Local Supermajorities

Fabrizio Cariani

[draft of August 27, 2014; comments welcome; please do not cite/quote without permission]

Abstract

This paper explores two non-standard supermajority rules in the context of judgment aggregation over multiple logically connected issues. These rules set the supermajority threshold in a local, context sensitive way—partly as a function of the input profile of opinions. To motivate the interest of these rules, I prove two results. First, I characterize each rule in terms of a condition I call 'Block Preservation'. Block preservation says that if a majority of group members accept a judgment set, then so should the group. Second, I show that one of these rules is, in a precise sense, a judgment aggregation analogue of a rule for connecting qualitative and quantitative belief that has been recently defended by Hannes Leitgeb. The structural analogy is due to the fact that Leitgeb sets thresholds for qualitative beliefs in a local, context sensitive way—partly as a function of the given credence function.

Introduction

This paper explores two non-standard supermajority rules in the context of judgment aggregation over multiple logically connected issues.[1] Informally, aggregating opinions by a supermajority rule means that a sentence φ is accepted in a group $\mathbf{G}$ just in case the proportion of individuals in $\mathbf{G}$ that accept φ meets or exceeds a threshold t. The rules I discuss here set the threshold in a very local, context-sensitive way—partly as a function of the input profile of opinions.

It is well-known that, if t is set *absolutely* (i.e., independently of features of the context of the aggregation problem) and is less than 1, there is no guarantee that the output of a supermajority rule will be consistent.

> *Example 1:* five individuals aggregate their beliefs on atomic sentences $A_1, ..., A_5$ and their disjunction $A_1 \vee ... \vee A_5$. Each individual i believes A_i, and rejects all the other members of $\{A_1, ..., A_5\}$. Additionally, everyone accepts $(A_1 \vee ... \vee A_5)$. Then, there is unanimous support for the disjunction and .8 support *against* each disjunct.

In Example 1, every supermajority rule with $t \leq .8$ recommends acceptance of an inconsistent set of sentences, namely $\{\sim A_1, ..., \sim A_5, (A_1 \vee ... \vee A_5)\}$. For $t > .8$ (and as long as it is short of 1), a variant of Example 1 involving more than five atomic sentences establishes a similar point. Example 1 is supermajoritarian generalization of the *doctrinal paradox*—the result that the majority rule can recommend acceptance of inconsistent sets of propositions even if each of the members of the group is consistent.[2]

Setting thresholds locally in the way I will suggest offers a way around this problem. Additionally, I hope to motivate the interest of these rules in two ways. First, I show (by characterizing each rule) that they have unique properties and they are immune from objections that apply to less context sensitive supermajority rules. Second, I show that one of these rules is, in a precise sense, a judgment aggregation analogue of a rule for connecting qualitative and quantitative belief that has been recently defended by Leitgeb (forthcoming).

---

[1] This work was spearheaded by List and Pettit (2002) but the literature now encompasses a variety of methods and applications. The key results that set up the present paper are found in Dietrich and List (2007a,b) and Nehring and Puppe (2007). See Grossi and Pigozzi (2014) for an updated introduction with an ample reference list: §6.3 of this book—which discusses Lang *et al.* (2011, 2012) and Nehring and Pivato (2011)—is especially relevant to the present discussion.

[2] Kornhauser and Sager (1986). In some form of other, the doctrinal paradox is discussed in all of the judgment aggregation references provided here. For important work distinguishing the doctrinal paradox from the discursive dilemma, see Mongin (2012).

1 Setup and Motivation

The following is a familiar formal framework for aggregating qualitative judgments.[3] Let $\mathbf{G}$ denote a group (a set of individuals). Suppose that we must aggregate the opinions of members of $\mathbf{G}$ on sentences drawn from an *agenda* $\mathbf{A}$. Think of the agenda as the set of issues on which the group must form an opinion. Although we will need to revise this in a later section, the prevailing formalism treats agendas syntactically, as sets of sentences of a formal language.

Start, then, with a sentential language $\mathcal{L}$. We will need the concepts of *logical consistency*, *deductive closure* and *completeness* of a set of $\mathcal{L}$-sentences S. In the present paper, we lift them (without presentation) from standard truth-functional logic. A set of sentences S is maximally consistent (relative to $\mathbf{A}$) iff S is consistent and no proper superset of S is consistent. It is minimally inconsistent iff S is inconsistent and every proper subset of S is consistent. Say that the *negation** of φ is $\sim\varphi$, unless the main connective of φ is negation (i.e. for some ψ, $\varphi = \sim\psi$), in which case the negation* of φ is ψ.

An *agenda* $\mathbf{A}$ is a finite set of sentences closed under negation*. A *judgment set* J is a subset of the agenda. A profile $\vec{J}$ is a sequence of judgment sets–one judgment set per group member. Each agent is associated with a judgment set: for instance the judgment set corresponding to the opinions of individual i is J_i. A *rational* profile is a sequence of maximally consistent judgment sets. Let $\mathbf{RP}$ be the set of all rational profiles. An aggregation rule $\mathbf{F}$ is a function from rational profiles to judgment sets.

To this relatively standard setup, I add one more definition:

$$\texttt{support}(\varphi, \vec{J}) := |\{i \in \mathbf{G} \mid \varphi \in J_i\}|/|\mathbf{G}|$$

Informally, $\texttt{support}(\varphi, \vec{J})$ outputs the proportion of φ-supporters in $\mathbf{G}$ (given the opinions recorded in $\vec{J}$). For example if, given a profile $\vec{J}$, one fifth of the $\mathbf{G}$-members accept φ and the rest reject it, $\texttt{support}(\varphi, \vec{J}) = .2$.

