# Democracy and Epistemic Justification

Lectures at the International Summer School on Philosophy and Probability
University of Konstanz

15 - 21 September 2002

Christian List

Nuffield College
Oxford
&
Department of Government
London School of Economics

christian.list@nuffield.oxford.ac.uk
http://www.nuff.ox.ac.uk/users/list/





## Overview

**Part 1: Social choice theory and the Condorcet jury theorem**

- Introduction
- Epistemic and procedural accounts of justification
- May's theorem as an illustration of the procedural account (two alternatives)
- The Condorcet jury theorem as an illustration of the epistemic account (two alternatives)

**Part 2: A Bayesian perspective**

- A Bayesian criterion of justification
- Two definitions of special majority voting

**Part 3: Decisions over multiple alternatives**

- Some remarks on procedural and epistemic perspectives
- A generalization of the Condorcet jury theorem to decisions over more than two alternatives
- The problem of cycling and its probability

**Part 4: Decisions over multiple interconnected propositions**

- Some remarks on procedural and epistemic perspectives
- A paradox and an impossibility result
- The probability of the occurrence of the paradox
- A generalization of the Condorcet jury theorem to decisions over multiple interconnected propositions

## Part 1: Social choice theory and the Condorcet jury theorem

**Social choice theory** models collective decisions as processes of aggregating individual inputs into collective outputs.

individual preferences / votes / beliefs
across the individual members of a group/collectivity

aggregation procedure,
e.g. voting system

collective preferences / decisions / beliefs

The **group or collectivity** is usually defined as a fixed set of individuals,

$$N = \{1, 2, ..., n\}.$$

An **aggregation procedure** is usually defined as a function

$$F : \text{domain of inputs} \rightarrow \text{co-domain of outputs},$$

where

- the domain of inputs contains **vectors of items** (such as individuals votes), i.e. one item for each individual (these vectors are also called **profiles**);
- the co-domain of outputs contains **single items** (such as an overall collective winner), i.e. a single item for the group/collectivity as a whole.



## An example of an aggregation problem

Consider a society $N$ that has to decide which one of two candidates, Bush or Gore, to select as their president.

**Input** to the aggregation:
a vector of votes $<x_1, x_2, ..., x_n>$,
where each $x_i \in$ {Bush, Gore}
**Output** of the aggregation:
a single decision $x$, where $x \in$ {Bush, Gore}

Examples of aggregation procedures:

**Simple majority voting:**
Define $F$ as follows. For each $<x_1, x_2, ..., x_n>$, let

$$N_{\text{Bush}} := \{i \in N : x_i = \text{Bush}\},$$
$$N_{\text{Gore}} := \{i \in N : x_i = \text{Gore}\},$$

$$F(<x_1, x_2, ..., x_n>) := \begin{cases} \text{Bush if } |N_{\text{Bush}}| > |N_{\text{Gore}}| \\ \text{Gore if } |N_{\text{Gore}}| > |N_{\text{Bush}}| \end{cases}$$

For simplicity, we do not consider ties here, assuming, say, that $n$ is odd. (Incidentally, this is not the aggregation procedure used in U.S. presidential elections.)

**Dictatorship:**
Define $F$ as follows. Appoint a fixed individual $d \in N$, for instance $d = 17$, as a dictator. For each $<x_1, x_2, ..., x_n>$, let

$$F(<x_1, x_2, ..., x_n>) := x_d.$$

**From a purely logical perspective** any well-defined function $F$ : domain of inputs $\rightarrow$ co-domain of outputs can serve as an aggregation procedure.

But **from a normative perspective** some aggregation procedures seem more attractive than others.

**Question:**
- When there are multiple logically possible aggregation procedures for a given class of aggregation problems, which procedure should we use?
- Under what conditions is the outcome of some collective decision process justified?



## Epistemic and procedural accounts of when a decision is justified

**(e.g. Cohen 1986; List and Goodin 2001 includes a more extensive bibliography)**

**Epistemic.** A decision is right if it is "correct" by some external (decision-procedure-independent) standard.

**Example.** Consider a jury decision.
- The external standard is the fact as to whether or not the defendant is truly guilty, e.g. whether or not he/she has truly committed a particular crime.
- The aim of the decision is to convict the defendant if and only if the defendant is guilty. (Additional assumption: convicting the innocent is worse than acquitting the guilty.)

**Note.** An epistemic justification of a decision procedure is an argument that the given procedure is good at reaching decisions that are "correct" by the relevant procedure-independent external standard.

**Procedural.** A decision is right if it has emerged through a procedure that has certain procedural properties (and the rightness of the decision is solely constituted by the fact that it has emerged through that procedure, irrespective of any procedure-independent facts).

**Example.** Relevant properties of a procedure might be:
- giving all individuals an equal opportunity to influence the outcome of the decision;
- being formally "unbiased" (in some suitable sense) with respect to different possible outcomes of the decision.

**Note.** A procedural justification of a decision procedure is an argument that the given procedure satisfies certain procedural properties. A property counts as *procedural* only if it does not make reference to any decision-procedure-independent facts about what the best outcome of a decision is.

## A procedural justification of an aggregation procedure ... ... as illustrated by the classical social-choice-theoretic approach

**Note:** In social-choice-theoretic terms, a procedural justification of an aggregation procedure has two components: a **normative** one and a **logical** one.

**Normative.** For a given aggregation problem, what are the minimal conditions that an acceptable aggregation procedure should satisfy?

The minimal conditions are usually taken to be **procedural** ones in the following sense: they refer only to the inputs and outputs of the aggregation procedure, but not to any additional parameters representing external facts or states of the world.

**Logical.** Given a set of such minimal conditions, what is the class of all possible aggregation procedures satisfying these conditions?

Ideally, the aim is to find a set of conditions which uniquely determines an aggregation procedure. On the procedural account, the output of such an aggregation procedure would then be the "right" decision, regardless of what that output is.

**Democracy.** Does the procedural account of justification always favour democratic aggregation procedures (like simple majority voting and unlike a dictatorship)?

– This depends entirely on the choice of the relevant minimal conditions.



## A procedural justification of simple majority voting

**Problem.** A group of individuals has to make a decision over two alternatives (e.g. two election candidates, or two alternatives like “conviction” or “acquittal”). Why, for instance, is simple majority voting a better aggregation procedure than a dictatorship?

**Minimal Conditions.**
- Universal Domain
- Anonymity
- Neutrality
- Strict Monotonicity

**Theorem (May 1952).** *Simple majority voting is the unique aggregation procedure for decisions over two alternatives satisfying universal domain, anonymity, neutrality and strict monotonicity.*

## An epistemic justification of a decision procedure

**Note:** An epistemic justification of a decision procedure (e.g. an aggregation procedure) has two components: an ontological one, and an epistemological one.

The **ontological** component is an argument that:

**Ontological Condition.** There exists a decision-procedure-independent fact as to what the best outcome is.

The **epistemological** component is an argument that:

**Epistemological Condition.** The outcome of the given decision procedure tracks the relevant decision-procedure-independent fact as to what the best outcome is.

What does "tracking the truth" mean?

**A criterion of truth-tracking inspired by Nozick's account of knowledge:**

An outcome $x$ of a decision procedure *tracks the truth* if the following four conditions hold:

(i) $x$ is correct (by the relevant external standard);
(ii) $x$ is chosen;
(iii) if $x$ were correct, then $x$ would be chosen;
(iv) if $x$ were not correct, then $x$ would not be chosen.

We call conditions (iii) and (iv) *the Nozickian truth-tracking criteria*. They state that outcome $x$, which is correct in the actual world, is not only accidentally chosen, but that, for any possible world, $x$ is chosen if and only if $x$ is correct.



**Democracy.** Does the epistemic account of justification always favour democratic decision procedures?

- This is not *prima facie* clear.

