# The (Im)Possibility of a Paretian Rational

Klaus Nehring[1]
Department of Economics, University of California, Davis

November 3, 2005

[1]I am grateful for the hospitality of the Institute for Advanced Study, Princeton, and the Department of Economics, Princeton University, where this research was done. I benefited from the audiences of seminars at the Institute for Advanced Study and Heidelberg University, and from discussions with Elchanan Ben-Porath, Alessandra Casella, Juergen Eichberger, Hans Gersbach, Christian List, Eric Maskin, Philip Pettit, and Tomas Sjostrom.

**Abstract**

We consider situations in which a group takes a collective decision by aggregating individual's judgments on a set of criteria according to some agreed-upon decision functions. Assuming the criteria and the decision to be binary, we demonstrate that, except when the aggregation rule is dictatorial or the decision rule is particularly simple, such reason-based social choice must violate the Pareto principle at some profile of individual judgments. In the second part of the paper, the normative implications of this impossibility result are discussed. We argue that the normative case for the Pareto Principle is strong in situations of "shared self-interest", but weak in situations of "shared responsibility".

**Keywords:** judgment aggregation, Pareto principle, discursive dilemma, group choice, responsibility

**JEL Classification:** D70, D71

## 1. INTRODUCTION

A key aspect of the economic conception of the rational decision making is its outcome orientation which entails a central role of the Pareto criterion: if every agent prefers one social outcome over another, that outcome is socially superior and should be brought about. Agreement on outcome choices is good enough whether or not it is based on an agreement on the substantive reasons for these choices; as evidenced by the (neo-)classical maxim "De gustibus non est disputandum", the justification of preferences and decisions by reasons is often viewed as impossible or besides the point.

By contrast, in the legal and political realms, the ability to support collective decisions by reasons appears to matter a lot: debates play a central role in politics as do notions of precedent and doctrine in law. Accordingly, in the normative thinking on politics in recent years, there has been renewed interest in the possibility of legitimizing political decisions by achieving a (partial) consensus on the reasons for these decisions through democratic deliberation (see Habermas (1984,1989), Cohen (1986,1989), Coleman-Ferejohn (1986) and many others). Likewise, the importance of arriving at legal decisions for the right reasons has been articulated forcefully in the recent law and economics literature (Kornhauser-Sager (1986), Chapman (1998)). Even more recently, these ideas have inspired a still smallish but vital literature on "judgment aggregation" that fits broadly into the field of social choice theory; the seminal paper is List-Pettit (2002).

The starting point for the present paper is the observation that reason-based social choice can easily come into conflict with the Pareto criterion. Consider, for example, the following variation on the "Discursive Dilemma"[1] which we shall refer to as the "Dilemma of the Paretian Rational", or simply Paretian Dilemma for short. Suppose that a three-member panel of judges in a tort case has to decide whether a defendant has to pay damages to the plaintiff. Legal doctrine requires that damages are due if and only if the following three premises are established: 1) the defendant had a duty to take care, 2) the defendant behaved negligently, 3) his negligence caused damage to the plaintiff. The pattern of judges' opinions is given in the following Table:

[1] Originally, the Discursive Dilemma is due to Kornhauser and Sager (1986) under the name of "Doctrinal Paradox".

| Judge | Duty | Negligence | Causation | Damages |
|---|---|---|---|---|
| I | Yes | Yes | No | No |
| II | No | Yes | Yes | No |
| III | Yes | No | Yes | No |
| Panel | Yes | Yes | Yes | Yes \ No ? |

Table 1: The Dilemma of the Paretian Rational

The panel has agreed to decide on each premise by majority vote. Given the actual pattern of opinions, it turns out that a majority agrees with each premise; reasoning from these premises according to legal doctrine, the panel comes to the conclusion that, yes, the Plaintiff should be awarded damages, contrary to their *unanimous* individual opinions that damages should be denied. There is an element of paradox here: how can an alternative (here: the awarding of damages) be judged superior by the group if each agent on his own judges it inferior? Note in particular that if the judges' opinions are common knowledge, one cannot justify this reversal on the basis of superior information possessed by the group, since, by assumption, the group choice relies on commonly known information only.

In this paper, we shall explore how robust this conflict between reason-based choice and the Pareto principle really is. At first sight, the conflict seems to depend critically on the aggregation rule employed: for example, if the outcome decision hinges on the validity of a conjunction of $K$ different premises, the dilemma is overcome by requiring supermajorities exceeding $1 - \frac{1}{K}$ for the acceptance of any premise. With 3 premises as above, each premise must thus be supported by strictly more than $\frac{2}{3}$ of the voters.[2] Yet, contrary to what may be suggested by this simple example, in many other cases the Paretian Dilemma cannot be overcome so easily: as soon as the decision function becomes a bit more complicated, all well-behaved (anonymous or non-dictatorial) aggregation rules overturn the unanimous outcome judgment for some profile of judgments: a "Paretian Rational" (advocate of reason-based social choice) is then "impossible".

[2]It is easy to see how this works: if all premises are accepted by the group under such a rule, this means that strictly less than $\frac{1}{3}$ of the voters reject each premise; by consequence, there must exist at least one voter who accepts all premises and therefore supports the outcome decision.

To illustrate the flavor of the results to come, suppose that the outcome decision is positive if and only if at least $L$ out of $K$ criteria are satisfied, and assume that $1 < L < K$. If the outcome depends on three premises ($K = 3$ and $L = 2$), majority voting on premises is consistent. Yet as soon as the outcome depends on four or more premises, *any* non-dictatorial aggregation rule is susceptible to Pareto conflicts! Generalizing this example, the main results of this paper, Theorems 5 and 7, classify all (monotone) decision functions in terms of their propositional structure according to which types of Pareto consistent[3] aggregation rules they admit. In particular, we show that whenever the outcome decision is "indecomposable" in that no single premise is either necessary or sufficient for the outcome decision, and whenever that decision depends on at least five premises, any non-dictatorial aggregation rule gives rise to potential Pareto conflicts.

In line with the emerging literature on judgment aggregation, we conceptualize reason-based choice as inference from independently aggregated premise judgments. This formulation has a variety of applications beyond its original judicial setting. An example closer to some readers' own professional experience may be the refereeing of a paper for a journal. Suppose, for instance, that a journal's editorial policy deems a paper acceptable if and only if it is valid, novel, and of significant interest. The paper is evaluated by aggregating the referees' opinions on each criterion according to some fixed procedure, for example by majority count. If the configuration of the referees' views is isomorphic to that in Table 1, the editor faces a non-trivial decision as to whether to follow through with the premise-based acceptance of the paper, or to endorse the referee's unanimous bottom line rejection.[4]

To move closer to economics proper, the "panel" might be a committee of the Federal Drug Administration that has to decide on whether to grant a patent for a new, genetically engineered drug on the basis of whether the drug is safe, effective and

[3] An aggregation rule is Pareto consistent if there is no profile of judmgents at which a Paretian Discursive Dilemma arises; see Section 2 for a formal definition.

[4] Of course, in reality editors typically play a more active role. On the other hand, the assumption of a mechanical aggregation procedure may be useful as a benchmark assumption, both in that it does not seem unreasonable as a default option for editors in practice – and an important part of the justification of their decisions to the paper's authors –, and to the extent that a more active role of the editor can be attributed, at least in part, to the difficulties of using a mechanical aggregation procedure.

ethical (either in terms of the research procedure on which it is based or in terms of the mechanism of its working). As a variant, this decision may be made by the population at large, not by an expert panel.[5] The point of these examples is first to show that the notion of reason-based choice makes intuitive sense and has practical relevance far beyond the legal realm, including in contexts within the subject matter of economics narrowly construed such as the regulation of industries. Second, the examples have been chosen to suggest that the normative appeal of the Pareto criterion depends on the context of application, and that it is *not obvious* when it bites and when not. We presume that for most readers, this appeal increases as one goes down the list of examples presented.

Whether or not the Pareto criterion is normatively applicable may in general depend on a number of features of decision situation at hand, for example on whether or not agents' judgments are commonly known, whether the agents are assumed to be "rational" or not, whether the judgments concern beliefs, tastes or values, etc.. Here, we shall not try to settle or analyze this complex issue exhaustively, but hope to provoke interest in further analysis. Beyond that, we shall argue that the applicability of the Pareto criterion turns in particular on the relation between the agents and the outcome decision. More specifically, we shall argue that, under appropriate conditions, if agents have a, possibly shared, *self-interest* in the decision, the Pareto criterion is normatively compelling, while it lacks normative force if agents have a *shared responsibility* for the decision.

Both cases arise in settings in which the notion of reason-based group choice seems relevant. At one end of the spectrum, judicial decision-making is arguably a paradigmatic situation of shared responsibility; at the other, many political and economic group decisions must be understood as partly or exclusively driven by self-interest. The results of this paper imply that, in the latter case, the notion of reason-based social choice must be revised substantially if it is to remain viable at all.[6]

**Related Literature.**—

The Dilemma of the Paretian Rational is related to and goes beyond the original

[5] In the case of a new drug, this is obviously unrealistic, but it need not be in variants, for instance, in the case of legalizing marijuana.

[6] See the discussions at the end of section 4 and subsection 7.1 for elaboration of this point.

Discursive Dilemma introduced by Kornhauser and Sager (1986). The latter's robustness has been demonstrated in a series of impossibility theorems by List-Pettit (2002), Pauly-van Hees (2003), Dietrich (2004a and 2004b), Nehring-Puppe (2005b), van Hees (2004) and Dokow-Holzman (2005). These results show the existence of unavoidable inconsistencies between those social decision procedures that directly aggregate outcome judgments and those that arrive at outcome judgments indirectly by way of logical inference from the aggregation of judgments on the premises. While such inconsistencies raise important questions regarding the nature of "good" decision procedures, they do not seem to be genuinely paradoxical. Indeed, why should the taking account of the reasons for a collective decision not modify the decision itself, compared to what that decision would have been taken without regard to these reasons? According to the Discursive Dilemma, reason-basedness makes a difference, but a difference that a primarily outcome-oriented view may well be able to live with. By contrast, the conflict between premise-based choice and *unanimous* outcome preferences exhibited in the Dilemma of the Paretian Rational challenges a normatively fundamental principle. In particular, while the overriding of a majority or supermajority of agents may well be normatively defensible in cases in which the agents are self-interested, the overriding of *unanimous* outcome preferences is much harder to justify in such cases, if it can be justified at all.

The above-mentioned literature on the Discursive Dilemma can be viewed as a special case of the general problem of aggregating interdependent binary "valuations" formulated originally in an insufficiently known paper by Wilson (1972) as a generalization of Arrow's preference aggregation problem (Arrow 1963) and pursued further in papers by Fishburn-Rubinstein (1986), Nehring-Puppe (2002, 2004, 2005)[7] and Dokow-Holzman (2005). The special feature of the above literature on the Discursive Dilemma is the existence of one binary judgment (the outcome judgment) that is functionally determined by the others, a feature shared by the present paper. However, this literature assumes that the outcome judgment is aggregated independently from the aggregation of the premise judgements. As these determine the outcome

[7]In these papers, which were motivated by the problem of characterizing strategy-proof voting rules on restricted preference domains, the aggregation problem was termed "voting by properties".

judgment uniquely, the merit of this independence assumption can be questioned.[8] Since the present paper assumes only that the logically independent premise judgments are aggregated independently from each other, it is immune to this critique. While all of the quoted contributions can be viewed as analogues or generalizations of Arrow's impossibility theorem, the main results of the present paper have no obvious analogue in the preference aggregation literature.

While the robustness and frequent inescapability of the Paretian Dilemma has not been recognized before in the literature, Dietrich and List (2004) have explored possible violations of the Pareto principle as the consequence of using aggregation rules that delegate the group decision on different premises to disjoint sets of "experts", a phenomenon akin to Sen's (1970) "Liberal Paradox".[9] We derive weak conditions for the occurrence of such an "expert paradox" in section 5, and show that frequently even the mere differential weighting of experts across premises leads to Pareto conflicts. Of course, in the case of more complex decision functions, the expert paradox is subsumed by the wholesale impossibility of a Paretian rational asserted by our main result.

Independently, Mongin (2005) shows the non-existence of non-dictatorial aggregation rules that are simultaneously consistent for a "rich" set of decision rules with at least three premises; for example, while the Paretian Dilemma in the case of the conjunctive decision function in Table 1 can be overcome by adopting a 2/3 supermajority rule, this rule, if applied to a disjunctive decision function would give rise to Pareto inconsistencies.[10] A conclusion similar to Mongin's is observed here in Section 3. However, such a conclusion seems both less surprising and much less troubling than the frequent impossibility of any non-dictatorial Pareto inconsistent aggregation rule for a given group decision function uncovered here: not only does the group (as imagined in the standard judgment aggregation scenario) face one decision at a time, it is intuitively also very natural to adapt the aggregation rule to the nature of the decision problem at hand.

Finally, a distinction between shared self-interest and shared responsibility with the

[8] **See Mongin (2005) for a forceful criticism.**

[9] An example isomorphic to Table 1 is independently given in List (2004b) without further analysis.

[10] I became aware of Mongin's work after this paper, including the discussion in section 3, had essentially been completed.

intended meaning does not seem to have been proposed before in the economics and social-choice literatures. Perhaps the closest contribution is Philip Pettit's "Groups with Minds of their Own" (Pettit 2001b) who argues for the importance of reason-based collective choice in the constitution of a distinct group agency.

**Outline of the Paper.—**

After setting up the framework and notation in Section 2, we characterize in section 3 the class of separable aggregation rules that are Pareto consistent with any given monotone decision function. This result provides the technical foundation for the remainder of the analysis. In section 4, we derive the main result of the paper, the classification of decision functions according to which kind of Pareto consistent aggregation rules they admit. In the following section 5, we use the characterization of section 3 to study when the use of different aggregation rules for different premises leads to Pareto conflicts.

We then ask in section 6 whether the Paretian Dilemma is the result (or even artefact) of the propositional structure of the individual and group judgments. In cases in which all premise judgments are judgments of belief, the natural alternative to a propositional framework is a Bayesian one. While this entails both mathematical and conceptual shifts, we suggest by way of examples that the broad picture does not seem to change fundamentally, leaving a more definitive conclusion to future research. We also point out some interesting connections with the impossibility results in the literature on Bayesian aggregation starting with Hylland-Zeckhauser (1969).

In section 7, we then turn to the normative implications of our results. We argue that while in contexts of "shared self-interests", the normative support for the Pareto criterion remains strong, it breaks down in contexts of "shared responsibility". Section 8 concludes, and the Appendix collects all proofs.

