

Factoring Out the Impossibility of Logical Aggregation*

Philippe Mongin[†]

October 2007. First version: June 2005

Abstract

According to a theorem recently proved in the theory of logical aggregation, any nonconstant social judgment function that satisfies independence of irrelevant alternatives (IIA) is dictatorial. We show that the strong and not very plausible IIA condition can be replaced with a minimal independence assumption plus a Pareto-like condition. This new version of the impossibility theorem likens it to Arrow's and arguably enhances its paradoxical value.

1 Introduction

In political science and legal theory, the so-called *doctrinal paradox* refers to the observation that if a group of voters casts separate ballots on each proposition of a given agenda, and the majority rule is applied to each of these votes separately, the resulting set of propositions may be logically inconsistent. A mathematical theory of *logical judgment aggregation* has recently grown out

*The author thanks for helpful comments B. Hill, F. Dietrich, R. Holzman, C. List, K. Nehring, two anonymous referees, as well as the participants in the LGS4 conference (June 2005), where this paper was first presented, the Cowles Foundation Workshop on Aggregation of Opinions (September 2006), and seminars at HEC, MIT, Brown University, Collegio Carlo Alberto, GREQAM, and CORE (2005-07). The author also thanks ICER for their hospitality when he was preparing this final version.

[†]HEC School of Management & Centre National de la Recherche Scientifique, 1 rue de la Libération, F-78350 Jouy-en-Josas. E-mail: mongin@hec.fr

of this straightforward point. Its method is to introduce a mapping from profiles of individual judgments to social judgments, where judgments are formalized as sets of formulas in some logical language, and then investigate the effect of imposing axiomatic conditions on this mapping. Among the results obtained is a striking impossibility theorem that abstractly generalizes the doctrinal paradox (Pauly and van Hees, 2006; see also Dietrich, 2006). This theorem states that a mapping defined on a universal domain is dictatorial - i.e., collapses the social judgment set into the set of a given individual whatever the profile - if and only if it is nonconstant and satisfies *independence of irrelevant alternatives* (IIA). In the logical context, the latter condition says that if a formula enters the judgment sets of exactly the same individuals in two profiles, it is a member of either both or none of the two social sets obtained from the mapping. More informally, one can decide whether a formula belongs to a social set just by considering which individual sets it belongs to, regardless of the other formulas that these sets may contain. An earlier variant relied on a stronger condition that is intuitively related to the neutrality axiom of social choice theory (List and Pettit, 2002).

Given the Arrowian undertones of the impossibility theorem, there is something puzzling about it. Arrow's social welfare function maps any conceivable profile of individual weak orderings to a social weak ordering, and his IIA says in effect that one can decide what the social preference is between two alternatives just by considering the individual preferences over these alternatives. The two aggregative schemes are closely related, and so are the two independence conditions. However, Arrow's conclusion that the social welfare function is dictatorial - in the sense of reproducing one individual's strict preference - depends not only on IIA, but on a unanimity condition. The standard proof based on Arrow (1963) uses the weak Pareto condition. Wilson's (1972) extension dispenses with it but needs a premiss to exclude antidictatorship, i.e., the rule which amounts to reversing one individual's strict preference. The impossibility of logical judgment aggregation seems to obtain without any assumption of the kind. This should come as a surprise to anybody acquainted with the social-choice theoretic tradition, not the least to theoretical economists.

The present paper offers a new theorem that will make the impossibility conclusion less mysterious. Still granting universal domain, it derives dictatorship from an IIA condition that is restricted to the atomic components of the language, hence much weaker than the existing one, plus an unrestricted unanimity condition in the style of the Pareto principle, without

which dictatorship would not follow. Beside likening the axioms of logical aggregation to those of social choice theory, and thus making them easier to understand, we reinstate a natural distinction, since unanimity-preservation and independence are conceptually different properties for an aggregator to satisfy. Under technical assumptions to be spelled out below, both the earlier result and ours deliver equivalent restatements of dictatorship, so what these assumptions allow us to do is to *factor out* the components of the strong IIA condition of the initial theory.

There are two more reasons why the present analysis may be instructive. Firstly, to those who wonder how a unanimity-preservation property can be hidden under the guise of an independence property, the mechanics of the theorem offers something of an explanation. Roughly speaking, the logical framework is responsible for the unusual blending; elsewhere, we have compared the impossibility of logical aggregation with related probabilistic possibilities and impossibilities, and reached the same critical conclusion (Mongin, 2005). Secondly, at the normative level, the strong IIA condition is hardly acceptable, whereas, as will be explained, restricting it to the atomic components of the language deflates a telling objection against it. To go beyond the stage of its initial paradox, the theory of logical judgment aggregation must uncover normatively defensible premisses. If, as we argue, our weakening of independence is one of them, the responsibility for the dictatorial conclusion must rest with the unanimity-preservation condition. The end of the paper elaborates on this interpretation, which likens the impossibility of logical aggregation to other “spurious unanimity” results.

The theory defines the *agenda* to be the set of logical formulas representing the propositions, or issues, on which both the individuals and society take sides. Our theorem has been obtained at the price of slightly expanding the agenda beyond what is required in order to get dictatorship from the strong IIA condition. We will relate this special restriction to the recently developed criteria for recognizing a dictatorial agenda.

2 The logical framework of aggregation

In the logical framework, a judgment consists in either accepting or rejecting a formula stated in some logical language. Like most previous writers, we will be concerned with the language of propositional logic. Accordingly, our set of formulas \mathcal{L} is constructed from the propositional connectives \neg , \vee , \wedge , \rightarrow ,

\leftrightarrow (“not”, “or”, “and”, “implies”, “is equivalent to”) and a set \mathcal{P} of distinct propositional variables (p.v.) p_1, \dots, p_m, \dots serving as the atomic components of the construction. We do not make the finiteness restriction that is often encountered in the literature. We say that a p.v. p *appears* in a formula φ to mean that p is a subformula of φ (e.g., p appears in $\neg p \wedge q$). Roman letters will be reserved for p.v., while greek letters will denote formulas of any kind. *Literals* \tilde{p} are those formulas which are either p.v. ($\tilde{p} = p$) or negations of p.v. ($\tilde{p} = \neg p$). The notation $\neg\tilde{p}$ means $\neg p$ in the former case and p in the latter; it accords with the general convention adopted here that double negations cancel. When we wish to emphasize that the literal value for p is fixed, we write \bar{p} instead of \tilde{p} . Two literals \tilde{p} and \tilde{q} are *essentially distinct* if $p \neq q$.

We define the *complexity* $c(\varphi)$ of a formula $\varphi \in \mathcal{L}$ to be the non-negative integer that is inductively defined as follows: if $\varphi = p$, $c(\varphi) = 0$; if $\varphi = \neg\psi$, $c(\varphi) = c(\psi)$; if $\varphi = \varphi_1 \vee \varphi_2$, $\varphi = \varphi_1 \wedge \varphi_2$, or $\varphi_1 \longleftrightarrow \varphi_2$, $c(\varphi) = \max\{c(\varphi_1), c(\varphi_2)\} + 1$. (This definition of complexity differs from the standard one, which states that if $\varphi = \neg\psi$, $c(\varphi) = c(\psi) + 1$.)