- Suppose, for instance, there exists an oracle of Delphi such that the oracle robustly determines the "true" best outcome of any decision problem. Then the oracle is the epistemically favoured decision procedure.

- Suppose, for instance, there exists a single individual or small group of individuals, the benevolent omniscient philosopher(s) (maybe as in Plato's *Republic*), that is particularly good at determining the "true" best outcome of any decision problem. Then maybe (?) an "epistocracy", i.e. the dictatorship of the epistemically privileged individual or group is the epistemically favoured decision procedure.

## A epistemic justification of simple majority voting

**Problem.** Consider a decision problem, where there are two alternatives, $x$ (e.g. convict) and $\neg x$ (e.g. acquit).

Is simple majority voting an epistemically favoured procedure?

**Definitions.**

| | |
|---|---|
| *x-correct* : | event that alternative $x$ is correct |
| *¬x-correct* : | event that alternative $x$ is not correct |
| $N_x$ : | random variable whose value is the number of individuals voting for $x$ |
| $N_{\neg x}$ : | random variable whose value is the number of individuals voting for $\neg x$ |

**Assumptions.**

- There are $n$ individuals. For simplicity, assume that $n$ is odd.
- Precisely one of *x-correct* or *¬x-correct* holds. (Which of these two alternatives is correct is called the *state of the world*.)
- Each individual has probability $p$ of voting for the correct alternative, i.e. for each $i \in N$,

  $P$(individual $i$ votes for $x$ | *x-correct*) = $p$

  $P$(individual $i$ votes for $\neg x$ | *x-correct*) = 1-$p$

  $P$(individual $i$ votes for $\neg x$ | *¬x-correct*) = $p$

  $P$(individual $i$ votes for $x$ | *¬x-correct*) = 1-$p$

  ($p$ is called individual $i$'s *competence*.)
- The votes of different individuals are independent from each other, given the state of the world.



**Proposition (Standard result; e.g. Grofman, Owen and Feld 1983).**

(i) $P(N_x > n/2 \mid x\text{-}correct) = \sum_{h>n/2} \binom{n}{h} p^h (1-p)^{n-h}$.

(ii) *If* $p > {}^1/_2$, *then* $P(N_x > n/2 \mid x\text{-}correct) > P(N_{\neg x} > n/2 \mid x\text{-}correct)$.

(*The result holds also if we substitute* $\neg x$ *for* $x$.)

**Theorem (Condorcet jury theorem; e.g. ibid.).** *If* $p > {}^1/_2$, *then* $P(N_x > n/2 \mid x\text{-}correct)$ *converges to* 1 *as* $n$ *increases*.

**Mechanism.** The convergence result is a consequence of the law of large numbers.

$p$ is the expected frequency of individuals voting for $x$ ($np$ is the expected number).

**For small *n***, the actual frequency $N_x/n$ may differ substantially from $p$.

**Law of large numbers** => As $n$ increases, the actual frequency $N_x/n$ will approximate $p$ increasingly closely.

Hence if $p > 1/2$, then $N_x/n$ is increasingly likely to exceed 1/2 as $n$ increases, and hence $x$ is increasingly likely to be the winner under simple majority voting.

We can make two observations.

**First observation.** If we use simple majority voting as the decision procedure (and accept the assumptions of the Condorcet jury model), the following hold:

(i) The probability that $x$ is chosen, given that $x$ is correct, converges to 1 as $n$ tends to infinity.
(ii) The probability that $x$ is not chosen, given that $x$ is not correct, converges to 1 as $n$ tends to infinity.

These statements might be interpreted as probabilistic versions of the Nozickian truth-tracking criteria (of course, this raises difficult questions about how to capture counterfactual conditionals). The interpretation suggests that, *in the limit*, simple majority voting satisfies the Nozickian truth-tracking criteria.

However, for any finite $n$, the probabilities in (i) and (ii) are less then one. In particular, for any finite $n$, we cannot say with probability one that

if $x$ were correct, then $x$ would be chosen;
if $x$ were not correct, then $x$ would not be chosen.

Simple majority voting *only approximately* tracks the truth. It is therefore a fallible procedure. I will later discuss the question of what degree of belief we can justifiably attach to the outcomes of simple majority voting.



**Second Observation.** The procedure of simple majority voting and the Condorcet jury theorem apply only to a very special aggregation problem:

precisely **one** decision over precisely **two** alternatives

This special aggregation problem can be generalized in (at least) two ways:

- in the **number of alternatives**

i.e. **one** decision over ***k*** (>2) alternatives (e.g. $k$ election candidates or $k$ policy options)

- in the **number of simultaneous decisions**

i.e. ***k*** (>1) simultaneous decisions over **two** alternatives each (where some of the decisions logically constrain others)

Below I will discuss both generalizations.

# Part 2: A Bayesian perspective

**(e.g. List 2002, which also includes a more extensive bibliography)**

As we have seen, the Condorcet jury theorem shows that:

(i) The probability that *x* is chosen, given that *x* is correct, converges to 1 as *n* tends to infinity.
(ii) The probability that *x* is not chosen, given that *x* is not correct, converges to 1 as *n* tends to infinity.

As we have also seen, for any finite *n*, the probabilities in (i) and (ii) are less then certainty.

**Problem.** Usually, we are trying to estimate an unobserved parameter: whether or not alternative *x* is correct. Our observed parameter is whether or not alternative *x* is chosen.

So we are interested in the following two Bayesian probabilities:

the probability that *x* is correct, given that *x* is chosen;
and the probability that *x* is not correct, given that *x* is not chosen.

We can rephrase this in the language of testing a hypothesis:

The **hypothesis** is that *x* is correct.

The **evidence** may include the fact as to whether or not *x* has been chosen in a given decision procedure and how many votes were cast in favour of *x* and how many against *x*.



## An example: two situations of jury decisions (‘guilty’ vs. ‘innocent’)

**Assumptions.** In each case,

- there are two states of the world: the defendant is guilty or innocent;
- each juror has a probability *p* of making a correct judgment;
- the judgments of different jurors are independent from each other, given the state of the world;
- we attach a fixed prior probability *r* to the proposition that the defendant is guilty.

We assume that *p* and *r* are the same in both situations.

**Situation A**
12 jurors vote for ‘guilty’ and 0 jurors vote for ‘innocent’

**Situation B**
507 jurors vote for ‘guilty’ and 493 jurors vote for ‘innocent’.

**Question.** In which situation is the given majority verdict more likely to be correct? In other words, which of the following two probabilities is greater:

(a) the probability that the defendant is guilty, given that we have situation A (12:0 for ‘guilty’); or
(b) the probability that the defendant is guilty, given that we have situation B (507:493 for ‘guilty’)?

**Note.**
In situation A, there is a 100% majority for ‘guilty’ (12 out of 12 jurors).
In situation B, there is a 50.7% majority for ‘guilty’ (507 out of 1000 jurors).

## Special majority voting

**(see List 2002)**

In many decisions, there is an asymmetry between a positive decision (e.g. convict) and a negative decision (e.g. acquit).

Unless we have very strong support for the hypothesis that a positive decision is correct, we may wish to be conservative, and prefer to err (if at all) on the negative side.

Thus we may distinguish between

- false positives: choosing $x$ when $x$ is not correct;
- and false negatives: choosing $\neg x$ when $x$ is correct.

**Special majority voting** is often used to implement such a conservative attitude.

**The "proportion" definition (the standard definition).**

**A special majority rule with parameter *q***
(where $^1/_2 \leq q \leq 1$, typically $q > 1/2$).

> A positive decision is reached if and only if the number of individuals supporting a positive decision divided by the total number of individuals $n$ exceeds $q$.

(The limiting case $q = {}^1/_2$ is the case of a simple majority.)