## 2. FRAMEWORK AND NOTATION

A group of $n$ agents $i \in I$ is faced with making a binary 1-0- ("Yes"-"No") decision. The group has agreed to make this choice on the basis of a set of $K \geq 2$ binary decision criteria ("premises") $\{1, .., K\}$ that will also be denoted simply by $K$. A *judgment*

$J$ is a set of premises, with the interpretation that $k \in J$ means that premise $k$ is accepted by the individual or group, while $k \in K \backslash J$ means that premise $k$ is rejected. Individual agent's judgments are denoted by $J_i$, the group judgment by $J_I$. An *aggregation rule* $F : \left(2^K\right)^I \to 2^K$ maps profiles of individual judgments $(J_i)_{i \in I}$ to a group judgment $J_I = F\left((J_i)_{i \in I}\right)$. In turn, the group judgment determines the group choice via a *group decision function* $\Phi : 2^K \to \{0, 1\}$. The composition $\Phi \circ F$ describes a "social choice function" that maps profiles of individual judgments to final outcomes. Throughout and w.l.o.g., all premises are assumed to be essential: that is, for all $k \in K$, there exists $J \in 2^K$ such that $\Phi(J) = 0$ and $\Phi(J \cup \{k\}) = 1$.

The goal of this paper is to determine under which conditions the reason-based group choice $\Phi\left(F\left((J_i)_{i \in I}\right)\right)$ agrees with the unanimous outcome choice of the individuals. Aggregation rules $F$ that ensure such agreement at all profiles $(J_i)_{i \in I}$ will be called Pareto consistent with the decision function $\Phi$.

**Definition 1** *The aggregation rule $F$ is **Pareto consistent with** $\Phi$ if, for all profiles $(J_i)_{i \in I}$, $\Phi\left(F\left((J_i)_{i \in I}\right)\right) = 0$ whenever $\Phi(F(J_i)) = 0$ for all $i \in I$, and $\Phi\left(F\left((J_i)_{i \in I}\right)\right) = 1$ whenever $\Phi(F(J_i)) = 1$ for all $i \in I$.*

Note that, in order to interpret an overriding of agents' unanimous outcome judgments $\Phi(F(J_i))$ as a genuine violation of the Pareto principle, agents must be assumed to care about the group decision on outcomes only. In scenarios in which the agents care directly about the group decision on the premises as well, this interpretation would cease to be appropriate; for example, in a juridical context, individual judges may care about the precedent set by the group decision on a particular premise.[11] In such cases, there may well be a trade-off between respecting agents' outcome judgments versus respecting their premise judgments, and an overriding of even a unanimous outcome judgment may well be Pareto efficient.

Throughout, we will assume that the group judgments on different premises are determined independently; that is, there exists a family of premise-wise aggregation

[11] Note, however, that the interest in precedents can be attributed to an interest in *future* outcome judgments; this apparent counterexample to the Pareto interpretation of Defintion 1 would thus cease to apply in a richer context in which future outcome decisions are explicitly modeled.

rules $F_k : 2^I \to \{0,1\}$ such that

$$F\left((J_i)_{i\in I}\right) = \{k \leq K : F_k\left(\{i \in I : J_i \ni k\}\right) = 1\};$$

such $F$ will be called *separable*. In many contexts, separability will be compelling due to the logical independence of the premises; in the literature on judgment aggregation, separable rules are also referred to as "premise-centered" aggregation rules. Throughout, and wrapped into the notion of separability, we will assume that each $F_k$ is monotone, i.e. $F_k(W) = 1$ and $W' \supseteq W$ imply $F_k(W') = 1$ , and respects unanimity, i.e. $F_k(\varnothing) = \varnothing$ and $F_k(I) = 1$.

It is convenient to represent a rule $F_k$ in terms of its families of "winning coalitions" $\mathcal{W}_k := F_k^{-1}(1)$, and to write $F = (\mathcal{W}_k)_{k\leq K}$. It is also often useful to consider the set of coalitions that are winning for the negation of $a_k$ $\mathcal{W}_k^0 := \{W : W^c \in F_k^{-1}(0)\}$; these are the sets of agents whose rejection of $a_k$ entails a rejection of $a_k$ by the group. In the important special case of anonymous rules, these rules can be parametrized in terms of a vector of quotas $(q_k)$ as $(\mathcal{W}_{q_k})$ by defining, for any $q \in [0,1]$, $\mathcal{W}_q := \{W \in 2^I : \#W > qn$ or $\#W = n$ and $q = 1\}$. For example, proposition-wise majority voting is defined for an odd number of agents by setting $\mathcal{W}_k = \mathcal{W}_{\frac{1}{2}}$ for all $k$.

**Representing Judgments and Decision Functions by Propositions.—**
Intuition is aided greatly by interpreting judgments and decision functions as propositions. This involves a certain amount of technical detail for now, but will pay off substantially later on. Thus, associate with each premise $k \in K$ an atomic proposition $a_k$, and let $\Pi$ denote the set of all complex propositions built from the atomic propositions $\{a_1, ..., a_K\}$ through the logical connectives "and", "or", "not" denoted by $\wedge, \vee$, and $\neg$, respectively. The judgment $J$ corresponds to the complex proposition

$$\pi_J := \left(\bigwedge_{j\in J} a_j\right) \wedge \left(\bigwedge_{j\in J^c} \neg a_j\right),$$

recording the affirmation of all accepted and the negation of all rejected premises. In the converse direction, any complex proposition $\pi \in \Pi$ describes a unique *set* of judgments $\mathcal{J}(\pi) \in 2^{\left(2^K\right)}$ in the obvious way. Formally, $\mathcal{J}(\pi)$ is pinned down by the follow-

ing three stipulations: i) $\mathcal{J}(\{a_k\}) = \{J \in 2^K : J \ni k\}$, ii) $\mathcal{J}(\pi \wedge \pi') = \mathcal{J}(\pi) \cap \mathcal{J}(\pi')$, and iii) $\mathcal{J}(\neg\pi) = \mathcal{J}(\pi)^c$. For example, with $K = 3$, one has $\mathcal{J}(a_1 \wedge a_2 \wedge a_3) = \{J \in 2^K : J \ni a_1, J \ni a_2, \text{ and } J \ni a_3\} = \{\{a_1, a_2, a_3\}\}$; more generally, these stipulations imply that $\mathcal{J}(\pi_J) = \{J\}$ for any judgment $J \in 2^K$. Note also the following two straightforward facts: first, any set of judgments $\mathcal{J} \in 2^{\left(2^K\right)}$ is described by some proposition; that is, for any $\mathcal{J} \in 2^{\left(2^K\right)}$, there exists $\pi \in \Pi$ such that $\mathcal{J} = \mathcal{J}(\pi)$. Second, complex propositions describe the same set of judgments if and only if they are equivalent; that is, $\mathcal{J}(\pi) = \mathcal{J}(\pi')$ for any $\pi, \pi' \in \Pi$ if and only if $\pi$ is logically equivalent to $\pi'$.

The propositional characterization of sets of judgments will be used in particular to characterize the "acceptance region" $\mathcal{J}^+ := \Phi^{-1}(1)$ and the "rejection region" $\mathcal{J}^- := \Phi^{-1}(0)$ associated with a decision function $\Phi$; these two regions denote the sets of all judgments leading to positive and negative outcome decisions, respectively. For example, the acceptance region of the decision function $\Phi_{2,3}$ on $2^3$ defined by $\Phi_{2,3}(J) = 1$ if and only if $\#J \geq 2$ can be described by the proposition $(a_1 \wedge a_2) \vee (a_1 \wedge a_3) \vee (a_2 \wedge a_3)$, which simply says that at least two premises are accepted.

**Monotone Decision Functions.—**

As a matter of significant technical and expositional simplification, we will maintain the assumption that the group decision function $\Phi$ is *monotone* in the sense that $\Phi(J) = 1$ and $J' \supseteq J$ imply $\Phi(J') = 1$; intuitively, a decision function is monotone if the the acceptance of any premise never reverses a positive decision. Much of the analytical benefit of focusing on monotone decision functions derives from the existence of the following canonical propositional representation as a disjunction of conjunctions that is directly useful in the study of Pareto consistency.

To this behalf, let $\min \mathcal{J}^+$ denote the family of inclusion-minimal judgments in $\mathcal{J}^+$, and write $\min \mathcal{J}^+ = \{J_m^+\}_{m \in M^+}$ for an appropriate index set $M^+$; similarly, let $\max \mathcal{J}^- = \{J_m^-\}_{m \in M^-}$ denote the family of inclusion-maximal judgments in $\mathcal{J}^-$ for an appropriate index set $M^-$. Clearly, if the underlying decision function is monotone, a judgment $J$ belongs to the acceptance region $\mathcal{J}^+$ if and only if it contains some judgment $J' \in \min \mathcal{J}^+$. The acceptance region $\mathcal{J}^+$ associated with any *monotone* group decision function $\Phi$ can therefore be described as a disjunction of conjunctions

by the proposition

$$\pi_{\Phi} := \bigvee_{m \in M^{+}} \left( \bigwedge_{j \in J_m^{+}} a_j \right); \tag{1}$$

that is, $\mathcal{J}(\pi_{\Phi}) = \mathcal{J}^{+}$, with $\pi_{\Phi}$ being the unique disjunction of conjunctions of atomic propositions $\pi$ such that $\mathcal{J}(\pi) = \mathcal{J}^{+}$.[12]

Similarly, the rejection region $\mathcal{J}^{-}$ is described by the negation $\neg\pi_{\Phi}$; in turn, $\neg\pi_{\Phi}$ is logically equivalent to a unique disjunction of conjunctions of *negated* atomic propositions given by

$$\bigvee_{m \in M^{-}} \left( \bigwedge_{j \in \left(J_m^{-}\right)^{c}} \neg a_j \right). \tag{2}$$

In the example of the decision function $\Phi_{2,3}$ introduced above, (2) is given by $(\neg a_1 \wedge \neg a_2) \vee (\neg a_1 \wedge \neg a_3) \vee (\neg a_2 \wedge \neg a_3)$ : the outcome decision is negative if and only if at least two of the three premises are rejected.

Since a decision rule $\Phi$ is uniquely characterized by the canonical representation $\pi_{\Phi}$ of its acceptance region given by (1), we will use the symbol $\Phi$ to refer to both the decision function and the proposition $\pi_{\Phi}$ characterizing it. Likewise, we will use the symbol $\neg\Phi$ for the canonical representation (2) of the rejection region $\mathcal{J}^{-}$. It is very natural to think of a decision function in propostional terms, since a judgment $J$ leads to a positive outcome decision if and only if the decision proposition $\pi_{\Phi}$ is logically entailed by the judgment proposition $\pi_J$.

Monotonicity seems plausible for many applications. Consider, for example, the case of a tenuring decision. The decision function $\Phi$ is the agreed upon tenuring standard. Each premise can be viewed as a "Lancasterian characteristic" of the candidate's record; monotonicity assumes that characteristics are unambiguously desirable or undesirable. The conjunctions in (1) represent the minimal combinations of desirable characteristics that warrant tenure. The disjunction operator captures the fact that some characteristics can substitute for others; for example excellence in teaching can

[12] If one lets $\Pi_{mon}$ denote the class of all complex propositions that are logically equivalent to a disjunction of conjunctions of atomic propositions as in (1). Clearly, any $\pi \in \Pi_{mon}$ defines a monotone decision rule with acceptance region $J(\pi)$. Thus, $\Pi_{mon}$ is the class of propositions associated with some monotone decision rule. It is also the smallest class of propositions containing the atomic propositions and closed under conjunction, disjunction, and logical equivalence.

make up for a narrow research record. In any case, the assumption of monotonicity is made for analytical and technical convenience only. The central tool of the analysis, the Pareto Intersection Property stated in Proposition 2, would remain applicable in modified form, and the overall drift of the results would likely remain the same; if anything, the balance would further tilt toward impossibility.

## 3. PARETO CONSISTENT AGGREGATION RULES: CHARACTERIZATION

The following result characterizes the class of separable aggregation rules that are Pareto consistent with it any given decision function $\Phi$.

**Proposition 2** ***(Pareto Intersection Property)*** *The separable aggregation rule $F = (\mathcal{W}_k)$ is Pareto consistent with the monotone decision function $\Phi$ if and only if it satisfies the following pair of conditions:*

*i) For any $J \in \min \mathcal{J}^+$ and any selection $W_k \in \mathcal{W}_k$ for $k \in J$, $\cap_{k\in J} W_k \neq \varnothing$.*

*ii) For any $J \in \max \mathcal{J}^-$ and any selection $W_k \in \mathcal{W}_k^0$ for $k \in J^c$, $\cap_{k\in J^c} W_k \neq \varnothing$.*

The idea behind the proof of Proposition 2 is the following. Pareto consistency amounts to the requirement that judgment profiles that are wholly contained in the acceptance region $\mathcal{J}^+$ (respectively the rejection region $\mathcal{J}^-$) map to a collective judgment in the same region. Thus, the characterization of Pareto consistent separable aggregation rules can be deduced from the characterization of separable aggregation rules on restricted domains of 'feasible' judgments $\mathcal{D}$ $(= \mathcal{J}^+, \mathcal{J}^-)$. Such a characterization has been provided in Nehring-Puppe 2004 (henceforth simply NP) in terms of an "Intersection Property" which captures the combinatorial structure of the domain $\mathcal{D}$ in terms of its "critical families"; to these correspond here the elements of the sets $\min \mathcal{J}^+$ and $\max \mathcal{J}^-$, which will be referred to as "critical judgments". In view of (1) and (2), these can be read off immediately from the canonical disjunctive representation of $\Phi$ and its negation.

In the anonymous case, the Pareto Intersection Property (henceforth: PIP) takes a particularly simple form in that the set of Pareto consistent rules can be characterized by a system of linear inequalities; for details, see NP, section 3.3.