Standard supermajority rules with absolute threshold t ($>.5$) are defined schematically as follows:

Absolute Supermajority$_t$: $\mathbf{SM}_t(\vec{J}) = \{\varphi \mid \texttt{support}(\varphi, \vec{J}) \geq t\}$

Notice that definitions like this characterize a function relative to a fixed group $\mathbf{G}$, agenda $\mathbf{A}$ and language $\mathcal{L}$. So, technically, in order to earn the right to speak of *the* supermajority rule with threshold t, we must fix the values of these parameters ($\mathbf{G}, \mathbf{A}$ and $\mathcal{L}$). In the following, unqualified general claims about

[3] For the sources of this framework see the references in fn. 1.

aggregation rules are intended to apply to groups of size at least 3 and agendas that contain at least one minimally inconsistent set of size no less than 3.

Example 1 (and its variations) shows that for every threshold $t < 1$, there is an agenda $\mathbf{A}$ and a group $\mathbf{G}$ such that $\mathbf{SM}_t$ (as defined for $\mathbf{A}$ and $\mathbf{G}$) is not guaranteed to preserve consistency. That is, for every threshold $t < 1$, there is an agenda $\mathbf{A}$ such that $\mathbf{SM}_t$ (as defined on $\mathbf{A}$) lacks the property:

Consistency Preservation: $\forall \vec{J} \in \mathbf{RP}; \mathbf{F}(\vec{J})$ is consistent.[4]

Note that the name 'consistency preservation' is apt because of the assumption that every judgment set in any profile in **RP** is consistent.

Despite this *prima facie* disappointing result, there is a persistent hunch that supermajoritarian rules could offer a general way out of the doctrinal paradox. The hunch is that, in any given case, we could *choose* a threshold that manages to yield consistent outputs. For instance, in Example 1, we could choose .95 so as to guarantee consistency. There might, then, be subtler, non-absolute ways of determining a supermajority threshold that *are* compatible with consistency-preservation.

List (forthcoming)—cashing in on results from Dietrich and List (2007a)—notes that there indeed is a subtle, non-absolute, way of defining supermajorities that guarantees consistency. Instead of picking a one-size-fits-all supermajority threshold, we can consider a threshold that varies with the logical properties of the agenda $\mathbf{A}$. In particular, consistency is guaranteed by any threshold that exceeds $(k-1)/k$ where k is the size of the largest minimally inconsistent subset of $\mathbf{A}$. Here we will choose $t_{\mathbf{A}} = ((k-1)/k) + \epsilon$, where $\epsilon < 1/|\mathbf{G}|$. So here is a rule that deploys this threshold:

Agenda Supermajority: $\mathbf{AS}(\vec{J}) = \{\varphi \mid \texttt{support}(\varphi, \vec{J}) \geq t_{\mathbf{A}}\}$

In Example 1, $k = 6$, because the largest minimally inconsistent subset is

$$\{\sim A_1, ..., \sim A_5, (A_1 \vee ... \vee A_5)\}$$

and hence $t_{\mathbf{A}}$ is just over $5/6$. Selecting the threshold as a function of the agenda enforces consistency preservation.

While **AS** has some attractive features, such as consistency preservation, it is too prudent an aggregation rule.

[4]*Notational Convention*: In presenting aggregation conditions, **F** occurs as a free variable. A more explicit way of giving the same definition would be to say that the set of Consistency Preserving rules is: $\{\mathbf{F} \mid \forall \vec{J} \in \mathbf{RP}; [\mathbf{F}(\vec{J})$ is consistent $].\}$

Example 2: Suppose that group **G** deliberates about the agenda **A**$=\{A, B, A \& B, negations\}$. Suppose that **G** is split in two factions: one faction accepts A, B and $A\&B$. The other accepts $\sim A$, $\sim B$ and $\sim(A\&B)$. 60% of **G** belong to the first faction; the remaining 40% belong to the second faction.

In Example 2, **AS** outputs the empty set (we may colorfully think of this as a case in which it recommends collective suspension of judgment on every issue in the agenda).

In some applications of judgment aggregation models, this outcome is counter-intuitive. Even though the agenda has enough complexity to raise the supermajority threshold above .5, group members' patterns of acceptance only seem to distinguish two alternative states. In other words, each judge treats the agenda as if it were the simpler agenda $\{(A\&B), \sim(A\&B)\}$. It would be intuitive in such cases to go with the majoritarian verdict.

By increasing the complexity of the agenda, we can get even more striking situations of this kind. Consider again the setup in Example 1 with five judges and five atomic sentences. As I noted earlier, in this case we have $t_{\mathbf{A}}$ just over 5/6. This means that the following distribution of judgments would also result in an empty collective judgment set:

$$\begin{array}{rl} \text{Judges 1-4:} & A_1, A_2, A_3, A_4, A_5, (A_1 \vee \ldots \vee A_5) \\ \text{Judge 5:} & \sim A_1, \sim A_2, \sim A_3, \sim A_4, \sim A_5, \sim(A_1 \vee \ldots \vee A_5) \end{array}$$

Under **AS**, even 4/5 uniform agreement over an entire judgment set may not be enough to warrant acceptance. It is easy to imagine applications of judgment aggregation theory in which these constraints are too strict.

We can turn these observations about particular cases into a general requirement.

(Weak) Block Preservation: $\forall \vec{J} \in \mathbf{RP}; \forall \mathbf{M} \subseteq \mathbf{G}$;
IF:

(i) **M** is a majority of the members of **G**

(ii) and $(\forall i, k \in \mathbf{M}, J_i = J_k)$

THEN: $[\forall i \in \mathbf{M}, \mathbf{F}(\vec{J}) = J_i]$.

Informally: if a majority of group members submit the same judgment set (i.e. the same opinion on every issue), then the collective ought to agree. Block

Preservation is, in my view, a reasonable (and in any case interesting) property of the majority rule.

