no premise other than p (and $\neg p$). In (3), a and b are root propositions. Any root proposition $p \in X$ must be a premise to itself: otherwise p would have no premises at all, violating "no underdetermination".[19] So the collective judgment on any root proposition p is (by III) formed solely on the basis of people's judgments on p via some voting method – majority voting if we stick closely to the standard premise-based procedure – while decisions on non-root propositions may depend on external premises.

When interpreting $\mathcal{R}$ as a premisehood relation, additional requirements on $\mathcal{R}$ may be appropriate. Surely, symmetry should not be required (unlike in Examples 1-3). Indeed, one might require that $\mathcal{R}$ is *anti-symmetric* on X^+ (so that no distinct propositions in X^+ are premises to each other) or, more strongly, *acyclic* on X^+ (so that in X^+ there is no cycle $p_1 \mathcal{R} p_2 \mathcal{R} p_3 ... \mathcal{R} p_m \mathcal{R} p_1$ where the p_i's are pairwise distinct and $m \geq 2$).

For some agendas X, specifying $\mathcal{R}$ is non-trivial: it is not obvious which propositions should count as reasons for/against which others. One might for instance draw on the *syntax* of the propositions in X: if $X^+ = \{a, a \to b, b\}$, one might argue that $a \mathcal{R} b$ because $a \to b \in X$, and not $b \mathcal{R} a$ because $b \to a \notin X$. Finding objective criteria for relevance would be an interesting research goal on its own.[20]

[18] Whether this is so may depend on whether "$\to$" is a material or subjunctive implication.

[19] Unless p is a tautology or contradiction: then even the empty set settles p.

[20] For different purposes, *relevance logicians* (e.g. Parikh 1999) propose syntactic and other criteria for when a proposition is relevant to another. Although this enterprise is controversial and its notion of relevance may differ from ours, one might use such criteria in defining $\mathcal{R}$.

4 Justifying the "no underdetermination" condition

I now give a technical and a conceptual motivation for the "no underdetermination" requirement on a relevance relation.

The technical reason is that "no underdetermination" is crucial for the existence of non-degenerate III aggregation rules. The condition of *judgment-set unanimity preservation*, whereby $F(A,...,A) = A$ for all unanimous profiles $(A,...,A)$ in the domain, is very mild (unlike the unrestricted *proposition*wise unanimity condition mentioned in Section 7).

Theorem 1 *Let $\mathcal{R}$ be an arbitrary binary relation on X. There exists a judgment-set unanimity preserving III aggregation rule with universal domain if and only if $\mathcal{R}$ satisfies "no underdetermination".*

This defence of the "no underdetermination" condition needs no collective completeness or consistency condition, not even a non-dictatorship condition. So, given underdetermination, not even rules with incomplete, or inconsistent, or dictatorial outputs can satisfy the conditions.

Proof. First, suppose "no underdetermination" is violated for $p \in X$. Then there are sets $A, A' \in \mathcal{J}$ such that $p \in A$ and $p \notin A'$ but $A \cap \mathcal{R}(p) = A' \cap \mathcal{R}(p)$ (i.e. A and A' disagree on p but agree on the relevant propositions). If we apply an III aggregation rule F with universal domain to the two unanimous profiles $(A,...,A)$ and $(A',...,A')$, the resulting judgment sets $F(A,...,A)$ and $F(A',...,A')$ agree on p by III. So, as A and A' disagree on p, $F(A,...,A) \neq A$ or $F(A',...,A') \neq A'$, violating judgment-set unanimity preservation.

Second, suppose $\mathcal{R}$ satisfies "no underdetermination". I show that (for instance) the unanimity rule with universal domain, given by $F(A_1,...,A_n) = A_1 \cap ... \cap A_n$, satisfies III (it obviously also preserves judgment-set unanimity). Consider any $p \in X$ and $(A_1,...,A_n), (A'_1,...,A'_n) \in \mathcal{J}^n$ such that $A_i \cap \mathcal{R}(p) = A'_i \cap \mathcal{R}(p)$ for all i. For all i, we have $A_i \cap \{r, \neg r : r \in \mathcal{R}(p)\} = A'_i \cap \{r, \neg r : r \in \mathcal{R}(p)\}$. By "no underdetermination", this set entails p or entails $\neg p$; in the first case, $p \in A_i$ and $p \in A'_i$, and in the second case $p \notin A_i$ and $p \notin A'_i$. So in any case $p \in A_i \Leftrightarrow p \in A'_i$. Hence $p \in \cap_i A_i \Leftrightarrow p \in \cap_i A'_i$, i.e. $p \in F(A_1,...,A_n) \Leftrightarrow F(A'_1,...,A'_n)$, as desired. ■

I now turn to a conceptual defence of "no underdetermination". This condition can be violated for $p \in X$ only if p is self-irrelevant, a rather special assumption. I can see only one (albeit prominent) case in which self-irrelevance has a clear motivation: the case of premise-based collective decision making. Here the decision on $p \in X$ should depend on people's *reasons* (*grounds*) for accepting or rejecting p (as in Examples 4). A person's reasons for accepting (rejecting) p can be viewed as a set E of sentences that, if specified exhaustively, logically entails p ($\neg p$).[21] But not

[21] If the set of reasons E did *not* entail p, it wouldn't be exhaustive, i.e. some "reasons" have been forgotten. For instance, if in inferring p from E the persons implicitly uses that $(\wedge_{e \in E} e) \rightarrow p$, i.e. that p follows from the members of E, then $(\wedge_{e \in E} e) \rightarrow p$ should be added as a reason to E. The so-enlarged set of reasons now logically entails p. (In other situations, $(\wedge_{e \in E} e) \rightarrow p$ is not the missing reason, i.e. E has to be enlarged differently.)

all sets of sentences E that entail p ($\neg p$) have to count as "sets of reasons" for accepting (rejecting) p; $\{p\}$ might not count as a set of reasons for accepting p. For instance, $\{a\}$ and $\{b\}$ might count as the sets of reasons for accepting a disjunction $a \vee b$; and $\{\neg a, \neg b\}$ might count as the only set of reasons for rejecting $a \vee b$. Let $\mathcal{E}(p)$ be the set of all sets of reasons for accepting or rejecting p. In the example, $\mathcal{E}(a \vee b) = \{\{a\}, \{b\}, \{\neg a, \neg b\}\}$. Plausibly, $\mathcal{E}(p)$ should contain sufficiently many sets of reasons so that p cannot be accepted or rejected without any set of reasons $E \in \mathcal{E}$. For instance, we cannot remove $\{a\}$ from $\mathcal{E}(a \vee b)$: otherwise someone who accepts a but not b would accept $a \vee b$ for *no* set of reasons $E \in \mathcal{E}(a \vee b)$. Given our assumption of reason-based aggregation, every reason for or against p (i.e. every member of any set of reasons $E \in \mathcal{E}(p)$) should be considered relevant to p: that is, $\cup_{E \in \mathcal{E}(p)} E \subseteq \mathcal{R}(p)$ (one might even argue that $\cup_{E \in \mathcal{E}(p)} E = \mathcal{R}(p)$). In this case, "no underdetermination" holds.[22]

5 Possibility or impossibility?

Are there appealing III aggregation rules, and how do they look? General answers to this question are harder to give than for proposition-wise independence. The reason is that criteria for the (in)existence of (non-degenerate) III aggregation rules typically concern not just logical interconnections (as for proposition-wise independence) but also relevance interconnections. More precisely, we need criteria on the *interplay* between logic and relevance. One such criterion is "no underdetermination", which is (by Theorem 1) necessary and sufficient for a limited possibility: "limited" because collective incompleteness is allowed (but collective consistency, agreement preservation, and non-dictatorship could have been required in Theorem 1, as the proof shows).

Below I derive one more possibility theorem – with creteria for the possibility of priory rules – and four impossibility theorems. I deliberately sacrifice some generality (of the criteria) for simplicity and elegance.[23]

6 Priority rules

In this section, I adopt Example 4's interpretation of relevance as premisehood; and I assume again that $\mathcal{R}$ is negation-invariant. Do there exist appealing premise-based

[22] What if some reasons $e \in \cup_{E \in \mathcal{E}(p)} E$ are outside the agenda X (i.e. not part of the decision problem), so that we cannot have $\cup_{E \in \mathcal{E}(p)} E \subseteq \mathcal{R}(p)$? One might *either* argue that such agendas are simply misspecified (in the context of reason-based aggregation): if $a \vee b \in X$ then X should include $a \vee b$'s reasons. *Or* one might defend "no underdetermination" for such agendas: since "$r \in X$ is relevant to $p \in X$" more precisely means "the individuals' judgments on r are relevant information for deciding p", if X excludes some of p's reasons then other propositions in X (perhaps p itself) become relevant to p as people's judgments on them are information on people's non-available reasons. If $X = \{a \vee b, \neg(a \vee b)\}$, the individuals' judgments on $a \vee b$ are relevant information for deciding $a \vee b$ as they reflect (partially) people's non-available reasons; hence we have $(a \vee b)\mathcal{R}(a \vee b)$ (but not so if X contains all reasons of $a \vee b$).

[23] I make no conjecture on the nature of *minimal* criteria for the (im)possibilities considered below, except that such conditions would not have a unified or structured form but the form of disjunctions of several cases. The reason is that the conditions must capture the joint and non-separable behaviour of relevance and logical connections, which is left general and uncontrolled in the framework.

(i.e. III) aggregation rules? I now introduce *priority rules* (generalising List 2004) and give simple criteria for when they can be used.

An impossibility threat comes not only from logical interconnections between root propositions (or other propositions), but also from transitivity violations of relevance $\mathcal{R}$. To see why, let $p \in X$ and suppose the premises of p's premises – call them the "pre-premises" – are *not* premises of p. The decision on p is settled by the decisions on p's premises (by "no underdetermination"), which in turn depend on how people judge the pre-premises (by III). This forces the decision on p to be some function f of how people judge the pre-premises. But by III the decision on p must be a function of how people judge p's premises (not pre-premises). So f depends on people's pre-premise judgments only indirectly: only through people's premise judgments as entailed by their pre-premise judgments – a strong restriction on f that suggests that impossibility is looming.