**Examples:**

- In many jury decisions, special majorities of 5/6 or even unanimity are often required for conviction.
- To change the Basic Law of the Federal Republic of Germany, 2/3 majorities in both chambers of parliament, Bundestag and Bundesrat, are required.



## Condorcet's insight

**Condorcet's Formula.** The probability that $x$ is correct, given that $x$ is supported by a majority of a given size is:

$$p^{h-k} / (p^{h-k}+(1-p)^{h-k})$$

where: $h$ number of individuals in majority
$k$ number of individuals in minority

**Implication.** The probability that the majority is correct, given the size of the majority, depends only on the absolute margin between the majority and the minority (i.e. $h$-$k$), but neither on the absolute size of the majority (i.e. $h$), nor on the ratio of the majority to the total size of the electorate (i.e. $h/n$).

I now derive a more general version of Condorcet's formula from Bayes's law and the Condorcet jury theorem.

**Complication.** The probability that the majority is correct, given the size of the majority, depends also on the prior probability $r$ that the specific option supported by that majority is the correct one.

**Example.** In a jury decision, the prior probability $r$ associated with a "guilty" verdict might be the (very small) probability that a randomly chosen member of the population is guilty of the relevant charge.

In Condorcet's formula, implicityly $r = 1/2$.

**Further formalism.**
$r = P(x\text{-}correct)$ : prior probability that option $x$ is correct

**Proposition (List 2002).** *Suppose $h > n/2$. Then*

$$P(x\text{-}correct \mid N_x = h) = \frac{r\, p^m}{r\, p^m + (1-r)\,(1-p)^m},$$

*where $m = 2h-n$ (in Condorcet's terms, $m = h-k$ with $k = n-h$).*



**Proof.** Suppose $h > n/2$. By Bayes's law,

$P(x\text{-}correct \mid N_x = h)$
$= P(x\text{-}correct)\, P(N_x = h \mid x\text{-}correct) \,/\, P(n_x = h)$.

As $N_x$ is binomially distributed,

$$P(N_x = h \mid x\text{-}correct) = \binom{n}{h} p^h\, (1-p)^{n-h}.$$

By elementary probability theory,

$$P(N_x = h) = P(N_x{=}h \mid x\text{-}correct)P(x\text{-}correct) + P(N_x{=}h \mid \neg x\text{-}correct)(1-P(x\text{-}correct)).$$

As $N_{\neg x}$ is binomially distributed,

$$P(N_x = h \mid \neg x\text{-}correct) = \binom{n}{h} (1-p)^h\, p^{n-h}.$$

Hence

$$P(N_x = h) = \binom{n}{h} p^h\, (1-p)^{n-h}\, P(x\text{-}correct) + \binom{n}{h} (1-p)^h\, p^{n-h}\, (1-P(x\text{-}correct)).$$

Therefore

$$P(x\text{-}correct \mid N_x = h) = \frac{P(x\text{-}correct)\binom{n}{h} p^h\,(1-p)^{n-h}}{\binom{n}{h} p^h\,(1-p)^{n-h}\, P(x\text{-}correct) + \binom{n}{h}(1-p)^h\, p^{n-h}\,(1-P(x\text{-}correct))}$$

$$= \frac{r\, p^{h-k}}{r\, p^{h-k} + (1-r)\,(1-p)^{h-k}},$$

where $r = P(x\text{-}correct)$ (prior probability that $x$ is correct)
$k = n-h$ (number of individuals in the minority).

## Determining the special majority required for a given Bayesian justification criterion

**A Bayesian justification criterion.** An agent considers a decision to be epistemically justified *only if* the agent's subjective probability that the decision is correct exceeds a certain fixed threshold $c$.

**"beyond any reasonable doubt" criterion.** Choose $c$ close to 1.

**Idea.** For any prior probability $r$ (that a positive decision is correct) and any individual competence $p$, we can use the proposition above to determine the margin between majority and minority required for a positive decision to meet the Bayesian justification criterion.

$P(x\text{-}correct \mid N_x - N_{not\text{-}x} = m)$ :

probability that alternative $x$ is correct, given that it is supported by a majority with a margin of $m$

By the above proposition, $P(x\text{-}correct \mid N_x - N_{not\text{-}x} = m)$ depends only on $m$, $p$ and $r$.

**Proposition (List 2002).** *Let c be a fixed threshold* $(0 \leq c \leq 1)$. *Then* $P(x\text{-}correct \mid N_x - N_{not\text{-}x} = m) \geq c$ *if and only if*

$$m \geq \frac{log(\frac{r-cr}{c-cr})}{log(^1/_p - 1)}.$$



Sample calculations of the values of $m$, for different values of $p$, $r$ and $c$

**Table. Values of *m* corresponding to different values of *p*, *r* and *c***

| | | $r = 0.001$ | $r = 0.01$ | $r = 0.25$ | $r = 0.4$ | $r = 0.5$ | $r = 0.6$ | $r = 0.75$ |
|---|---|---|---|---|---|---|---|---|
| $p = 0.51$ | $c = 0.5$ | 173 | 115 | 28 | 11 | 0 | 0 | 0 |
| | $c = 0.75$ | 201 | 143 | 55 | 38 | 28 | 18 | 0 |
| | $c = 0.99$ | 288 | 230 | 143 | 125 | 115 | 105 | 88 |
| | $c = 0.999$ | 346 | 288 | 201 | 183 | 173 | 163 | 146 |
| $p = 0.55$ | $c = 0.5$ | 35 | 23 | 6 | 3 | 0 | 0 | 0 |
| | $c = 0.75$ | 40 | 29 | 11 | 8 | 6 | 4 | 0 |
| | $c = 0.99$ | 58 | 46 | 29 | 25 | 23 | 21 | 18 |
| | $c = 0.999$ | 69 | 58 | 40 | 37 | 35 | 33 | 29 |
| $p = 0.6$ | $c = 0.5$ | 18 | 12 | 3 | 1 | 0 | 0 | 0 |
| | $c = 0.75$ | 20 | 15 | 6 | 4 | 3 | 2 | 0 |
| | $c = 0.99$ | 29 | 23 | 15 | 13 | 12 | 11 | 9 |
| | $c = 0.999$ | 35 | 29 | 20 | 19 | 18 | 17 | 15 |
| $p = 0.75$ | $c = 0.5$ | 7 | 5 | 1 | 1 | 0 | 0 | 0 |
| | $c = 0.75$ | 8 | 6 | 2 | 2 | 1 | 1 | 0 |
| | $c = 0.99$ | 11 | 9 | 6 | 5 | 5 | 4 | 4 |
| | $c = 0.999$ | 13 | 11 | 8 | 7 | 7 | 6 | 6 |
| $p = 0.9$ | $c = 0.5$ | 4 | 3 | 1 | 1 | 0 | 0 | 0 |
| | $c = 0.75$ | 4 | 3 | 1 | 1 | 1 | 1 | 0 |
| | $c = 0.99$ | 6 | 5 | 3 | 3 | 3 | 2 | 2 |
| | $c = 0.999$ | 7 | 6 | 4 | 4 | 4 | 3 | 3 |

**(Source: List 2002)**

**Proportions corresponding to absolute margins.** Given an electorate of size $n$, a margin $m$ between the majority and the minority is equivalent to a proportion

$$q = {}^{1}/_{2}({}^{m}/_{n} + 1)$$

of the electorate.

$P(x\text{-}correct \mid N_x/n = q)$ :

probability that alternative $x$ is correct, given that it is supported by a proportion of $q$ of an electorate of size $n$

**Proposition (List 2002).** *Let c be a fixed threshold* ($0 \leq c \leq 1$). *Then* $P(x\text{-}correct|N_x/n = q) \geq c$ *if and only if*

$$q \geq \quad {}^{1}/_{2}\left(\frac{log(\frac{r\text{-}cr}{c\text{-}cr})}{n\ log({}^{1}/_{p} - 1)} + 1\right).$$

**Note.** For any fixed values of $p$, $r$ and $c$, the value of $q$ tends to ${}^{1}/_{2}$ as the number of individuals $n$ increases.