Absolute supermajorities (with the exception of the majority rule itself) fail to satisfy it. Moreover, and crucially, Agenda Supermajority (**AS**) violates it, as long as the agenda contains a minimally inconsistent set of size 3 (or greater). For that degree of complexity in the agenda will raise the threshold over 2/3 and thus prevent acceptance of 60/40 blocks (as in Example 2).

These observations raise two natural questions:

(Q1) are there rules that satisfy consistency preservation *and* block preservation?

(Q2) are there rules *that are supermajoritarian in inspiration*, but satisfy consistency preservation and block preservation?

Q1 is easily answered: the computer science literature has drawn attention to a vast class of *distance-based* rules (Konieczny *et al.*, 2004, Pigozzi 2006, Miller and Osherson 2009, Chandler 2013). Many distance-based rules—for instance, the rule based on summing up Hamming distances—are both consistency preserving and block preserving. No doubt many other distance-based rules are as well. Since I do not aim to provide a fully general answer to (Q1) and since it is independently interesting to explore rules that do not presuppose a distance metric, I set distance-based rules aside and omit detailed presentation of them.

Q2 is more interesting, because its answer depends on what it means to be "supermajoritarian in inspiration". On one way of making sense of this concept, the answer must be negative. Say that a *family of thresholds* (on **A**) is a function τ such that for each $\varphi \in \mathbf{A}$, $\tau(\varphi)$ is a threshold ($.5 < \tau(\varphi) \leq 1$). Let **T** be the set of all families of thresholds (on **A**). We define the class of quota rules

$$\textbf{Quota Rule: } \exists \tau \in \mathbf{T}; \forall \vec{J} \in \mathbf{RP}; \mathbf{F}(\vec{J}) = \{\varphi \mid \texttt{support}(\varphi, \vec{J}) \geq \tau(\varphi)\}$$

Informally, **F** is a quota rule if there is a family of thresholds τ such that φ is accepted by **F** (on $\vec{J}$) iff φ is accepted by a proportion of $\tau(\varphi)$ individuals (or more).

Now, let $\mathbf{T}_{uni} = \{\tau \in \mathbf{T} \mid \text{for any } \varphi, \psi \in \mathbf{A}, \tau(\varphi) = \tau(\psi)\}$. Say that a quota rule (based on agenda **A**) is *uniform* iff it satisfies the above condition with $\mathbf{T}_{uni}$ in place of **T**. Informally, uniform quota rules are such that every sentence in the agenda is assigned the same threshold.

We could try making sense of "**F** is supermajoritarian in inspiration" as meaning "**F** is a uniform quota rule." However, it is an easy consequence of the main theorem in List and Dietrich (2007a) that no uniform quota rule is consistent and block preserving. In particular, their Corollary 2(a) implies that a uniform quota rule **F** is consistent if and only if its threshold is $t_{\mathbf{A}}$ or greater. And we know that supermajority rules based on such thresholds are not block preserving.

2 Local Supermajorities

The key to an affirmative answer to Q2 is to adopt a more permissive conception of what it is to be "supermajoritarian in inspiration". A natural idea is to swap the quantifiers in the definition of Quota Rules, thus making the family of thresholds dependent on the profile.

> **Relaxed Quota Rule:** $\forall \vec{J} \in \mathbf{RP}; \exists \tau \in \mathbf{T}; \mathbf{F}(\vec{J}) = \{\varphi \mid \mathtt{support}(\varphi, \vec{J}) \geq \tau(\varphi)\}$

Say that a relaxed quota rule is *uniform* just in case it satisfies this definition with $\mathbf{T}_{uni}$ in place of **T**.

It turns out that the property of being a relaxed quota rule is weak enough that it can be combined with block preservation and consistency preservation. In fact, this combination can be instantiated by a variety of non-equivalent rules. I present two here.

Each rule is defined in two steps. Here is the first rule:[5]

> **Auxiliary Definition:** $\mathtt{cons}(\vec{J}, t) = \{\varphi \mid \mathtt{support}(\varphi, \vec{J}) \geq t\}$, if this set is consistent, $\emptyset$ otherwise.
>
> **Local Supermajority:**
>
> $$\mathbf{LS}^{-}(\vec{J}) = \bigcup_{.5 \leq t \leq 1} \mathtt{cons}(\vec{J}, t)$$

Informally, one way of thinking about $\mathbf{LS}^{-}$ is that it sets its supermajority threshold only after 'peeking' at the profile $\vec{J}$. A crucial property of cons-sets is that: if $\mathtt{cons}(\vec{J}, t)$ is non-empty, and $t < t'$, $\mathtt{cons}(\vec{J}, t) \supseteq \mathtt{cons}(\vec{J}, t')$. For this reason, the effect of taking the union of the cons-sets is the same as the effect of choosing as a threshold one of the maximally inclusive thresholds (in general, there is no guarantee that there will be *one* maximally inclusive threshold).

Although $\mathbf{LS}^{-}$ is consistency preserving, some may object to it on the grounds that its output is not guaranteed to be deductively closed (or, as I will simply say: *closed*).

[5] This rule is equivalent to the ranking rule of Lang *et al.* (2011,2012) and the leximin rule of Nehring and Pivato (2011).

Closure: $\forall \vec{J} \in$ **RP**; $F(\vec{J})$ is deductively closed.

Here is an example that shows that **LS**$^-$ is not closed:

Example 3: Let $\mathbf{G} = \{1, ..., 7\}; \mathbf{A} = \{A, B, (A \,\&\, B), negations\}$.

- $J_1 = J_2 = J_3 = \{A, B, (A \,\&\, B)\}$.
- $J_4 = J_5 = \{A, {\sim}B, {\sim}(A \,\&\, B)\}$
- $J_6 = J_7 = \{{\sim}A, B, {\sim}(A \,\&\, B)\}$

Let $\vec{J} = \langle J_1, J_2, J_3, J_4, J_5, J_6, J_7 \rangle$.