It is debatable whether premisehood (more generally, relevance) is inherently a transitive concept. If $\mathcal{R}$ is assumed transitive – whether for conceptual reasons or just to remove one impossibility source – interesting candidates for III aggregation arise, as explained now. List (2004) introduces *sequential priority rules* in judgment aggregation (generalising sequential rules in standard social choice theory). Here the propositions of a (finite) agenda are put in a priority order $p_1, p_2, ...$ and decided sequentially, where earlier decisions logically constrain later ones. As is easily seen, such a rule is III if relevance is given by $p_j \mathcal{R} p_{j'} \Leftrightarrow j \leq j'$, a linear order on X^+. I now introduce similar rules relative to an arbitrary (possibly quite partial) relevance relation. Informally, these rules decide the propositions in the order of relevance: each $p \in X$ is decided by logical entailment from previously accepted relevant propositions *except* if the latter propositions do not settle p, in which case p is decided via some local decision method (e.g. via majority voting on p). Formally, a *priority rule* is an aggregation rule F with universal domain such that there is for every proposition $p \in X^+$ a ("local") aggregation rule D_p for the binary agenda $\{p, \neg p\}$ (where D_p has for this agenda universal domain and consistent and complete outcomes) with

$$F(A_1, ..., A_n) \cap \{p, \neg p\} = \begin{cases} \{\tilde{p} \in \{p, \neg p\} : F(A_1, ..., A_n) \cap \mathcal{R}(p) \backslash \{p, \neg p\} \vdash \tilde{p}\} & \text{if this set is non-empty} \\ D_p(A_1 \cap \{p, \neg p\}, ..., A_n \cap \{p, \neg p\}) & \text{otherwise} \end{cases} \quad (4)$$

for all profiles $(A_1, \ldots, A_n) \in \mathcal{J}^n$. So the pair $p, \neg p$ is decided locally via D_p unless the previous decisions $F(A_1, ..., A_n) \cap \mathcal{R}(p) \backslash \{p, \neg p\}$ are logically constraining (hence "priority" to the previous decisions). In practice, first every root proposition $p \in X^+$ and $\neg p$ are decided by a local vote using D_p. Then every non-root proposition $p \in X^+$ to which only root propositions (and possibly p and $\neg p$) are relevant is decided: *either* by entailment from the previous decisions on relevant root propositions *or* (if neither p nor $\neg p$ is entailed) by a local vote using D_p. And so on.

The local rule D_p may be chosen as the same rule for all $p \in X^+$ (e.g. majority rule). Or D_p may vary: D_p might assign more weight to individuals with expertise on p (e.g. to physicists if p is "Nuclear energy is safe"), or to individuals personally affected by the decision on p (e.g. to the citizens of towns X and Y if p is "A road between X and Y should be built"). Such "expert rights" or "liberal rights" are (unlike those in Dietrich and List 2004) *conditional rights*: they can be overruled by previous decisions on relevant propositions. If the group can be partitioned into experts on different fields, and a proposition q's premises fall each into exactly one of the fields, the decision on each premise $p \in X^+$ of q could be delegated entirely to the experts on p, i.e. D_p uses only these experts' judgments. This generalises List's (2005) *distributed premise-based procedure*. The premises from different fields might form different subtrees preceding q.

Once we specify the family $(D_p)_{p \in X^+}$ of local rules, the recursive formula (4) defines a unique priority rule $F_{(D_p)_{p \in X^+}} := F$, provided that relevance $\mathcal{R}$ is *well-founded* on X^+.[24]

The following theorem shows that, for transitive relevance, $F_{(D_p)_{p \in X^+}}$ (a) satisfies III, and (b) generates consistent outcomes if certain logical independencies hold within X. Result (a) is surprising in one respect: one might have expected that III can be violated for non-self-relevant $p \in X$ due to the second case in (4). Let me motivate result (b). $F_{(D_p)_{p \in X^+}}$ could generate inconsistent outcomes if there are logical dependencies between root propositions, or more generally between any propositions $p_i \in X, i \in I$, that are *mutually irrelevant* (i.e. for no distinct $i, i' \in I$ $p_i \mathcal{R} p_{i'}$). To see why, notice that no p_i's precede other p_i's in the priority order (by irrelevance), whence the decisions on the p_i's ignore each other. But even if the (mutually irrelevant) p_i's are logical independent, inconsistent outcomes may still arise if there are logical interconnections between the sets $\mathcal{R}(p_i), i \in I$, as is easily imagined. This is why result (b) requires certain logical independencies between the sets $\mathcal{R}(p_i), i \in I$. To define these logical independencies, some terminology is needed. As usual, negation-closed sets $A_i, i \in I$, are called *logically independent* if $\cup_{i \in I} B_i$ is consistent for all consistent sets $B_i \subseteq A_i, i \in I$. Logical independence fails whenever $A_i \cap A_{i'} \neq \emptyset$ for some $i \neq i'$, because the sets B_i and $B_{i'}$ can pick different members of a pair $p, \neg p \in A_i \cap A_{i'}$.[25] This "easy" way to render $\cup_{i \in I} B_i$ inconsistent is excluded in the following weaker definition. I call negation-closed sets $A_i, i \in I$, *logically quasi-independent* if $\cup_{i \in I} B_i$ is consistent for all consistent sets $B_i \subseteq A_i, i \in I$, such that any pair $p, \neg p$ in an intersection $A_i \cap A_{i'}$ $(i \neq i')$ has a member that is both in B_i and in $B_{i'}$. (So $A_i \cap A_{i'}$ has the same intersection with B_i as with $B_{i'}$).

The theorem moreover requires relevance $\mathcal{R}$ to be *vertically finite*: there is no infinite sequence $(p_k)_{k=1,2,...}$ in X^+ that is ascending (i.e. each p_k is relevant to and distinct from p_{k+1}) or descending (i.e. each p_{k+1} is relevant to and distinct from p_k). In short, the network of inter-relevances is nowhere "infinitely deep", but possibly "infinitely broad". This exclusion of "infinite relevance chains" is a debatable condition on the concept of relevance;[26] without it the theorem would not hold.

[24] $\mathcal{R}$ is *well-founded* on X^+ if every non-empty set $S \subseteq X^+$ has an $\mathcal{R}$-minimal element s (i.e. for no $r \in S \backslash \{s\}$ $r \mathcal{R} s$); or, more intuitively, if there is no infinite sequence $(p_k)_{k=1,2,...}$ in X^+ such that each p_{k+1} is relevant to and distinct from p_k. The priority rule $F \equiv F_{(D_p)_{p \in X^+}}$ uniquely exists because, for every $(A_1, ..., A_n) \in \mathcal{J}^n$, $F(A_1, ..., A_n)$ is the union of the sets $f(p) := F(A_1, ..., A_n) \cap \{p, \neg p\}$, $p \in X^+$, where the function f is uniquely defined on X^+ by recursion on $\mathcal{R}$ using the well-founded recursion theorem (e.g. Fenstad 1980). If $\mathcal{R}$ is *not* well-founded on X^+, there could exist no or many priority rules with local rules $(D_p)_{p \in X^+}$.

[25] Unless p is a tautology or contradiction.

[26] For instance, one might argue (like Gärdenfors 2006) that every proposition can, in principle, be explained in terms of more fundamental premises; this creates infinite descending relevance chains. On the other hand, realistic agendas might still be vertically finite: they might not *include* all arbitrarily fundamental premises.

Theorem 2 *Let relevance $\mathcal{R}$ be transitive, vertically finite and negation-invariant.*

(a) *Every priority rule satisfies III.*[27]

(b) *Every priority rule generates consistent judgment sets if, for all mutually irrelevant propositions $p_i \in X, i \in I$, the sets $\mathcal{R}(p_i), i \in I$, are logically quasi-independent.*

The logical quasi-independence condition reduces to a logical independence condition if no mutually irrelevant propositions share any relevant proposition – but often the relevance (premisehood) relation is not of this special kind. Consider for instance case 4 in Figure 1.[28] Or consider a scientific board using a priority rule to derive collective judgments on several scientific propositions: then mutually irrelevant propositions, e.g. "Species X survives in Hawaii" and "Species Y survives in Australia", might well share premises, e.g. "The ozone hole exceeds size Z" or general biological or chemical hypotheses.

Proof. Let $\mathcal{R}$ and X be as specified. I leave it to the author to verify that $\mathcal{R}$'s vertical finiteness implies (in fact, is equivalent to) the following: every non-empty set $S \subseteq X^+$ has an $\mathcal{R}$-maximal element s (i.e. for no $r \in S \backslash \{s\}$ $s\mathcal{R}r$) and an $\mathcal{R}$-minimal element s (i.e. for no $r \in S \backslash \{s\}$ $r\mathcal{R}s$). In short:

$$\max_{\mathcal{R}} S \neq \emptyset \text{ and } \min_{\mathcal{R}} S \neq \emptyset, \text{ for all } \emptyset \neq S \subseteq X^+. \tag{5}$$

In particular, $\mathcal{R}$ is well-founded on X^+. Let $F \equiv F_{(D_p)_{p \in X^+}}$ be a priority rule.

(a) To show III, I prove that all $p \in X^+$ have the following property: for all $(A_1, ..., A_n), (A'_1, ..., A'_n) \in \mathcal{J}^n$, if $A_i \cap \mathcal{R}(p) = A'_i \cap \mathcal{R}(p)$ for all i then

$$F(A_1, ..., A_n) \cap \{p, \neg p\} = F(A'_1, ..., A'_n) \cap \{p, \neg p\}. \tag{6}$$

Suppose for a contradiction that the property fails for some $p \in X^+$. By (5) there is a $p \in X^+$ that is $\mathcal{R}$-minimal such that the property fails. So there are $(A_1, ..., A_n), (A'_1, ..., A'_n) \in \mathcal{J}^n$ with $A_i \cap \mathcal{R}(p) = A'_i \cap \mathcal{R}(p)$ for all i such that (6) is false. By p's minimality property and $\mathcal{R}$'s transitivity,

$$F(A_1, ..., A_n) \cap \mathcal{R}(p) \backslash \{p, \neg p\} = F(A'_1, ..., A'_n) \cap \mathcal{R}(p) \backslash \{p, \neg p\}. \tag{7}$$

Let $Y := \{\tilde{p} \in \{p, \neg p\} : \text{the set (7) entails } \tilde{p}\}$.