Sample calculations of the values of $q$, for different values of $p$, $r$ and $c$

**Table. Values of $q$ corresponding to different values of $p$, $n$ and $c$, with $r = 0.001$**

| | | $n = 12$ | $n = 50$ | $n = 100$ | $n = 300$ | $n = 500$ | $n = 1000$ | $n = 10000$ |
|---|---|---|---|---|---|---|---|---|
| $p = 0.51$ | $c = 0.5$ | n/a | n/a | n/a | 78.9% | 67.3% | 58.7% | 50.9% |
| | $c = 0.75$ | n/a | n/a | n/a | 83.5% | 70.1% | 60.1% | 51.1% |
| | $c = 0.99$ | n/a | n/a | n/a | 98% | 78.8% | 64.4% | 51.5% |
| | $c = 0.999$ | n/a | n/a | n/a | n/a | 84.6% | 67.3% | 51.8% |
| $p = 0.55$ | $c = 0.5$ | n/a | 85% | 67.5% | 55.9% | 53.5% | 51.75% | 50.2% |
| | $c = 0.75$ | n/a | 90% | 70% | 56.7% | 54% | 52% | 50.2% |
| | $c = 0.99$ | n/a | n/a | 79% | 59.7% | 55.8% | 52.9% | 50.3% |
| | $c = 0.999$ | n/a | n/a | 84.5% | 61.5% | 56.9% | 53.5% | 50.4% |
| $p = 0.6$ | $c = 0.5$ | n/a | 68% | 59% | 53% | 51.8% | 50.9% | 50.1% |
| | $c = 0.75$ | n/a | 70% | 60% | 53.4% | 52% | 51% | 50.1% |
| | $c = 0.99$ | n/a | 79% | 64.5% | 54.9% | 52.9% | 51.5% | 50.2% |
| | $c = 0.999$ | n/a | 85% | 67.5% | 55.9% | 53.5% | 51.8% | 50.2% |
| $p = 0.75$ | $c = 0.5$ | 79.2% | 57% | 53.5% | 51.2% | 50.7% | 50.4% | 50.1% |
| | $c = 0.75$ | 83.4% | 58% | 54% | 51.4% | 50.8% | 50.4% | 50.1% |
| | $c = 0.99$ | 95.9% | 61% | 55.5% | 51.9% | 51.1% | 50.6% | 50.1% |
| | $c = 0.999$ | n/a | 63% | 56.5% | 52.2% | 51.3% | 50.7% | 50.1% |
| $p = 0.9$ | $c = 0.5$ | 66.7% | 54% | 52% | 50.7% | 50.4% | 50.2% | 50.1% |
| | $c = 0.75$ | 66.7% | 54% | 52% | 50.7% | 50.4% | 50.2% | 50.1% |
| | $c = 0.99$ | 75% | 56% | 53% | 51% | 50.6% | 50.3% | 50.1% |
| | $c = 0.999$ | 79.2% | 57% | 53.5% | 51.2% | 50.7% | 50.4% | 50.1% |

**(Source: List 2002)**

## Observations.

- Under the “proportion” definition, for any fixed Bayesian justification criterion, the value of $q$ required for implementing that criterion depends on the size of the electorate and converges to $^1/_2$ as the size of the electorate increases.
- In a large electorate, $q$ approximates $^1/_2$. The epistemic justifiability of special majority voting under the “proportion” definition, with $q$ significantly larger than $^1/_2$, is therefore questionable.
- What matters from a Bayesian perspective is not the proportion of the electorate supporting a positive decision, but rather the absolute margin between the majority and the minority.
- Regardless of how large the electorate is and regardless of how large the majority is in proportional terms, the margin between majority and minority required for implementing a given Bayesian justification criterion remains the same.



## Special majority voting – an alternative definition

**The "absolute margin" definition (see List 2002).**

**A special majority rule with parameter *m* (where $m \geq 0$)**

> A positive decision is reached if and only if the difference between the number of individuals supporting a positive decision and the number of individuals supporting a negative decision exceeds $m$.

(the limiting case $m = 0$ is the case of a simple majority)

## Why can special majority voting under the standard definition be counterproductive?

$P(N_x/n \geq q \mid x\text{-}correct)$ :

probability that alternative $x$ will be supported by a proportion of at least $q$ of an electorate of size $n$, given that $x$ is correct.

**Proposition (List 2002).**

- *If* $^1/_2 < p < q$, *then* $P(N_x/n \geq q \mid x\text{-}correct)$ *converges to* 0 *as n increases*.
- *If* $p > q$, *then* $P(N_x/n \geq q \mid x\text{-}correct)$ *converges to* 1 *as n increases*.

**Note.** Under the "proportion" definition, in a large electorate, the probability that the correct alternative will be selected may thus be very low, unless the competence of individuals is very high (i.e. unless $p>q$). Under simple majority voting, by contrast, $p>1/2$ is sufficient for the probability of the correct alternative being selected to converge to 1.

**However**, special majority voting under the "proportion" definition is an effective method of avoiding 'false positives', i.e. decisions in favour of alternative $x$ when alternative $x$ is not correct (e.g. convicting the innocent).

**But** the avoidance of 'false positives' may come at the expense of almost never obtaining 'true positives' either.

**In short**, this suggests that special majority voting under the "proportion" definition violates the first Nozickian truth-tracking condition, **even in the limit**.



## Why is special majority voting under the alternative definition epistemically sound?

$P(N_x - N_{not\text{-}x} \geq m \mid x\text{-}correct)$ :

probability that $x$ will be supported by a majority with a margin of at least $m$ between the majority and the minority, given that $x$ is correct

**Proposition (List 2002).** *For any* $m > 0$, *if* $p > {}^1/_2$, *then* $P(N_x - N_{not\text{-}x} \geq m \mid x\text{-}correct)$ *converges to* 1 *as n increases*.

**Note.**

- Under the "absolute margin" definition, alternative $x$, if correct, is very likely to obtain the required special majority support in a large electorate, so long as individual competence exceeds 1/2 – *regardless of how large the parameter m is*.
- The previous results already imply that the probability of 'false positives' under the "absolute margin" definition of special majority majority can be made as small as we like, simply by choosing a sufficiently large parameter $m$.
- But, under the "absolute margin" definition of special majority voting, the avoidance of 'false positives' does *not* come at the expense of almost never obtaining 'true positives'.

**In short,** special majority voting under the "absolute margin" definition not only satisfies the Bayesian justification criterion, but it also approximates (and in the limit satisfies) the Nozickian truth-tracking criteria.

## Conclusions

I have argued that …

- If we care about special majorities *because and only because* we care about truth-tracking, then **(in this model)** the "absolute margin" definition is the appropriate definition of special majority voting.

- Adopting that definition would require modifications of those legal documents that prescribe the use of special majority voting.

  For example,

  (1) instead of requiring a *two thirds* majority in a parliament, a margin of, say, 222 votes would be required. If the size of the parliament changes, the "absolute margin" criterion would remain the same.

  (2) For referenda, a special majority criterion would be the required margin between the majority and the minority, independently of the total population size and the size of the majority in proportional terms.

- **(In this model),** if we nonetheless want to defend special majority voting in a large electorate under the "proportion" definition **(maybe for good reasons)**, possibly with $q$ significantly greater than $^1/_2$, then our justification cannot be an epistemic one. If we justify special majority voting in terms of minority protection or legitimacy considerations, then this is an instance of a procedural justification.