In this case we have $\texttt{support}(A, \vec{J}) = \texttt{support}(B, \vec{J}) = 5/7$; $\texttt{support}(A\&B, \vec{J}) = 3/7$; $\texttt{support}({\sim}(A \,\&\, B), \vec{J}) = 4/7$. Given this: $\mathbf{LS}^-(\vec{J}) = \{A, B\}$, which entails, but obviously does not include, $(A\&B)$. In one sentence: though **LS**$^-$ guarantees consistent outputs, it does not guarantee closed ones. If closure is desirable, **LS**$^-$ is objectionable.

One way to dodge this objection is to close under entailment *after* the aggregation. Though I do not think closing in this way is particularly objectionable, there is an alternative way of enforcing closure that will prove to have independent interest.

Auxiliary Definition: $\texttt{closed}(\vec{J}, t) = \{\varphi | \texttt{support}(\varphi, \vec{J}) \geq t\}$, if this set is consistent *and closed*, $\emptyset$ otherwise.

Closed Local Supermajority:

$$\mathbf{LS}^+ \,(\vec{J}) = \bigcup_{.5 \leq t \leq 1} \texttt{closed}(\vec{J}, t)$$

The key modification is the italicized addition in the Auxiliary Definition. **LS**$^+$ guarantees consistency and closure.

I do not claim that **LS**$^-$ and **LS**$^+$ are ideal ways of aggregating opinions. Instead, I acknowledge that it is clear they have puzzling features. For instance, both rules can output the empty set on some input judgments. Classic instances of the doctrinal paradox are an example.

Example 4: Suppose $\mathbf{G} = \{Charlotte, Lara, Chandra\}$. Charlotte accepts $\{A, B, A\&B\}$; Lara: $\{A, {\sim}B, {\sim}(A\&B)\}$. Chandra: $\{{\sim}A, B, {\sim}(A\& B)\}$.

Depending on the purposes of the aggregation and the background features of the aggregation context, this empty output might (or might not be) a bad verdict (in any case, this is a verdict that $\mathbf{LS}^-/\mathbf{LS}^+$share with **AS**).

Another, more serious objection is that it is easy to imagine situations in which Local Supermajority incentivizes a certain kind of insincere voting. In Example 4, Chandra might have an incentive to submit the judgment $\{\sim A, \sim B, \sim(A \,\&\, B)\}$. If she were to submit this 'insincere' judgment set, $\mathbf{LS}^-$ would output $\{A, \sim B, \sim(A \,\&\, B)\}$. Chandra might 'trade' losing on $\sim B$ in order to get the collective outcome to agree with her on $\sim(A \,\&\, B)$. It might be deemed bad if an aggregation rule encourages this kind of insincerity and it is well-known in the literature that the possibility of strategic voting is tightly connected to failures of Independence constraints (Dietrich and List 2007b, Nehring and Puppe 2007).[6]

In defense of the present enterprise, however, I wish to make three points. First, while these objections are well taken, the point of studying of aggregation rules is not to find *the right rule* in some abstract *a-priori* sense. It is rather to understand the rationale behind different rules (or classes of rules), by thinking about the aggregation conditions that characterize them. As I will show, Local Supermajorities allow us to explore aggregation conditions that have a distinctive intuitive motivation, but are otherwise difficult to identify.

Second, Local Supermajorities are not the only rules that allow strategic voting. Given that the possibility of strategic voting is tightly connected with failures of Independence, the aforementioned distance-based rules also allow for strategic voting (as do many others). Moreover, Independence constraints are both normatively suspicious and involved in almost every impossibility result.

Third, it is easy to imagine circumstances in which aggregating by $\mathbf{LS}^-$ or $\mathbf{LS}^+$ is relatively innocent. The kind of insincere voting that is allowed by these rules can only be exploited by an individual who knows the entire profile of judgments. But there are many applications of judgment aggregation tools in which we can reasonably rule out this possibility. An example is the case of deference to a group of epistemic experts that motivated Pettit's discussion of Supermajority rules (2006). Suppose you are trying to form an opinion by deference on the basis of the judgment of several independent and equally informed experts. If Jane, an expert, has no stakes in what you get to believe (or does not *know* what the others are testifying), there is no direct threat that she would misrepresent her opinion.[7] The general point here is that whether the possibility of insincere

[6] Independence is the requirement that the collective judgment on φ depend only on the individual judgments on φ.

[7] There might even be cases of collective decision-making in which the rule cannot easily be exploited. Imagine a society whose members submit their judgments to a social planner (in ignorance of the judgments of others). The social planner is then be tasked with choosing an

voting counts against a rule is an application-dependent problem.

3 Basic Analysis of Local Supermajorities

In this section, I provide a simple characterization of $\mathbf{LS}^{-}$and $\mathbf{LS}^{+}$. The characterization of $\mathbf{LS}^{-}$ turns crucially on the following strengthening of block preservation.[8]

Strong Block Preservation: $\forall \vec{J} \in \mathbf{RP}; \forall S, T \subseteq \mathbf{A}; \forall t > .5;$
IF:

(i) S is consistent

(ii) $\forall \psi \in S$, $\texttt{support}(\psi, \vec{J}) \geq t$,

(iii) if T and S are inconsistent, $\exists \psi \in T, \texttt{support}(\psi, \vec{J}) < t$

THEN: $[S \subseteq \mathbf{F}(\vec{J})]$

This entails, but is not entailed by the original block preservation condition. It states a sufficient condition for the sentences in a set S to belong to the aggregated output: namely, that each sentence in S be supported by at least $t \cdot |\mathbf{G}|$ individuals while set inconsistent with S is uniformly supported by at least $t \cdot |\mathbf{G}|$ individuals.