Case 1: $Y \neq \emptyset$. Then, by the first case in (4), $F(A_1, ..., A_n) \cap \{p, \neg p\} = Y$, and for the same reason $F(A'_1, ..., A'_n) \cap \{p, \neg p\} = Y$. This implies (6), contradicting the choice of p.

Case 2: $Y = \emptyset$. Then, by the second case in (4), $F(A_1, ..., A_n) \cap \{p, \neg p\} = D_p(A_1 \cap \{p, \neg p\}, ..., A_n \cap \{p, \neg p\})$ and $F(A'_1, ..., A'_n) \cap \{p, \neg p\} = D_p(A'_1 \cap \{p, \neg p\}, ..., A'_n \cap \{p, \neg p\})$. These two sets are distinct (as (6) is violated), and so for some i $A_i \cap \{p, \neg p\} \neq A'_i \cap \{p, \neg p\}$. So, as $A_i \cap \mathcal{R}(p) = A'_i \cap \mathcal{R}(p)$, $\mathcal{R}(p)$ does not contain both of $p, \neg p$, hence contains none of $p, \neg p$ by negation-invariance. So the set (7) equals $F(A_1, ..., A_n) \cap \mathcal{R}(p)$, which contains a member of each pair $r, \neg r \in \mathcal{R}(p)$, and hence entails p or $\neg p$ by "no underdetermination". This contradicts that $Y = \emptyset$.

[27] This still holds if the vertical finiteness condition is weakened to well-foundedness on X^+.

[28] If $\mathcal{R}$ is wished to be transitive, close the plotted inter-relevances under transitivity.

(b) Assume the condition. For all $p \in X$, put $\mathcal{R}^p := \mathcal{R}(p) \cup \{p, \neg p\}$ and $\mathcal{R}_p := \mathcal{R}(p) \backslash \{p, \neg p\}$. Let $(A_1, ..., A_n) \in \mathcal{J}^n$. The (desired) consistency of $A := F(A_1, ..., A_n)$ follows from the following claims.

Claim 1: $X = \cup_{p \in \max_{\mathcal{R}} X^+} \mathcal{R}^p$; in particular, $A = \cup_{p \in \max_{\mathcal{R}} X^+}(A \cap \mathcal{R}^p)$.

Claim 2: for any mutually irrelevant propositions $p_i \in X, i \in I$, the sets $\mathcal{R}^{p_i}, i \in I$, are logically quasi-independent; in particular, the sets $\mathcal{R}^p, p \in \max_{\mathcal{R}} X^+$, are logically quasi-independent.

Claim 3: $A \cap \mathcal{R}^p$ is consistent for all $p \in X^+$ (hence for all $p \in \max_{\mathcal{R}} X^+$).

Proof of Claim 1. For a contradiction, suppose $X \backslash \cup_{p \in \max_{\mathcal{R}} X^+} \mathcal{R}^p \neq \emptyset$. Then, by negation-invariance, $X^+ \backslash \cup_{p \in \max_{\mathcal{R}} X^+} \mathcal{R}^p \neq \emptyset$. Hence by (5) there is a $q \in \max_{\mathcal{R}} \left(X^+ \backslash \cup_{p \in \max_{\mathcal{R}} X^+} \mathcal{R}^p\right)$. By $q \notin \cup_{p \in \max_{\mathcal{R}} X^+} \mathcal{R}^p$, $q \notin \max_{\mathcal{R}} X^+$. So q is relevant to some $q' \in X^+ \backslash \{q\}$. We have $q' \notin X^+ \backslash \cup_{p \in \max_{\mathcal{R}} X^+} \mathcal{R}^p$, as q is maximal in $X^+ \backslash \cup_{p \in \max_{\mathcal{R}} X^+} \mathcal{R}^p$. So $q' \in \cup_{p \in \max_{\mathcal{R}} X^+} \mathcal{R}^p$. Hence, as $\mathcal{R}$ is transitive, q is relevant to some $p \in \max_{\mathcal{R}} X^+$, a contradiction as $q \notin \cup_{p \in \max_{\mathcal{R}} X^+} \mathcal{R}^p$.

Proof of Claim 2. Consider mutually irrelevant $p_i \in X, i \in I$, and consistent sets $B_i \subseteq \mathcal{R}^{p_i}, i \in I$, such that any pair $p, \neg p$ in an intersection $\mathcal{R}^{p_i} \cap \mathcal{R}^{p_{i'}}$ $(i \neq i')$ has a member that is in B_i and in $B_{i'}$. I show that $\cup_{i \in I} B_i$ is consistent. W.l.o.g. let each B_i contain a member of each pair $p, \neg p \in \mathcal{R}^{p_i}$ (otherwise extend the B_i's to consistent sets $\bar{B}_i \subseteq \mathcal{R}^{p_i}$ with the property; the present proof shows the consistency of $\cup_{i \in I} \bar{B}_i$, hence of $\cup_{i \in I} B_i$). As the sets $\mathcal{R}(p_i), i \in I$, are logically quasi-independent, (*) $\cup_{i \in I}(B_i \cap \mathcal{R}(p_i))$ is consistent. By "no underdetermination", (**) each $B_i \cap \mathcal{R}(p_i)$ entails a $\tilde{p}_i \in \{p_i, \neg p_i\}$. Each B_i is either $(B_i \cap \mathcal{R}(p_i)) \cup \{\tilde{p}_i\}$ or $(B_i \cap \mathcal{R}(\tilde{p}_i)) \cup \{\neg \tilde{p}_i\}$; so, as the latter set is inconsistent by (**) whereas B_i is consistent, $B_i = (B_i \cap \mathcal{R}(p_i)) \cup \{\tilde{p}_i\}$. Hence $\cup_{i \in I} B_i = \cup_{i \in I}((B_i \cap \mathcal{R}(p_i)) \cup \{\tilde{p}_i\})$, which is consistent by (*) and (**).

Proof of Claim 3. Suppose the contrary: there is a $p \in X^+$ for which $A \cap \mathcal{R}^p$ is inconsistent. By (5), there is a $p \in X^+$ that is $\mathcal{R}$-maximal subject to $A \cap \mathcal{R}^p$ being inconsistent. By an argument similar to that for Claim 1,

$$\mathcal{R}_p = \cup_{q \in \max_{\mathcal{R}}(X^+ \cap \mathcal{R}_p)} \mathcal{R}^q; \text{ hence } A \cap \mathcal{R}_p = \cup_{q \in \max_{\mathcal{R}}(X^+ \cap \mathcal{R}_p)}(A \cap \mathcal{R}^q). \tag{8}$$

By Claim 2, the sets $\mathcal{R}^q, q \in \max_{\mathcal{R}}(X^+ \cap \mathcal{R}_p)$, are logically quasi-independent. Hence, as each $A \cap \mathcal{R}^q$ in (8) is consistent (by the maximality of p), the set $A \cap \mathcal{R}_p$ is consistent. So, as $A \cap \{p, \neg p\}$ is related via (4) to $A \cap \mathcal{R}_p$ $(= A \cap \mathcal{R}(p) \backslash \{p, \neg p\})$, the union $(A \cap \mathcal{R}_p) \cup (A \cap \{p, \neg p\}) = A \cap \mathcal{R}^p$ is also consistent. This contradicts the choice of p. ■

7 A defensible (restricted) unanimity condition

It is natural to require (as in later theorems) a unanimity preservation condition. But it would be against the relevance-based approach to require *global* unanimity preservation, i.e. to require for *all* $p \in X$ that a unanimity for p implies social acceptance of p. Indeed, a unanimity for p can be *spurious*: different individuals i can hold p for different reasons, that is (in the relevance terminology) they may hold different explanations $E_i \subseteq \{r, \neg r : r\mathcal{R}p\}$ of p.[29] I will not require spurious unanimities to be respected. This follows the frequent view that spurious unanimities have less normative force. It also follows our relevance-based approach, since propositions relevant

[29] On spurious unanimities, see for instance Mongin (2005-b) and Bradley (forthcoming).

to p should not suddenly be treated as irrelevant if a unanimity accepts p. Instead, I will impose a unanimity condition restricted to a fixed set $\mathcal{P} \subseteq X$ of "privileged" propositions:

Agreement Preservation. For every profile $(A_1, ..., A_n)$ in the domain and every privileged proposition $p \in \mathcal{P}$, if $p \in A_i$ for all individuals i then $p \in F(A_1, ..., A_n)$.

I assume that $\mathcal{P}$ is chosen such that a unanimity for a $p \in \mathcal{P}$ cannot be spurious, i.e. such that each $p \in \mathcal{P}$ can be explained in just one way:[30]

$$\mathcal{P} \subseteq \{p \in X : p \text{ has a single } \mathcal{R}\text{-explanation}\}. \tag{9}$$

By default (i.e. if $\mathcal{P}$ is not explicitly defined otherwise), I assume that (9) holds with a "=". This maximal choice of $\mathcal{P}$ is often natural, though not necessary for the theorems below. Another potentially natural choice is to include in $\mathcal{P}$ only propositions $p \in X$ to which just p itself is relevant; so $\{p\}$ is p's only explanation, i.e. p has no "external" explanation.[31]

In the "classical" case that each $p \in X$ is just self-relevant, $\mathcal{P}$ by default equals X,[32] i.e. agreement preservation applies globally. So the "classical" relevance notion renders not only III equivalent to standard independence but also agreement preservation equivalent to standard (proposition-wise) unanimity preservation. If $X^+ = \{a, b, a \wedge b\}$, with (negation-invariant) relevance given by (3), $\mathcal{P}$ may contain $a \wedge b$ (which has a single explanation: $\{a, b\}$) but not $\neg(a \wedge b)$ (which has three explanations: $\{\neg a, b\}, \{a, \neg b\}, \{\neg a, \neg b\}$). So a unanimity for $\neg(a \wedge b)$ can be spurious and need not be respected.

Example 1 (continued) For the preference agenda, agreement preservation is equivalent to the weak Pareto principle, in virtue of defining $\mathcal{P}$ as

$$\mathcal{P} := \{\neg xRy : x, y \in Q, x \neq y\} = \{yPx : x, y \in Q, x \neq y\}, \tag{10}$$

the set of *strict* ranking propositions yPx. I call (10) the *Arrowian* set of privileged propositions. Note that (under the Arrowian relevance relation) each $\neg xRy = yPx \in \mathcal{P}$ has indeed a single explanation: $\{\neg xRy, yRx\}$.