**Fact 3** ***(Anonymous Pareto Intersection Property)*** *If $(\mathcal{W}_k)$ is anonymous and Pareto consistent with the monotone decision function $\Phi$, there exists a system of quotas $(q_k)$ such that*

*i) $\mathcal{W}_k = \mathcal{W}_{q_k}$ for all $k \leq K$,*

*ii) for any $J \in \min \mathcal{J}^+$, $\sum_{k \in J} q_k \geq \#J - 1$,*

*iii) for any $J \in \max \mathcal{J}^-$, $\sum_{k \in J^c} q_k \leq 1$,*

*iv) for all $H \in \mathcal{H}$, $nq_k$ is not an integer other than $0$ or $n$.*

*Conversely, if $(q_k)$ is a vector of quotas satisfying ii), iii) and iv), the aggregation rule $(\mathcal{W}_{q_k})$ is anonymous and Pareto consistent.*

The conditions ii) and iii) are the counterparts to the two set-theoretic conditions making up the PIP. The role of condition iv) is to ensure that the dual committees $\mathcal{W}^0_{q_k}$ are equal to the committees $\mathcal{W}_{1-q_k}$; this becomes important in situations in which all anonymous choice-functions require some quota $q_k$ to be equal to $\frac{1}{2}$; clause iv) implies in this case that $n$ must be odd. This makes intuitive sense in that majority voting is well-defined only for an odd number of individuals.

To see how the PIP works in the anonymous case, consider a system of quotas $(q_k)$ violating part iii) of Fact 3, and consider a critical judgment $\overline{J} \in \max \mathcal{J}^-$ such that $\sum_{k \in \overline{J}^c} q_k > 1$. For $k \in \overline{J}^c$, pick rational numbers $\mu_k < q_k$ such that $\sum_{k \in \overline{J}^c} \mu_k = 1$ and consider any profile $(J_i)$ such that exactly the fraction $\mu_k$ of agents has the judgment $\overline{J} \cup \{k\}$.[13] By choice of $\overline{J}$, each judgment of the form $\overline{J} \cup \{k\}$ is in the acceptance region $\mathcal{J}^+$; hence, at the defined profile, the group favors unanimously a positive outcome decision. On the other hand, each premise $k \in \overline{J}^c$ is favored by only the fraction $\mu_k$ of agents, which is not enough to reach the quota $q_k$; hence each premise $k \in \overline{J}^c$ is rejected by the group, and only the premises in $\overline{J}$ are accepted. But since $\overline{J}$ is in the rejection region $\mathcal{J}^-$, the group reaches a negative outcome decision, violating Pareto consistency.

To further illustrate the content and power of the Pareto Intersection Property, consider first the simplest example in which $\Phi$ is the conjunction of $K$ atomic propo-

[13] The meticulous reader will note that this explanation appeals to an appropriately chosen number of agents $n$, while Fact 3 is valid for any fixed $n$.

sitions. Here the PIP requires that

$$\sum_{k \leq K} q_k \geq K - 1.$$

This means that, in particular, a strict supermajority rule with uniform supermajority $q_k = q$ is Pareto consistent if and only if

$$q \geq 1 - \frac{1}{K}.$$

More generally, consider the class of decision functions $\Phi_{L,K}$ equivalent to the satisfaction of at least $L$ out of $K$ atomic propositions (criteria). The critical jugments in $\min \mathcal{J}^+$ are exactly the judgments of cardinality $L$. Similarly, the critical jugments in $\max \mathcal{J}^-$ are exactly the judgments of cardinality $L - 1$; their complements have thus cardinality $K - L + 1$. The PIP therefore requires that

$$\sum_{k \in J} q_k \geq L - 1,$$

for all $J$ with $\#J = L$, and

$$\sum_{k \in J^c} q_k \leq 1$$

for all $J$ with $\#J = L - 1$. Adding up these inequalities implies that

$$\frac{1}{K} \sum_{k \leq K} q_k \geq 1 - \frac{1}{L}, \tag{3}$$

as well as

$$\frac{1}{K} \sum_{k \leq K} q_k \leq \frac{1}{K - L + 1}, \tag{4}$$

respectively. Equations (3) and (4) are jointly satisfiable if and only if

$$1 - \frac{1}{L} \leq \frac{1}{K - L + 1},$$

i.e. iff $L = 1$, $L = K$, or $L = 2$ and $K = 3$. In particular, if $K \geq 4$ and $1 < L < K$, all anonymous aggregation rules are Pareto inconsistent. According to Theorem 7

below, this conclusion extends to all non-dictatorial aggregation rules.

This class of examples also shows that the set of Pareto consistent aggregation rules depends heavily on the decision function $\Phi$, if it is non-degenerate at all. Indeed, for the system of quotas $(q_k)$ to simultaneously be Pareto consistent with both the complete conjunction *and* the complete disjunction of $K$ premises, the quotas need to satisfy

$$\sum_{k \leq K} q_k \geq K - 1 \text{ and } \sum_{k \leq K} q_k \leq 1;$$

this is possible at all only if $K = 2$![14]

## 4. PARETO CONSISTENT AGGREGATION RULES: EXISTENCE

Rather than considering one decision function at a time, we now take a broader view and classify decision functions according to the kind of Pareto-consistent aggregation rules they admit. Along with anonymity, we will consider the following properties: dictatorship, local dictatorship, veto power, and neutrality. These are most crisply defined in terms of the winning coalitions characterizing separable aggregation rules.

**Definition 4** *1. A separable aggregation rule $F = (\mathcal{W}_k)_{k \in K}$ is **dictatorial** if there exists an individual $i \in I$ such that $\{i\} \in \mathcal{W}_k \cap \mathcal{W}_k^0$ for all $k \leq K$;*

*2. $F$ is **locally dictatorial** if there exists an individual $i \in I$ and a premise $a_k$ such that $\{i\} \in \mathcal{W}_k \cap \mathcal{W}_k^0$;*

*3. $F$ exhibits **veto power** if there exists an individual $i \in I$ and a premise $a_k$ such that $\{i\} \in \mathcal{W}_k \cup \mathcal{W}_k^0$;*

*4. $F$ is **neutral** if, for all premises $a_k$ and $a_\ell$, $\mathcal{W}_k = \mathcal{W}_\ell = \mathcal{W}_\ell^0$.*

Thus, neutrality requires both that the winning coalitions for any premise are the same as those for its complement, and that these winning coalitions are identical across premises. Clearly, in the anonymous case, neutrality amounts to majority

[14] Using the Pareto Intersection Property, this can be generalized to show that if the same (possibly non-anonymous) aggregation rule is Pareto consistent with the both the disjunction and the conjunction of the same three premises, it must be dictatorial; a similar result has recently been obtained independently by Mongin (2005).

voting on each premise; as remarked above, for this to be well-defined, the number of individuals must be odd. Note that, by contrast, non-anonymous and neutral aggregation rules exist even if the number of individuals is even. Also note that an aggregation rule exhibits no veto power if and only if any premise is accepted/rejected whenever at least $n-1$ individuals accept/reject it.

The existence of a possibility result depends on the complexity and structure of the group decision function $\Phi$. The complexity of the group decision function can be measured by the maximal cardinality of its critical judgments $\kappa_\Phi$ which is formally defined as

$$\kappa_\Phi := \max\left[\max\{\#J : J \in \min \mathcal{J}^+\}, \max\{\#J^c : J^c \in \max \mathcal{J}^-\}\right].$$

Decision rules with the smallest complexity measure $\kappa_\Phi = 2$ will be called **simple**; they turn out to be exactly the decision functions for which majority voting on properties is Pareto consistent.

**Theorem 5** *The following four statements are equivalent:*

1. *Majority voting on properties is Pareto consistent with* $\Phi$;
2. *There exists a neutral and non-dictatorial separable aggregation rule that is Pareto consistent with* $\Phi$;
3. $\Phi$ *is simple* ($\kappa_\Phi = 2$);
4. $\Phi$ *has one of the following three forms, where* $a, b, c, d$ *are not necessarily distinct:*

$$\begin{aligned} \Phi &= ab \vee cd, \ \textit{or} \\ \Phi &= ab \vee cd \vee bc, \ \textit{or} \\ \Phi &= ab \vee cd \vee bc \vee ad. \end{aligned}$$

To illustrate the equivalence between the third and fourth statements, $\Phi = ab \vee cd$ has complexity $\kappa_\Phi = 2$ since $\neg\Phi \Longleftrightarrow \overline{a}\overline{c} \vee \overline{a}\overline{d} \vee \overline{b}\overline{c} \vee \overline{b}\overline{d}$.[15] By contrast, $\Phi = ab \vee cd \vee e$

[15] Recall the definition of $\neg\Phi$ in section 2. To compactify notation of specific propositions, we frequently denote the negation of an atomic proposition by $\overline{a}$ instead of $\neg a$, and abbreviate a con-

has $\kappa_\Phi = 3$ since $\neg\Phi$ has the canonical representation $\neg\Phi \Longleftrightarrow \overline{ace} \vee \overline{a}\overline{d}\overline{e} \vee \overline{b}\overline{c}\overline{e} \vee \overline{b}\overline{d}\overline{e}$.

By Theorem 5, majority voting is Pareto consistent only in very special circumstances. Yet as illustrated by the example of conjunctive and disjunctive decision functions, Pareto consistency can sometimes be achieved by appropriate supermajority rules that treat a premise and its negation asymmetrically; whether or not this is possible more generally depends on the qualitative structure of the decision function, as will be shown now. We will distinguish three types of decision functions – "indecomposable", "fully decomposable" and "partly decomposable"– and consider them in turn.

The decision function/proposition $\Phi$ is **indecomposable** if no premise is by itself either necessary or sufficient for the satisfaction of $\Phi$. Formally, $\Phi$ is indecomposable if each conjunct of the canonical representations of both $\Phi$ and $\neg\Phi$ combines at least two premises. For example, among the family of propositions $\Phi_{L,K}$ described above, $\Phi_{L,K}$ is indecomposable if and only if $1 < L < K$.

To complete the picture, we need to consider intermediate cases between those of simple conjunctions/disjunctions on the one hand and indecomposable ones on the other. Suppose thus that $\Phi$ is decomposable (not indecomposable). We note the following elementary fact.

**Fact 6** *A monotone proposition $\Phi$ is decomposable if and only if there exists a premise $a_k$ and a monotone proposition $\Phi'$ with $n-1$ arguments such that*

$$\Phi(a_1, ..., a_n) = \Phi'(a_1, .., a_{k-1}, a_{k+1}, .., a_n) \vee a_k$$

*or*

$$\Phi(a_1, ..., a_n) = \Phi'(a_1, .., a_{k-1}, a_{k+1}, .., a_n) \wedge a_k.$$

If the monotone proposition $\Phi'$ described by this Fact is in turn decomposable, one is able to further simplify the representation of $\Phi$ by peeling off further premises until one arrives at an indecomposable proposition $\Phi^*$ or has used up all arguments, in which case $\Phi^*$ can be viewed as "empty". It is easy to see that $\Phi^*$ is uniquely

junction such as $a \wedge b \wedge$ c as $abc$.

defined.[16] We will refer to $\Phi^*$ as the "core" of $\Phi$. If the core is empty, $\Phi$ is **fully decomposable**; if the core $\Phi^*$ is non-empty but $\Phi$ is decomposable, then $\Phi$ is **partly decomposable**.

**Theorem 7** *i) If $\Phi$ is monotone and indecomposable, it admits a non-dictatorial and Pareto consistent separable aggregation rule $F$ if and only if $\Phi$ is simple. If $\Phi$ is simple, then a separable aggregation rule $F$ is Pareto consistent if and only if it is neutral.*

*ii) If $\Phi$ is monotone and fully decomposable, then it admits an anonymous separable aggregation rule without veto power that is Pareto consistent.*

*iii) If $\Phi$ is monotone and partly decomposable with a simple core, then it admits an anonymous separable aggregation rule that is Pareto consistent.*

*On the other hand, all Pareto consistent separable aggregation rules exhibit veto power.*

*iv) If $\Phi$ is monotone and partly decomposable with a non-simple core, then it admits a non-dictatorial separable aggregation rule that is Pareto consistent.*

*On the other hand, all Pareto consistent separable aggregation rules are locally dictatorial.*

**Corollary 8** *If $\Phi$ is monotone, then it admits an anonymous Pareto consistent separable aggregation rule $F$ if and only if its core $\Phi^*$ is empty or simple.*

The key step to proving the first part of the Theorem is to show that in the indecomposable case any Pareto consistent aggregation rule must be neutral; by Theorem

[16] Formally, one arrives at a uniquely defined "canonical decomposition" of $\Phi$ of the following form, for an appropriate enumeration of the premises $\{a_1, ..a_K\}$:

$$\Phi = \left( \bigvee_{k=K,..,k_1} a_k \right) \vee \left( \left( \bigwedge_{k=k_1,..,k_2} a_k \right) \wedge ... \left( \bigvee_{k=k_{m-1},..,k_m} a_k \vee \Phi^* \left(1, .., k_m - 1\right) \right) \right),$$

where $\Phi^*$ is indecomposable or empty (iff $k_m = 1$), with $m \geq 0$ and $K + 1 \geq k_0 \geq k_1 \geq k_2 \geq ...k_{m-1} \geq k_m - 1 \geq 0$.

In this notation, $\Phi$ is indecomposable if $m = 0$ and $k_0 = K$. Also, if $\Phi$ is partly composable and $k_1 = k_0 = K + 1$, this means that the leading disjuctive term is empty, and that therefore the expression for $\Phi$ starts properly with a disjunctive term; similarly, if $k_{m-1} = k_m - 1$, the last disjunctive term is empty.

5 above, $\Phi$ must therefore be simple; conversely, as we know again from Theorem 5, for simple $\Phi$, neutrality ensures Pareto consistency.

The proof of part ii) of Theorem 7 is non-trivial due to the fact that the quotas $q_k$ associated with the specified aggregation rules may *need* to be non-constant, i.e. there may not exist $q : \frac{1}{2} \leq q < 1$ such that $q_k \in \{q, 1-q\}$. This happens, for example, for the decision function $\Phi = f \vee e \vee (d \wedge (c \vee ba))$. Indeed, in such cases the required supermajorities may be more extreme than a straightforward analogy to pure con- or disjunctions would suggest. In the example $\Phi = f \vee e \vee (d \wedge (c \vee ba))$, the least extreme supermajority $\max_k \max(q_k, 1-q_k)$ must be at least $\frac{9}{10}$,[17] while a conjunction of 6 premises would require a uniform supermajority of only $\frac{5}{6}$ to be Pareto consistent.

To illustrate why partial decomposability entails a veto as asserted in part iii), consider the partly decomposable decision function $\Phi = a \vee bc \vee bd \vee cd$. Suppose that $F$ is an anonymous, Pareto consistent separable aggregation rule. From part i) of the Theorem, it is clear that $F$ must require majority voting over the core premises $b, c$, and $d$. Consider the following profile of judgments among $n = 2m + 1$ agents: $m$ agents hold the judgment $\overline{a}b\overline{c}d$, another $m$ agents hold the judgment $\overline{a}bc\overline{d}$, and a single agent $i$ holds the judgment $a\overline{b}\overline{c}\overline{d}$. In particular, all agents affirm $\Phi$. On the other hand, a majority rejects each of the three core-premises $b, c$, and $d$. Thus, for the group to affirm $\Phi$ as required by Pareto consistency, it must affirm $a$. But since $i$ is the only agent affirming $a$, this means that $i$ must have a veto against rejecting $a$.