The claim is that $\mathbf{LS}^{-}$ is the only uniform relaxed quota rule that is consistency preserving and strongly block preserving.

Proof:
Suppose $\mathbf{A}$ and $\vec{J}$ are respectively an arbitrary agenda and an arbitrary profile. First, one can easily check that $\mathbf{LS}^{-}$ satisfies the three characterizing properties.

Second, let $\mathbf{F}$ be an arbitrary rule that satisfies the three characterizing properties. We prove that $\forall \varphi \in \mathbf{A}$, $\varphi \in \mathbf{LS}^{-}(\vec{J}) \Longleftrightarrow \varphi \in \mathbf{F}(\vec{J})$.

$[\Rightarrow]$: suppose $\varphi \in \mathbf{LS}^{-}(\vec{J})$. Then there is a threshold $t > .5$ such that $\texttt{cons}(\vec{J}, t) = \mathbf{LS}^{-}(\vec{J})$ and so $\varphi \in \texttt{cons}(\vec{J}, t)$ and all of these are consistent. We show that $\texttt{cons}(\vec{J}, t)$ is a set of sentences that

appropriate threshold, and derive some range of collective choices. Under such a system, there is relatively little threat that individuals would vote insincerely.

[8]The principle is related to the principle of supermajoritarian efficiency in Nehring and Pivato (2011), which they describe informally as the requirement to overrule "if necessary, small supermajorities in favor of an equal or larger number of supermajorities of equal or greater size" (p.9).

meets the conditions in the antecedent of strong block preservation, instantiating the quantifiers with profile $\vec{J}$, agenda $\mathbf{A}$ and threshold t. Because $\mathbf{F}$ satisfies Strong Block Preservation, establishing this implies $\texttt{cons}(\vec{J}, t) \subseteq \mathbf{F}(\vec{J})$ and so $\varphi \in \mathbf{F}(\vec{J})$.

Obviously: $\texttt{cons}(\vec{J}, t) \subseteq \mathbf{A}$. For conditions (i) and (ii), it is also obvious that $\texttt{cons}(\vec{J}, t)$ is consistent (otherwise it would be empty) and $\texttt{support}(\psi, \vec{J}) \geq t$ for $\psi \in \texttt{cons}(\vec{J}, t)$. Finally, to establish (iii), there cannot be any set T inconsistent with $\texttt{cons}(\vec{J}, t)$ such that for all $\psi \in T$, $\texttt{support}(\psi, \vec{J}) \geq t$. If there was, $T \subseteq \texttt{cons}(\vec{J}, t)$, which is impossible because $\texttt{cons}(\vec{J}, t)$ is consistent. Since the antecedent of strong block preservation is met, $\texttt{cons}(\vec{J}, t) \subseteq \mathbf{F}(\vec{J})$, so $\varphi \in \mathbf{F}(\vec{J})$.

[$\Leftarrow$]: suppose $\varphi \in \mathbf{F}(\vec{J})$. Because $\mathbf{F}$ is a uniform relaxed quota rule, given $\vec{J}$, there is a threshold $t_{\vec{J}}$ such that

$$(\#)\forall\psi\ [\psi \in \mathbf{F}(\vec{J}) \text{ iff } \texttt{support}(\psi, \vec{J}) \geq t_{\vec{J}}]$$

Fix some such $t_{\vec{J}}$. Note that by (#) and the fact that $\mathbf{F}(\vec{J})$ is consistent, $\mathbf{F}(\vec{J}) = \texttt{cons}(\vec{J}, t_{\vec{J}})$. But we know $\texttt{cons}(\vec{J}, t_{\vec{J}}) \subseteq \mathbf{LS}^-(\vec{J})$, so $\mathbf{F}(\vec{J}) \subseteq \mathbf{LS}^-(\vec{J})$, and hence $\varphi \in \mathbf{LS}^-(\vec{J})$.

$\mathbf{LS}^+$ is also a uniform relaxed quota rule. The key difference is that $\mathbf{LS}^+$ is not strongly block preserving. In Example 3, any rule that satisfies strong block preservation entails that $\{A, B\}$ is accepted, and in fact $\mathbf{LS}^-(\vec{J}) = \{A, B\}$. By contrast, $\mathbf{LS}^+(\vec{J}) = \emptyset$ (the key observation being that the consistent, closed and non-empty judgment set $\{A, B, A \,\&\, B\}$ is not supported by a majority).

There is, however, a weakening of strong block preservation that applies to $\mathbf{LS}^+$.

Intermediate Block Preservation: $\forall \vec{J} \in \mathbf{RP}; \forall S \subseteq \mathbf{A}; \forall t > .5;$
IF:

(i) S is consistent and closed

(ii) $\forall \psi \in S$, $\texttt{support}(\psi, \vec{J}) \geq t$

(iii) $\forall \psi \in \mathbf{A}$ s.t. ψ is not entailed by S, $\texttt{support}(\psi, \vec{J}) < t$

THEN $[S \subseteq \mathbf{F}(\vec{J})]$

This condition is intermediate in strength between the initial block preservation condition and the strong one. On the one hand, it entails that any judgment set J that is supported by a majority of individuals is accepted by the group. On the other, it is silent on what happens in cases like Example 3.

To characterize $\mathbf{LS}^+$, then we say that it is the only uniform relaxed quota rule that satisfies consistency preservation, deductive closure, and intermediate block preservation.