The axiomatic system of propositional logic defines an inference relation $S \vdash \varphi$ holding between sets of formulas S and formulas φ . The properties of this relation will be taken for granted. The further notions of a theorem (to be denoted by $\vdash \varphi$), a contradiction, a consistent (an inconsistent, an independent) set of formulas, are defined from \vdash in the usual way. We write “ $S \not\vdash \varphi$ ” instead of “not $S \vdash \varphi$ ”. Recall that if literals $\tilde{p}, \tilde{q}, \tilde{r} \dots$ are essentially distinct, no logical connection holds between them; in particular, $S = \{\tilde{p}, \tilde{q}, \tilde{r} \dots\}$ is consistent and independent. A formula that is not equivalent to a literal, i.e., φ such that for all \tilde{p} , $\varphi \longleftrightarrow \tilde{p}$, will be called *genuinely molecular*.

Beyond propositional logic itself, the present framework is relevant to those more expressive languages in which the latter can be embedded isomorphically. This is a wide class, which includes, e.g., the modal propositional logics with epistemic interpretations which have become known to game theorists (see Bacharach et alii, 1997).

The *agenda* is the nonempty subset $\Phi \subseteq \mathcal{L}$ of formulas representing the actual propositions on which the individuals and society pass a positive or negative judgment. Accordingly, the overall set of formulas to be considered is $\Phi^* = \Phi \cup \{\neg\varphi : \varphi \in \Phi\}$. It is convenient (and innocuous) to assume that double negations cancel in Φ^* and that all formulas in Φ are positive, i.e., do not have \neg as their top-level connective (e.g., p , $\neg p \vee q$ are positive, while $\neg p$, $\neg(p \vee q)$ are not). Define $\Phi_0 = \Phi \cap \mathcal{P}$ and $\Phi_0^* = \Phi_0 \cup \{\neg p : p \in \Phi_0\}$; with

the conventions just made, Φ_0^* is exactly the set of literals of Φ^* . Given the purported application, it is reasonable to assume that $\varphi \in \Phi$ is neither a theorem, nor a contradiction.

For any two $\tilde{p}, \tilde{q} \in \Phi_0^*$, let us say that \tilde{p} is directly related to \tilde{q} , to be denoted by $\tilde{p}D\tilde{q}$, if there is $X \subset \Phi^*$ such that (i) $X \cup \{\tilde{p}\}$ and $X \cup \{\neg\tilde{q}\}$ are consistent, and (ii) $X \cup \{\tilde{p}\} \vdash \tilde{q}$; we call X a supporting set for $\tilde{p}D\tilde{q}$. In words, one literal is directly related to another if the second can be inferred from the first with the help of auxiliary premisses that must be consistent with the first and not by themselves sufficient to entail the second. It immediately follows that D does not hold between \tilde{p} and $\neg\tilde{p}$, that a supporting set X must contain at least one genuinely molecular formula, and that D satisfies contraposition, i.e., $\tilde{p}D\tilde{q}$ iff $\neg\tilde{q}D\neg\tilde{p}$.

Let us now say that \tilde{p} is related to \tilde{q} , to be denoted by $\tilde{p}R\tilde{q}$, if there are $k \geq 2$ and $\tilde{p}_1, \dots, \tilde{p}_k \in \Phi_0^*$ such that $\tilde{p}_1 = \tilde{p}$, $\tilde{p}_k = \tilde{q}$, and $\tilde{p}_j D \tilde{p}_{j+1}$, $1 \leq j \leq k - 1$. In words, one literal is related to another if the two can be linked by a finite sequence of literals each of which is directly related to the next. When R (D) goes both ways, we simply say that the literals are related (directly related).

For example, if $\Phi = \{p, q, \neg p \vee q, p \vee \neg q, \neg p \vee \neg q\}$, p and q , as well as $\neg p$ and $\neg q$, are directly related, while p (q) is directly related to $\neg q$ (resp. $\neg p$), and p (q) is related to $\neg p$ (resp. $\neg q$). The missing relations would hold if $p \vee q$ were an element of Φ . We now introduce specific conditions on agendas.

AGENDA CONDITIONS ON Φ :

(i) *Closure Under Propositional Variables*: if $\varphi \in \Phi$, and $p \in \mathcal{P}$ appears in φ , then $p \in \Phi_0$.

(ii) *Relatedness*: For all $\tilde{p}, \tilde{q} \in \Phi_0^*$, $\tilde{p}R\tilde{q}$.

(iii) *Specific Asymmetry*: For some $\tilde{p}, \tilde{q} \in \Phi_0^*$ and $X \subset \Phi^*$, X is a supporting set for $\tilde{p}D\tilde{q}$ without being a supporting set for $\tilde{q}D\tilde{p}$.

(iv) *Privileged Triple*: For some $\tilde{p}, \tilde{q}, \tilde{r} \in \Phi_0^*$, there is a set $Y \subset \Phi^*$ that is consistent with $\{\tilde{p}, \neg\tilde{q}, \neg\tilde{r}\}$, $\{\neg\tilde{p}, \tilde{q}, \neg\tilde{r}\}$ and $\{\neg\tilde{p}, \neg\tilde{q}, \tilde{r}\}$, and inconsistent with $\{\neg\tilde{p}, \neg\tilde{q}, \neg\tilde{r}\}$.

The last two conditions impose lower bounds 2 and 3, respectively, on $|\Phi_0|$, since they hold only if the literals mentioned in them are essentially distinct. Assuming distinctness, (iii) holds of $\tilde{p}, \tilde{q} \in \Phi_0^*$ if $\neg\tilde{p} \vee \tilde{q} \in \Phi^*$, and (iv) holds of $\tilde{p}, \tilde{q}, \tilde{r} \in \Phi_0^*$ if $\tilde{p} \vee \tilde{q} \vee \tilde{r} \in \Phi^*$. There are many other ways in which these conditions can be met. For instance, (iii) holds of the same pair if $(\neg\tilde{p} \vee \tilde{q}) \wedge \tilde{r}, \tilde{r} \in \Phi^*$, and (iv) holds of the same triple if $\tilde{p} \vee \tilde{q} \vee \tilde{r} \vee \tilde{s} \vee \tilde{z}, \neg\tilde{s}, \dots, \neg\tilde{z} \in \Phi^*$. It is easy to check that the conditions are logically

independent. For instance, to show that (iv) does not imply (iii), take the agenda $\Phi = \{p, q, r, p \longleftrightarrow (q \longleftrightarrow r)\}$, which satisfies (iv), but not (iii).