8 Semi-vetodictatorship and semi-dictatorship

Hereafter, we consider an III and agreement preserving aggregation rule $F : \mathcal{J}^n \to \mathcal{J}$, relative to some fixed relevance relation $\mathcal{R}$ and some fixed set of privileged propositions $\mathcal{P}$. I give conditions (on logical and relevance links) that force F to be degenerate: a (semi-)dictatorship or (semi-)vetodictatorship.

First, how should these degenerate rules be defined? The relevance-based framework allows us to generalise the standard social-choice-theoretic definitions. Recall

[30] "Agreement" means "non-spurious unanimity". Hence the term "agreement preservation".

[31] Interesting normative questions can be raised about the choice of $\mathcal{P}$. For instance, Nehring (2005) suggests in his analysis of the Pareto/unanimity condition that unanimities are normatively binding if they reflect "self-interested" judgments, or if they carry "epistemic priority".

[32] Unless X contains contradictions: these have *no* explanation, hence are not in $\mathcal{P}$.

that an (Arrowian) "dictator" is an individual who can socially enforce his *strict* preferences between options, but not necessarily his *indifferences.* Similarly, a "vetodictator" can prevent ("veto") any *strict* preference, but not necessarily any indifference. Put in our terminology, a dictator (vetodictator) can enforce (veto) any *privileged* proposition of the preference agenda (given (10)). The following definitions generalise this to arbitrary agendas.

Definition 1 *An individual i is*

(a) *a* dictator *(respectively,* semi-dictator*) if, for every privileged proposition $p \in \mathcal{P}$, we have $p \in F(A_1, ..., A_n)$ for all $(A_1, ..., A_n) \in \mathcal{J}^n$ such that $p \in A_i$ (respectively, such that $p \in A_i$ and $p \notin A_j$, $j \neq i$);*

(b) *a* vetodictator *(respectively,* semi-vetodictator*) if, for every privileged proposition $p \in \mathcal{P}$, i has a* veto *(respectively,* semi-veto*) on p, i.e. a judgment set $A_i \in \mathcal{J}$ not containing p such that $p \notin F(A_1, ..., A_n)$ for all $A_j \in \mathcal{J}$, $j \neq i$ (respectively, for all $A_j \in \mathcal{J}$, $j \neq i$, containing p).*

In the standard models without a relevance relation, *conditional entailment* between propositions (first used by Nehring and Puppe 2002/2005) has proven useful to understand agendas. Roughly, $p \in X$ conditionally entails $q \in X$ if p together with other propositions in X entails q (with a non-triviality condition on the choice of "other" propositions). I cannot use conditional entailments here, as they reflect only *logical* links between propositions. Rather, I now define *constrained entailments*, a related notion that reflects both logical and relevance links. It will turn out that certain paths of constrained entailments lead to degenerate aggregation rules.[33]

Definition 2 *For propositions $p, q \in X$, if $\{p\} \cup Y \vdash q$ for a set $Y \subseteq \mathcal{P}$ consistent with every explanation of p and with every explanation of $\neg q$, I say that p* constrained entails q (in virtue of Y)*, and I write $p \vdash_* q$ or $p \vdash_Y q$.*[34]

The amount of constrained entailments in X is crucial for whether impossibilities arise. Trivially, every *unconditional* entailment is also a constrained entailment (namely in virtue of $Y = \emptyset$). Intuitively, if there are more inter-relevances between propositions, $\mathcal{P}$ becomes smaller, and propositions have more and larger explanations; so the requirements on Y in constrained entailments become stronger; hence there are fewer constrained entailments, and more room for possibilities of aggregation.

The preference agenda X (Example 1, with Arrowian $\mathcal{R}$ and $\mathcal{P}$) displays many constrained entailments (hence impossibilities). For instance, $xRy \vdash_{\{yPz\}} xPz$ (if x, y, z are pairwise distinct options), as yPz is in $\mathcal{P}$ and is consistent with each explanation

[33] Nehring and Puppe (2002/2005) use paths of *conditional* entailment to define their *totally blocked* agendas. For such agendas, they obtain strong dictatorship by imposing that F satisfies proposition-wise independence, an unrestricted unanimity condition, and a monotonicity condition. I impose relevance-based conditions on F (III and agreement preservation); these conditions, like Arrow's conditions, imply less than strong dictatorship.

[34] Many alternative notions of constrained entailment turn out to be non-suitable: they do not preserve interesting properties along paths of constrained entailments. The present definition is the weakest one to preserve semi-winning coalitions. The requirement that $Y \subseteq \mathcal{P}$ allows one to apply agreement preservation. In view of different results to those derived here, it might be fruitful to impose additional requirements on Y, e.g. that Y be consistent also with explanations of $\neg p$ and/or of q.

of xRy ($\{xRy, yRx\}$ and $\{xRy, \neg yRx\}$) and the only explanation of $\neg xPz = zRx$. By contrast, *no* non-trivial constrained entailments arise in our example $X^+ = \{a, b, a \wedge b\}$ with (negation-invariant) relevance given by (3): for instance, it is not the case that $a \vdash_{\{\neg(a \wedge b)\}} \neg b$ since $\neg(a \wedge b) \notin \mathcal{P}$; and it is not the case that $a \vdash_{\{b\}} a \wedge b$, as $\{b\}$ is inconsistent with the explanation $\{a, \neg b\}$ of $\neg(a \wedge b)$. As a result, our impossibilities will not apply to this agenda – and cannot, as the premise-based procedure for odd n (see Example 4) satisfies all conditions.

To obtain impossibility results, richness in constrained entailments is not sufficient. At least one constrained entailment $p \vdash_* q$ must hold in a "truly" constrained sense. By this I mean more than that p does not *un*conditionally entail q, i.e. more than that p is consistent with $\neg q$: I mean that every explanation of p is consistent with every explanation of $\neg q$.

Definition 3 *For propositions $p, q \in X$, p* truly constrained entails *q if $p \vdash_* q$ and moreover every explanation of p is consistent with every explanation of $\neg q$.*

For instance, if relevance is an equivalence relation (as in Examples 1-3) that partitions X into *pairwise logically independent* subagendas[35] (as for the preference agenda) then all constrained entailments across equivalence classes are truly constrained. Also, $p \vdash_* q$ is truly constrained if $p \nvdash q$ and moreover p and q are root propositions (see Example 4).

Our impossibility results rest on the following path conditions.

Definition 4 (a) *For propositions $p, q \in X$, if X contains propositions $p_1, ..., p_m$ ($m \geq 2$) with $p = p_1 \vdash_* p_2 \vdash_* ... \vdash_* p_m = q$, I write $p \vdash\vdash q$; if moreover one of these constrained entailments is truly constrained, I write $p \vdash\vdash_{\text{true}} q$.*

(b) *A set $Z \subseteq X$ is* pathlinked (in X) *if $p \vdash\vdash q$ for all $p, q \in Z$, and* truly pathlinked (in X) *if moreover $p \vdash\vdash_{\text{true}} q$ for some (hence all) $p, q \in Z$.*

While pathlinkedness forces to a limited form of *neutral* aggregation (see Lemma 3), true pathlinkedness forces to the following degenerate aggregation rules.

Theorem 3 *If the set $\mathcal{P}$ of privileged propositions is inconsistent and truly pathlinked, there is a semi-vetodictator.*

Theorem 4 *If the set $\{p, \neg p : p \in \mathcal{P}\}$ of privileged or negated privileged propositions is truly pathlinked, there is a semi-dictator.*

In the present (and all later) theorems, the qualification "truly" can be dropped if relevance is restricted to taking a form for which pathlinkedness (of the set in question) implies true pathlinkedness, for instance if $\mathcal{R}$ is restricted to being an equivalence relation that partitions X into logically independent subagendas.[36]

[35] That is, if X_1, X_2 are distinct subagendas, $A \cup B$ is consistent for all consistent $A \subseteq X_1, B \subseteq X_2$.

[36] The argument for the latter is as follows. By Lemma 1, all constrained entailments *within* any of the subagendas are unconditional entailments. This implies that the pathlinked set in question contains propositions linked by a path containing a constrained entailment *across* subagendas. The latter is truly constrained by an earlier remark.

Under the conditions of Theorems 3 and 4, there may be more than one semi-(veto)dictator, and moreover there need not exist any (veto)dictator.[37]

There are many applications. The preference agenda (Example 1) is discussed later. If in Example 4 we let $\mathcal{P}$ be the set of root propositions, and if these root propositions are interconnected in the sense of Theorem 4 (3), then some individual is semi-(veto)decisive on all "fundamental issues"; and hence, premise-based or prioritarian aggregation rules take a degenerate form (at least with respect to the local decision methods D_p for root propositions $p \in X$). Let me discuss Example 2 in more detail.

Example 2 (continued) For many instances of this aggregation problem (of judging values of and constraints between variables), the conditions of Theorems 3 and 4 hold, so that semi-(veto)dictatorships are the only solutions. To make this point, let relevance be again given by (2), and let the privileged propositions be given by

$$\mathcal{P} = \{V = v : V \in \mathbf{V} \& v \in Rge(V)\} \cup \{c, \neg c : c \in \mathbf{C}\}. \tag{11}$$

Also, let $|\mathbf{V}| \geq 2$ (to make it interesting), and assume[38]

$$\{\neg c : c \in \mathbf{C}\} \notin \mathcal{J}^*. \tag{12}$$

First, consider Theorem 3. Obviously, $\mathcal{P}$ is inconsistent, as $\mathbf{C} \neq \emptyset$ by (12). Often, $\mathcal{P}$ is also truly pathlinked. The latter could be shown by establishing that

(a) $\mathcal{P}_1 := \{V = v : V \in \mathbf{V} \& v \in Rge(V)\}$ is truly pathlinked, and

(b) for all $c \in \mathbf{C}$ there are $p, q, r, s \in \mathcal{P}_1$ with $c \vdash_* p$, $q \vdash_* c$, $\neg c \vdash_* r$, $s \vdash_* \neg c$.