The assumptions behind Theorems 5 and 7 are special in a number of ways whose relaxation would not substantially alter the upshot of these results. First, if one would drop the monotonicity assumption on the decision function $\Phi$, the logic of the analysis would remain the same; the PIP would simply have to be formulated in a more general way. However, the analysis would become substantially more complicated and most likely more messy due to the absence of a canonical propositional representation. Nonetheless, the thrust of the results, i.e. the confinement of possibility results to decision functions with quite special and simple structure, would in all likelihood remain the same; if anything, in the absence of monotonicity, the balance would

[17] This is achieved by the following quotas obtained by Linear Programming: $q_f = q_e = 0.1$, $q_d = 0.8$, $q_c = 0.2$, $q_a = q_b = 0.6$ .

probably tilt even further towards Pareto inconsistency, in that it will frequently be impossible to ensure even one-sided respect for unanimity on positive (or negative) outcome decisions only.

Second, we have assumed that the outcome decision can take on two values only. If the outcome decision is multiple-valued (with $\Phi : J \mapsto y \in Y$), the weakest natural notion of Pareto consistency requires that if all agents agree on some outcome decision (which can thus be viewed as implicitly the unanimously "most preferred"), the group choice must agree with this decision. Under this definition, the role of the sets $\mathcal{J}^-$ and $\mathcal{J}^+$ is now taken by the partition of $\{0,1\}^K$ into the inverse images $\mathcal{J}_y := \Phi^{-1}(y)$, with $y \in Y$. On the one hand, multi-valuedness leads to additional flexibility in chopping up the domain of the decision function in appropriate pieces; in the extreme case in which the $\mathcal{J}_y'$s are all singletons, Pareto consistency loses its bite. On the other hand, with more than two outcome decisions in the range of $\Phi$, at least one region must be non-monotonic, which will typically make it hard for Pareto consistency to be satisfied. For example, if the decision function $\Phi : J \mapsto \{0, ..., \#K\}$ is given by $\Phi(J) = \#J$, non-dictatorial Pareto consistent aggregation is possible only if $\#K \leq 2$.[18] Generalizing this example, it seems likely that the second effect of multi-valuedness will dominate in most applications.

Third, we have assumed a universal domain of premise judgments in the sense that any combination of premise judgments is allowed. If certain combinations of premise judgments are excluded, for example due to logical or semantic entailment relations, this complicates the exact characterization of the possibilities but leaves the broad picture intact. Roughly speaking, if the entailment structure among premises is such that the space of judgments is two-dimensional (i.e. can be embedded in the product of two trees in an appropriate sense, cf. Nehring-Puppe 2003), possibility results predominate. On the other hand, as soon as the space of judgments is at least three-dimensional, possibility results will obtain only in fairly special and restrictive circumstances.[19]

[18] This follows from the fact that if $\#K > 2$, the sets $\mathcal{J}_y$ for $1 \leq y \leq \#K-1$ are totally blocked; due to Theorem 1 of Nehring-Puppe (2005a), this implies that the aggregation rule must be dictatorial.

[19] To be a bit more precise and specific, consider the case of majority voting on interrelated premises. For this to be consistent, the space of judgments must be a median space; see NP, Theorem 4. For majority voting on premises to be Pareto-consistent in addition, the sets $\mathcal{J}^+$ and $\mathcal{J}^-$ must themselves be median-spaces. This is easily possible in the case of at most two dimensions,

In all of this, we have maintained the assumption that the premise judgments are to be aggregated independently ("separability"). In view of the assumed logical independence of the premises, this assumption will be plausible in many applications. (This is in marked contrast with the original Discursive Dilemma, where the normative appeal of aggregating outcome judgements independently from the judgments on the premises that determine the outcome truth-functionally can be questioned). Separability is especially plausible if the judgments on different premises are *epistemically independent* in that the relevant evidential considerations behind these judgments are independent.[20] For instance, epistemic independence seems highly plausible in the introductory adjudication example of Table 1, where the evidence germane to the three issues of duty, negligence, and causation is clearly quite distinct.

Of course, *if* one is willing to give up separability, it is possible to find rules that satisfy Pareto consistency. One such class of rules are the "maxmin rules" defined as follows. Let $K'=\{1,..,K,K+1\}$ denote the set of *all* judged propositions comprising all premises and the conclusion as the $K+1$-th proposition. For any judgment $J$, let $J' \supseteq J$ denote the associated judgment on propositions, with

$$J' = \begin{cases} J & \text{if } \Phi(J)=0 \\ J \cup \{K+1\} & \text{if } \Phi(J)=1 \end{cases}.$$

For any $k \in K'$, let $s_{k,J}$ denote the support for judging $k$ according to $J$, i.e.

$$s_{k,J} = \begin{cases} \#\{i : k \in J_i'\} & \text{if } k \in J \ni k \\ \#\{i : k \notin J_i'\} & \text{if } k \notin J \end{cases}.$$

An aggregation rule $F$ is a *maxmin rule* if it maximizes the weakest support, i.e. if at all profiles it satisfies

$$F\left((J_i)_{i\in I}\right) \in \arg\max_{J\subseteq K}\left(\min_{k\in K'} s_{k,J}\right).$$

but only in special and restricted ways otherwise.

[20] One way to make the notion of epistemic independence more rigorous formally would be to postulate a setting in which the aggregated premises judgments are themselves determined truth-functionally by judgments about "basic" premises; in this setting, epistemic independence could be equated with disjointness of the basic premises determining each aggregated premise.

Maxmin rules are the analogue of maxmin (Simpson-Kramer) rules in the context of preference aggregation. As in that context, maxmin rules are Condorcet consistent; that is, they agree with proposition-wise majority voting whenever this is consistent. More importantly for the present discussion, they are Pareto-consistent as well, for any proposition $\Phi$, since for any group judgment in violation of a unanimous judgment on any proposition, premise or conclusion, its minimal support $\min_{k \in K'} s_{k,J}$ equals 0, while the minimal support of any individual's judgment $\min_{k \in K'} s_{k,J_i}$ is at least 1.

However, the maxmin rule's violation of separability entails potentially problematic spill-overs across premises. Consider, for instance, the following profile of judgments of a 7-member panel in the adjudication of a tort claim based on two conjunctive premises.

| Judge | Duty | Negligence | Damages |
|---|---|---|---|
| I,II | Yes | Yes | Yes |
| III,IV | No | Yes | No |
| V,VI,VII | Yes | No | No |
| Majority | Yes | Yes | No |
| *Minimax* | *Yes* | *No* | *No* |

and

| Judge | Duty | Negligence | Damages |
|---|---|---|---|
| I,II | Yes | Yes | Yes |
| III,IV | No | Yes | No |
| V | Yes | No | No |
| VI,VII | No | No | No |
| Majority | No | Yes | No |
| *Minimax* | *No* | *Yes* | *No* |

Table 2: The Maxmin Rule at two profiles

Given the top profile of preferences displayed, majority voting on premises entails a negative outcome decision; since this agrees with the outcome decision preferred a majority of judges, it also the maxmin outcome decision. Suppose, though, that

judges VI and VII reconsider, and come to the conclusion that the defendant did have a duty to take care, after all. While the 5:2 majority view on outcomes remains unchanged, now a 5:2 majority of judges affirms a failure of Duty. The maxmin rule breaks the inconsistency of propositionswise majority voting at its weakest link, which is the 4:3 majority in favor of Negligence.

Thus, the group judgment on Negligence has changed even though all of the judges views on this issue remain unchanged. Whether or not such ripple effects are deemed acceptable will depend on the context. In the adjudication example, this seems quite problematic, in that for an outcome judgment to be *well-justified* by group judgments on the underlying premises, these group judgments presumably should be grounded in the individual's judgments on that issue, or at least some epistemically related one. Since the judges views on the issue Duty have no plausible evidential bearing on the issue of Negligence, one feels that the maxmin judgment on Negligence at the second profile is merely a post-hoc rationalization (in the ordinary language, pejorative sense) of the maintained negative outcome decision on Damages.

Finally, it should be borne in mind that the maxmin rule is a rather special in satisfying Pareto consistency almost by construction. In general, natural non-separable aggregation rules may easily violate Pareto consistency; for example, if the group judgment is aggregated by maximizing the total support $\sum_{k \in K'} s_{k,J}$ rather than the minimal support (in analogy to the Kemeny or "median" rule for preference maximization), one can show that, depending on the nature of the monotone decision function, Pareto consistency may or may not hold. Thus, *if* Pareto consistency is deemed normatively inviolable, and if separability is abandoned for that reason, it would likely constitute a rather powerful criterion for selecting among non-separable aggregation rules. Obviously, this constitutes an interesting and potentially rich field for future investigations.

## 5. JUDGMENT DELEGATION TO EXPERTS

In the above analysis, we have related the existence and properties of Pareto consistent separable aggregation rules to the structure of the decision function $\Phi$. We shall now show that Pareto consistency also imposes substantial restrictions on aggregation

rules that do not depend on the structure of the decision function. Broadly speaking, the aggregation rules governing different premises must be "sufficiently similar" to each other to ensure Pareto consistency.

Dissimilarity in this sense originates in particular from the delegation of the judgments on various premises to distinct subgroups of "experts", as illustrated by the following example due to Dietrich and List (2004) who point out its formal analogy to Sen's (1970) "Liberal Paradox".

**Example 9** *Let* $\Phi = a \wedge b$ *and* $I = \{1, 2\}$, *and assume that agent* 1 *decides* $a$ *and agent* 2 *decides* $b$ *(that is,* $\{1\} \in \mathcal{W}_a \cap \mathcal{W}_a^0$ *and* $\{2\} \in \mathcal{W}_b \cap \mathcal{W}_b^0$*). Suppose that* $J_1 = \{a\}$ *and* $J_2 = \{b\}$. *Then both agents reject* $\Phi$; *nonetheless, since both affirm the premises they are authorized to judge,* $F(J_1, J_2) = \{a, b\}$; *the group therefore accepts* $\Phi$, *a Pareto inconsistency.*

The general point contained in the example is that the sets of agents that have decisive influence on the group choice of a premise must always overlap. Formally, for $k \leq K$, define the family of decisive coalitions $\mathcal{D}_k := \mathcal{W}_k \cap \mathcal{W}_k^0$.

**Theorem 10** *For any monotone* $\Phi$ *and any* $F$ *that is Pareto-consistent with* $\Phi$ :

$$\textit{for all } j, k \leq K \textit{ and all } W \in \mathcal{D}_j, W' \in \mathcal{D}_k : W \cap W' \neq \varnothing.$$

Theorem 10 has two more specific corollaries. For any $k$, let $E_k$ denote the set of "essential agents" or "experts" whose judgment counts in the group judgment on premise $k$ : $E_k := \{i \in I :$ there exists $W \in \mathcal{W}_k^c$ such that $W \cup \{i\} \in \mathcal{W}_k\}$. Note that $E_k \in \mathcal{D}_k$ by definition. Thus we have

**Corollary 11** *For any monotone* $\Phi$ *and any Pareto-consistent* $F$ :

$$\textit{for all } j, k \in K : \; E_j \cap E_k \neq \varnothing. \tag{5}$$

Thus, generalizing Example 9 above, whenever two premises are effectively judged by disjoint sets of experts, a potential Pareto inconsistency arises.

More can be said if more is known about the structure of the aggregation rules on the individual premises. Suppose in particular that the aggregation rule $F = (\mathcal{W}_k)$ is

**premise-wise neutral** in the sense that the acceptance and rejection of each premise is treated symmetrically, i.e., for all $k \leq K$, $\mathcal{W}_k = \mathcal{W}_k^0$. For example, $F$ is premise-wise neutral if each premise is judged by weighted majority voting, where individual agents' weights may differ across premises. Indeed, in this case, the aggregation rules used for the different premises must be *identical* for Pareto consistency to obtain; there is no room at all for a differential weighting of agents according to their expertise![21]

**Proposition 12** *If the separable aggregation rule $F = (\mathcal{W}_k)$ is premise-wise neutral and Pareto consistent, it is neutral.*

A result with a drift similar to that of Corollary 11 has been obtained before by Dietrich and List (2004). Corollary 11 goes beyond their results by deriving the necessary overlap of experts from a condition that is explicitly formulated in terms of the structure of the proposition $\Phi$ (monotonicity) rather than implicitly as in their "connectedness"; moreover, since not all monotone propositions are connected, Corollary 11 cannot be derived from their result.[22] Dietrich and List (2004) have no counterpart to the more general Theorem 10 or to Proposition 12.[23]

## 6. ALTERNATIVE LOGICS OF JUDGMENT AGGREGATION

As pointed out at the end of section 4, the thrust of our results is robust to the particular assumptions made: the monotonicity assumption on the decision function, the binary accept-or-reject character of the group decision, and the logical independence of premises. Nonetheless, all of these variations maintain the propositional, hence discrete, structure of individual and group judgments that is the hallmark of the existing literature on judgment aggregation. It is thus natural to wonder whether

[21] Besides its intrinsic interest, Proposition 12 is a key step in proving that when $\Phi$ is indecomposable, Pareto consistency requires neutrality, as asserted by part i) of Theorem 7 above.

[22] An example of a monotone but not connected proposition is $\Phi = abc \vee abd \vee ace \vee bcf$; by Lemma 17 in the Appendix, $\neg\Phi = \overline{a}\overline{b} \vee \overline{a}\overline{c} \vee \overline{b}\overline{c} \vee \overline{a}\overline{f} \vee \overline{b}\overline{e} \vee \overline{c}\overline{d}$. $\Phi$ is not connected, because none of the conjuncts making up $\Phi$ (resp. $\neg\Phi$) contains both $d$ and $e$ (resp. both $\overline{d}$ and $\overline{e}$). Conversely, not all connected propositions are monotone.

[23] On the other hand, Dietrich and List (2004) obtain results that have no counterpart here.

the "Impossibility of a Paretian Rational" is derived more from that propositional structure rather than from reason-basedness itself.