Given its emphasis on impossibility results, the literature on logical aggregation has justifiably focused on small-sized agendas. We follow this trend by illustrating the lower bound of our conditions taken together, i.e., $|\Phi_0| = 3$.

$$\mathcal{A}_1 = \{p, q, r, p \vee q \vee r, \neg p \vee q, \neg p \vee r, \neg r \vee q, \neg p \vee \neg q\}.$$

$$\mathcal{A}_2 = \{p, q, r, p \vee q \vee r, p \vee \neg q \vee \neg r, \neg p \vee q \vee \neg r, \neg p \vee \neg q \vee r\}.$$

These two agendas turn out to be *minimal* for (i)-(iv), i.e., they do not include any other agenda satisfying these conditions (we skip the tedious proof). Here is another minimal agenda:

$$\mathcal{A}_3 = \{p, q, r, p \vee q \vee r, p \wedge q \wedge r\}.$$

In the general case $|\Phi_0| \geq 3$, we single out for future discussion the following two definitely non-minimal agendas:

$$\mathcal{A}_k^+ = \{p_1, \dots, p_m, \dots\} \cup \{\text{all 2-disjunctions } \widetilde{p}_{j_1} \vee \widetilde{p}_{j_2}\} \cup \{\overline{p}_{j_1} \vee \dots \vee \overline{p}_{j_k}\},$$

and

$$\mathcal{A}_K^+ = \{p_1, \dots, p_m, \dots\} \cup \{\text{all } K\text{-disjunctions } \widetilde{p}_{j_1} \vee \dots \vee \widetilde{p}_{j_K}\},$$

where $k, K \in [3, |\Phi_0|]$ are fixed numbers.

The theory of logical aggregation investigates aggregative rules for *judgment sets*, where logical restrictions are imposed on what counts as such a set. Technically, it is any nonempty subset $B \subseteq \Phi^*$ that is *consistent*, as well as *maximal* in the following relative sense: for any $\varphi \in \Phi^*$, either φ or $\neg\varphi$ belongs to B . This maximal consistency property implies the weaker one that B is *deductively closed* in the same relative sense, i.e., for all $\varphi \in \Phi^*$ and $C \subset B$, if $C \vdash \varphi$, then $\varphi \in B$. It follows in particular that $\varphi \in B \Leftrightarrow \psi \in B$ when $\vdash \varphi \leftrightarrow \psi$, and that $\varphi \in B$ when $\vdash \varphi$. Deductive closure and its consequences raise the “logical omniscience problem” that is widely discussed in epistemic logic (see, e.g., Bacharach et alii, 1997). Strong as these requirements are, they are in a sense milder than maximality, which compels the individuals and society to take a stand on any formula of Φ^* , however large this set may be.

A standard tool of propositional logic, *Lindenbaum’s extension lemma* states that any consistent set of formulas $S \subset \mathcal{L}$ can be extended to a maximal

consistent set S^+ of formulas of \mathcal{L} (see, e.g., Chellas, 1980). It is routine to generalize the lemma to maximality as it is defined here. This variant will recur in the proofs to deliver judgment sets of a predetermined form.

Each of the n individuals will be endowed with a judgment set, and so will society. Formally, given some set D of judgment sets, a *social judgment function* (SJF) is a mapping

$$F : D^n \rightarrow D, \quad (A_1, \dots, A_n) \mapsto A.$$

We put $N = \{1, \dots, n\}$ and often write A, A', \dots instead of $F(A_1, \dots, A_n), F(A'_1, \dots, A'_n), \dots$. A SJF is *dictatorial* if there is $j \in N$ - the *dictator* - such that for all $(A_1, \dots, A_n) \in D^n$,

$$A_j = F(A_1, \dots, A_n).$$

Equivalently, a dictatorial SJF is a projection of the product D^n on one of its components. (Incidentally, Arrow's dictatorship does not have the projection property, since it retains only the strict part of the dictator's ordering.)

Some of the proofs below will involve considering local dictators and antidictators. Let us say that j is a *dictator on* $\tilde{p} \in \Phi_0^*$ if

$$\forall (A_1, \dots, A_n) \in D^n, \quad \tilde{p} \in F(A_1, \dots, A_n) \Leftrightarrow \tilde{p} \in A_j.$$

and an *antidictator on* $\tilde{p} \in \Phi_0^*$ if

$$\forall (A_1, \dots, A_n) \in D^n, \quad \tilde{p} \in F(A_1, \dots, A_n) \Leftrightarrow \neg \tilde{p} \in A_j.$$

If Agenda Condition (i) holds, the clause that j is a dictator on all $\tilde{p} \in \Phi_0^*$ (more briefly: on Φ_0^*) fully defines a SJF, for which j is a dictator in the sense first stated (this follows from the maximal consistency of judgment sets). By contrast, if Agenda Condition (i) holds, the clause that j is an antidictator on Φ_0^* defines a F , but it cannot be said without further assumption on the agenda that j is an overall antidictator, since negating all $\varphi \in A_j$ may lead to contradictions.

3 Axiomatic conditions on social judgment functions

We now introduce axiomatic conditions on F . It will be a maintained assumption that F satisfies *Universal Domain*, i.e., that D is the set of all logically possible judgment sets.

Axiom 1 (*Systematicity*)

$$\forall \varphi, \psi \in \Phi^*, \forall (A_1, \dots, A_n), (A'_1, \dots, A'_n) \in D^n,$$

$$[\varphi \in A_i \Leftrightarrow \psi \in A'_i, i \in N] \Rightarrow [\varphi \in A \Leftrightarrow \psi \in A'].$$

Axiom 2 (*Independence of Irrelevant Alternatives, IIA*)

$$\forall \varphi \in \Phi^*, \forall (A_1, \dots, A_n), (A'_1, \dots, A'_n) \in D^n,$$

$$[\varphi \in A_i \Leftrightarrow \varphi \in A'_i, i \in N] \Rightarrow [\varphi \in A \Leftrightarrow \varphi \in A'].$$

Axiom 3 (*Independence of Irrelevant Propositional Alternatives, IIPA*)

$$\forall \tilde{p} \in \Phi_0^*, \forall (A_1, \dots, A_n), (A'_1, \dots, A'_n) \in D^n,$$

$$[\tilde{p} \in A_i \Leftrightarrow \tilde{p} \in A'_i, i \in N] \Rightarrow [\tilde{p} \in A \Leftrightarrow \tilde{p} \in A'].$$

(In view of the maximal consistency of judgment sets, we might have written “ $\varphi \in \Phi$ ” and “ $p \in \Phi_0$ ” in the statements of IIA and IIPA.)

The three conditions are listed from the logically strongest to the weakest. Before we discuss them normatively, we clarify the connection between IIPA and IIA. Under what condition is the former *strictly* weaker than the latter? An absolutely minimal condition would be that Φ (or equivalently, Φ^*) contains a genuinely molecular formula. With modest auxiliary assumptions, we show that this condition is indeed sufficient.

Proposition 1 *Assume that $n \geq 2$ and Agenda Condition (i) holds. If there is a genuinely molecular formula $\varphi \in \Phi^*$, IIPA does not imply IIA.*

Proof. The proof goes by induction on the complexity of $\varphi \in \Phi^*$.