Part (a) might even hold in the sense of, for all $V = v, V' = v' \in \mathcal{P}_1$ with $V \neq V'$, a truly constrained entailment $V = v \vdash_* V' = v'$ (rather than an indirect path $V = v \vdash\vdash V' = v'$); indeed, there might be a set of constraints $C \subseteq \mathbf{C}$ and a set of value assignments $D \subseteq \mathcal{P}_1$ such that $V = v \vdash_{C \cup D} V' = v'$ (hence, under the constraints in C, the set of value assignments $\{V = v\} \cup D$ implies that $V' = v'$).

Part (b) might hold for the following reasons. Consider a constraint $c \in \mathbf{C}$. Plausibly, $V = v \vdash_D \neg c$ for some $V = v \in \mathcal{P}_1$ and $D \subseteq \mathcal{P}_1$; here, $\{V = v\} \cup D$ is a set of value assignments violating the constraint c. It is also plausible that $c \vdash_D V = v$ for some $V = v \in \mathcal{P}_1$ and some $D \subseteq \mathcal{P}_1$; here, the value assignments in D imply, under the constraint c, that $V = v$. Moreover, we could have $V = v \vdash_D c$ for some $V = v \in \mathcal{P}_1$ and $D \subseteq \mathcal{P}_1$: this is so if the set of value assignments $\{V = v\} \cup D$

[37] Suppose $\mathcal{R}(p) = \{p\}$ for all $p \in X$ (only *self*-relevance allowed), $\mathcal{P} = X$ (all propositions privileged), and $|X| < \infty$. Then constrained entailment reduces to standard conditional entailment, and pathlinkedness of X reduces to Nehring and Puppe's (2002/2005) *total blockedness* condition whereby there is a path of *conditional* entailments between any $p, q \in X$. Dokow and Holzman (2005) show that *parity rules* F, defined on $\mathcal{J}^n$ by $F(A_1, ..., A_n) = \{p \in X : |\{i \in M : p \in A_i\}|$ is odd$\}$ for an odd-sized subgroup $M \subseteq N$, take values in $\mathcal{J}$ for certain agendas X that are totally blocked (hence pathlinked, in fact *truly* pathlinked) and satisfy an algebraic condition. Such a parity rule is also III and agreement preserving, and hence provides the required counterexample because every $i \in M$ a semi-dictator and a semi-vetodictator, but not a dictator and not a vetodictator (unless $|M| = 1$).

[38] Condition (12) requires that *at least one* constraint between variables holds, i.e. that the variables are not totally independent from each other. This assumption is natural in cases where the question is not *whether* but only *how* the variables affect each other, as it is the case for macroeconomic variables.

violates all constraints in $\mathbf{C}$ except c, hence entails c by (12). Finally, we could have $\neg c \vdash_{\tilde{C} \cup D} V = v$ for some $V = v \in \mathcal{P}_1$ and $D \subseteq \mathcal{P}_1$, and some set $\tilde{C}$ of negated constraints; indeed, suppose $\{\neg c\} \cup \tilde{C}$ contains the negations of all except of one constraint in $\mathbf{C}$, hence entails the remaining constraint by (12); under this remaining constraint, the value assignments in D could imply that $V = v$.

Now consider Theorem 4. The special form (11) of $\mathcal{P}$ in fact implies that the conditions of Theorem 4 hold whenever those of Theorem 3 hold (hence in many cases, as argued above). Specifically, let $\mathcal{P}$ be truly pathlinked. To prove that also $\{p, \neg p : p \in \mathcal{P}\}$ is truly pathlinked, it suffices to show that, for all $V = v \in \mathcal{P}$, there is a $p \in \mathcal{P}$ with $\neg(V = v) \vdash\vdash p$ and $p \vdash\vdash \neg(V = v)$. Consider any $V = v \in \mathcal{P}$, and choose any $p \in \mathbf{C}$ ($\subseteq \mathcal{P}$). As $\mathcal{P}$ is pathlinked and by (11) contains $\neg p$ and $V = v$, we have $\neg p \vdash\vdash V = v$ and $V = v \vdash\vdash \neg p$; hence (using Lemma 4 below) $\neg(V = v) \vdash\vdash p$ and $p \vdash\vdash \neg(V = v)$, as desired.

I now derive lemmas that will help both prove the theorems and understand constrained entailment. I first give a sufficient condition for when a constrained entailment reduces to an unconditional entailment.

Lemma 1 *For all $p, q \in X$ with $\mathcal{R}(p) \subseteq \mathcal{R}(\neg q)$ or $\mathcal{R}(\neg q) \subseteq \mathcal{R}(p)$, $p \vdash_* q$ if and only if $p \vdash q$.*

Proof. Let p, q be as specified. Obviously, $p \vdash q$ implies $p \vdash_\emptyset q$. Suppose for a contradiction that $p \vdash_* q$, say $p \vdash_Y q$, but $p \nvdash q$. Then $\{p, \neg q\}$ is consistent. So there is an $B \in \mathcal{J}$ containing p and $\neg q$. Then

- the set $B \cap \{r, \neg r : r\mathcal{R}p\}$ is an explanation of p;
- the set $B \cap \{r, \neg r : r\mathcal{R}\neg q\}$ is an explanation of $\neg q$.

One of these two sets is a superset of the other one, as $\mathcal{R}(p) \subseteq \mathcal{R}(\neg q)$ or $\mathcal{R}(\neg q) \subseteq \mathcal{R}(p)$; call this superset A. As $p \vdash_Y q$, $A \cup Y$ is consistent. So, as $A \vdash p$ and $A \vdash \neg q$, $\{p, \neg q\} \cup Y$ is consistent. It follows that $\{p\} \cup Y \nvdash q$, in contradiction to $p \vdash_Y q$. ■

The next fact helps in choosing the set Y in a constrained entailment.

Lemma 2 *For all $p, q \in X$, if $p \vdash_* q$ then $p \vdash_Y q$ for some set Y containing no proposition relevant to p or to $\neg q$.*

Proof. Let $p, q \in X$, and assume $p \vdash_* q$, say $p \vdash_Y q$. The proof is done by showing that $p \vdash_{Y \backslash (\mathcal{R}(p) \cup \mathcal{R}(\neg q))} q$. Suppose for a contradiction that not $p \vdash_{Y \backslash (\mathcal{R}(p) \cup \mathcal{R}(\neg q))} q$. Then

(*) $\{p, \neg q\} \cup Y \backslash (\mathcal{R}(p) \cup \mathcal{R}(\neg q))$ is consistent.

I show that

(**) $p \vdash p'$ for all $p' \in Y \cap \mathcal{R}(p)$ and $\neg q \vdash q'$ for all $q' \in Y \cap \mathcal{R}(\neg q)$,

which together with (*) implies that $\{p, \neg q\} \cup Y$ is consistent, a contradiction since $p \vdash_Y q$. Suppose for a contradiction that $p' \in Y \cap \mathcal{R}(p)$ but $p \nvdash p'$. Then there is a $B \in \mathcal{J}$ containing p and $\neg p'$. The set $A := B \cap \{r, \neg r : r\mathcal{R}p\}$ does not entail $\neg p$, hence is an explanation of p (as $\mathcal{R}$ is a relevance relation). So $A \cup Y$ is consistent (as $p \vdash_Y q$), a contradiction since $A \cup Y$ contains both p' and $\neg p'$. For analogous reasons, for all $q' \in Y \cap X^l$ it cannot be that $\neg q \nvdash q'$. ■

Now I introduce notions of decisive and semi-decisive coalitions, and I show that semi-decisiveness is preserved along paths of constrained entailments.

Definition 5 *A possibly empty coalition $C \subseteq N$ is* decisive *(respectively,* semi-decisive*) for $p \in X$ if its members have judgment sets $A_i \in \mathcal{J}$, $i \in C$, containing p, such that $p \in F(A_1, ..., A_n)$ for all $A_i \in \mathcal{J}$, $i \in N \backslash C$ (respectively, for all $A_i \in \mathcal{J}$, $i \in N \backslash C$, not containing p).*

While a decisive coalition for p can (by appropriate judgment sets) always socially enforce p, a semi-decisive coalition can do so provided all other individuals reject p. Let $\mathcal{W}(p)$ and $\mathcal{C}(p)$ be the sets of decisive and semi-decisive coalitions for $p \in X$, respectively.

Lemma 3 *For all $p, q \in X$, if $p \vdash_* q$ then $\mathcal{C}(p) \subseteq \mathcal{C}(q)$. In particular, if $Z \subseteq X$ is pathlinked, all $p \in Z$ have the same semi-decisive coalitions.*[39]

Proof. Suppose $p, q \in X$, and $p \vdash_* q$, say $p \vdash_Y q$, where by Lemma 2 w.l.o.g. $Y \cap \mathcal{R}(p) = Y \cap \mathcal{R}(\neg q) = \emptyset$. Let $C \in \mathcal{C}(p)$. So there are sets $A_i \in \mathcal{J}$, $i \in C$, containing p, such that $p \in F(A_1, ..., A_n)$ for all $A_i \in \mathcal{J}$, $i \in N \backslash C$, containing $\neg p$. By Y's consistency with every explanation of p, it is possible to change each A_i, $i \in C$, into a set (still in $\mathcal{J}$) that contains every $y \in Y$ and has the same intersection with $\mathcal{R}(p)$ as A_i; this change preserves the required properties, i.e. it preserves that $p \in A_i$ for all $i \in C$ (as $\mathcal{R}$ is a relevance relation), and preserves that $p \in F(A_1, ..., A_n)$ for all $A_i \in \mathcal{J}$, $i \in N \backslash C$, containing $\neg p$ (by $Y \cap \mathcal{R}(p) = \emptyset$ and III). So we may assume w.l.o.g. that $Y \subseteq A_i$ for all $i \in C$. Hence, by $\{p\} \cup Y \vdash q$, all A_i, $i \in C$, contain q.