> *Proof:*
> First, we note that $\mathbf{LS}^+$ satisfies all of the characterizing conditions. As before, we prove that for all $\varphi \in \mathbf{A}$, $\varphi \in \mathbf{LS}^+(\vec{J}) \Leftrightarrow \varphi \in \mathbf{F}(\vec{J})$, where $\mathbf{F}$ is an arbitrary aggregation function satisfying the characterizing conditions.
>
> $[\Rightarrow]$: suppose $\varphi \in \mathbf{LS}^+(\vec{J})$. Then there is a threshold t such that $\texttt{closed}(\vec{J}, t) = \mathbf{LS}^+(\vec{J})$ and so $\varphi \in \texttt{closed}(\vec{J}, t)$. To prove $\varphi \in \mathbf{F}(\vec{J})$, we check that $\texttt{closed}(\vec{J}, t)$ satisfies the conditions in the antecedent of Intermediate Block Preservation. First, $\texttt{closed}(\vec{J}, t)$ must be a deductively closed and consistent subset of the agenda (thus satisfying condition (i)). Condition (ii) is satisfied by construction of $\texttt{closed}(\vec{J}, t)$. To check condition (iii), suppose there was a ψ, $\texttt{closed}(\vec{J}, t) \nvdash \psi$ with $\texttt{support}(\psi, \vec{J}) \geq t$. This would imply $\psi \in \texttt{closed}(\vec{J}, t)$, which contradicts the claim that ψ is not entailed by $\texttt{closed}(\vec{J}, t)$. By intermediate block preservation $\varphi \in \mathbf{F}(\vec{J})$.
>
> $[\Leftarrow]$: suppose $\varphi \in \mathbf{F}(\vec{J})$. We know that $\mathbf{F}(\vec{J})$ is consistent and closed. Because $\mathbf{F}$ is a uniform relaxed quota rule, given $\vec{J}$, there is a threshold $t_{\vec{J}}$ such that
>
> $$\varphi \in \mathbf{F}(\vec{J}) \text{ iff } \texttt{support}(\varphi, \vec{J}) \geq t_{\vec{J}}$$
>
> Fix some such $t_{\vec{J}}$. Because $\mathbf{F}(\vec{J})$ is consistent and closed, $\mathbf{F}(\vec{J}) = \texttt{closed}(\vec{J}, t_{\vec{J}}))$. Since, by definition, $\texttt{closed}(\vec{J}, t_{\vec{J}})) \subseteq \mathbf{LS}^+(\vec{J})$, $\varphi \in \mathbf{LS}^+(\vec{J})$.

It seems possible that to simplify both characterization results, especially in light of the fact that the block preservation conditions are (i) fairly complex conditions and (ii) are only used in one direction of the respective proofs. But for the present purposes, they are useful because they allow us to think of $\mathbf{LS}^-$ and $\mathbf{LS}^+$ in terms of what kinds of blocks of opinion they preserve.

4 $\mathbf{LS}^+$ and the Stability Theory of Belief

Much literature has noted the close formal analogy between the problem of judgment aggregation and the problem of connecting graded belief and qualitative belief (Levi 2004, Douven and Romejin 2007, Chandler 2013, Briggs *et* *al.* forthcoming). More generally, there is a natural correspondence between acceptance rules and a subset of the set of all aggregation rules.

In this section, I (1) make this correspondence precise, and (2) note that, under the correspondence, $\mathbf{LS}^+$ matches an acceptance rule recently developed by Hannes Leitgeb (forthcoming).

To make the connection precise, we must revise some elements of our aggregation setup. First, it is slightly more convenient to discuss acceptance rules in a setup with propositions (understood as sets of worlds), rather than sentences of a formal language. Suppose we have possible worlds drawn from a set W. Perhaps, we can start with the language $\mathcal{L}$ and identify possible worlds as maximally consistent sets of literals of $\mathcal{L}$. For every sentence φ, there is a set of worlds at which that sentence is true (if we identify worlds with maximally consistent sets of literals, these would be the maximally consistent sets that sententially entail φ). Propositions are just such sets. To mark this shift to propositional talk, I am going to start using variables p, q, r, ranging over sets of worlds. I do not reintroduce analogues for the other definitions (e.g. `support`, *aggregation rule*, etc.).

A more substantive amendment concerns the structure of the agenda. In the standard aggregation setup, we just require agendas to be closed under the operation of *negation**. In propositional talk, the parallel move would be to require agendas to be closed under complementation relative to the set W of all worlds. In the acceptance literature, however, it is presupposed that agendas that have quite more structure. In particular, to set up our required correspondence we require agendas to be boolean algebras.[9]

The question then is how to connect judgment aggregation and a theory of acceptance. I address this in two steps. First, we note that a profile of judgments determines a probability function over **A**.

> **Step 1: Determination.** For each rational profile $\vec{J}$, let $P^{\vec{J}}(\cdot)$ be the function whose value for a given $p \in \mathbf{A}$ is $\texttt{support}(p, \vec{J})$.

Under the assumption that **A** is a boolean algebra and that $\vec{J} \in \mathbf{RP}$, $P^{\vec{J}}(\cdot)$ is a probability function–it obeys the axioms of the probability calculus (this can be easily verified by going through each of the Kolmogorov axioms).

[9] In many cases, the extra richness of the underlying space of propositions does not make a real difference. An agenda **A** that involves only sentential atoms and truth-functional connections among them can easily be extended to agendas over the entire boolean algebra $\mathbf{A}^*$ in such a way that any rational profile of opinions on **A** determines by entailment a rational profile of opinions on $\mathbf{A}^*$.

Now, say that an *acceptance rule* **R** (on **A**) is a function that takes a probability function (defined on **A**) and outputs a rational judgment set. The next step is to note that we can define aggregation rules by composing the *determination step* with different *acceptance rules*. That is, we hold fixed the determination step, and then note:

> **Step 2: Composition.** For every acceptance rule **R**, the function $\mathbf{F}(\vec{J}) = \mathbf{R}(P^{\vec{J}})$ that is the result of composing $P^{\vec{J}}$ and **R** is an aggregation rule.