Suppose that Φ^* contains φ with $c(\varphi) = 1$. (From the assumption that φ is genuinely molecular, this is the first complexity value to consider.) If $\varphi \in \Phi$, either (a) $\varphi = \tilde{p}_1 \vee \tilde{p}_2$, or (b) $\varphi = \tilde{p}_1 \wedge \tilde{p}_2$, or (c) $\varphi = \tilde{p}_1 \longleftrightarrow \tilde{p}_2$. In case (a), take (A_1, \dots) and (A'_1, \dots) such that

$$\tilde{p}_1, \neg \tilde{p}_2 \in A_1, \neg \tilde{p}_1, \tilde{p}_2 \in A'_1, \text{ and } A_i = A'_i, i \in N \setminus \{1\}.$$

With this choice of profiles, the properties of judgment sets imply that $\varphi \in A_i \Leftrightarrow \varphi \in A'_i, i \in N$. Now, suppose that 1 is a dictator on \tilde{p}_1 and an

antidictator on \tilde{p}_2 ; this defines a SJF because of Condition (i), and IIPA holds. It follows that $\tilde{p}_1, \tilde{p}_2 \in A$, and by deductive closure, $\varphi \in A$, while $\neg\tilde{p}_1, \neg\tilde{p}_2 \in A'$, and by consistency, $\varphi \notin A'$ - a contradiction with IIA. The same choice of profiles and SJF takes care of (b). If (c), take (A_1, A_2, \dots) and (A'_1, A'_2, \dots) such that

$$\begin{aligned}\tilde{p}_1, \tilde{p}_2 &\in A_1, \neg\tilde{p}_1, \tilde{p}_2 \in A_2, \\ \neg\tilde{p}_1, \neg\tilde{p}_2 &\in A'_1, \neg\tilde{p}_1, \tilde{p}_2 \in A'_2, \\ A_i &= A'_i, i \in N \setminus \{1, 2\},\end{aligned}$$

and suppose that 1 is an antidictator on p_1 and 2 a dictator on p_2 . This SJF satisfies IIPA, but not IIA.

Now, $\varphi = \neg(\tilde{p}_1 \vee \tilde{p}_2)$ is equivalent to $\neg\tilde{p}_1 \wedge \neg\tilde{p}_2$, which is of type (b), $\varphi = \neg(\tilde{p}_1 \wedge \tilde{p}_2)$ to $\neg\tilde{p}_1 \vee \neg\tilde{p}_2$, which is of type (a), and $\varphi = \neg(\tilde{p}_1 \longleftrightarrow \tilde{p}_2)$ to $\neg\tilde{p}_1 \longleftrightarrow \tilde{p}_2$, which is of type (c). Since judgment sets respect logical equivalence, we have in effect shown that $\text{IIPA} \not\Rightarrow \text{IIA}$ for all Φ^* , not just for all Φ , containing a genuinely molecular φ of complexity 1.

Suppose that the conclusion was reached for all Φ^* containing a genuinely molecular formula of complexity at most c , and take $\varphi \in \Phi^*$ with $c(\varphi) = c+1$. If $\varphi \in \Phi$, there are again three cases, i.e., (a') $\varphi = \varphi_1 \vee \varphi_2$, (b') $\varphi = \varphi_1 \wedge \varphi_2$, and (c') $\varphi = \varphi_1 \longleftrightarrow \varphi_2$, with $c(\varphi_1), c(\varphi_2) \leq c$. In case (a'), suppose that either $\vdash \varphi_1 \rightarrow \varphi_2$ or $\vdash \varphi_2 \rightarrow \varphi_1$; then, either $\vdash \varphi \longleftrightarrow \varphi_2$ or $\vdash \varphi \longleftrightarrow \varphi_1$, and the conclusion that $\text{IIPA} \not\Rightarrow \text{IIA}$ follows from the inductive hypothesis and the fact that judgment sets respect logical equivalence. A similar argument takes care of (b') when either $\vdash \varphi_1 \rightarrow \varphi_2$ or $\vdash \varphi_2 \rightarrow \varphi_1$ holds. Hence, in the first two cases, we may assume that $\{\varphi_1, \varphi_2\}$ is logically independent. This assumption permits repeating in terms of φ_1 and φ_2 the construction made for (a) $\varphi = \tilde{p}_1 \vee \tilde{p}_2$ and (b) $\varphi = \tilde{p}_1 \wedge \tilde{p}_2$.

In case (c'), it is impossible that both $\vdash \varphi_1 \rightarrow \varphi_2$ and $\vdash \varphi_2 \rightarrow \varphi_1$ hold, since this would mean that φ is a theorem. Now, if either $\not\vdash \varphi_2 \rightarrow \varphi_1$ or $\not\vdash \varphi_1 \rightarrow \varphi_2$, we can repeat in terms of φ_1 and φ_2 the construction made for (c) $\varphi = \tilde{p}_1 \longleftrightarrow \tilde{p}_2$ (possibly by interchanging the indexes 1 and 2).

The remaining formulas in Φ^* are dealt with as in the first inductive step. Logical equivalence reduces them to positive formulas such as $\neg\varphi_1 \wedge \neg\varphi_2$, $\neg\varphi_1 \vee \neg\varphi_2$, and $\neg\varphi_1 \longleftrightarrow \varphi_2$, which leads us back to (b'), (a') and (c'), since $c(\neg\varphi_1) = c(\varphi_1)$ and $c(\neg\varphi_2) = c(\varphi_2)$. ■

Now, to the normative assessment of the axioms. *Systematicity* requires that two formulas be treated alike by society if they draw the support of

exactly the same people, even if these formulas refer to semantically unrelated items. Although this was the condition assumed in the first place, it is quite obviously unattractive. Take a two-individual society in which 1 judges that the budget should be balanced, 2 disagrees, and the social judgment endorses 1. Then, if 1 also judges that marijuana should be legalized, and 2 disagrees again, the social judgment should endorse 1 again. Economists will recognize that this is a neutrality condition in the style of those of social choice theory and that it is no more appealing here than it is there. Samuelson (1977) once described neutrality as transparently close to dictatorship and gratuitous; we refer the reader to the witty example he devised to reject this condition.

Instead of permitting variations in both the profile (A_1, \dots, A_n) and formula φ , *Independence of Irrelevant Alternatives* fixes the formula and allows only the profile to vary; in this way, it avoids the confounding of semantic contents that spoils the earlier condition. It singles out the requirement contained in Systematicity that the social judgment on φ should depend only on the individual judgments on φ . Exactly as for Arrow's condition, the best normative defence for this restriction is that it prevents some possible manipulations (see Dietrich, 2006, and Dietrich and List, forthcoming). However, the condition remains open to a charge of irrationality. One would expect society to pay attention not only to the individuals' judgments on φ , but also to their *reasons* for accepting or rejecting this formula, and these reasons may be represented by other formulas than φ in the individual sets. Before deciding that two profiles call for the same acceptance or rejection, society should in general take into account more information than is supposed in the condition.