To establish that $C \in \mathcal{C}(q)$, I consider any sets $A_i \in \mathcal{J}$, $i \in N \backslash C$, all containing $\neg q$, and I show that $q \in F(A_1, ..., A_n)$. We may assume w.l.o.g. that $Y \subseteq A_i$ for all $i \in N \backslash C$, by an argument like the one above (using that Y is consistent with any explanation of $\neg q$, $\mathcal{R}$ is a relevance relation, $Y \cap \mathcal{R}(\neg q) = \emptyset$, and III). As $\{\neg q\} \cup Y \vdash \neg p$, all A_i, $i \in N \backslash C$, contain $\neg p$. Hence $p \in F(A_1, ..., A_n)$. Moreover, $Y \subseteq F(A_1, ..., A_n)$ by $Y \subseteq \mathcal{P}$. So, as $\{p\} \cup Y \vdash q$, $q \in F(A_1, ..., A_n)$, as desired. ■

I now prove Theorems 3 and 4, after stating a last (obvious) lemma.

Lemma 4 (contraposition) *For all $p, q \in X$ and all $Y \subseteq \mathcal{P}$, $p \vdash_Y q$ if and only if $\neg q \vdash_Y \neg p$.*

Proof of Theorem 3. Let $\mathcal{P}$ be inconsistent and truly pathlinked. I first prepare the proof by establishing three simple claims.

Claim 1. (i) The set $\mathcal{C}(p)$ is the same for all $p \in \mathcal{P}$; call it $\mathcal{C}_0$. (ii) The set $\mathcal{C}(\neg p)$ is the same for all $p \in \mathcal{P}$.

Part (i) follows from Lemma 3. Part (ii) follows from it too because, by Lemma 4, $\{\neg p : p \in \mathcal{P}\}$ is like $\mathcal{P}$ pathlinked, q.e.d.

Claim 2. $\emptyset \notin \mathcal{C}_0$ and $N \in \mathcal{C}_0$.

By agreement preservation, $N \in \mathcal{C}_0$. Suppose for a contradiction that $\emptyset \in \mathcal{C}_0$. Consider any judgment set $A \in \mathcal{J}$. Then $F(A, ..., A)$ contains all $p \in \mathcal{P}$, by $N \in \mathcal{C}_0$ if $p \in A$, and by $\emptyset \in \mathcal{C}_0$ if $p \notin A$. Hence $F(A, ..., A)$ is inconsistent, a contradiction, q.e.d.

[39] Constrained entailments preserve semi-decisiveness but usually not decisiveness.

By Claim 2, there is a minimal coalition C in $\mathcal{C}_0$ (with respect to inclusion), and $C \neq \emptyset$. By $C \neq \emptyset$, there is a $j \in C$. Write $C_{-j} := C \backslash \{j\}$. As $\mathcal{P}$ is truly pathlinked, there exist $p \in \mathcal{P}$ and $r, s \in X$ such that $p \vdash\vdash r$, $r \vdash_* s$ truly, and $s \vdash\vdash p$.

Claim 3. $\mathcal{C}(r) = \mathcal{C}(s) = \mathcal{C}_0$; hence $C \in \mathcal{C}(r)$ and $C_{-j} \notin \mathcal{C}(s)$.

By Lemma 3, $\mathcal{C}(p) \subseteq \mathcal{C}(r) \subseteq \mathcal{C}(s) \subseteq \mathcal{C}(p)$. So $\mathcal{C}(r) = \mathcal{C}(s) = \mathcal{C}(p) = \mathcal{C}_0$, q.e.d.

Now let Y be such that $r \vdash_Y s$, where by Lemma 2 w.l.o.g. $Y \cap \mathcal{R}(r) = Y \cap \mathcal{R}(\neg s) = \emptyset$. By $C \in \mathcal{C}(r)$, there are judgment sets $A_i \in \mathcal{J}$, $i \in C$, containing r, such that $r \in F(A_1, ..., A_n)$ for all $A_i \in \mathcal{J}$, $i \in N \backslash C$, not containing r. I assume w.l.o.g. that

$$\text{for all } i \in C_{-j},\ Y \subseteq A_i, \text{ hence (by } \{r\} \cup Y \vdash s) \ s \in A_i, \tag{13}$$

which I may do by an argument like that in the proof of Lemma 3 (using that Y is consistent with any explanation of q, $\mathcal{R}$ is a relevance relation, $Y \cap \mathcal{R}(r) = \emptyset$, and III). By (13) and as $C_{-j} \notin \mathcal{C}(s)$ (see Claim 3), there are sets $B_i \in \mathcal{J}$, $i \in N \backslash C_{-j}$, containing $\neg s$, such that, writing $B_i := A_i$ for all $i \in C_{-j}$,

$$\neg s \in F(B_1, ..., B_n). \tag{14}$$

I may w.l.o.g. modify the sets B_i, $i \in N \backslash C_{-j}$, into new sets in $\mathcal{J}$ as long as their intersections with $\mathcal{R}(\neg s)$ stays the same (because the new sets then still contain $\neg s$ as $\mathcal{R}$ is a relevance relation, and still satisfy (14) by III). First, I modify the set B_i for $i = j$: as $r \vdash_* s$ truly, $B_j \cap \{t, \neg t : t \in \mathcal{R}(\neg s)\}$ (an explanation of $\neg s$) is consistent with any explanation of r, hence with $A_j \cap \{t, \neg t : t \in \mathcal{R}(r)\}$, so that I may assume that $A_j \cap \{t, \neg t : t \in \mathcal{R}(r)\} \subseteq B_j$; which implies that

$$B_i \cap \mathcal{R}(r) = A_i \cap \mathcal{R}(r) \text{ for all } i \in C. \tag{15}$$

Second, I modify the sets B_i, $i \in N \backslash C$: I assume (using that $Y \cap \mathcal{R}(\neg s) = \emptyset$ and Y's consistency with any explanation of $\neg s$) that

$$\text{for all } i \in N \backslash C,\ Y \subseteq B_i, \text{ hence (as } \{\neg s\} \cup Y \vdash \neg r) \ \neg r \in B_i. \tag{16}$$

The definition of the sets A_i, $i \in C$, and (16) imply, via (15) and III, that

$$r \in F(B_1, ..., B_n). \tag{17}$$

By (14), (17), and the inconsistency of $\{r, \neg s\} \cup Y$, the set Y is not a subset of $F(B_1, ..., B_n)$. So there is a $y \in Y$ with $y \notin F(B_1, ..., B_n)$. We have $\{j\} \in \mathcal{C}(\neg y\}$ for the following two reasons.

- B_j contains $\neg y$; otherwise $y \in B_i$ for all $i \in N$, so that $y \in F(B_1, ..., B_n)$ by $y \in \mathcal{P}$.
- Consider any sets $C_i \in \mathcal{J}$, $i \neq j$, not containing $\neg y$, i.e. containing y. I show that $\neg y \in A := F(C_1, ..., C_{j-1}, B_j, C_{j+1}, ..., C_n)$. For all $i \neq j$, $C_i \cap \{t, \neg t : t \in \mathcal{R}(y)\}$ is consistent with y, hence is an explanation of y (as $\mathcal{R}$ satisfies "no underdetermination"); for analogous reasons, $B_i \cap \{t, \neg t : t \in \mathcal{R}(y)\}$ is an explanation of y. These two explanations must be identical by $y \in \mathcal{P}$. So $C_i \cap \mathcal{R}(y) = B_i \cap \mathcal{R}(y)$. Hence, by $y \notin F(B_1, ..., B_n)$ and III, $y \notin A$. So $\neg y \in A$, as desired.

By $\{j\} \in \mathcal{C}(\neg y)$ and Claim 1, $\{j\} \in \mathcal{C}(\neg q)$ for all $q \in \mathcal{P}$. So j is a semi-vetodictator. ■

Proof of Theorem 4. Let $\{p, \neg p : p \in \mathcal{P}\}$ be truly pathlinked. I will reduce the proof to that of Theorem 3. I start again with two simple claims.

Claim 1. The set $\mathcal{C}(q)$ is the same for all $q \in \{p, \neg p : p \in \mathcal{P}\}$; call it $\mathcal{C}_0$.

This follows immediately from Lemma 3, q.e.d.

Claim 2. $\emptyset \notin \mathcal{C}_0$ and $N \in \mathcal{C}_0$.

By agreement preservation, $N \in \mathcal{C}(p)$ for all $p \in \mathcal{P}$; hence $N \in \mathcal{C}_0$. This implies, for all $p \in \mathcal{P}$, that $\emptyset \notin \mathcal{C}(\neg p)$; hence $\emptyset \notin \mathcal{C}_0$, q.e.d.

Now by an analogous argument to that in the proof of Theorem 3, but based this time on the present Claims 1 and 2 rather than on the two first claims in Theorem 3's proof, one can show that there exists an individual j such that $\{j\} \in \mathcal{C}(\neg q)$ for all $q \in \mathcal{P}$. So, by the present Claim 1 (which is stronger than the first claim in Theorem 3's proof),

$$\{j\} \in \mathcal{C}(q) \text{ for all } q \in \mathcal{P}. \tag{18}$$

So j is a semi-dictator, for the following reason. Let $q \in \mathcal{P}$ and let $(A_1, ..., A_n) \in \mathcal{J}^n$ be such that $q \in A_j$ and $q \notin A_i$, $i \neq j$. By (18) there is a set $B_j \in \mathcal{J}$ containing q such that $q \in F(B_1, ..., B_n)$ for all $B_i \in \mathcal{J}$, $i \neq j$, not containing q. Since q has only one explanation (by $q \in \mathcal{P}$), the two explanations $A_j \cap \{t, \neg t : t \in \mathcal{R}(q)\}$ and $B_j \cap \{t, \neg t : t \in \mathcal{R}(q)\}$ are identical. So $A_j \cap \mathcal{R}(q) = B_j \cap \mathcal{R}(q)$. Hence, using III and the definition of B_j, $q \in F(A_1, ..., A_n)$, as desired. ■

9 Dictatorship and strong dictatorship

In fact, the semi-dictator of Theorem 4 is in many cases (including the preference aggregation problem) a dictator, and in some cases even a strong dictator in the sense of the following definition that generalises the classical notion of strong dictatorship in social choice theory.

Definition 6 *An individual i is a* strong dictator *if $F(A_1, ..., A_n) = A_i$ for all $(A_1, ..., A_n) \in \mathcal{J}^n$.*

So a strong dictator imposes his judgments on all rather than just privileged propositions. I will give simple criteria for obtaining (weak or strong) dictatorship, in terms of the following irreversibility property.