## Part 3: Decisions over multiple alternatives

### Generalizing the Condorcet jury theorem to decisions over multiple alternatives (see List and Goodin 2001)

**Problem.** Consider a decision problem, where there are $k$ alternatives, $x_1, x_2, ..., x_k$.

**Difficulty.** There exist several aggregation procedures for decisions over $k$ alternatives, where each is a consistent extension of simple majority voting for two alternatives. Here are some examples (see also Mueller 1989):

**Plurality voting.** Each individual submits one vote. Choose the alternative who receives the largest number of votes.

**The Condorcet winner criterion.** Each individual submits a complete preference ranking over all alternatives. Choose an alternative that beats (or at least ties with) all others in pairwise elections using majority rule.

**Borda count.** Each individual submits a complete preference ranking over all alternatives. Give each of the $k$ alternatives a score of 1 to $k$ based on the alternative's ranking in an individual's preference ordering; i.e., the alternative ranked first receives $k$ points, the second one $k$-l, .., the lowest-ranked alternative one point. Choose an alternative with a maximal score.

**Hare system.** Each individual indicates the alternative he/she ranks highest of the $k$ alternatives. Remove from the list of alternatives the one (or in case of ties, ones) ranked highest by the fewest individuals. Repeat the procedure for the remaining $k$-1 alternatives. Continue until only (at most) one alternative remains. Choose this alternative (if any).

## A procedural critique of plurality voting (some brief remarks)

**(for a good discussion see Riker 1982)**

**Note:** Social-choice-theorists of a procedural variety often reject plurality voting.

In non-technical terms, plurality voting may select an alternative which is **dis**preferred to another alternative by a majority of individuals.

**Example:**

| | 25 individuals | 20 individuals | 15 individuals |
|---|---|---|---|
| 1st preference | $x_1$ | $x_2$ | $x_3$ |
| 2nd preference | $x_2$ | $x_3$ | $x_2$ |
| 3rd preference | $x_3$ | $x_1$ | $x_1$ |

**Note:**

- $x_1$ receives more first choice votes (25) than $x_2$ (20) and $x_3$ (15). Hence $x_1$ wins under plurality voting.
- 35 against 25 individuals prefer each of $x_2$ and $x_3$ to $x_1$. Hence $x_1$, the plurality winner, is pairwise-majority-dispreferred to every other alternative.
- $x_2$ beats every other alternative in pairwise majority comparisons. Hence $x_2$ is the Condorcet winner.

It is often argued that, if there is at all such a thing as a "will of the people", then this "will" is much better captured by the Condorcet winner than by the plurality winner.



## An epistemic justification of plurality voting (List and Goodin 2001)

**Definitions.**

for each $j \in \{1, 2, ..., k\}$:

$x_j$-*correct* : event that alternative $x_j$ is correct

$N_1, N_2, ..., N_k$ : random variables whose values are the numbers of individuals voting for $x_1, x_2, ..., x_k$, respectively.

**Assumptions.**

- There are $n$ individuals. For simplicity, assume that $n$ is indivisible by $k$.
- Precisely one of $x_1$-*correct*, $x_2$-*correct*, ..., $x_k$-*correct* holds. (Which of these $k$ alternatives is correct is called the *state of the world*.)
- Suppose the state of the world is $x_j$. Then each individual $i$ has probabilities $p_1, p_2, ..., p_k$ of voting for alternatives $x_1, x_2, ..., x_k$, respectively, where $\sum_j p_j = 1$ and, for each $h \neq j$,

  $$P(\text{individual } i \text{ votes for } x_j \mid x_j\text{-}correct) > P(\text{individual } i \text{ votes for } x_h \mid x_j\text{-}correct)$$

  Informally, the probability, $pj$, of voting for the "correct" outcome, $j$, exceeds each of the probabilities, $ph$, of voting for any of the "wrong" outcomes, $x_h \in x_j$.
  (The vector $<p_1, p_2, ..., p_k>$ is called individual $i$'s *competence vector*.)
- The votes of different individuals are independent from each other, given the state of the world.

**Proposition (List and Goodin, 2001).**

- $P(N_j > N_h \text{ for all } h \neq j \mid x_j\text{-}correct)$

$$= \sum_{<n_1, n_2, \ldots, n_k> \in W_j} \frac{n!}{n_1!\, n_2! \ldots n_k!} p_1^{n_1} p_2^{n_2} \ldots p_k^{n_k},$$

where $W_j = \{<n_1, n_2, \ldots, n_k> : \sum_i n_i = n$
$\& \; n_j > n_h \text{ for all } h \neq j\}$.

- *If* $p_j > p_h$ for all $h \neq j$, *then*
  $P(N_j > N_h \text{ for all } h \neq j \mid x_j\text{-}correct)$
  $> P(N_{j^*} > N_h \text{ for all } h \neq j^* \mid x_j\text{-}correct)$ for any $j^* \neq j$.

**Theorem (List and Goodin, 2001).** *If* $p_j > p_h$ for all $h \neq j$, *then* $P(N_j > N_h$ for all $h \neq j \mid x_j$-*correct*) *converges to* 1 *as* $n$ *increases*.



**Mechanism.** Again, the convergence result is a consequence of the law of large numbers.

$p_1, p_2, \ldots, p_k$ are the expected frequencies of individuals voting for $x_1, x_2, \ldots, x_k$, respectively ($np_1, np_2, \ldots, np_k$ are the expected numbers).

**For small $n$**, the actual frequencies $N_1/n, N_2/n, \ldots, N_k/n$ may differ substantially from the expected ones.

**Law of large numbers** => As $n$ increases, the actual frequencies $N_1/n, N_2/n, \ldots, N_k/n$ will approximate $p_1, p_2, \ldots, p_k$ increasingly closely.

Hence if $p_j > p_h$ for all $h \neq j$, then $N_j/n$ is increasingly likely to exceed $N_h/n$ for all $h \neq j$ as $n$ increases, and hence $x_j$ is increasingly likely to be the winner under plurality voting.

**Probability that alternative $x_1$ is chosen, given that alternative 1 is correct:**

| ***k*** | ***p*₁, *p*₂, ..., *p*ₖ** | ***n*** | | | | | |
|---|---|---|---|---|---|---|---|
| | | **11** | **51** | **101** | **301** | **601** | **1001** |
| **2** | **0.51, 0.49** | 0.527 | 0.557 | 0.580 | 0.636 | 0.688 | 0.737 |
| | **0.6, 0.4** | 0.753 | 0.926 | 0.979 | ≈1 | ≈1 | ≈1 |
| **3** | **0.34, 0.33, 0.33** | 0.268 | 0.338 | 0.358 | 0.407 | 0.449 | 0.489 |
| | **0.4, 0.35, 0.25** | 0.410 | 0.605 | 0.692 | 0.834 | 0.918 | 0.965 |
| | **0.5, 0.3, 0.2** | 0.664 | 0.937 | 0.987 | ≈1 | ≈1 | ≈1 |
| **4** | **0.26, 0.25, 0.25, 0.24** | 0.214 | 0.266 | 0.296 | 0.361 | 0.420 | 0.476 |
| | **0.4, 0.3, 0.2, 0.1** | 0.512 | 0.770 | 0.873 | 0.980 | 0.998 | ≈1 |
| | **0.5, 0.3, 0.1, 0.1** | 0.708 | 0.939 | 0.987 | ≈1 | ≈1 | ≈1 |
| **5** | **0.21, 0.2, 0.2, 0.2, 0.19** | 0.157 | 0.214 | 0.243 | 0.308 | | |
| | **0.3, 0.2, 0.2, 0.2, 0.1** | 0.360 | 0.653 | 0.812 | 0.980 | | |
| | **0.35, 0.2, 0.15, 0.15, 0,15** | 0.506 | 0.883 | 0.974 | ≈1 | | |

**(Source: List and Goodin 2001)**

**Observation.** If we use plurality voting as the decision procedure (and accept the assumptions of the generalized Condorcet jury model), the following hold:

(i) The probability that $x_j$ is chosen, given that $x_j$ is correct, converges to 1 as $n$ tends to infinity.