The new condition of *Independence of Irrelevant Propositional Alternatives* (IIPA) amounts to reserving IIA to p.v. In the doctrinal paradox and related examples, trouble arises from the assumption that the majority rule applies to molecular formulas and p.v. alike - i.e., that this independent and even neutral rule dictates on the whole of Φ . In contrast, when restricted to Φ_0 , independence becomes more acceptable *because the formulas now represent primary data*. One can object to IIA being applied to $p \vee q$, where p represents the proposition that (1) the budget should be balanced and q represents the proposition that (2) Marijuana should be legalized, on the ground that there are two propositions involved, and that society should know how each individual feels about either of them, and not simply about their disjunction. No similar objection arises when IIA is applied to either p or q in isolation because the reasons for accepting or rejecting the intended

propositions are beyond the expressive possibilities of the language.

Of course, as pointed out earlier, the language of propositional logic can be embedded in more powerful ones that will refine its analysis of propositions. The formulas replacing the initial p.v. in these richer frameworks will be built from new elementary signs, in terms of which it will often become possible to express some of the reasons for accepting or rejecting the intended propositions. To illustrate, suppose that we are also interested in the proposition that (3) it should be the case that either the budget be balanced or marijuana be legalized. Propositional logic will represent it by a p.v. r , on a par with p and q , without being able to convey the fact that the reasons for (3) connect with the reasons for (1) and (2). By contrast, deontic logic (see Chellas, 1980) will represent (3) by $O(a \vee b)$, where O is the obligation operator, and if its axiom system includes:

$$\vdash O(a \vee b) \leftrightarrow Oa \vee Ob,$$

the reasons that (1) and (2) provide for accepting (3) will be explicit *within the language*. This example seems to suggest that we were wrong to reject IIA while accepting IIPA. The argument against the former works against the latter when we take the perspective of a more powerful logic.

The quick answer to this objection is that the more powerful logic in question will also have its independent building blocks, and the argument made in defence of applying independence to the p.v. will carry through to these new items. For instance, a and b , which represent “the budget *is* balanced” and “marijuana *is* legalized” respectively, are building blocks of the language of deontic logic; it is to them that the initial argument should now be applied. Since there is no such thing as a most powerful logic, there will always be room for the argument to apply.

The last conditions to be introduced are directly inspired by the Pareto principle. As for independence, we state two versions graded by logical strength:

Axiom 4 (*Unanimity Preservation, UP*) For all $\varphi \in \Phi$, and all $(A_1, \dots, A_n) \in D^n$,

$$\varphi \in A_i, i \in N \Rightarrow \varphi \in A.$$

Axiom 5 (*Unanimity Preservation for Propositional Alternatives, UPPA*) For all $p \in \Phi_0$, and all $(A_1, \dots, A_n) \in D^n$,

$$p \in A_i, i \in N \Rightarrow p \in A.$$

We defer the normative assessment of UP and UPPA until the end of the paper.

4 The impossibility of logical judgment aggregation

By itself, IIPA is compatible with a large number of non-dictatorial rules some of which satisfy UPPA, while others do not. To illustrate the latter class, consider j to be antidictatorial on Φ_0^* under Agenda Condition (i); thus, we have a well defined F that satisfies IIPA, though not UPPA. To illustrate the former class, consider the following variant of majority voting:

(R) *Apply majority voting to the $p \in \Phi_0$, and take the deductive closure of the resulting set of literals. If n is odd, let the n individuals vote; if n is even, let only the first $(n-1)$ ones vote.*

We require $n \geq 3$ to make this rule non-dictatorial. Under (i), it defines a F that satisfies IIPA and UPPA by construction.

For suitably defined agendas, UPPA cannot be strengthened into UP without precipitating dictatorship. This is the main result of the paper.

Theorem 2 *Assume that $n \geq 2$ and Agenda Conditions (i), (ii), (iii) and (iv) hold. If F satisfies Independence of Irrelevant Propositional Alternatives and Unanimity Preservation, F is dictatorial.*

We start by illustrating one aspect of the Agenda Conditions, which is the cardinality restriction $|\Phi_0| \geq 3$ implied by (iv). If $|\Phi_0| = 1$ or $|\Phi_0| = 2$, IIPA and UP together do *not* imply dictatorship. The counterexample is **(R)**, which happens to satisfy UP in these two cases. This is a trivial point when $|\Phi_0| = 1$, but it must be proved when $|\Phi_0| = 2$.

■ Assume that (i) holds, $|\Phi_0| = 2$, and $n \geq 3$. We check that **(R)** satisfies UP in the different cases for $\varphi \in \Phi$. Suppose that $\varphi = \tilde{p} \vee \tilde{q} \in A_i$, $i \in N$. Given the maximal consistency of judgment sets, there must be a majority for \tilde{p} or a majority for \tilde{q} , and the conclusion that $\varphi \in A$ follows from applying **(R)**. Formulas $\varphi = \tilde{p} \wedge \tilde{q}$ and $\varphi = \tilde{p} \longleftrightarrow \tilde{q}$ are handled similarly, and there is no other case to consider up to logical equivalence. ■

Interestingly, this proof remains unchanged if we make the initial assumptions that $|\Phi_0| \geq 3$ and that Φ is restricted to 2-disjunctions up to logical equivalence. This shows that the size of formulas, and not only the number

of p.v., matters to the truth of the theorem. A major change occurs when Agenda Condition (iv) is applied on top of (i). To see that, take three distinct $p, q, r \in \Phi_0$, and assume for simplicity that $p \vee q \vee r \in \Phi$. It is not always the case anymore that unanimous endorsement of a disjunction is preserved by **(R)**. Specifically, for odd n , take the profile in which $p \vee q \vee r \in A_i, i \in N$, and

$$p, \neg q, \neg r \in A_i, i \in I; \neg p, q, \neg r \in A_j, j \in J; \neg p, \neg q, r \in A_k, k \in K,$$

with $|I| = |K| = (n - 1)/2$ and $|J| = 1$. Since $\neg p, \neg q$, and $\neg r$ obtain a majority against p, q, r respectively, **(R)** and the consistency of A imply that $p \vee q \vee r \notin A$, a contradiction with UP. A profile like this one illustrates how Condition (iv) blocks the implication from UPPA to UP and makes room for the dictatorial conclusion of Theorem 1.

There remain some limiting cases to cover. If $n = 1$, the theorem trivially holds, and if $n = 2$ and $|\Phi_0| = 1$, it does not hold, as the following SJF shows:

$$\begin{aligned} F(\{p\}, \{p\}) &= \{p\}, F(\{\neg p\}, \{\neg p\}) = \{\neg p\}, \\ F(\{p\}, \{\neg p\}) &= F(\{\neg p\}, \{p\}) = p. \end{aligned}$$

Perhaps surprisingly, the theorem holds in the case $n = |\Phi_0| = 2$. This will be seen indirectly from the proof of the theorem and directly from an argument given at the end of the section.

The proof is organized into three lemmas and a concluding step.