Definition 7 *For $p, q \in X$, p* irreversibly *constrained entails q if $p \vdash_Y q$ for a set Y for which $\{q\} \cup Y \nvdash p$.*

So a constrained entailment $p \vdash_* q$ is irreversible if the constrained entailment is not a "constrained equivalence", i.e. if p and q do not conditionally entail each other (for at least one choice of Y). If X is the preference agenda (with Arrowian $\mathcal{R}$ and $\mathcal{P}$), *all* constrained entailments between (distinct) propositions are irreversible. For instance, $xRy \vdash_* xRz$ is irreversible (for distinct options x, y, z), since $xRy \vdash_{\{yPz\}} xRz$, where $\{xRz, yPz\} \nvdash xRy$.

By the next result, the semi-dictatorship of Theorem 4 becomes a dictatorship if we only slightly strengthen the pathlinkedness condition: in *at least one* path, *at least one* constrained entailment should be irreversible.

Definition 8 (a) *For propositions $p, q \in X$, I write $p \vdash\vdash_{\text{irrev}} q$ if X contains propositions $p_1, ..., p_m$ ($m \geq 2$) with $p = p_1 \vdash_* p_2 \vdash_* ... \vdash_* p_m = q$, where at least one of these constrained entailments is irreversible.*

(b) *A pathlinked set $Z \subseteq X$ is* irreversibly pathlinked (in X) *if $p \vdash\vdash_{\text{irrev}} q$ for some (hence all) $p, q \in Z$.*

Theorem 5 *If the set $\{p, \neg p : p \in \mathcal{P}\}$ of privileged or negated privileged propositions is truly and irreversibly pathlinked, some individual is a dictator.*

As an application, I obtain the full Arrow theorem by proving that, if X is the preference agenda, $\{p, \neg p : p \in \mathcal{P}\}$ is truly and irreversibly pathlinked.[40]

Corollary 1 *(Arrow's Theorem) For the preference agenda (with Arrowian $\mathcal{R}$ and $\mathcal{P}$), some individual is a dictator.*

Proof. Let X be the preference agenda, with $\mathcal{R}$ and $\mathcal{P}$ Arrowian. I show that (i) $\mathcal{P}$ is pathlinked, and (ii) there are $r, s \in \mathcal{P}$ with true and irreversible constrained entailments $r \vdash_* \neg s \vdash_* r$. Then, by (i) and Lemma 4, $\{\neg p : p \in \mathcal{P}\}$ is (like $\mathcal{P}$) pathlinked, which together with (ii) implies that $\{p, \neg p : p \in \mathcal{P}\}$ is truly and irreversibly pathlinked, as desired.

(ii): For any pairwise distinct options $x, y, z \in Q$, we have $xPy \vdash_{\{yPz\}} xRz$ ($= \neg zRx$), and $xRz \vdash_{\{zPy\}} xPy$, in each case truly and irreversibly.

(i): Consider any $xPy, x'Py' \in \mathcal{P}$. I show that $xPy \vdash\vdash x'Py'$. The paths to be constructed depend on whether $x \in \{x', y'\}$ and whether $y \in \{x', y'\}$. As $x \neq y$ and $x' \neq y'$, the following list of cases is exhaustive. Case $x \neq x', y' \& y \neq x', y'$: $xPy \vdash_{\{x'Px, yPy'\}} x'Py'$. Case $y = y' \& x \neq x', y'$: $xPy \vdash_{\{x'Px\}} x'Py = x'Py'$. Case $y = x' \& x \neq x', y'$: $xPy \vdash_{\{yPy'\}} xPy' \vdash_{\{x'Px\}} x'Py'$. Case $x = x' \& y \neq y', x'$: $xPy \vdash_{\{yPy'\}} xPy'$. Case $x = y' \& y \neq x', y'$: $xPy \vdash_{\{x'Px\}} x'Py \vdash_{\{yPx\}} x'Px$. Case $x = x' \& y = y'$: $xPy \vdash_\emptyset xPy$. Case $x = y' \& y = x'$: taking any $z \in Q \backslash \{x, y\}$, $xPy \vdash_{\{yPz\}} xPz \vdash_{\{yPx\}} yPz \vdash_{\{zPx\}} yPx$. ■

The proof of Theorem 5 uses two further lemmas. For any set $\mathcal{S}$ of coalitions $C \subseteq N$, I define $\overline{\mathcal{S}} := \{C' \subseteq N : C \subseteq C' \text{ for some } C \in \mathcal{S}\}$.

Lemma 5 *For all $p, q \in X$,*

(a) *$p \vdash_* q$ irreversibly if and <u>only</u> if $\neg q \vdash_* \neg p$ irreversibly;*

(b) *if $p \vdash_* q$ irreversibly then $\overline{\mathcal{C}(p)} \subseteq \mathcal{C}(q)$.*

Proof. Let $p, q \in X$. Part (a) follows from Lemma 4 and the fact that, for all $Y \subseteq \mathcal{P}$, $\{q\} \cup Y \nvdash p$ if and only if $\{\neg p\} \cup Y \nvdash \neg q$.

[40] This property of $\{p, \neg p : p \in \mathcal{P}\}$ strengthens Nehring's (2003) finding that the preference agenda is totally blocked, which gave him already a weaker version of Arrow's theorem. Part (i) of our proof is analogous to Nehring's proof and also to proofs for the *strict* preference agenda by Dietrich and List (forthcoming-b) and Dokow and Holzman (2005).

Regarding (b), suppose $p \vdash_* q$ irreversibly, say $p \vdash_Y q$ with $\{q\} \cup Y \nvdash p$. We can assume w.l.o.g. that $Y \cap \mathcal{R}(p) = Y \cap \mathcal{R}(\neg q) = \emptyset$, since otherwise we could replace Y by $Y' := Y \backslash (\mathcal{R}(p) \cup \mathcal{R}(\neg q))$, for which still $p \vdash_{Y'} q$ (by the proof of Lemma 2) and $\{q\} \cup Y' \nvdash p$. To show $\overline{\mathcal{C}(p)} \subseteq \mathcal{C}(q)$, consider any $C' \in \overline{\mathcal{C}(p)}$. So there is a $C \in \mathcal{C}(p)$ with $C \subseteq C'$. Hence there are $A_i \in \mathcal{J}$, $i \in C$, containing p, such that $p \in F(A_1, ..., A_n)$ for all $A_i \in \mathcal{J}$, $i \in N \backslash C$, containing $\neg p$. Like in earlier proofs, I may suppose w.l.o.g. that, for all $i \in C$, $Y \subseteq A_i$ (using that Y is consistent with all explanations of p, $\mathcal{R}$ is a relevance relation, III, and $Y \cap \mathcal{R}(p) = \emptyset$); hence, by $\{p\} \cup Y \vdash q$, $q \in A_i$ for all $i \in C$. Further, as $\{\neg p, q\} \cup Y$ is consistent (by $\{q\} \cup Y \nvdash p$), there are sets $A_i \in \mathcal{J}$, $i \in C' \backslash C$, such that $\{\neg p, q\} \cup Y \subseteq A_i$ for all $i \in C' \backslash C$.

I have to show that $q \in F(A_1, ..., A_n)$ for all $A_i \in \mathcal{J}$, $i \in N \backslash C'$, containing $\neg q$. Consider such sets A_i, $i \in N \backslash C'$. Again, we may assume w.l.o.g. that for all $i \in N \backslash C'$, $Y \subseteq A_i$ (as Y is consistent with all explanations of $\neg q$, $\mathcal{R}$ is a relevance relation, III, and $Y \cap \mathcal{R}(\neg q) = \emptyset$), which by $\{\neg q\} \cup Y \vdash \neg p$ implies that $\neg p \in A_i$ for all $i \in N \backslash C'$. In summary then,

$$A_i \supseteq \begin{cases} \{p, q\} \cup Y & \text{for all } i \in C \\ \{\neg p, q\} \cup Y & \text{for all } i \in C' \backslash C \\ \{\neg p, \neg q\} \cup Y & \text{for all } i \in N \backslash C'. \end{cases}$$

So $p \in F(A_1, ..., A_n)$ (by the choice of the sets A_i, $i \in C$) and $Y \subseteq F(A_1, ..., A_n)$ (by $Y \subseteq \mathcal{P}$). Hence, as $\{p\} \cup Y \vdash q$, $q \in F(A_1, ..., A_n)$, as desired. ■

In the following characterisation of decisive coalitions it is crucial that $p \in \mathcal{P}$.

Lemma 6 *If $p \in \mathcal{P}$, $\mathcal{W}(p) = \{C \subseteq N :$ all coalitions $C' \supseteq C$ are in $\mathcal{C}(p)\}$.*

Proof. Let $p \in \mathcal{P}$ and $C \subseteq N$. If $C \in \mathcal{W}(p)$ then clearly all coalitions $C' \supseteq C$ are in $\mathcal{C}(p)$. Conversely, suppose all coalitions $C' \supseteq C$ are in $\mathcal{C}(p)$. As $C \in \mathcal{C}(p)$, there are sets A_i, $i \in C$, containing p, such that $p \in F(A_1, ..., A_n)$ for all sets A_i, $i \in N \backslash C$, not containing p. To show that $C \in \mathcal{W}(p)$, consider any sets A_i, $i \in N \backslash C$ (containing or not containing p); I show that $p \in F(A_1, ..., A_n)$. Let $C' := C \cup \{i \in N \backslash C : p \in A_i\}$. By $C \subseteq C'$, $C' \in \mathcal{C}(p)$. So there are sets B_i, $i \in C'$, containing p, such that $p \in F(B_1, ..., B_n)$ for all sets B_i, $i \in N \backslash C'$, not containing p. As p has a single explanation, we have for all $i \in C'$ $A_i \cap \{r, \neg r : r \in \mathcal{R}(p)\} = B_i \cap \{r, \neg r : r \in \mathcal{R}(p)\}$, hence $A_i \cap \mathcal{R}(p) = B_i \cap \mathcal{R}(p)$. So, by III and the definition of the sets B_i, $i \in C'$, and since $p \notin A_i$ for all $i \in N \backslash C'$, $p \in F(A_1, ..., A_n)$, as desired. ■

Proof of Theorem 5. Let $\{p, \neg p : p \in \mathcal{P}\}$ be truly and irreversibly pathlinked. By Theorem 4, there is a semi-dictator i. I show that i is a dictator.