In particular, one may argue that even though individual judgments may plausibly be modelled as propositional, it may be artificial to force social judgments into the same all-or-nothing mold when agents disagree. Such disagreement could reasonably give rise to an element of doubt that should be reflected in intermediate degrees of acceptance or "truth". Thus Pauly-van Hees (2003) and van Hees (2004) study the Discursive Dilemma in the context of multi-valued logic. In Appendix 1, we show that under standard interpretations of the logical connectives "and" and "or", allowing for multi-valuedness leaves the set of rationalizable social choice functions completely unchanged and hence does nothing to mitigate the conflict between reason-basedness and Pareto consistency.

In those cases in which all premise judgments represent beliefs, one may want to depart even further from the propositional setting and assume that all judgments come in the form of probabilities. While a complete and fully satisfactory understanding of the Bayesian version is beyond the scope of the present paper, we shall argue that conflicts between reason-basedness and the Pareto principle arise again naturally and take a broadly similar shape. Assuming stochastic independence of the aggregated events, the following model is formulated in a rather special way to make the analogy to the propositional set-up as tight as possible. It is not meant as an exhaustive discussion of the issues arising in a Bayesian setting.[24]

To arrive at the Bayesian counterpart, suppose now that a group decision is to be taken on the basis of individual agent's probability judgments $\mathbf{p}_i^k$ on $K$ subjectively stochastically independent contingencies $E_k$. Thus each agent's beliefs $\mathbf{p}_i$ are described by a product measure $\otimes_k \mathbf{p}_i^k$ on the state space $\{0,1\}^K$, with $E_k = \{1\} \times \{0,1\}^{K \setminus k}$, where $\mathbf{p}_i^k$ is uniquely specified by the number $p_i^k = \mathbf{p}_i^k(E_k)$, the subjective probability of agent $i$ that the $k-th$ contingency $E_k$ materializes. The individual assessments are aggregated by a separable aggregation rule $H = (H_k)$ into

[24] Much of the specialness of the following formulation is more apparent than real. In particular, the basic points of the following discussion would easily generalize to conditional independence structures. As demonstrated by the explosive growth of "Bayes' nets" and "graphical models" in Bayesian theory and applications over the last 15 years, these are of extremely wide applicabilty and fundamental importance; see, for example, Pearl (1988) and Cowell et al. (1999).

a social product probability measure $\mathbf{p}_I = \otimes_k \mathbf{p}_I^k$, where $\mathbf{p}_I^k = H_k\left(\left(\mathbf{p}_i^k\right)_{i\in I}\right)$; since the component state-spaces are binary, we will write more simply $p_I^k = H_k\left(\left(p_i^k\right)_{i\in I}\right)$, viewing $H_k$ as a mapping from $[0,1]^I$ to $[0,1]$.

Again, the group needs to make a Yes-No-decision on the basis of the aggregated group probabilities. In a Bayesian setting, it is natural to assume that the group uses an expected utility criterion described by an agreed-upon group utility function $u : 2^K \to \mathbf{R}$, where $u(\omega)$ is the (possibly negative) utility gain in state $\omega$ of having chosen "Yes" rather than "No". A given utility function $u$ induces the decision function $\Phi_u$, with

$$\Phi_u(\mathbf{p}) = 1 \text{ if and only if } \sum_{\omega\in 2^K} u(\omega)\,\mathbf{p}(\omega) > 0.$$

Of particular interest are utility functions of the form $u = 1_S - \tau$, where $S$ is an event in $2^K$; in this case, the decision function $\Phi_{(1_S-\tau)}$ simplifies to

$$\Phi_{(1_S-\tau)}(\mathbf{p}) = 1 \text{ if and only if } \mathbf{p}(S) > \tau.$$

That is, the decision is "Yes" if and only if the group assessment of the event $S$ exceeds some threshold value $\tau$. An aggregation rule $H$ is **Pareto consistent with respect to** $u$ if, for all profiles $(p_i)_{i\in I}$, $\Phi_u\left(\otimes_k H_k\left(\left(p_i^k\right)_{i\in I}\right)\right) = 1$ (respectively $= 0$) whenever $\Phi_u(\mathbf{p}_i) = 1$ (respectively $= 0$) for all $i \in I$. Note the key role of epistemic independence in making the outcome decision uniquely determined by the beliefs over the marginal events $E_k$.

One can now ask, in complete analogy to the questions at the heart of sections 3 to 5, which aggregation rules $H$ are Pareto consistent for a given $u$, and, second, whether for given $u$ there exist aggregation rules with specified desirable properties at all. It is likely going to be substantially harder to solve these questions than before, since the class of component aggregation rules $H_k : [0,1]^I \to [0,1]$ is obviously much larger than the class of rules $F_k : \{0,1\}^I \to \{0,1\}$, and a counterpart to the Pareto Intersection Property seems unlikely to exist.

Pareto consistency becomes an issue already in the simplest of problems, for example in a Bayesian counterpart to the classical conjunction problem.

**Example 13** *Suppose two expected-value maximizing agents share the profits from a potential investment equally. The success of this investment depends on the joint realization of two independent events* $E_1$ *and* $E_2$*. The investment is successful if and only if both materialize; in this case, the investment recoups the initial outlays tenfold; in the alternative, it is completely wasted. Thus we have* $u = 1_S - \tau$*, where* $S = E_1 \cap E_2 = \{(1,1)\}$ *and* $\tau = \frac{1}{10}$*.*

*Consider now the following profile of probability judgments illustrated in table 1 below. Agent 1 believes that the first contingency will materialize with 90% probability, but the second only with 10% probability; the investment will therefore succeed with 9% probability, implying a negative expected return. Agent 2 likewise believes that the investment will succeed with 9% probability, but for different reasons. While she thinks that the second contingency will materialize with 90% probability, she gives only a 10 % chance to the first.*

*By contrast, aggregating the probability judgments for the two contingencies directly suggests a group probability of 50% for each in view of the symmetry of the individual 90% and 10%=100%-90% estimates. (A natural aggregation rule* $H_k$ *to deliver this besides the arithmetic mean rule, and one that is arguably more attractive, is to let the group* odds ratio *for event* $E_k$ *against* $E_k^c$ *be the geometric mean of the individual odds ratios.) This entails a 25% probability for the investment to succeed, hence a clear decision to invest.*

| | $p^1(E_1)$ | $p^2(E_2)$ | $p(E_1 \cap E_2)$ | Decision |
|---|---|---|---|---|
| Agent 1 | 0.9 | 0.1 | 0.09 | Don't Invest |
| Agent 2 | 0.1 | 0.9 | 0.09 | Don't Invest |
| Group {1,2} | 0.5 | 0.5 | 0.25 | Invest |

Table 3: A Bayesian Version of the Paretian Dilemma

The example illustrates that well-motivated aggregation rules $H$ can be Pareto inconsistent, just as the premise-wise majority rule was in the original Paretian Dilemma. Just as in that case, we do not claim that this is the only reasonable aggregation rule, nor that the Paretian Dilemma cannot be avoided by choice of a

different rule. Indeed, in this particular example Pareto consistency could be achieved for instance by letting the group probability be the *geometric* mean of individual probabilities, $H_k = H^{geo}$ for all $k$, where

$$H^{geo}\left(\left(p_i^k\right)\right) = \left(\prod_{i\in I} p_i^k\right)^{\frac{1}{n}}$$

Note that at the profile given in Table 4, this leads to group probabilities of 30% for each contingency, and thus 9% for the investment to succeed. [25]

But serependipity cannot always succeed in the case of decisions based on more complex events. This follows from the following Proposition.

**Proposition 14** *There exist events $S$ such that no anonymous separable aggregation rule $H$ is Pareto consistent with $\Phi_{(1_S-\tau)}$, for any $\tau \in (0,1)$.*

Proposition 14 is verified by constructing an example with 6 marginal events, setting $S = (E_1 \cap E_2) \cup (E_3 \cap E_4) \cup (E_5 \cap E_6)$, a probability threshold $\tau \in (0,1)$ and two values for probability estimates $\alpha, \beta$ with $\alpha > \beta$, and such that $\alpha^2 + 2\beta^2 > \tau > 2\alpha\beta + \beta^2$.

| Judge\Event | $E_1$ | $E_2$ | $E_3$ | $E_4$ | $E_5$ | $E_6$ | $S$ | Decision |
|---|---|---|---|---|---|---|---|---|
| I | $\alpha$ | $\alpha$ | $\beta$ | $\beta$ | $\beta$ | $\beta$ | $\alpha^2+2\beta^2$ | Yes |
| II | $\beta$ | $\beta$ | $\alpha$ | $\alpha$ | $\beta$ | $\beta$ | $\alpha^2+2\beta^2$ | Yes |
| III | $\beta$ | $\beta$ | $\beta$ | $\beta$ | $\alpha$ | $\alpha$ | $\alpha^2+2\beta^2$ | Yes |
| Panel | $p_I^1$ | $p_I^2$ | $p_I^3$ | $p_I^4$ | $p_I^5$ | $p_I^6$ | ? | ? |

and

[25] A potential criticism of this aggregation rule is its asymmetric treatment of the positive and negative realizations of the contingencies; for example, due to this asymmetry, this aggregation rule would fail to be Pareto consistent for decision problems of the form $\Phi_{(1_S-\tau)}$ if $S$ is a disjunction rather than conjunction of two independent events.

We note that with only two independent events and an *odd* number of agents, the latter problem can be overcome in turn by using instead the event-wise median of the individual probabilities. This follows from results of Peters et al. (1992) . On the other hand, the median is Pareto inconsistent if $E$ is the conjunction of more than two events, as can be seen by interpreting Example associated with Table 1 in the introduction probabilistically.

| Judge\Event | $E_1$ | $E_2$ | $E_3$ | $E_4$ | $E_5$ | $E_6$ | $S$ | Decision |
|---|---|---|---|---|---|---|---|---|
| I | $\alpha$ | $\beta$ | $\beta$ | $\alpha$ | $\beta$ | $\beta$ | $2\alpha\beta + \beta^2$ | No |
| II | $\beta$ | $\beta$ | $\alpha$ | $\beta$ | $\beta$ | $\alpha$ | $2\alpha\beta + \beta^2$ | No |
| III | $\beta$ | $\alpha$ | $\beta$ | $\beta$ | $\alpha$ | $\beta$ | $2\alpha\beta + \beta^2$ | No |
| Panel | $p_I^1$ | $p_I^2$ | $p_I^3$ | $p_I^4$ | $p_I^5$ | $p_I^6$ | ? | ? |

Table 4: Impossibility of a Paretian Rational in the Bayesian Setting.

By anonymity and separability of $H$, the group probabilities $p_k^1$ on each marginal event $E_k$ must be the same at both profiles, and therefore the group decision must be the same as well. Yet since the agents agree on a different outcome decision at the two profiles, Pareto consistency must be violated in one of them.

**Connections with the Bayesian Literature** While there does not seem to exist a direct counterpart to the above observations such as Proposition 14 in the literature, the existence of potential conflicts between Bayesian group rationality and the Pareto axiom is well-known, starting with the classic contribution of Hylland-Zeckhauser (1969). In all of these contributions, the conflict results from a simultaneous disagreement about probabilities and utilities, in contrast to Proposition 14, where the impossibility results from a disagreement about probabilities only.

There exists also more a directly related literature that is concerned with the purely epistemic aggregation of probability judgments only; see in particular the classic survey by Genest and Zidek (1985). In the discussion, two aggregation rules play a dominant rule, the "linear" and the "logarithmic" "opinion pools". In the linear opinion pool, the group probability of each event is the (possibly weighted) *arithmetic* average of individual probabilities; by contrast, in the logarithmic opinion pool, the group probability of each *state* is proportional to the (possibly weighted) *geometric* average of individual probabilities. While the linear opinion pool respects unanimous probability judgments for all events by construction, the logarithmic opinion pool does not.[26] On the other hand, the logarithmic opinion pool preserves stochastic indepen-

[26]In example 13, for instance, the logarithmic opinion pool yields the 25% estimate for the probability of the event $E_1 \cap E_2$ that has been suggested heuristically above.

dence of individual probability judgments, while the linear one does not. Neither rule has emerged as the dominant one.[27]

The main contribution of this section vis-a-vis this Bayesian literature is to show that potential conflicts between reason-based judgment aggregation and the Pareto criterion are not tied to a specific, demanding model of decision-theoretic rationality at the group level such as the Bayesian one, but comes with the notion of reason-based judgment aggregation as such. In particular, such conflicts do not depend on the aggregation of an entire coherent preference or likelihood ordering respectively probability measure, but arise already in the context of the simplest binary decision problems.

## 7. ON THE NORMATIVE STATUS OF THE PARETO AXIOM

The results of this paper have demonstrated the great robustness of the Dilemma of the Paretian Rational in propositional judgment aggregation and beyond. Even in the simplest cases, it is always present as a possibility, in that some prima facie sensible aggregation rules fail to be to consistent with the Pareto criterion. In many cases, such failures occur even for aggregation rules that are natural and well-motivated, such as premise-wise majority voting in the propositional case or the "geometric mean odds" rule in the Bayesian case illustrated in Example 13. Indeed, if the group decision problem is sufficiently complex, the Dilemma runs deep enough so that all separable aggregation rules are vulnerable to conflicts with the Pareto principle.

This raises the normative issue over which principle should give way, when necessary, or how they should be qualified or traded-off against each other. As its resolution turns out to be quite complex and touches on some controversial considerations such as, in the Bayesian case, the normativity of the common prior assumption, unavoid-

[27] Obviously, the linear opinion pool could have been used to "overcome" the Pareto consistency problem. It is not clear, however, whether such overcoming is necessary (cf. section 7) nor appropriate. In particular, if individual probability measures are aggregated linearly, the group probability of the decision-relevant event $S$ is no longer a function of the group probabilities $p_I^k$ of the marginal events $E_k$, and thus no longer grounded in them. This would appear to seriously compromise the reason-basedness of the group choice. A more complete treatment would have to discuss the pros and cons of the linear opinion pool from the present perspective in greater detail.

Two "parallel" axiomatizations of these aggregation rules based on separability considerations can be found in Fishburn-Rubinstein (1986); Paretian considerations play no role there.

ably the discussion will be somewhat rough and incomplete, and a fully satisfactory analysis is left to future work. The more modest goal of the present discussion is to establish the existence of a live conflict between the two principles. In particular, we will argue that the "Dilemma" does not show that the notion of reason-basedness is ill-founded from the start, and that, indeed, in some situations it is the Pareto principle that should give way.