Lemma 3 *Assume IIPA, UP and Agenda Condition (ii). Then, the following property, to be called Limited Systematicity, holds: for all $\tilde{p}, \tilde{q} \in \Phi_0^*$ and all $(A_1, \dots, A_n), (A'_1, \dots, A'_n) \in D^n$,*

$$[\tilde{p} \in A_i \Leftrightarrow \tilde{q} \in A'_i, i \in N] \Rightarrow [\tilde{p} \in A \Leftrightarrow \tilde{q} \in A']$$

Proof. We will prove the following weaker form of the claim: for all $\tilde{p}, \tilde{q} \in \Phi_0^*$ s.t. $\tilde{p}R\tilde{q}$, and all $(A_1, \dots, A_n), (A'_1, \dots, A'_n) \in D^n$,

$$(*) [\tilde{p} \in A_i \Leftrightarrow \tilde{q} \in A'_i, i \in N] \Rightarrow [\tilde{p} \in A \Rightarrow \tilde{q} \in A'].$$

When this is established, the claim will result from applying (ii).

By definition of $\tilde{p}R\tilde{q}$, there exist a sequence of literals $\tilde{p}_1, \dots, \tilde{p}_k$, with $\tilde{p} = \tilde{p}_1$ and $\tilde{q} = \tilde{p}_k$, and a sequence of sets X_1, \dots, X_{k-1} , s.t. $X_j \cup \{\tilde{p}_j\} \vdash \widetilde{p_{j+1}}$, $1 \leq j \leq k - 1$, and the X_j are consistent with \tilde{p}_j and $\neg \widetilde{p_{j+1}}$.

We derive (*) by induction on the length k of the sequence of literals. Suppose that $k = 1$. Given $(A_1, \dots, A_n), (A'_1, \dots, A'_n) \in D^n$ that satisfy the antecedent of (*), there exist $(B_1, \dots, B_n) \in D^n$ with the following properties:

$$\tilde{p} \in A_i \Leftrightarrow \tilde{p} \in B_i, \tilde{q} \in A'_i \Leftrightarrow \tilde{q} \in B_i, i \in N,$$

and

$$X_1 \subset B_i, i \in N.$$

The definitional properties of X_1 ensure that $X_1 \cup \{\tilde{p}, \tilde{q}\}$ and $X_1 \cup \{\neg\tilde{p}, \neg\tilde{q}\}$ are consistent, so that the extension lemma delivers B_i of the prescribed sort. Now, suppose that $\tilde{p} \in A$; then $\tilde{p} \in B$ from IIPA, and since UP entails that $X \subset B$, deductive closure entails that $\tilde{q} \in B$, hence $\tilde{q} \in A'$ from IIPA again.

Assume that the property has been proved for all $\tilde{p}, \tilde{q} \in \Phi_0^*$ s.t. $\tilde{p}R\tilde{q}$ with connecting sequences of length at most $k - 1$. Take $\tilde{p}R\tilde{q}$ with a connecting sequence of length k . Given $(A_1, \dots, A_n), (A'_1, \dots, A'_n) \in D^n$ satisfying the antecedent of (*), the extension lemma now applied to $X_{k-1} \cup \{\widetilde{p_{k-1}}, \tilde{q}\}$ and $X_{k-1} \cup \{\neg\widetilde{p_{k-1}}, \neg\tilde{q}\}$ ensures that for some $(B_1, \dots, B_n) \in D^n$,

$$\tilde{p} \in A_i \Leftrightarrow \widetilde{p_{k-1}} \in B_i, \tilde{q} \in A'_i \Leftrightarrow \tilde{q} \in B_i, i \in N,$$

and

$$X_{k-1} \subset B_i, i \in N.$$

Suppose that $\tilde{p} \in A$; then $\widetilde{p_{k-1}} \in B$ from the inductive hypothesis, and $\tilde{q} \in A'$ follows by the same steps as in the case $k = 1$. ■

Lemma 4 *Assume UP and Agenda Condition (iii). Then, for the literals \tilde{p} and \tilde{q} of this condition, the following property, to be called Specific Monotonicity, holds: for all $(A_1, \dots, A_n), (A'_1, \dots, A'_n) \in D^n$,*

$$(**) [\tilde{p} \in A_i \Rightarrow \tilde{q} \in A'_i, i \in N] \Rightarrow [\tilde{p} \in A \Rightarrow \tilde{q} \in A'].$$

Proof. Given $(A_1, \dots, A_n), (A'_1, \dots, A'_n) \in D^n$ satisfying the antecedent of (**), we consider $(B_1, \dots, B_n) \in D^n$ such that

$$\tilde{p} \in A_i \Leftrightarrow \tilde{p} \in B_i, \tilde{q} \in A'_i \Leftrightarrow \tilde{q} \in B_i, i \in N,$$

and

$$X \subset B_i, i \in N,$$

where X is as stipulated in (iii). The B_i exist because the sets $X \cup \{\tilde{p}, \tilde{q}\}$, $X \cup \{\neg\tilde{p}, \neg\tilde{q}\}$, $X \cup \{\neg\tilde{p}, \tilde{q}\}$ permit applying the extension lemma, the first two being consistent because X is a supporting set for $\tilde{p}D\tilde{q}$, and the third because X is not a supporting set for $\tilde{q}D\tilde{p}$. Now, if $\tilde{p} \in A$, then $\tilde{p} \in B$ from IIPA, and since UP implies that $X \subset B$, $\tilde{q} \in B$ follows from deductive closure, whence $\tilde{q} \in A'$ from IIPA again. ■

Lemma 5 *Assume Limited Systematicity, UP and Agenda Condition (iv). Then, there are $i^* \in N$, $\tilde{p} \in \Phi_0^*$ and $(B_1, \dots, B_n) \in D^n$ s.t.*

$$(***) \tilde{p} \in B_{i^*}, \neg\tilde{p} \in B_i, i \neq i^*, \text{ and } \tilde{p} \in F(B_1, \dots, B_n).$$

Proof. Take $\tilde{p}, \tilde{q}, \tilde{r}$ as in (iv) and consider any sequence of profiles $(A_j^i)_{i,j=1,\dots,n} \in D^n$ with the following pattern:

$$\begin{array}{ccccccc} \tilde{p} \in A_1^1 & \neg\tilde{p} \in A_2^1 & \dots & & \neg\tilde{p} \in A_n^1 & & \\ \tilde{p} \in A_1^2 & \tilde{p} \in A_2^2 & \neg\tilde{p} \in A_3^2 & \dots & \neg\tilde{p} \in A_n^2 & & \\ \dots & \dots & \dots & & \dots & & \\ \tilde{p} \in A_1^n & \tilde{p} \in A_2^n & \tilde{p} \in A_3^n & \dots & \tilde{p} \in A_n^n & & \end{array}.$$

Denote by $(A^i)_{i=1,\dots,n}$ the associated sequence of social judgment sets. Define i^* to be the index of the first line i with $\tilde{p} \in A^i$, or equivalently the index of the first individual i s.t. both $\tilde{p} \in A_i^i$ and $\tilde{p} \in A^i$ hold. UP ensures that this number exists. We will prove that there is a profile $(B_1, \dots, B_n) \in D^n$ satisfying (***) for the individual i^* so defined.