For example, it is easy to see that the standard propositions-wise majority rule (**MA**) is characterized by the acceptance rule **LO** (for *Lockean*) that says:

$$\mathbf{LO}(P) = \{p \mid P(p) > .5\}$$

in the sense that for all profiles $\vec{J}$, $\mathbf{MA}(\vec{J}) = \mathbf{LO}(P^{\vec{J}})$. More generally, it is easy to see that supermajority rules with threshold t are the result of composing the determination step with an acceptance rule ($\mathbf{LO}_t$) that says that p is accepted iff $P(p) \geq t$.

Note that not every aggregation rule can be represented as the composition of the determination step and an acceptance rule. For example, there is no way of capturing the dictatorship of judge i in this way: the reason is that the probability function $P^{\vec{J}}$ loses information that is explicitly represented in a profile of judgments—for example, information about the identity of the judges.

However, and remarkably, $\mathbf{LS}^{-}$and $\mathbf{LS}^{+}$ can be represented in this way. In particular, $\mathbf{LS}^{+}$corresponds to one of the (broadly) Lockean rules defined in Leitgeb (forthcoming). Leitgeb is driven by a parallel insight to the one that drove the definition of $\mathbf{LS}^{-}/\mathbf{LS}^{+}$. He notes that some of the impossibilities concerning acceptance rules disappear if we choose our acceptance threshold as dependent on the probability function P.

The thresholds in Leitgeb's acceptance rule are defined in terms of the notion of a P-stable proposition. The motivation and definition of this notion are complex, but Leitgeb provides a simple characterization we can borrow in this context (p. 10). Let P be a probability function:

$$q \text{ is } P\text{-}stable \text{ iff either } P(q) = 1 \text{ or for all } w \in q,\ Pr(\{w\}) > P(W - q)$$

One of Leitgeb's principal results is that any acceptance rule according to which the acceptance threshold (for P) is $P(q)$ with q a P-stable proposition makes the set of accepted propositions consistent and deductively closed. I appeal to this result in the proofs below.

Given a probability measure P, there may be multiple P-stable propositions. However, as long as the agenda is finite (as we required), there always is a strongest P-stable proposition. Let $h_{\mathbf{A}}$ be a function that, given a probability function P (on $\mathbf{A}$), outputs the strongest P-stable proposition. We can then define the following acceptance rule:

$$\mathbf{ST}(P) = \{q \mid P(q) \geq P(h_{\mathbf{A}}(P))\}$$

The key claim of this section is:

$$\mathbf{LS}^{+}(\vec{J}) = \mathbf{ST}(P^{\vec{J}})$$

Informally, closed local supermajority is (equivalent to) the result of converting the profile of opinions into a probability function and applying a maximally inclusive version of Leitgeb's acceptance rule.

Proof:
[$\subseteq$]: suppose $p \in \mathbf{LS}^{+}(\vec{J})$. We want to show $p \in \mathbf{ST}(P^{\vec{J}})$.

Let $q = \bigcap[\mathbf{LS}^{+}(\vec{J})]$. Since $\mathbf{A}$ is a finite boolean algebra, $q \in \mathbf{A}$. Because $\mathbf{LS}^{+}$'s output is deductively closed (with respect to $\mathbf{A}$), $q \in \mathbf{LS}^{+}(\vec{J})$. We show that q is the strongest $P^{\vec{J}}$-stable proposition.

(I) q is $P^{\vec{J}}$-stable: Suppose it is not, then $P^{\vec{J}}(q) < 1$ and there is $w \in q$, s.t. $P^{\vec{J}}(\{w\}) \leq P^{\vec{J}}(W - q)$. Then let $r := (q - \{w\}) \cup (W - q)$; because of the properties of w, and $W - q$, $P^{\vec{J}}(r) \geq P^{\vec{J}}(q)$, i.e. $\texttt{support}(r, \vec{J}) \geq \texttt{support}(q, \vec{J})$. Since q is in $\mathbf{LS}^{+}(\vec{J})$ and r is more likely than q (according to $P^{\vec{J}}$), $r \in \mathbf{LS}^{+}(\vec{J})$. However, this is incompatible with $q = \bigcap[\mathbf{LS}^{+}(\vec{J})]$. Because $w \in q$, and $q = \bigcap[\mathbf{LS}^{+}(\vec{J})]$, $w \in r$, but that contradicts r's definition.

(II) q is strongest among the $P^{\vec{J}}$-stable propositions: suppose there is an r that is $P^{\vec{J}}$-stable and $r \subsetneq q$. Let w be the world that witnesses $r \subsetneq q$ (so $w \in q, w \notin r$). Because r is $P^{\vec{J}}$-stable, $\{s \mid P^{\vec{J}}(s) \geq P^{\vec{J}}(r)\}$ must be consistent and closed. But then $\texttt{closed}(\vec{J}, P^{\vec{J}}(r))$ would be consistent and closed. Moreover, because $r \subsetneq q$, $P^{\vec{J}}(r) \leq P^{\vec{J}}(q)$. This, together with the fact that $\texttt{closed}(\vec{J}, P^{\vec{J}}(r))$ is consistent and closed, implies:

$$\texttt{closed}(\vec{J}, P^{\vec{J}}(q)) \subsetneq \texttt{closed}(\vec{J}, P^{\vec{J}}(r))$$

This implies (i) $\texttt{closed}(\vec{J}, P^{\vec{J}}(r)) \subseteq \mathbf{LS}^+(\vec{J})$; in addition we have (ii) $r \in \texttt{closed}(\vec{J}, P^{\vec{J}}(r))$; (iii) $q = \bigcap[\mathbf{LS}^+(\vec{J})]$ and (iv) $w \in q, w \notin r$. But these four claims are contradictory.