To ensure that the following discussion rests on a sufficiently well-specified premises, we will assume that all agents' judgments are commonly known ("Complete Information"). Thus, all deliberation has already taken place, and there is nothing left to discuss; any remaining difference in agents' judgments reveals an "agreement to disagree" in the sense of Aumann (1976). On the one hand, the Complete Information assumption is in line with the absence of a formally described information structure from the model. On the other hand, under asymmetric information, unanimity of interim judgments would fail to be the normatively appropriate Pareto requirement, robbing the Paretian Dilemma of normative relevance for reasons that having nothing to do with the reason-basedness of social choice[28,29].

The assumption of Complete Information is not consistent with some interpretations of the judgment aggregation problem that can be found in the literature. In particular, it rules out an evaluation of aggregation rules in terms of their truth-tracking properties in the manner of the Condorcet Jury Theorem[30], for this literature attributes the difference in rational agents' judgments to differences in their information.

Similarly, the assumption of Complete Information also undercuts much of the motivation behind delegating the judgment on different premises to different experts as studied in section 5. For such delegation presumably hinges on the putative experts

[28]To see the issue more clearly in a voting context, suppose that the agents agree that a positive outcome decision should be taken if the probability of its "success" is sufficiently high ($= q^*$), and that this threshold probability exceeds the agents' common prior. If the agents' private signals are sufficiently weak, it may well be, that no agent would favor *on the basis of his own information* that the project be undertaken, even though all agents would agree to do so once their information is pooled .

[29]In a Bayesian setting, this has been argued compellingly by Holmstrom-Myerson (1983); they propose instead a notion of Interim Pareto Dominance, which requires not just unanimity but common knowledge of unanimity.

[30]For works along these lines in the judgment aggregation literature, see for example Bovens-Rabinowicz (2004) and List (2004a).

being "better informed" than the non-experts; but under Complete Information, such asymmetries of information are ruled out by definition.

### 7.1. Group Choice based on Shared Self-Interest

We will argue that the correct resolution of the conflict between Paretianism and Rationalism will depend on the nature of the group choice problem. Suppose first that the decision at hand is a matter of shared self-interest, as in the profit sharing example above (Example 13). In that example, both agents' interests are perfectly aligned; the two agents merely disagree in their assessment of the underlying uncertainty. Since under the stated assumptions both agents believe that the investment will make negative expected profits, it seems compelling to argue that the agents' joint decision should be to reject the investment project, in agreement with the Pareto criterion.

How solid is the case for the Pareto criterion here on reflection? In particular, could one reasonably argue that the group probability derived from premise-based aggregation, taken to be 25% here, is the "right" probability all and all from an impartial point of view, and that the agents should base their joint decision on it? Such an argument can be supported by appealing to the normativity of the common prior assumption, according to which in view of Aumann's (1976) celebrated result disagreement among agents entails the existence of at least one imperfectly rational agent.[31]

A crucial consideration here is the assumption of Complete Information. For if the agents "agree to disagree" with each other, they will "agree to disagree" with the reason-based group judgment as well (if that judgment differs from their own) since that judgment is a commonly known logical consequence of the individual judgments given the aggregation rule. Thus arguing that the 25% belief is the right belief, everything considered, amounts to saying that *each agent*, given his current information, should have beliefs different from those that he in fact has. This line of reasoning has the structure of a classical paternalistic argument, according to which an overriding of individual preferences is warranted because of rationality deficits in their formation.

Opinions will differ as to whether and when such paternalism is legitimate or even

[31] For a rejection of the Pareto criterion that appeals to the normativity of the common prior assumption, see Gilboa et al. (2004).

mandated. Among the arguments against paternalism and in favor of maintaining the Pareto principle, two deserve to be mentioned in particular; we shall refer to them as the "pluralist" and the "liberal" defenses of the Pareto principle, respectively. The pluralist defense is simply to reject the normativity of the common prior assumption, so that disagreement among agents is no evidence *per se* of agent's irrationality, and therefore does not motivate by itself any paternalistic overriding of the Pareto principle [32]. The liberal defense of the Pareto principle might accept the common prior assumption as normative, but would argue that in matters of self-interest agents should have a right to make a mistake. So, in view of these two defenses, the argument for the Pareto criterion remains strong, and we expect that most economists would subscribe to it, even though there will be some dissenters.

Assuming, then, that the Pareto principle is accepted in situations of (possibly shared) self-interest, what do the results of this paper imply for the notion reason-based group choice? One response would be to argue that "reasons" are simply irrelevant to group choice among self-interested agents, and that the group choice should be determined by agents' preferences over outcomes in line with received social choice theory via standard criteria of distributional justice, optimal voting or fair bargaining.

But this need not be the only possible response. For example, in discussions of deliberative democracy, it is often suggested that the ability to support collective choices by collectively affirmed reasons enhances their "legitimacy" (see, for example, Pettit (2001a)). Presumably, considerations of legitimacy would sometimes justify the choice of an outcome favored by a minority as the better supported one. In this manner, reason-basedness can function as a form of minority protection. However, there seems to be no point in overruling a *unanimous* "majority" on grounds of deficient legitimacy: no need for legitimation without contestation, one might say. Thus, our impossibility results imply that a satisfactory account of reason-based social choice as legitimizing requires a departure from the separable premise-aggregation assumed here. The maxmin rules introduced at the end of section 4 have some appeal as "legitimizing" aggregation rules; their separability failure may be – would have to be – a price worth paying.

[32] See Morris (1995), for example, for a forceful critique of the common prior assumption.

### 7.2. Group Choice based on Shared Responsibility

Are there situations in which the Pareto axiom lacks normative force even though the premises of these defenses are accepted? We will now try to show that there are indeed such situations, and the Pareto principle does not apply when the group decision is one of "shared responsibility" rather than "shared self-interest".

To illustrate the idea, suppose that the above profit-sharing example is modified so that now the two agents are trustees of a fund that has been endowed for the benefit of an under-age heir. In contrast to a profit-sharing example, the members of the group have now no (legitimate) *personal stake* in the decision. Their judgments merely serve as informational inputs from which the group decision is to be derived in an appropriate way; this implies in particular that the individuals' own preferences over outcomes carry no normative weight on their own. In particular, unanimity of individual "preferences" *qua preferences* does not normatively entail a corresponding group preference, which is to say that the Pareto principle is normatively defeasible *in principle*.

Yet it would be rash to conclude from this argument alone that the Pareto principle can in fact be legitimately violated, for it might be supported indirectly through what may be called the "Unanimity Principle". By "Unanimity Principle", we mean the requirement that the group judgement on any particular proposition or event must agree with the individual judgments on this proposition or event whenever these are unanimous. While the Pareto Principle is concerned with the socially rational choice of *outcomes*, the Unanimity Principle is concerned with the socially rational aggregation of judgments *qua judgments*, irrespective of their outcome implications. In the introductory adjudication example, for instance, the Unanimity principle asserts that a unanimous negation of the conjunctive proposition ("had duty *and* was negligent *and* conduct was causal") obliges the panel to render the same judgment on this proposition; by the assumed agreed-upon content of legal doctrine this judgment "happens to" entail a negative outcome decision on the damages; in this way, the Unanimity Principle may lend indirect support for the Pareto Principle. Similarly, in the Bayesian investment example, the Unanimity principle would imply that the group must assign a 9% probability to the conjoint event $E_1 \cap E_2$, which, as a result of the shared utility-function over final outcomes, would entail a rejection of the

investment project.

We now argue that the Unanimity Principle should *not* be considered a general principle of the rational judgment aggregation. In particular, we submit that the Unanimity Principle does not hold in cases in which the premise judgments are epistemically independent and prior to the conclusion. (In the case of propositional judgment aggregation, epistemic independence has already been briefly discussed above; in the Bayesian case, epistemic independence can be identified with stochastic independence.)[33]

If premises are epistemically independent and thus prior to the conclusion, *all* relevant information about the outcome decision is contained in the agents' premise judgments; their outcome judgements make thus no independent contribution, and should arguably not carry any independent weight. Hence if the group aggregation rule entails an outcome judgment that contradicts the unanimous outcome judgment of the agents, this contradiction by itself is no reason to doubt the wisdom of the aggregation rule. Indeed, under epistemic independence of premises it is easy to understand how a group aggregation rule can *rightly* override a unanimous outcome judgment: for while the latter depends crucially on how the judgments on different premises are "correlated" across agents, such correlation arguably should not matter given the assumed epistemic independence.[34]

It is worth emphasizing that while the argument against in the Unanimity Principle is particularly strong under epistemic independence of the premises, its validity is not confined to this case. Indeed, the Unanimity Principle seems to be open to justifiable violations quite generally as long as there is *some* evidential spillover from other,

[33] Arguably, the former entails the latter: if premises are epistemically independent, they must be viewed as *epistemically prior* to the entailed outcome judgments. For instance, arguing by *modus tollens,* if in the Bayesian investment example some agent had evidence for the likelihood of the conjoint event $E_1 \cap E_2$ that is not derived from direct evidence on the underlying events $E_1$ and $E_2$ (for example based on the track record of similar earlier investment decisions), this should be reflected in a negative subjective correlation between the marginal events $E_1$ and $E_2$ in his probability judgments, in contradiction to the assumed subjective independence between them.

[34] This argument seems cogent if agents are assumed to be fully rational. If not, there may be room for learning about the likely rationality departures of an agent's judgement on some premise from his judgement pattern on other premises compared to that of the other agents. Since the point of our argument here is merely to claim that the Unanimity Principle is *sometimes* defeasible, we leave this stone unturned.

related judgments.

To fully understand the difference between situations of shared interests and situations of shared responsibility, it is instructive to compare the role of the pluralist and liberal defenses of the Pareto criterion in both contexts. Since both defenses pertain to the propositional or probabilistic judgments underlying agents' preferences, they can be applied as meaningfully to the Unanimity principle as to the Pareto principle. However, in situations of shared responsibility, they lose plausibility or force. The breakdown of the liberal defense is straightforward, in that it is part and parcel of the very notion of responsibility that agents are not simply free to make mistakes in their exercise of such responsibility.

The breakdown of the pluralist defense is more subtle. In contrast to situations of shared interest, in situations of shared responsibility the normative underwriting of a group judgment that differs from the unanimous judgement of individual agents does not mean that the agents' are paternalized in the sense that their belief is *replaced* by a normatively superior one. Instead, the aggregate belief should now be thought of as belonging to "the group", as distinct from the individual agents; it is normatively valid as an "optimal" summary of the information contained in individual judgments for the purposes of coming to the best-justified collective decision. The group has to be thought of as epistemically in a position *different* from any individual (neither superior nor inferior) in virtue of its impartial (agent-neutral) vantage point; claiming that a particular belief is best for the group does therefore not entail that it must be best for any agent individually,.

**A Paradox of Responsible Choice ?.—**

The notion of a distinct group belief suggests the following normative implication. Suppose that only one of the two agents is tasked with the investment decision as the *single* trustee while the other is merely an outside onlooker. Then arguably the single trustee agent would rightly discharge his responsibility to the heir by basing the decision on *his own* best estimate of the success of the project, in this case with the consequence of not undertaking the investment decision.[35]

[35]Note that if one was to argue that the individual would be obliged to take an impartial viewpoint towards his own beliefs in discharging his responsibility, this viewpoint would hardly coincide with

There is an element of paradox here, in that the responsible decision for a group may differ from what responsibility would require of each member acting individually. Note that it is not possible to resolve this difference by claiming that the group has different information from the individuals; after all, everything is assumed to be commonly known here. Instead, it appears that a resolution of this apparent paradox must involve the notion that being held (or holding oneself) responsible as a group engenders a group agency that cannot be reduced directly to the agency of the underlying individuals – which is very much in line with the rationale for introducing the notion of a group belief in the first place [36].

## 8. CONCLUSION: SOME DIRECTIONS FOR FUTURE RESEARCH

The results of this paper have demonstrated the existence of a robust tension between reason-based choice and the Pareto principle. In the specific propositional framework assumed in the bulk of the paper, this tension becomes in fact a stark conflict in that under all but the simplest decision functions, this conflict is unavoidable. Future research should explore the robustness of our findings under different assumptions, especially different assumptions on the structure of the judgments to be aggregated. Our preliminary observations in section 6, however, suggest already that the overall bottom line is likely to generalize, and that at best only rather special premise aggregation rules will agree with the Pareto principle.

We have also probed into the normative implications of this conflict, and identified circumstances that render the Pareto principle either normative compelling or defeasible, characterizing them in terms of "shared self-interest" versus "shared responsibility". In contexts of shared self-interest, the articulation of reasons for a collective choice may still play a role, for example in legitimizing group decisions. However, these reasons will need to relate to the group decision in fairly restricted ways so as not to fall prey to the Paretian Dilemma. We have given here just one

that of the group, since the group will have onlookers to. The viewpoint would have to be that of "all" rational believers, but who are these "all"?

[36] The connection between reason-based social choice and non-reducible group agency is central to the work of Pettit; see in particular Pettit (2001b). While Pettit does not appeal to the notion of "shared responsibility" as it is used here, there is probably a non-trivial degree of overlap that remains to be clarified.

example (the minmax rules) of how this might be done. Almost everything remains to be done.

There also clearly needs to be more work trying to establish their proper underlying conceptual structure and valid reach of the notion of shared responsibility. For example, what kind of entities can the group be meaningfully responsible to? Do these need to consist of one ore more human beings, or could it be entities such as "our organization", "the state" or even "the biosphere"? How should one conceptualize mixed situations in which agents share a common responsibility, but self-interest plays a legitimate role as well? These and many other interesting questions remain to be answered.

## APPENDIX

### A1. An Irrelevance Result for Multi-valued Logics

A *multi-valued aggregation rule* $G$ maps profiles of judgments $(J_i)$ to multi-valued truth assignments $t = (t_k)$; it is separable if $G = (G_k)$, where $G_k : \{0,1\}^I \to [0,1]$ is monotone and respects unanimity ($G_k(\varnothing) = 0, G_k(I) = 1$). A *decision function* $\Phi_\tau : [0,1]^K \to \{0,1\}$ makes the social choice dependent on the continuous truth value of the proposition $\Phi$, which in turn is determined from the truth values of the premises on the basis of an appropriate multi-valued logic. Pareto consistency is defined as before via the induced (two-valued) social choice function $\Phi_\tau \circ G : \{0,1\}^{K \times I} \to \{0,1\}$.