If $i^* = 1$, it is enough to take the $(A_j^1)_{j=1,\dots,n}$ line. If $i^* \geq 2$, we define three sets of individuals:

$$I = \{1, \dots, i^* - 1\}, J = \{i^*\}, K = \{i^* + 1, \dots, n\},$$

the last of which may be empty, and we assume by way of contradiction that $\neg\tilde{p} \in B$ for all profiles s.t. $\tilde{p} \in B_{i^*}$, $\neg\tilde{p} \in B_i$, $i \neq i^*$. In particular, $\neg\tilde{p} \in B$ if $(B_1, \dots, B_n) \in D^n$ has this property as well as the following ones:

$$\begin{array}{l} \neg\tilde{q} \in B_i, i \in I \cup J; \tilde{q} \in B_i, i \in K, \\ \tilde{r} \in B_i, i \in I; \neg\tilde{r} \in B_i, i \in J \cup K, \\ Y \subset B_i, i \in N, \end{array}$$

where the set Y is as stipulated by (iv). The three consistency properties listed in this condition are exactly those needed for the B_i to be well defined,

given the by now familiar extension argument. Limited Systematicity entails, on the one hand, that $\neg q \in B$ in view of the $(A_j^{i^*})_{j=1,\dots,n}$ line, and on the other hand, that $\neg r \in B$ in view of the $(A_j^{i^*-1})_{j=1,\dots,n}$ line. Then, $\neg p, \neg q, \neg r \in B$, and the conclusion drawn from UP that $Y \subset B$ contradicts the fourth, inconsistency property of Y stated in (iv). ■

Proof. (End) When Agenda Conditions (ii) and (iii) hold together, Limited Systematicity can be applied to strengthen IIPA and Specific Monotonicity into the following *Strong Monotonicity* property: for all $(A_1, \dots, A_n), (A'_1, \dots, A'_n) \in D^n$ and all $\tilde{p} \in \Phi_0^*$,

$$(**') [\tilde{p} \in A_i \Rightarrow \tilde{p} \in A'_i, i \in N] \Rightarrow [\tilde{p} \in A \Rightarrow \tilde{p} \in A'].$$

If Agenda Condition (iv) also holds, we can compare any \tilde{p} and any (A_1, \dots, A_n) with those \bar{p} and (B_1, \dots, B_n) which satisfy (**), and applying Limited Systematicity and Strong Monotonicity, conclude that i^* is a dictator on Φ_0^* . From Agenda Condition (i) and the maximal consistency of JS, i^* is also a dictator. ■

The reader will have noticed the social-theoretic undertones of this proof. Lemma 5 establishes a “semi-decisiveness property” for a particular item that Lemmas 3 and 4 extend to a “decisiveness property” for any item whatever.

A more direct proof works for the case $n = 2$. For brevity, we state it for the agenda \mathcal{A}_k^+ containing all 2-disjunctions, but it is potentially as general as the previous one.

Proof. Suppose that F satisfies IIPA and UP, but is non-dictatorial. It follows that there are two profiles $(A_1, A_2), (A'_1, A'_2) \in D^2$, s.t. $A \neq A_1, A' \neq A'_2$. From Agenda Condition (i) and maximal consistency, two judgment sets differ from each other iff they differ in terms of some p.v. Thus, there are two literals \tilde{p}, \tilde{q} s.t.

$$\neg\tilde{p} \in A_1, \tilde{p} \in A, \text{ and } \neg\tilde{q} \in A'_2, \tilde{q} \in A'.$$

If $p \neq q$, we can find $(B_1, B_2) \in D^2$ obeying the following equivalences:

$$\tilde{p} \in A_i \Leftrightarrow \tilde{p} \in B_i, \tilde{q} \in A'_i \Leftrightarrow \tilde{q} \in B_i, i = 1, 2.$$

Deductive closure entails that $\neg\tilde{p} \vee \neg\tilde{q} \in B_i, i = 1, 2$, so that $\neg\tilde{p} \vee \neg\tilde{q} \in B$ from UP. However, IIPA requires that both $\tilde{p} \in B$ and $\tilde{q} \in B$, which violates consistency. If $p = q$, take $r \neq p$ and modify the equivalences defining the B_i as follows:

$$\tilde{p} \in A_i \Leftrightarrow \tilde{p} \in B_i, \tilde{p} \in A'_i \Leftrightarrow \tilde{r} \in B_i, i = 1, 2.$$

■

This proof uses only two p.v., which shows that Theorem 1 extends to the limiting case $|\Phi_0| = n = 2$ with only Agenda Conditions (i), (ii) and (iii). Indeed, the main proof needed three p.v. and Condition (iv) only in Lemma 5, in order to construct a profile in which some i is “semi-decisive” on some \tilde{p} , but such a profile is trivially available when there are just two individuals.

The decomposition of IIA into UP and IIPA depends on strengthening the Agenda Conditions somewhat, as will be explained in the next section.

5 Comparisons and further comments

Among the existing impossibility theorems, Pauly and van Hees’s (2006) is the easiest to compare with Theorem 1. These authors conclude that F is dictatorial from the two conditions that F satisfies IIA and is non-constant, starting from an agenda Φ with the following features: $|\Phi_0| \geq 2$, Agenda Condition (i) holds, as well as

(ii’) $\tilde{p} \wedge \tilde{q} \in \Phi$ for all literals \tilde{p}, \tilde{q} .

Pauly and van Hees’s choice of conjunctions rather than disjunctions is immaterial, since both become available when one takes the negation closure Φ^* . Thus, their agenda is equivalent to $\mathcal{A}_k^+ \setminus \{\overline{p_{j_1}} \vee \dots \vee \overline{p_{j_k}}\}$, and it satisfies all our conditions except for (iv). If we add this restriction, their theorem and ours will hold at the same time, thus delivering the promised decomposition.

Corollary 6 *Suppose that $\Phi = \mathcal{A}_k^+$, or more generally, that $\mathcal{A}_k^+ \setminus \{\overline{p_{j_1}} \vee \dots \vee \overline{p_{j_k}}\} \subset \Phi$ and Φ satisfies Agenda Condition (iv). Then, F is non-constant and satisfies Independence of Irrelevant Alternatives if and only if it is equivalent to Independence of Irrelevant Propositional Alternatives conjoined with Unanimity Preservation.*

Starting from a finite \mathcal{P} , Dietrich (2006) has proved an enlightening variant of Pauly and van Hees’s theorem that weakens IIA by reserving it to the *atoms* of \mathcal{L} , i.e., to the conjunctions $\tilde{p}_1 \wedge \dots \wedge \tilde{p}_m$ obtained for all possible $\tilde{p}_1, \dots, \tilde{p}_m$, where $|\mathcal{P}| = m \geq 2$. Like other agendas in the literature that do not satisfy (i), this one is incomparable with both Pauly and van Hees’s and ours, except for the limiting case $m = 2$.