Claim. For all $q \in \{p, \neg p : p \in \mathcal{P}\}$, $\mathcal{C}(q)$ contains all coalitions containing i.

Consider any $q \in \{q, \neg q : q \in \mathcal{P}\}$ and any coalition $C \subseteq N$ containing i. By true pathlinkedness there exist $p \in \mathcal{P}$ and $r, s \in X$ such that $p \vdash\vdash r \vdash_* s \vdash\vdash q$, where $r \vdash_* s$ is a truly constrained entailment. By $\{i\} \in \mathcal{C}(p)$ and Lemma 3, $\{i\} \in \mathcal{C}(r)$. So, by Lemma 5(b), $C \in \mathcal{C}(s)$. Hence, by Lemma 3, $C \in \mathcal{C}(q)$, q.e.d.

By this claim and Lemma 6, $\{i\} \in \mathcal{W}(p)$ for all $p \in \mathcal{P}$. This implies that i is a dictator, by an argument similar to the one that completed the proof of Theorem 4. ■

Finally, for what agendas do we even obtain *strong* dictatorship? Surely not for the preference agenda, as it is well-known that Arrow's conditions only imply weak dictatorship.[41]

Trivially, if *all* propositions are privileged, every dictatorship is strong:

Corollary 2 *If $\mathcal{P} = X$, and X is truly and irreversibly pathlinked, some individual is a strong dictator.*

But the assumption $\mathcal{P} = X$ removes nearly all generality: agreement preservation becomes global unanimity preservation, and the relevance relation is forced to give each $p \in X$ a single explanation. However, strong dictatorship follows under a much less restrictive condition than $\mathcal{P} = X$. Call $p \in X$ *logically equivalent to* $A \subseteq X$ if A entails p and p entails all $q \in A$ (i.e. intuitively, if p is equivalent to the conjunction of all $q \in A$). For instance, $a \wedge b$ is equivalent to $\{a, b\}$ (where $a, b, a \wedge b \in X$).

Theorem 6 *If $\{p, \neg p : p \in \mathcal{P}\}$ is truly and irreversibly pathlinked and each proposition in X is logically equivalent to a set of negated privileged propositions $A \subseteq \{\neg p : p \in \mathcal{P}\}$, some individual is a strong dictator.*

Proof. Let the assumptions hold. By Theorem 5, there is a dictator i. To show that i is a strong dictator, I consider any $(A_1, ..., A_n) \in \mathcal{J}^n$, and I show that $A_i = F(A_1, ..., A_n)$. Obviously, it suffices to show that $F(A_1, ..., A_n) \subseteq A_i$. Suppose $q \in F(A_1, ..., A_n)$. By assumption, q is logically equivalent to some $A \subseteq \{\neg p : p \in \mathcal{P}\}$. For all $\neg p \in A$, we have $\neg p \in F(A_1, ..., A_n)$ (by $q \vdash \neg p$), hence $p \notin F(A_1, ..., A_n)$, and so $p \notin A_i$ (as $p \in \mathcal{P}$ and i is a dictator), implying that $\neg p \in A_i$. This shows that $A \subseteq A_i$. So $q \in A_i$ (since $A \vdash q$), as desired. ■

The preference agenda X, which has *not strongly* dictatorial solutions, indeed violates the extra condition in Theorem 6: some propositions in X (namely precisely the privileged propositions xPy) are not logically equivalent to any set of negated privileged propositions xRy.

Example 2 (continued) As argued earlier, $\{p, \neg p : p \in \mathcal{P}\}$ is truly pathlinked in many instances of this aggregation problem. The other conditions in Theorem 6 also often hold, so that strong dictatorship follows. The reasons are simple.

First, X is often rich in irreversible constrained entailments. For instance, if X contains value assignments $V = 3$ and $W = 3$ and the constraint $W > V$, then $V = 3 \vdash_* \neg(W = 3)$ irreversibly, since $V = 3 \vdash_{\{W>V\}} \neg(W = 3)$ but $\{\neg(W = 3), W > V\} \nvdash V = 3$; or, if X contains a constraint c that is strictly stronger than another constraint $c' \in \mathbf{C}$, then $c \vdash_* c$ irreversibly, since $c \vdash_\emptyset c'$ but $\{c'\} \cup \emptyset = \{c'\} \nvdash c$.

Second, if $\mathcal{P}$ is again given by (11), each proposition $q \in X$ is indeed logically equivalent to a set of negated privileged propositions $A \subseteq \{\neg p : p \in \mathcal{P}\}$: if q has the form $V = v$, one should take $A = \{\neg(V = v') : v' \in Rge(V) \backslash \{v\}\}$; otherwise q has the form $\neg p$ with $p \in \mathcal{P}$, and so one may take $A = \{q\}$.

[41] Lexicographic dictatorships satisfy all conditions but are only weak dictatorships.

10 Conclusion

The impossibility findings might be interpreted as showing how relevance $\mathcal{R}$ should *not* be specified. Indeed, in order to enable non-degenerate III aggregation rules, $\mathcal{R}$ must display sufficiently many inter-relevances. But such richness in inter-relevances may imply that collective decisions have to be made in a "holistic" manner: many semantically unrelated decisions must be bundled and decided simultaneously. Two propositions, say one on traffic regulations and one on diplomatic relations with Argentina, have to be treated simultaneously if the relevance relation (specified sufficiently richly to enable non-degenerate aggregation rules) displays some possibly indirect link between the two.[42] Large and semantically disparate decision problems are a hard challenge in practice.

As I began to discuss in Section 3, several types of relevance relations, hence informational constraints, are of interest to aggregation theory. Which III aggregation rules are there if relevance is, for instance, transitive? Or asymmetric? Or well-founded? In addition to such questions, it is worth exploring further the premise-based and prioritarian approaches. If a distance-based approach compatible with the informational constraint III could be developed, the theories of belief merging and judgment aggregation would meet. Developing different types of III aggregation procedures should go hand in hand with developing objective criteria for when to consider a proposition as relevant to another, i.e. for which informational restriction to impose. This second research goal has a normative dimension. Reaching both goals would enable us to give concrete recommendations for practical group decision-making.

11 References

Bradley, R (forthcoming) Taking advantage of difference in opinion, *Episteme*, forthcoming

Chapman, B (2002) Rational Aggregation, Polit Philos Econ 1(3): 337-354

Claussen, CA, Roisland, O (2005) Collective economic decisions and the discursive paradox, working paper, Central Bank of Norway Research Division

Dietrich, F (2006) Judgment aggregation: (im)possibility theorems, *J Econ Theory* 126(1): 286-298

Dietrich, F (forthcoming) A generalised model of judgment aggregation, *Soc Choice Welfare*

Dietrich, F (2005) The possibility of judgment aggregation on agendas with subjunctive implications, working paper, Maastricht Univ

Dietrich, F, List, C (2004) A liberal paradox for judgment aggregation, working paper, London School of Economics

Dietrich, F, List, C (forthcoming-a) Majority voting generalized: judgment aggregation by quota rules, *J Theor Polit*

[42] That is, if the two propositions are related in terms of the transitive, symmetric and reflexive closure of relevance $\mathcal{R}$. This closure partitions the totality of propositions (questions) into equivalence classes of irreducible decision problems. If there is a single equivalence class, *all* decisions (including those on traffic regulations and on diplomatic relations) have to be treated simultaneously.

Dietrich, F, List, C (forthcoming-b) Arrow's theorem in judgment aggregation, *Soc Choice Welfare*
Dietrich, F, List, C (2006) Judgment aggregation on restricted domains, working paper, Maastricht Univ
Dokow, E, Holzman, R (2005) Aggregation of binary evaluations, working paper, Technion Israel Institute of Technology
Eckert, D, Pigozzi, G (2005) Belief merging, judgment aggregation and some links with social choice theory, in *Belief Change in Rational Agents*, J. Delgrande et al. (eds.)
Fenstad, JF (1980) *General recursion theory: an axiomatic approach*, Springer
Genest, C, Zidek, JV (1986) "Combining Probability Distributions: A Critique and Annotated Bibliography", *Statistical Science* 1(1): 113-135
Gärdenfors, P (2006) An Arrow-like theorem for voting with logical consequences, *Econ Philos* 22(2): 181-190
Konieczny, S, Pino-Perez, R (2002) Merging information under constraints: a logical framework, *J Logic Comput* 12(5): 773-808
List, C (2004) A model of path-dependence in decisions over multiple propositions, *Am Polit Sci Rev* 98(3): 495-513
List, C (2005) Group knowledge and group rationality: a judgment aggregation perspective, *Episteme* 2(1): 25-38
List, C, Pettit, P (2002) Aggregating sets of judgments: an impossibility result. *Econ Philos* 18: 89-110
List, C, Pettit, P (2004) Aggregating sets of judgments: two impossibility results compared. *Synthese* 140(1-2): 207-235
Mongin, P (2005-a) Factoring out the impossibility of logical aggregation, working paper, CNRS, Paris
Mongin, P (2005-b) Spurious Unanimity and the Pareto Principle, *LSE Choice Group Working Papers* 1 (5)
Nehring, K (2003) Arrow's theorem as a corollary. *Econ Letters* 80(3): 379-382
Nehring, K (2005) The (im)possibility of a Paretian rational, working paper, Univ California at Davies
Nehring, K, Puppe, C (2002) Strategy-proof social choice on single-peaked domains: possibility, impossibility and the space between, working paper, Univ California at Davies
Nehring, K, Puppe, C (2005) The structure of strategy-proof social choice, part II: non-dictatorship, anonymity and neutrality, working paper, Karlsruhe Univ
Nehring, K., Puppe, C (2006) Consistent judgment aggregation: the truth-functional Case, working paper, Karlsruhe Univ
Parikh, R (1999) Beliefs, belief revision and splitting languages, *Logic, language and computation* 2: 266-278
Pauly, M, van Hees, M (2006) Logical constraints on judgment aggregation, *J Philos Logic* 35: 569-585
Rubinstein, A, Fishburn, P (1986) Algebraic aggregation theory, *J Econ Theory* 38: 63-77
van Hees, M (forthcoming) The limits of epistemic democracy, *Soc Choice Welfare*
Wilson, R (1975) On the Theory of Aggregation, *J Econ Theory* 10: 89-99