According to standard axiomatizations of multi-valued logic, the truth-value of a conjunction of propositions is equal to the minimum of the truth values of its constituents, and the truth value of a disjunction equal to the maximum of the truth-values of its constituents; this holds, for example, for Łukasiewicz's logic adopted in van Hees (2004) and for standard versions of fuzzy logic due to Zadeh (1965). Thus, in view of the canonical disjunctive representation of monotone propositions, the truth value of the proposition $\Phi = \bigvee_{m \in M^+} \left( \bigwedge_{k \in J_m^+} a_k \right)$ at the truth assignment $t = (t_k)$ is given by

$$\max_{m \in M^+} \min_{k \in J_m^+} t_k; \tag{6}$$

hence decision functions based on this logic take the form

$$\Phi_\tau(t) = 1 \text{ iff } \max_{m \in M^+} \min_{k \in J_m^+} t_k > \tau,$$

where $\tau \in (0,1)$ is an appropriate "truth threshhold".

While multi-valued group judgments capture the disagreement among agents in a direct and natural way, they do not expand the set of reason-based social choice functions, and therefore do nothing to overcome Pareto inconsistencies.

**Proposition 15** *Let $G = (G_k)$ be any separable multi-valued aggregation rule, $\Phi$ any monotone proposition and $\tau \in (0,1)$. Then there exists a separable aggregation rule*

$F = (F_k)$ *such that*

$$\Phi_\tau \circ G = \Phi \circ F.$$

This aggregation rule $F = (F_k)$ is simply given by by accepting premise $k$ at those profiles at which the truth-value of that premise under $G$ exceeds the critical threshold $\tau$; formally, set $F_k(W) = 1$ if and only if $G_k(W) > \tau$ for $W \in 2^I$ and $k \leq K$. For example, if the degree of truth $t_k$ at a profile is given by the fraction of agents affirming premise $k$, than $F_k$ is a quota-rule with quota $\tau$.

Proposition 15 is formally proved as follows. For any profile $(J_i)$, $\Phi_\tau \circ G((J_i)) = 1$ if and only if $\max_{m \in M^+} \min_{k \in J_m^+} G_k((J_i)) > \tau$, i.e. iff there exists $m \in M^+$ such that, for all $k \in J_m^+$, $G_k((J_i)) > \tau$, i.e. iff there exists $m \in M^+$ such that, for all $k \in J_m^+$, $F_k((J_i)) = k$. But this is equivalent to saying that $F((J_i)) \in \mathcal{J}^+$, that is $\Phi \circ F((J_i)) = 1$.

## A.2 Proofs

The following Fact establishes that the families of winning coalitions $\mathcal{W}_k$ and $\mathcal{W}_k^0$ are dual to each other; it will be referenced repeatedly in the following proofs; see NP for the straightforward verification.

**Fact 16** $\mathcal{W}_k^0 = \{W \in 2^I : \textit{for all } W' \in \mathcal{W}_k : W \cap W' \neq \varnothing\}$;
$\mathcal{W}_k = \{W \in 2^I : \textit{for all } W' \in \mathcal{W}_k^0 : W \cap W' \neq \varnothing\}$.

**Proof of Proposition 2.**

$F = (\mathcal{W}_k)_{k \in K}$ is Pareto consistent iff

i) for all profiles $(J_i)$ such that $\{J_i\} \subseteq \mathcal{J}^-$, $F((J_i)) \in \mathcal{J}^-$, and likewise

ii) for all profiles $(J_i)$ such that $\{J_i\} \subseteq \mathcal{J}^+$, $F((J_i)) \in \mathcal{J}^+$.

Thus i) and ii) define the consistency of the aggregation rules $(\mathcal{W}_k)_{k \in K}$ on the restricted domains $\mathcal{J}^-$ resp. $\mathcal{J}^+$. These have been characterized in NP, Theorem 3.

For $k \in K$, let $H_k := \{J \in \mathcal{J}^- : J \ni k\}$, and let $\mathcal{H}^- := \{H_k\}_{k \in K} \cup \{H_k^c\}_{k \in K}$.

A complement-free family $\mathcal{G} \subseteq \mathcal{H}^-$ is a (non-trivial) critical family if $\cap \mathcal{G} = \varnothing$ and $\cap_{G \in \mathcal{G}} (\mathcal{G} \backslash G) \neq \varnothing$. It is easily verified that the critical families are exactly the families of the form $\{\{H_k\}_{k \in J_m^+}\}$, where $J_m^+ \in \mathcal{J}^+$. Thus, by NP, Theorem 3, $F_{|(\mathcal{J}^-)^I}$ is consistent iff part i) of the Proposition is satisfied. Likewise, $F_{|(\mathcal{J}^+)^I}$ is consistent iff part ii) of the Proposition is satisfied. ■

**Proof of Theorem 5.**

As in the proof of Proposition 2, $F = (\mathcal{W}_k)_{k \in K}$ is Pareto consistent iff

i) for all profiles $(J_i)$ such that $\{J_i\} \subseteq \mathcal{J}^-$, $F((J_i)) \in \mathcal{J}^-$, and likewise

ii) for all profiles $(J_i)$ such that $\{J_i\} \subseteq \mathcal{J}^+$, $F((J_i)) \in \mathcal{J}^+$.

By Theorem 4 of NP, these two conditions hold for some (or for any) neutral and non-dictatorial separable aggregation rule if and only if the associated critical families have cardinality 2, i.e. in view of the argument in the proof of Proposition 2, if and only if $\kappa_\Phi = 2$. This establishes the equivalence of 1), 2) and 3).

As to the equivalence between 3) and 4), one verifies the implication 4) $\Rightarrow$ 3) by inspecting the canonical representation of $\neg\Phi$ in each case. We will show by

contradiction that $\kappa_\Phi = 2$ implies $n \leq 5$. The remaining claim follows "by hand" from similarly straightforward but tedious arguments that are omitted.

Suppose, thus, that $\kappa_\Phi = 2$ and that $n \geq 5$.

**Lemma 17** *$J \in \mathcal{J}^+$ iff for all $J' \in \mathcal{J}^-$ (or $J' \in \max \mathcal{J}^-$), $J \cap (J')^c \neq \varnothing$.*
*$J \in \mathcal{J}^-$ iff for all $J' \in \mathcal{J}^+$ (or $J' \in \min \mathcal{J}^+$), $J' \cap J^c \neq \varnothing$.*

**Verification.** $J \in \mathcal{J}^+$ iff not $J \subseteq J'$ for some $J' \in \max \mathcal{J}^-$ iff, for all $J' \in \max \mathcal{J}^-$, not $J \subseteq J'$, i.e. iff for all $J' \in \max \mathcal{J}^-$, $J \cap (J')^c \neq \varnothing$. The second part is completely analogous. □

**Lemma 18** *For all $j$, there exists $J \in \min \mathcal{J}^+$ such that $j \in J$ and there exists $J' \in \max \mathcal{J}^-$ such that $j \in (J')^c$.*

**Verification.** Suppose, contrary to the first part of the claim, that for all $J \in \min \mathcal{J}^+$, $j \notin J$. Then, by Lemma 17, for any $S$, $S \in \mathcal{J}^-$ iff $S \backslash j \in \mathcal{J}^-$, i.e. $j$ is inessential. Again, the second part is completely analogous. □

Since $n \geq 5$, in view of Lemma 18, $\min \mathcal{J}^+$ must contain at least two disjoint sets $J, J'$. For if this was not the case, clearly $\min \mathcal{J}^+$ must be of the form $\{\{a_\ell, a_k\}\}_{k \in K \backslash \ell}$, in which case $\max \mathcal{J}^- = \{\{a_\ell\}, \{a_k\}_{k \in K \backslash \ell}\}$, which implies $\kappa_\Phi \geq 3$.

Thus $\min \mathcal{J}^+$ contains three sets of the form

$$\{a_1, a_2\}, \{a_3, a_4\}, \{a_5, a_\ell\}, \tag{7}$$

where w.l.o.g. $\ell = 1$ or $\ell = 6$. Evidently $\{a_1\} \in \mathcal{J}^-$, since $\{a_1, a_2\} \in \min \mathcal{J}^+$.

Let $J$ denote any element of $\max \mathcal{J}^-$ containing $\{a_1\}$. By Lemma 17, $J^c$ must intersect all elements of $\min \mathcal{J}^+$, which implies

$$J^c \ni a_2, \ J^c \cap \{a_3, a_4\} \neq \varnothing \text{ and } J^c \cap \{a_5, a_\ell\} \neq \varnothing.$$

Hence $\# J^c \geq 3$, contradicting the assumption that $\kappa_\Phi = 2$. It follows that in fact $n \leq 4$ ■

**Proof of Fact 6.**

Consider a decomposable $\Phi$ with $a_k$ entailing $\Phi$. Then evidently $\{k\} \in \min \mathcal{J}^+$, and therefore, for all other $J$ in $\min \mathcal{J}^+$, $k \notin J$. Thus no other conjunct in the canonical representation of $\Phi$ contains $a_k$, which means that $\Phi(a_1, ..., a_n) = \Phi'(a_1, .., a_{k-1}, a_{k+1}, .., a_n) \vee a_k$.

The dual case of $\Phi$ entailing $a_k$ is analogous. $\square$

**Proof of Theorem 7.**

**Part i).** Step 1: Pareto consistency implies Neutrality

Suppose that $F$ given by $(\mathcal{W}_k)_{k \in K}$ is Pareto consistent.

Let $\trianglerighteq$ denote a linear ordering on $2^{2^I}$ that is monotone with respect to set inclusion, i.e. such that $\mathcal{W} \supseteq \mathcal{W}'$ implies $\mathcal{W} \trianglerighteq \mathcal{W}'$, with asymmetric component $\triangleright$, and partition the set of atomic propositions $K$ into three subsets $K^=, K^>$, and $K^<$ as follows:

$$K^= = \{k : \mathcal{W}_k = \mathcal{W}_k^0\}, K^> = \{k : \mathcal{W}_k \triangleright \mathcal{W}_k^0\}, K^< = \{k : \mathcal{W}_k \triangleleft \mathcal{W}_k^0\}.$$

We need the following Lemmata.

**Lemma 19** *For any two committees $\mathcal{W} \triangleright \mathcal{W}^0$, and $\mathcal{W}' \trianglerighteq (\mathcal{W}')^0$, there exist $W \in \mathcal{W}$ and $W' \in \mathcal{W}'$ such that $W \cap W' = \varnothing$.*

**Proof.** Suppose not. I.e., for all $W \in \mathcal{W}$ and $W' \in \mathcal{W}' : W \cap W' \neq \varnothing$.

Then by Fact 16,

$$\mathcal{W}^0 \supseteq \mathcal{W}' \text{ and } (\mathcal{W}')^0 \supseteq \mathcal{W}.$$

By monotonicity of $\trianglerighteq$ and the assumptions on $\mathcal{W}, \mathcal{W}'$ therefore

$$\mathcal{W} \triangleright \mathcal{W}^0 \trianglerighteq \mathcal{W}' \trianglerighteq (\mathcal{W}')^0 \trianglerighteq \mathcal{W},$$

a contradiction. $\square$

**Lemma 20** *For any $j, k$ and $J$ such that $\{j, k\} \subseteq J \in \min \mathcal{J}^+$ or $\{j, k\} \subseteq J^c$ and $J \in \min \mathcal{J}^-$, $j \in K^>$ if and only if $k \in K^<$.*

This follows immediately from Lemma 19 and the PIP. □

To complete the proof of Step 1, in view of Proposition 12 of section 5 (demonstrated below), it suffices to show that $K^> \cup K^< = \varnothing$.

First, we note that $K^> \in \mathcal{J}^-$. Indeed, suppose that instead $K^> \in \mathcal{J}^+$. Then there exists $J \subseteq K^>$ such that $J \in \min \mathcal{J}^+$. But since $\#J \geq 2$ by indecomposability, this contradicts Lemma 20.

Second, we verify that indeed $K^> \cup K^< \in \mathcal{J}^-$. Let $L$ denote any superset of $K^>$ in $\max \mathcal{J}^-$. Since by construction $L^c \subseteq K^= \cup K^<$ and $\#L^c \geq 2$ by indecomposability, $L^c \subseteq K^=$ by Lemma 20, i.e. $L \supseteq K^> \cup K^<$, from which the claim follows immediately.

Finally, we show that $K^> \cup K^< = \varnothing$. If this was not the case, then by indecomposability and Lemma 20, $K^> \neq \varnothing$. Fix any $k \in K^>$. By Lemma 18, there exists $J \in \min \mathcal{J}^+$ containing $k$ with $\#J \geq 2$. Since $J \nsubseteq K^> \cup K^<$ by the previous claim, $J \cap K^= \neq \varnothing$ , in contradiction to Lemma 20.

Step 2: If $F$ is efficient and neutral, $\kappa_\Phi \leq 2$.

Step 3: If neutral and $\kappa_\Phi \leq 2$, then $F$ is efficient.

The last two steps follow immediately from Theorem 5. □

**Part ii).**

Define quotas $(q_k)$ as follows: If in the canonical decomposition $a_k$ appears in a disjunction, let $q_k = \frac{1}{2^k}$. If, on the other hand, $a_k$ appears in a conjunction, let $q_k = 1 - \frac{1}{(2+\epsilon)^k}$, where $\epsilon$ is any irrational strictly positive number.

We claim by induction on $K$ that

i) for any $J \in \min \mathcal{J}^+$, $\sum_{k \in J} q_k \geq \#J - 1 + \frac{1}{(2+\epsilon)^k}$,

ii) for any $J \in \max \mathcal{J}^-$, $\sum_{k \in J^c} q_k \leq 1 - \frac{1}{(2+\epsilon)^k}$.

By Fact 3, this implies that $(\mathcal{W}_{q_k})$ is Pareto consistent.

Suppose that our claim holds for $K$. Then we will show from the recursive characterization of critical sets given by the following Lemma 21 that the claim holds in fact for $K+1$.

**Lemma 21** *a) Let* $\Phi'(a_1, ..., a_{K+1}) = \Phi(a_1, ..., a_K) \vee a_{K+1}$ *with associated sets* $\mathcal{J}^{-\prime}$ *and* $\mathcal{J}^{+\prime}$. *Then*

$$\begin{aligned} \max \mathcal{J}^{-\prime} &= \max \mathcal{J}^{-}, \text{ and} \\ \min \mathcal{J}^{+\prime} &= \left(\min \mathcal{J}^{+}\right) \cup \{a_{K+1}\}. \end{aligned}$$

*b) Let* $\Phi'(a_1, ..., a_{K+1}) = \Phi(a_1, ..., a_K) \wedge a_{K+1}$ *with associated sets* $\mathcal{J}^{-\prime}$ *and* $\mathcal{J}^{+\prime}$. *Then*

$$\begin{aligned} \max \mathcal{J}^{-\prime} &= \{J \cup \{a_{K+1}\} : J \in \max \mathcal{J}^{-}\} \cup \{a_1, ..., a_K\}, \text{ and} \\ \min \mathcal{J}^{+\prime} &= \{J \cup \{a_{K+1}\} : J \in \min \mathcal{J}^{+}\}. \end{aligned}$$

The proof of the Lemma is straightforward, hence omitted.