Several writers have recently provided necessary and sufficient conditions for an agenda Φ of finite cardinality to be dictatorial, granted that F satisfies some axiomatic conditions. The most general criterion, Dokow and

Holzman's (2005), requires F to satisfy IIA and UP, and makes it possible to recover several theorems at once, including Pauly and van Hees's when \mathcal{P} is finite, but not ours, which proceeds from the non-standard IIPA. However, since Dokow and Holzman's conditions on F imply IIPA and UP, which we have shown to imply dictatorship for agendas Φ^* satisfying (i)-(iv), such agendas qualify as impossibility domains in Dokow and Holzman's sense, and it is necessary that they meet their criterion. Moving from our syntactical to their semantic language, it is easy to check that $\mathcal{A}_1, \mathcal{A}_2, \mathcal{A}_3, \mathcal{A}_k^+, \mathcal{A}_K^+$ do satisfy the two properties of "total blockedness" and "not being an affine space" on Φ^* that the two scholars have shown to be necessary and sufficient for their F to be dictatorial. Beyond this formal linkage, there are heuristic analogies that would deserve to be tightened between the two properties and, respectively, our conditions (ii) and (iii) imposed on Φ_0^* ; as it turns out, (iv) has also an analogue in a third property that they derive. These connections will be clarified in another paper, which will also investigate the extent to which Agenda Conditions (ii), (iii), and (iv), are, in the presence of (i), not only sufficient, but necessary, for Theorem 2 to hold.

Elaborating on their earlier work on strategy-proof social choice, Nehring and Puppe (2006) have also provided necessary and sufficient conditions for an agenda satisfying (i) to make F dictatorial. They require that F be onto and satisfy a Monotonic Independence property that strengthens IIA in the direction of the Strong Monotonicity property introduced within the above proof. There is no straightforward implication from this pair of conditions to IIPA and UP, which means that the argument used to apply Dokow and Holzman's criterion to our Φ^* cannot be reproduced, but there are heuristic links with their "non-conjunctiveness" condition that may be usefully explored.

Despite being much larger than necessary, the agenda \mathcal{A}_K^+ is of some interest, both technical and conceptual. Developing the idea of local dictatorship in a new direction, we define $j \in N$ to be a *dictator in* $(A_1, \dots, A_n) \in D^n$ if $F(A_1, \dots, A_n) = A_j$, and F to be *dictatorial profile by profile* if for all $(A_1, \dots, A_n) \in D^n$, there is a dictator in (A_1, \dots, A_n) .

Proposition 7 *If $\Phi = \mathcal{A}_K^+$, with $K = n$, and F satisfies UP, then F is dictatorial profile by profile.*

Proof. Suppose to the contrary that there exists $(A_1, \dots, A_n) \in D^n$ s.t. $A \neq A_i$, for all $i \in N$. From (i) and maximality, for all i , there is \tilde{p}_i with the property that $\tilde{p}_i \in A_i$ and $\neg \tilde{p}_i \in A$. Deductive closure entails that $\bigvee_{i=1, \dots, n} \tilde{p}_i \in$

A_i for all $i \in N$, whence $\bigvee_{i=1, \dots, n} \tilde{p}_i \in A$ from UP. This implies that A is inconsistent. (The argument parallels that used to simplify the proof of Theorem 2 when $n = 2$.) ■

Using this proposition as a lemma, we can devise a quick proof of Theorem 2 for $\Phi = \mathcal{A}_K^+$ (such a proof was used in the 2005 version of the paper). More importantly, it comes as a surprise that profile by profile dictatorship follows from UP *without any application of IIPA*. A condition that goes such a long way towards dictatorship cannot be very appealing. Furthermore, UP is open to an objection similar to that raised against IIA. Unanimity over a disjunctive formula may be *spurious* in the sense of relying on conflicting endorsements of the reasons - here identified with the disjuncts - for endorsing that very formula. The issue of spurious unanimity has been discussed by Mongin (1997) in a probabilistic context, and recently studied afresh by Nehring (2005) in a more general framework. It is perhaps no surprise if Nehring's thorough investigation of the Pareto principle leads him to results that are evocative of Theorem 2.

If UP turns out to be dubious for the same reasons as IIA was, it seems natural to relax it in the same way, i.e., by restricting it to p.v. When this is also done, attractive - or at least well-regarded - aggregative rules emerge, such as **(R)** and further variants of majority voting.

6 Conclusion

This paper has derived the impossibility of logical judgment aggregation from assumptions that, for one, highlight connections with social choice theory, and for another, disentangle the initial paradox based on the essentially unique assumption of IIA. Independence has been reduced to such a weak requirement as to exonerate it from its responsibility in the unpleasant conclusion. Although unanimity-preservation now comes first on the black list, many economists may be reluctant to give up a condition that is so closely related to the Pareto principle. The remaining options are to blame either (a) the logical framework of judgment, or (b) universal domain, or (c) the demanding assumptions put on judgment sets. Logical judgment theorists have begun to explore weakenings of (b) and (c); however, some of the possibilities they have discovered - e.g., oligarchies - remain unattractive or exaggeratedly specific. Another avenue that recommends itself to the theoretical economist is to depart from (a) and its set of embodied constraints - typically, by moving

to the probabilistic setting.

7 References

- Arrow, K.J. (1963), *Social Choice and Individual Values*, New York, Wiley (1st ed.1951).
- Bacharach, M., L. A. Gérard-Varet, P. Mongin, H. Shin (1997) (eds), *Epistemic Logic and the Theory of Games and Decisions*, Dordrecht, Kluwer.
- Chellas, B. (1980), *Modal Logic*, Cambridge, Cambridge University Press.
- Dietrich, F. (2006), “Judgment Aggregation: (Im)possibility Theorems”, *Journal of Economic Theory*, 126, p. 286-298.
- Dietrich, F. and C. List (forthcoming), “Strategy-Proof Judgment Aggregation”, to appear in *Economics and Philosophy*.
- Dokow, E. and R. Holzman (2005), “Aggregation of Binary Evaluations”, mimeo, Technion-Israel Institute of Technology.
- List, C. and P. Pettit (2002), “Aggregating Sets of Judgments: An Impossibility Result”, *Economics and Philosophy*, 18, p. 89-110.
- Mongin, P. (1997), “Spurious Unanimity and the Pareto Principle”, THEMA Working Papers, Université de Cergy-Pontoise.
- Mongin, P. (2005), “Logical Aggregation, Probabilistic Aggregation, and Social Choice”, mimeo, invited lecture at the LGS4 Conference, Caen, June 2005.
- Nehring, K. (2005), “The (Im)possibility of a Paretian Rational”, mimeo, University of California at Davis.
- Nehring, K. and C. Puppe (2006), “Consistent Judgment Aggregation: The Truth-Functional Case”, mimeo, University of California at Davis and Universität Karlsruhe.
- Pauly, M. and M. van Hees (2006), “Logical Constraints on Judgment Aggregation”, *Journal of Philosophical Logic*, 35, p. 569-585.
- Samuelson, P.A. (1977), “Reaffirming the Existence of “Reasonable” Bergson-Samuelson Social Welfare Functions”, *Economica*, 44, p. 81-88.
- Wilson, R. (1972), “Social Choice Theory Without the Pareto Principle”, *Journal of Economic Theory*, 5, p. 478-486.