(ii) The probability that $x_j$ is not chosen, given that $x_j$ is not correct, converges to 1 as $n$ tends to infinity.

In analogy to the two-alternative case, this suggests that, *in the limit*, plurality voting satisfies the Nozickian truth-tracking criteria.



## Comparing truth-trackers

**Question:** Is plurality voting the unique aggregation procedure that tracks the truth in decisions over $k$ alternatives, or do other proposed aggregation procedures (which are often favoured on procedural grounds) track the truth too?

What about

- the Condorcet winner criterion,
- the Borda count,
- the Hare system?

**(On the Condorcet winner criterion and the Borda count, see also Young 1988.)**

**Note:** There is an informational difference between the plurality rule and the other cited rules: the plurality rule takes single most-preferred alternatives as its input, whereas the other rules take complete preference orderings over all alternatives as their input.

## An extension of the Condorcet jury model to the case of complete preference orderings (see also List and Goodin 2001)

**Definitions.**

$P_1, P_2, ..., P_{k!}$ : the $k$! logically possible strict preference orderings over the $k$ alternatives

$X^*_1, X^*_2, ..., X^*_{k!}$ : random variables whose values are the numbers of individuals submitting the orderings $P_1, P_2, ..., P_{k!}$, respectively.

**Assumptions.**

- Each individual has probabilities $p^*_1, p^*_2, ..., p^*_{k!}$ of submitting $P_1, P_2, ..., P_{k!}$ as his/her preference ordering, respectively (where $\sum_i p^*_i = 1$).
- The submission of preference orderings by different individuals are independent from each other, given the state of the world.

The joint distribution of $X^*_1, X^*_2, ..., X^*_{k!}$ is again a multinomial distribution.

**Note:** To compare the epistemic merits of plurality voting with those of the other aggregation procedures, we need to find a way to convert probability distributions over most preferred options (as specified in the model for plurality voting over $k$ alternatives) into probability distributions over complete preference orderings:

$$p^*_i := \frac{p_{i_1}}{1} * \frac{p_{i_2}}{1-p_{i_1}} * \frac{p_{i_3}}{1-(p_{i_1}+p_{i_2})} * ... * \frac{p_{i_{k-1}}}{1-(p_{i_1}+p_{i_2}+...+p_{i_{k-2}})} * 1.$$



## Probability that alternative $x_1$ is chosen, given that alternative 1 is correct, under various procedures, for $n$=51:

| ***k*** | | Plurality | Condorcet | Borda | Hare | Coombs |
|---|---|---|---|---|---|---|
| **2** | **0.51, 0.49** | 0.557 | 0.557 | 0.557 | 0.557 | 0.557 |
| **3** | **0.60, 0.30, 0.10** | 0.988 | 0.993 | 0.995 | 0.993 | 0.993 |
| **3** | **0.51, 0.25, 0.24** | 0.972 | 0.991 | 0.994 | 0.989 | 0.993 |
| **3** | **0.40, 0.30, 0.30** | 0.666 | 0.740 | 0.760 | 0.737 | 0.775 |
| **3** | **0.34, 0.33, 0.33** | 0.333 | 0.348 | 0.360 | 0.369 | 0.372 |
| **3** | **0.335, 0.3325, 0.3325** | 0.311 | 0.315 | 0.326 | 0.338 | 0.339 |

**(Source: List and Goodin 2001)**

## Conclusions

- Plurality voting, often rejected on procedural grounds, performs surprisingly well from an epistemic perspective.
- Several widely-discussed decision rules perform almost equally well epistemically. Given the assumptions of the Condorcet jury framework about individual competence, they all converge on the same outcome. This indicates that the cases in which these rules produce divergent outcomes are ones in which the assumptions of the Condorcet jury framework are violated.

## An implication of the *k*-option Condorcet jury mechanism for the probability of cycles (List 2001a)

**Condorcet's paradox:** Suppose there are three individuals, 1, 2, 3, with the following preferences over three alternatives, $x$, $y$, z:

Individual 1: $x > y > z$
Individual 2: $y > z > x$
Individual 3: $z > x > y$

There are majorities of 2 out 3 for $x > y$, $y > z$ and $z > x$, a cycle. There exists no Condorcet winner.

**Note:**

- Cyclical preference orderings violate the condition of **transitivity** (if $x > y$ and $y > z$, then $x > z$).
- The possibility of cycles implies that the Condorcet winner criterion is **not** well-defined for all logically possible profiles of individual preference orderings. The Condorcet winner criterion therefore violates the analogue of the condition of unrestricted domain in May's theorem.



## How frequent are cycles?

As soon as we specify a probability distribution over the different possible individual preference orderings, we can determine the probability of the occurrence of a cycle.

**Impartial culture assumption.** Each individual's probabilities $p^*_1, p^*_2, ..., p^*_{k!}$ of submitting $P_1, P_2, ..., P_{k!}$ as his/her preference ordering satisfy $p^*_1 = p^*_2 = ... = p^*_{k!} = 1/k!$.

**Probability of cycles (e.g. Gehrlein, 1983).**
Impartial culture assumption => the probability of the occurrence of a (top-)cycle and of the existence of a Condorcet winner *decreases* with increases in the number of individuals and with increases in the number of alternatives.

What happens if we deviate from the impartial culture assumption?

Consider three-option case as a simple illustration:

Suppose there are $n$ voters/jurors ($n$ odd).

6 logically possible strict orderings of the three options $x$, $y$, $z$:

| *label* | $P_{X1}$ | $P_{Y2}$ | $P_{Z1}$ | $P_{X2}$ | $P_{Y1}$ | $P_{Z2}$ |
|---|---|---|---|---|---|---|
| 1st | $z$ | $z$ | $y$ | $y$ | $x$ | $x$ |
| 2nd | $x$ | $y$ | $z$ | $x$ | $y$ | $z$ |
| 3rd | $y$ | $x$ | $x$ | $z$ | $z$ | $y$ |

Let $n(P_{X1})$, $n(P_{X2})$, $n(P_{Y1})$, $n(P_{Y2})$, $n(P_{Z1})$, $n(P_{Z2})$ be the total numbers of individuals submitting orderings $P_{X1}$, $P_{X2}$, $P_{Y1}$, $P_{Y2}$, $P_{Z1}$, $P_{Z2}$, respectively.

**Proposition.** (Miller 2000) *The anonymous profile* $<n(P_{X1}), n(P_{X2}), n(P_{Y1}), n(P_{Y2}), n(P_{Z1}), n(P_{Z2})>$ *generates a cycle under pairwise majority voting if and only if*

$$\Big( (n(P_{X1}) > n(P_{X2}) \; \& \; n(P_{Y1}) > n(P_{Y2}) \; \& \; n(P_{Z1}) > n(P_{Z2}))$$
$$\text{or } (n(P_{X1}) < n(P_{X2}) \; \& \; n(P_{Y1}) < n(P_{Y2}) \; \& \; n(P_{Z1}) < n(P_{Z2}))\Big)$$
$$\& \; |n(P_{X1}) - n(P_{X2})| < n'/2$$
$$\& \; |n(P_{Y1}) - n(P_{Y2})| < n'/2$$
$$\& \; |n(P_{Z1}) - n(P_{Z2})| < n'/2,$$

*where* $n' = |n(P_{X1}) - n(P_{X2})| + |n(P_{Y1}) - n(P_{Y2})| + |n(P_{Z1}) - n(P_{Z2})|$.