To verify the inductive argument, consider the case in which $\Phi'(a_1, ..., a_{K+1}) = \Phi(a_1, ..., a_K) \wedge a_{K+1}$; the alternative case of $\Phi'(a_1, ..., a_{K+1}) = \Phi(a_1, ..., a_K) \vee a_{K+1}$ is verified analogously. Consider $J \in \max \mathcal{J}^{-\prime}$. If $J = \{a_1, ..., a_K\}$, $J^c = \{a_{K+1}\}$, and thus

$$\sum_{k \in J^c} q_k = q_{K+1} = 1 - \frac{1}{(2+\epsilon)^{K+1}}.$$

If $J \neq \{a_1, ..., a_K\}$, then $J = J' \cup \{a_{K+1}\}$ for some $J' \in \max \mathcal{J}^{-}$, hence by induction

$$\sum_{k \in J^c} q_k = \sum_{k \in (J')^c} q_k \leq 1 - \frac{1}{(2+\epsilon)^{K}} \leq 1 - \frac{1}{(2+\epsilon)^{K+1}}.$$

On the other hand, consider $J \in \min \mathcal{J}^{+\prime}$. Then $J = J' \cup \{a_{K+1}\}$ for some $J' \in \min \mathcal{J}^{+\prime}$, hence by induction

$$\sum_{k \in J} q_k = \sum_{k \in J'} q_k + q_{K+1} \geq \left(\#J - 2 + \frac{1}{(2+\epsilon)^{K}}\right) + \left(1 - \frac{1}{(2+\epsilon)^{K+1}}\right) \geq \#J - 1 + \frac{1}{(2+\epsilon)^{K+1}}.$$

**Part iii).**

Define quotas $(q_k)$ as follows: If $a_k$ is a core-premise (i.e. if $k \leq k_m - 1$), set $q_k = \frac{1}{2}$. If $a_k$ is a non-core premise and appears in the canonical decomposition in a disjunction, set $q_k = 0$. If, on the other hand, $a_k$ appears in a conjunction, set $q_k = 1$.

We claim by induction on the number of non-core premises $K - k_m + 1$ that

i) for any $J \in \min \mathcal{J}^+$, $\sum_{k \in J} q_k \geq \#J - 1$,

ii) for any $J \in \max \mathcal{J}^-$, $\sum_{k \in J^c} q_k \leq 1$.

By Fact 3, this implies that $(\mathcal{W}_{q_k})$ is Pareto consistent. The remaining proof parallels that of Part ii) exactly.

By Theorem 5, the claim holds if $K - k_m + 1 = 0$. Suppose that our claim holds for $K - k_m + 1 = L$. We need to show that it holds in fact for $L + 1$ as well. Consider the case in which $\Phi'(a_1, ..., a_{L+1}) = \Phi(a_1, ..., a_L) \wedge a_{L+1}$; the alternative case of $\Phi'(a_1, ..., a_{L+1}) = \Phi(a_1, ..., a_L) \vee a_{L+1}$ is verified analogously. Consider $J \in \max \mathcal{J}^{-\prime}$. If $J = \{a_1, ..., a_L\}$,

$$\sum_{k \in J^c} q_k = q_{L+1} = 1.$$

If $J \neq \{a_1, ..., a_L\}$, then $J = J' \cup \{a_{L+1}\}$ for some $J' \in \max \mathcal{J}^-$, hence by induction

$$\sum_{k \in J^c} q_k = \sum_{k \in (J')^c} q_k \leq 1.$$

On the other hand, consider $J \in \min \mathcal{J}^{+\prime}$. Then $J = J' \cup \{a_{L+1}\}$ for some $J' \in \min \mathcal{J}^{+\prime}$, hence by induction

$$\sum_{k \in J} q_k = \sum_{k \in J'} q_k + q_{L+1} \geq (\#J - 2) + 1 = \#J - 1.$$

This proves the existence of anonymous separable aggregation rules that are Pareto consistent.

Conversely, take any Pareto consistent separable aggregation rule $F = (\mathcal{W}_k)$. By Lemma 21, there exists a critical set $J \in \min \mathcal{J}^+$ such that $J$ contains at least two core and at least one non-core premises, or there exists a critical set $J \in \max \mathcal{J}^-$ such that $J^c$ contains at least two core and at least one non-core premises. Assume w.l.o.g. the former, with $J \supseteq \{j_1, j_2, \ell\}$, where $j_1$ and $j_2$ are core premises and $\ell$ is a non-core premise.

By relativizing the critical sets to core premises and applying the argument behind Part i), $(\mathcal{W}_k)$ must be neutral on core premises. Hence in particular

$$\mathcal{W}_{j_1} = \mathcal{W}^0_{j_1} = \mathcal{W}_{j_2}. \tag{8}$$

Fix any $W$ and $i \in W$ such that $W \in \mathcal{W}_{j_1}$ while $W \backslash i \notin \mathcal{W}_{j_1}$, hence $W^c \cup \{i\} \in \mathcal{W}^0_{j_1}$. By (8), $W^c \cup \{i\} \in \mathcal{W}_{j_2}$. By the PIP therefore, for all $W' \in \mathcal{W}_\ell$,

$$W \cap (W^c \cup \{i\}) \cap W' = \{i\} \cap W' \neq \varnothing.$$

By Fact 16 this implies that $\{i\} \in \mathcal{W}^0_\ell$, which means that $i$ has veto power against $a_\ell$.

We note that further reasoning along these lines shows that there is a unique anonymous separable aggregation rule that ensures Pareto consistency. This rule consists of majority voting on core premises combined with unanimity voting on non-core premises or their negations.

**Part iv).**

By a construction analogous to that of Part iii) in which majority voting on core premises is replaced by dictatorship on core premises, one demonstrates the existence of non-dictatorial Pareto consistent separable aggregation rules

Conversely, it is immediate from the neutrality on core premises established in Part iii) that any Pareto consistent separable aggregation rule must be dictatorial on core premises, hence locally dictatorial. ■

**Proof of Theorem 10.**

We need the following lemmas.

**Lemma 22** *For any monotone $\Phi$ : For all $j, k \in K$, at least one of the following statements holds:*

1. *there exists $J \in \min \mathcal{J}^+$ such that $\{j, k\} \subseteq J$;*
2. *there exists $J \in \max \mathcal{J}^-$ such that $\{j, k\} \subseteq J^c$;*

3. *there exists* $J \in \min \mathcal{J}^+$, $J' \in \max \mathcal{J}^-$ *and* $\ell \notin \{j,k\}$ *such that* $\{j,\ell\} \subseteq J$ *and such that* $\{k,\ell\} \subseteq J^c$.

**Proof.** Fix $j,k \in K$. By Lemma 18, there exist $J \in \min \mathcal{J}^+$ and $J' \in \max \mathcal{J}^-$ such that $j \in J$ and $k \notin J'$. By Lemma 17, $J \cap (J')^c \neq \varnothing$. Let $\ell \in J \cap (J')^c$. If neither (1) nor (2) is satisfied, then $\ell \notin \{j,k\}$, whence (3) is satisfied. □

**Lemma 23** *Let $F$ be Pareto consistent and separable w.r.t. the monotone proposition $\Phi$. Then, for all $j,k \in K$, $\mathcal{W}_j \supseteq \mathcal{W}_k$, $\mathcal{W}_j \supseteq \mathcal{W}_k^0$, $\mathcal{W}_j^0 \supseteq \mathcal{W}_k$, or $\mathcal{W}_j^0 \supseteq \mathcal{W}_k^0$.*

Take any $j,k \in K$, and consider in turn the three cases described by Lemma 22. Suppose first that there exists $J \in \min \mathcal{J}^+$ such that $\{j,k\} \subseteq J$. In this case, by the PIP, for all $W \in \mathcal{W}_j$ and $W' \in \mathcal{W}_k$, $W \cap W' \neq \varnothing$. By Fact 16, therefore $\mathcal{W}_j^0 \supseteq \mathcal{W}_k$.

Similarly, if there exists $J \in \max \mathcal{J}^-$ such that $\{j,k\} \subseteq J^c$, one infers that $\mathcal{W}_j \supseteq \mathcal{W}_k^0$.

Finally, suppose that there exists $J \in \min \mathcal{J}^+$, $J' \in \max \mathcal{J}^-$ and $\ell \notin \{j,k\}$ such that $\{j,\ell\} \subseteq J$ and such that $\{k,\ell\} \subseteq J^c$. By what we have just shown for the first case, $\mathcal{W}_j^0 \supseteq \mathcal{W}_\ell$. Likewise, by what we have just shown for the second case, $\mathcal{W}_\ell \supseteq \mathcal{W}_k^0$. Hence by transitivity of set inclusion, $\mathcal{W}_j^0 \supseteq \mathcal{W}_k^0$. □

The Theorem is now an immediate result of Lemma 23 and the definition of $\mathcal{D}_j$. ■

**Proof of Proposition 12.**

Note that premise-wise neutrality implies that, for all $k, \mathcal{D}_k = \mathcal{W}_k = \mathcal{W}_k^0$. Hence, by Fact 16, Theorem 10 implies that, for all $j,k \in K$, $\mathcal{W}_j^0 \supseteq \mathcal{W}_k$ and $\mathcal{W}_k^0 \supseteq \mathcal{W}_j$, and thus by premise-wise neutrality again $\mathcal{W}_j = \mathcal{W}_k$. □

## REFERENCES


[1] Aumann, R. (1976), Agreeing to Disagree, *Annals of Statistics,* 4, 1236-1239.

[2] Bovens, L. and W. Rabinowicz (2004), "Voting Procedures for Complex Collective Decisions. An Epistemic Perspective," *Ratio Juris* 17, 241-258.

[3] Chapman B. (1998) "More Easily Done than Said: Rules, Reason and Rational Social Choice," *Oxford Journal of Legal Studies* 18, pp. 293-329.

[4] Cohen, J. (1986). "An Epistemic Conception of Democracy." *Ethics* 97: 26-38.

[5] Cohen, J. (1989). Deliberation and Democratic Legitimacy. *The Good Polity.* A. Hamlin and P. Pettit (eds), 17-34.

[6] Coleman, J. and J. Ferejohn (1986). "Democracy and Social Choice." *Ethics* 97: 6-25.

[7] Cowell, R., P. Dawid, S. Lauritzen, and D. Spiegelhalter (1999). *Probabilistic Networks and Expert Systems,* Springer, Heidelberg, New York.

[8] Dietrich, F. (2004a), "Judgement Aggregation: (Im)Possibility Theorems," forthcoming in *Journal of Economic Theory.*

[9] Dietrich, F. (2004b), "Judgement Aggregation in General Logics," mimeographed.

[10] Dietrich, F., and C. List (2004): "A Liberal Paradox for Judgment Aggregation," mimeo.

[11] Dokow, E. and R. Holzman (2005): "Aggregation of Binary Evaluations," mimeo.

[12] Genest, C. and J. Zidek (1986): "Combining Probability Distributions: A Critique and an Annotated Bibliography," *Statistical Science* 1, 114-148.

[13] Gilboa, I. , Samet, D. and D. Schmeidler (2004), "Utilitarian Aggregation of Beliefs and Tastes," *Journal of Political Economy*, 112, 932-938.

[14] Gutmann, A. and D. Thompson (1996). *Democracy and Disagreement.* Cambridge, Mass, Harvard University Press.

[15] Habermas, J. (1984, 1989). *A Theory of Communicative Action*, Vols 1 and 2, Cambridge, Polity Press.

[16] Holmstrom, B. and R. Myerson (1983), Efficient and Durable Decision Rules with Incomplete Information, *Econometrica,* 51, 1799-1819.

[17] Hylland, A. and Zeckhauser, R. (1979), "The Impossibility of Bayesian Group Decision Making with Separate Aggregation of Beliefs and Values," *Econometrica* 47, 1321-1336.

[18] Kornhauser, L. A., and L. G. Sager (1986) "Unpacking the Court," *Yale Law Journal* 82,

[19] List, C. and Pettit, P. (2002), "Aggregating Sets of Judgements: An Impossibility Result," *Economics and Philosophy* 18, 89-110.

[20] List, C. (2004a), "The Probability of Inconsistencies in Complex Collective Decisions," forthcoming in *Social Choice and Welfare.*

[21] List, C. (2004b), "The Discursive Dilemma and Public Reason," mimeo.

[22] Morris, S. (1995), "The Common Prior Assumption in Economic Theory," *Economics and Philosophy,* 11, 227-253.

[23] Mongin, P. (2005), "Factoring Out the Impossibility of Logical Aggregation," mimeo.

[24] Nehring, K. (2004), "The Veil of Public Ignorance," *Journal of Economic Theory* 119, 247-270.

[25] Nehring, K. and Puppe, C. (2003), "Efficient and Strategy-Proof Voting Rules: A Characterization", mimeo.

[26] Nehring, K. and Puppe, C. (2004), "The Structure of Strategy-Proof Social Choice: General Characterization and Possibility Results on Median Spaces," mimeo.

[27] Nehring, K. and Puppe, C. (2005a), "On the Possibility of Strategy-Proof Social Choice: Non-Dictatorship, Anonymity, and Neutrality" mimeo.

[28] Nehring, K. and Puppe, C. (2005b), "Consistent Judgment Aggregation: A Characterization," mimeo.

[29] Pauly, M. and van Hees, M. (2003), "Logical Constraints on Judgement Aggregation," Working Paper.

[30] Pearl, J. (1988): *Probabilistic Reasoning in Expert Systems: Networks of Plausible Inference.* Morgan Kaufmann, San Mateo, CA.

[31] Peters, H., H. van der Stel and T. Storcken (1992), Pareto Optimality, Anonymity, and Strategy-Proofness in Location Problems, *International Journal of Game Theory* 21, 221-235.

[32] Pettit, P. (2001a), "Deliberative Democracy and the Discursive Dilemma," *Philosophical Issues* 11, 268-299.

[33] Pettit, P. (2001b), "Groups with Minds of their Own," mimeo.

[34] Sen, A. K.: The Impossibility of a Paretian Liberal. *Journal of Political Economy* 78: 152-157 (1970).

[35] Van Hees, M. (2004): The Limits of Epistemic Democracy, mimeo.

[36] Zadeh, L.A. (1965): "Fuzzy sets," *Information and Control* 8, 338-353.