[$\supseteq$]: suppose $p \in \mathbf{ST}(P^{\vec{J}})$. Let $q = \bigcap \mathbf{ST}(P^{\vec{J}})$ and $t' = P^{\vec{J}}(q)$. Note that it immediately follows that $P^{\vec{J}}(p) \geq t'$. Furthermore, it follows that if $\texttt{closed}(\vec{J}, t')$ is non-empty, $p \in \texttt{closed}(\vec{J}, t')$ and so $p \in \mathbf{LS}^+(\vec{J})$.

All we have to show is that $\texttt{closed}(\vec{J}, t')$ is non-empty, which it is iff $\{p|\texttt{support}(p, \vec{J}) \geq t'\}$ (i.e. $\{p|P^{\vec{J}}(p) \geq t'\}$) is closed and consistent. This follows by Leitgeb's theorem that for every probability function P, including $P^{\vec{J}}$, $\{p \,|\, P(p) \geq P(q)\}$ is closed and consistent when q is P-stable.

This result is interesting in a few different ways. Most obviously, it forges a somewhat surprising connection between an acceptance rule and an aggregation rule that are motivated in rather different (though structurally similar) ways.

Relatedly, it allows us to bring the characterization results of §3 to bear on our understanding of acceptance rules. For example, recent aggregation literature (Lang 2011, 2012, Nehring and Pivato 2011) discusses techniques to define more complex rules that extend $\mathbf{LS}^-$(where $\mathbf{F}$ extends $\mathbf{LS}^-$iff for all $\vec{J}$, $\mathbf{LS}^-(\vec{J}) \subseteq \mathbf{F}(\vec{J})$ and for some $\vec{J}$ the inclusion is strict). An interesting conceptual question is whether these extensions can be given philosophical motivation in the application to acceptance rules.

Finally, the result opens up important questions in collective epistemology. To highlight this, let me outline a possible application of these results to models of deference to a group of experts that are epistemic peers of each other (Pettit 2006). We might consider two kinds of models: the *Bayesian* model suggests selecting a probability function P_{old} and then updating by conditionalization on the testimony of each of the experts to get to a new probability function P_{new}. If we want our model to output a set of propositions (the propositions one ought to come to accept as a result of deference), we might apply an acceptance rule to P_{new}. The *Aggregation* model suggests instead finding an appropriate aggregation rule $\mathbf{F}$, and applying to the input opinions of the expert. In principle, there is no guarantee that these models will agree. However, under the correspondence result proved in this section, there need not be any disagreement between an Aggregation model based on $\mathbf{LS}^+$ and a Bayesian model based on the appropriate version of Leitgeb's stability theory (provided that, given a profile $\vec{J}$, P_{new} and $P^{\vec{J}}$ are related in the appropriate way). This work, of course, remains to be done.

References

Briggs, R. Cariani, F. Easwaran, K. and Fitelson, B. (forthcoming) Individual coherence and group coherence. Forthcoming in *Essays in Collective Epistemology*, J. Lackey (ed.), Oxford University Press.

Chandler, J. (2013) Acceptance, aggregation and scoring rules, *Erkenntnis*, 78(1), 201-217.

Dietrich, F. and List C. (2007a) Judgment aggregation by quota rules. *Journal of Theoretical Politics*, 19(4).

Dietrich, F. and List C. (2007b) Strategy-proof judgment aggregation. *Economics and Philosophy*, 23, 269-300.

Douven, I. and Romeijn, J. (2007) The discursive dilemma as a lottery paradox. *Economics and Philosophy*, 23, 301-319.

Grossi, D. and Pigozzi, G. (2014) *Judgment aggregation. A primer*. Morgan & Claypool.

Konieczny S., Lang J. and Marquis, P. (2004) DA^2 merging operators. *Artificial Intelligence*, 157(1-2), 49-79.

Kornhauser, L.A. and Sager, L.G. (1986) Unpacking the court. *The Yale Law Journal* 96(1), 82-117.

Lang J., Pigozzi, G., Slovkovik, M. and van der Torre, L. (2011) Judgment aggregation rules based on minimization. In *TARK*, 238-246.

Lang J., Pigozzi, G., Slovkovik, M. and van der Torre, L. (2012) Judgment aggregation rules based on minimization - extended version, technical report, *Université Paris-Dauphine*.

Leitgeb, H. (forthcoming) The stability theory of belief. Forthcoming in *the Philosophical Review*.

Levi, I. (2004) List and Pettit. *Synthese*, 140(1/2), 237-242.

List, C. (forthcoming) When to defer to supermajority testimony and when not to. Forthcoming in *Essays in Collective Epistemology*, J. Lackey (ed.), Oxford University Press.

List, C. and Pettit, P. (2002) Aggregating sets of judgments; an impossibility result. *Economics and Philosophy*, 18(1), 89-110.

Mongin, P. (2012) The doctrinal paradox, the discursive dilemma, and logical aggregation theory. *Theory and Decision*, 73(3), 315-355.

Miller, M. and Osherson, D. (2009) Methods for distance-based judgment aggregation. *Social Choice and Welfare*, 32, 575-601.

NEHRING, K. and PIVATO, M. (2011) Majority rule in the absence of a majority, manuscript.

NEHRING, K. and PUPPE, C. (2007) The structure of strategy-proof social choice – Part I: General characterization and possibility results on median spaces. *Journal of Economic Theory*, 135, 269-305.

PIGOZZI, G. (2006) Belief merging and the discursive dilemma: an argument-based account to paradoxes of judgment aggregation. *Synthese*, 152, 285-298.

PETTIT, P. (2006) When to defer to majority testimony and when not to. *Analysis*, 66(3), 179-187.