Let $p_{X1}$, $p_{X2}$, $p_{Y1}$, $p_{Y2}$, $p_{Z1}$, $p_{Z2}$ be the probabilities that an individual submits the orderings $P_{X1}$, $P_{X2}$, $P_{Y1}$, $P_{Y2}$, $P_{Z1}$, $P_{Z2}$, respectively. In an *impartial culture* we have $p_{X1} = p_{X2} = p_{Y1} = p_{Y2} = p_{Z1} = p_{Z2}$.

Let $X_{X1}$, $X_{X2}$, $X_{Y1}$, $X_{Y2}$, $X_{Z1}$, $X_{Z2}$ be the random variables whose values are the numbers of voters/jurors with orderings $P_{X1}$, $P_{X2}$, $P_{Y1}$, $P_{Y2}$, $P_{Z1}$, $P_{Z2}$, respectively.

As before, the joint distribution of $X_{X1}$, $X_{X2}$, $X_{Y1}$, $X_{Y2}$, $X_{Z1}$, $X_{Z2}$ is a multinomial distribution.



**Proposition (List 2001a).** *Suppose*

$$[\, [p_{X1} < p_{X2} \text{ or } p_{Y1} < p_{Y2} \text{ or } p_{Z1} < p_{Z2}]$$
$$\& \; [p_{X1} > p_{X2} \text{ or } p_{Y1} > p_{Y2} \text{ or } p_{Z1} > p_{Z2}] \,]$$
$$\text{or } |p_{X1} - p_{X2}| > n'/2$$
$$\text{or } |p_{Y1} - p_{Y2}| > n'/2$$
$$\text{or } |p_{Z1} - p_{Z2}| > n'/2,$$

*where* $n' = |p_{X1} - p_{X2}| + |p_{Y1} - p_{Y2}| + |p_{Z1} - p_{Z2}|$. *Then the probability that there will be* ***no cycle*** *under pairwise majority voting tends to* 1 *as n tends to infinity*.

**Note:** The condition of this proposition is already satisfied if at least one of $p_{X1} < p_{X2}$, $p_{Y1} < p_{Y2}$, $p_{Z1} < p_{Z2}$ and at least one of $p_{X1} > p_{X2}$, $p_{Y1} > p_{Y2}$, $p_{Z1} > p_{Z2}$ are satisfied.

For instance, the condition is satisfied if
$p_{X1} = 1/6 - \varepsilon$, $p_{Y1} = 1/6 + \varepsilon$ and $p_{X2} = p_{Y2} = p_{Z1} = p_{Z2} = 1/6$.

**Conclusion.** Given suitable systematic, however slight, deviations from an impartial culture, the probability that there will be a cycle under pairwise majority voting vanishes as the size of the electorate increases.

# Part 4: Decisions over multiple interconnected propositions

## The "Doctrinal Paradox" or "Discursive Dilemma"

**(Kornhauser and Sager 1986; Kornhauser 1992; Chapman 1998; for a bibliography see http://www.nuff.ox.ac.uk/users/list/doctrinalparadox.htm.)**

proposition $P$ : the defendant did action X
proposition $Q$ : the defendant was contractually obliged not to do action X
proposition $R$ : the defendant is liable

legal doctrine: $(R \leftrightarrow (P \wedge Q))$

**The Doctrinal Paradox (Conjunctive Version)**

| | $P$ | $Q$ | $(R \leftrightarrow (P \wedge Q))$ | $R$ |
|---|---|---|---|---|
| Judge 1 | Yes | Yes | Yes | Yes |
| Judge 2 | Yes | No | Yes | No |
| Judge 3 | No | Yes | Yes | No |
| Majority | Yes | Yes | Yes | No |



**Problem.** Propositionwise majority voting over multiple connected propositions may generate an inconsistent collective set of judgments even when the sets of judgments of all individuals are consistent.

**Why is the study of decisions over multiple connected propositions relevant?**

Decisions over multiple connected propositions occur in many different contexts: e.g.

- a committee has to make a decision that involves the resolution of several premises
- a political party or interest group seeks to come up with an entire policy package, where such a package consists of several interconnected propositions
- a panel of experts is required to give an expert opinion on a set of multiple connected issues, where consistency across issues is relevant.

**Generalization of the doctrinal paradox**

The doctrinal paradox itself concerns only

- a specific profile of individual sets of judgments; and
- a specific aggregation procedure, i.e. propositionwise majority voting.

## Modelling Aggregation over Multiple Propositions (see List and Pettit 2002)

$N$ : set of individuals: 1, 2, 3, …, $n$ (at least two)
$X$ : set of propositions on which judgments are to be made, including
- ‘atomic’ propositions: e.g. $P$ and $Q$
  (e.g. “premises”, “conclusions”)
  (at least two of them)
- compound propositions: e.g. $((P \wedge Q) \rightarrow R)$, $(P \leftrightarrow \neg Q)$
  (e.g. “propositions stating logical interconnections between other propositions”)
  (at least one of them)

(formally, any proposition from prop. calculus)

For each individual $i$:
$\Phi_i$ : set of those propositions in $X$ which individual $i$ accepts
(“individual $i$’s set of judgments over the propositions in $X$”)

**Assumption.** The set of judgments of each individual satisfies certain minimal consistency conditions, namely
- completeness,
- consistency,
- deductive closure.



## Minimal <u>Procedural</u> Conditions on an Acceptable Aggregation Procedure

**Input**
$\Phi_1, \Phi_2, ..., \Phi_n$
(a *profile* of sets of beliefs or judgments across individuals)

aggregation ↓ procedure

**Output**
$\Phi$
(collective set of beliefs or judgments)

**Minimal Consistency Conditions**
“A group should be rational/consistent in the judgments it collectively endorses.”

We formalize this by requiring that the collective set of judgments should also satisfy:
- completeness,
- consistency,
- deductive closure.

## Minimal Responsiveness Conditions

"A collectivity should be responsive to the beliefs or judgments of individuals in forming collective beliefs or judgments."

We formalize this in terms of three conditions:

**UNIVERSAL DOMAIN (U).** An aggregation procedure should accept as admissible input any logically possible profile of individual sets of judgments.

**ANONYMITY (A).** All individuals should have equal weight in determining the collective set of judgments.

**SYSTEMATICITY (S).** The aggregation procedure should treat all propositions in an evenhanded way.

**An impossibility theorem:**

**Theorem (List and Pettit 2002).** There exists no aggregation procedure (generating complete, consistent and deductively closed collective sets of judgments) which satisfies universal domain, anonymity and systematicity.

**Consequence.** Any aggregation procedure will violate at least one of the six minimal conditions introduced above.



## A probabilistic approach

**Question.** How likely is the occurrence of the doctrinal paradox? Or, how likely is the occurrence of profiles of individual sets of judgments which would, under propositionwise majority voting, generate the paradox?

An answer to this question also casts light on an escape-route from the impossibility result via domain restriction: i.e. it tells us how likely it is that a profile of individual sets of judgments will fall into a [restricted] domain in which the impossibility result can be avoided.

## A Necessary and Sufficient Condition for the Occurrence of the Paradox

**Assumptions:**

- $n$ individuals
- three propositions, $P$, $Q$ and $R$
- all individuals accept the connection rule $(R \leftrightarrow (P \wedge Q))$
- each individual holds a consistent set of judgments over $P$, $Q$ and $R$

**Table 2: All logically possible consistent sets of judgments over *P*, *Q* and *R*, given $(R \leftrightarrow (P \wedge Q))$**

| Label | Judgment on $P$ | Judgment on $Q$ | Judgment on $R$ |
|---|---|---|---|
| $PQ$ | Yes | Yes | Yes |
| $P\neg Q$ | Yes | No | No |
| $\neg PQ$ | No | Yes | No |
| $\neg P\neg Q$ | No | No | No |

A collective inconsistency occurs if and only if there are

- majorities for each of $P$ and $Q$, and
- a majority against $R$.



## A Simple Probability-Theoretic Framework

**Assumptions:**

- Each individual has probabilities $p_{PQ}$, $p_{P\neg Q}$, $p_{\neg PQ}$, $p_{\neg P\neg Q}$ of holding the sets of judgments $PQ$, $P\neg Q$, $\neg PQ$, $\neg P\neg Q$, respectively (where $p_{PQ} + p_{P\neg Q} + p_{\neg PQ} + p_{\neg P\neg Q} = 1$);
- the judgments of different individuals are independent from each other.

**Definition.** An *impartial culture*: perfect equiprobability, i.e. $p_{PQ} = p_{P\neg Q} = p_{\neg PQ} = p_{\neg P\neg Q}$.

**Probability that there will be a collective inconsistency under propositionwise majority voting (given $(R \leftrightarrow (P \wedge Q))$), for various scenarios**

| | Scenario 1 $p_{PQ} = 0.25$ $p_{P\neg Q} = 0.25$ $p_{\neg PQ} = 0.25$ $p_{\neg P\neg Q} =$ 0.25 | Scenario 2 $p_{PQ} = 0.26$ $p_{P\neg Q} = 0.25$ $p_{\neg PQ} = 0.25$ $p_{\neg P\neg Q} =$ 0.24 | Scenario 3 $p_{PQ} = 0.3$ $p_{P\neg Q} = 0.25$ $p_{\neg PQ} = 0.25$ $p_{\neg P\neg Q} =$ 0.2 | Scenario 4 $p_{PQ} = 0.24$ $p_{P\neg Q} = 0.27$ $p_{\neg PQ} = 0.25$ $p_{\neg P\neg Q} =$ 0.24 | Scenario 5 $p_{PQ} = 0.49$ $p_{P\neg Q} = 0.2$ $p_{\neg PQ} = 0.2$ $p_{\neg P\neg Q} =$ 0.11 | Scenario 6 $p_{PQ} = 0.51$ $p_{P\neg Q} = 0.2$ $p_{\neg PQ} = 0.2$ $p_{\neg P\neg Q} =$ 0.09 | Scenario 7 $p_{PQ} = 0.55$ $p_{P\neg Q} = 0.2$ $p_{\neg PQ} = 0.2$ $p_{\neg P\neg Q} =$ 0.05 | Scenario 8 $p_{PQ} = 0.33$ $p_{P\neg Q} = 0.33$ $p_{\neg PQ} = 0.33$ $p_{\neg P\neg Q} =$ 0.01 |
|---|---|---|---|---|---|---|---|---|
| $n = 3$ | 0.0938 | 0.0975 | 0.1125 | 0.0972 | 0.1176 | 0.1224 | 0.1320 | 0.2156 |
| $n = 11$ | 0.2157 | 0.2365 | 0.3211 | 0.2144 | 0.3570 | 0.3432 | 0.2990 | 0.6188 |
| $n = 31$ | 0.2487 | 0.2946 | 0.4979 | 0.2409 | 0.5183 | 0.4420 | 0.2842 | 0.9104 |
| $n = 51$ | 0.2499 | 0.3101 | 0.5815 | 0.2405 | 0.5525 | 0.4414 | 0.2358 | 0.9757 |
| $n = 71$ | ≈ 0.2500 | 0.3216 | 0.6417 | 0.2393 | 0.5663 | 0.4327 | 0.1983 | 0.9930 |
| $n = 101$ | ≈ 0.2500 | 0.3362 | 0.7113 | 0.2375 | 0.5798 | 0.4201 | 0.1562 | 0.9989 |
| $n = 201$ | ≈ 0.2500 | 0.3742 | 0.8511 | 0.2317 | 0.6118 | 0.3882 | 0.0774 | ≈ 1.0000 |
| $n = 501$ | ≈ 0.2500 | 0.4527 | 0.9754 | 0.2149 | 0.6729 | 0.3271 | 0.0124 | ≈ 1.0000 |
| $n = 1001$ | ≈ 0.2500 | 0.5426 | 0.9985 | 0.1897 | 0.7366 | 0.2634 | 0.0008 | ≈ 1.0000 |
| $n = 1501$ | ≈ 0.2500 | 0.6097 | 0.9999 | 0.1676 | 0.7808 | 0.2192 | 0.0001 | ≈ 1.0000 |

**(Source: List 2001b)**

**Note.** Slight differences in $p_{PQ}$, $p_{P\neg Q}$, $p_{\neg PQ}$, $p_{\neg P\neg Q}$ trigger substantial differences in the resulting probability that a collective inconsistency will occur.

## Convergence Results

**Proposition 2 (List 2001b).** *Let the connection rule be* $(R \leftrightarrow (P \wedge Q))$.

**(a)** *Suppose* $(p_{PQ} + p_{P\neg Q} > 1/2)$ *and* $(p_{PQ} + p_{\neg PQ} > 1/2)$ *and* $(p_{PQ} < 1/2)$. *Then the probability of a collective inconsistency under propositionwise majority voting converges to* 1 *as n tends to infinity*.

**(b)** *Suppose* $(p_{PQ} + p_{P\neg Q} < 1/2)$ *or* $(p_{PQ} + p_{\neg PQ} < 1/2)$ *or* $(p_{PQ} > 1/2)$. *Then the probability of a collective inconsistency under propositionwise majority voting converges to* 0 *as n tends to infinity*.

Scenarios 2, 3, 5 and 8 in the table in the previous slide satisfy the conditions of proposition 2a.

Scenarios 4, 6 and 7 satisfy the conditions of proposition 2b.



**Mechanism.** Convergence results are a consequence of the law of large numbers.

$np_{PQ}$, $np_{P\neg Q}$, $np_{\neg PQ}$, $np_{\neg P\neg Q}$ are the expected numbers of the 4 different combinations of individual judgments across $n$ individuals.

**For small *n***, the actual numbers may differ substantially from the expected ones.

**Law of large numbers** ➔ As $n$ increases, the actual frequencies will approximate the expected ones increasingly closely.

If the probabilities $p_{PQ}$, $p_{P\neg Q}$, $p_{\neg PQ}$, $p_{\neg P\neg Q}$ satisfy a set of strict inequalities, the actual frequencies are increasingly likely to satisfy a matching set of strict inequalities.

Our necessary and sufficient condition for the occurrence of a collective inconsistency then directly implies the convergence results.

## An Epistemic Framework
## (see also Bovens and Rabinowicz 2001; Pettit and Rabinowicz 2001)

**Simple assumptions** (inspired by the classical Condorcet jury theorem):

- there is an external fact on whether each of $P$ and $Q$ is true (and, by implication, on whether $R$ is true) -- this is called the *state of the world*;
- each individual has probabilities $p$ and $q$ of making a correct judgment on $P$ and $Q$, respectively, where $p$, $q > 0.5$ ("competence");
- each individual's judgments on $P$ and $Q$ are independent from each other, given the state of the world;
- the judgments of different individuals are independent from each other, given the state of the world.

**Proposition.** *Let the connection rule be* $(R \leftrightarrow (P \wedge Q))$.

**(a)** *Suppose P and Q are true*.

- *Suppose* $0.5 < p, q < \sqrt{(0.5)}$. *Then the probability of a collective inconsistency under propositionswise majority voting converges to* 1 *as n tends to infinity*.
- *Suppose* $p, q > \sqrt{(0.5)}$. *Then the probability of a collective inconsistency under propositionwise majority voting converges to* 0 *as n tends to infinity*.

**(b)** *Suppose that not both P and Q are true and* $p, q > 0.5$. *Then the probability of a collective inconsistency under propositionswise majority voting converges to* 0 *as n tends to infinity*.



## Voting for the Premises Versus Voting for the Conclusion (i.e. examples of procedures <u>other than</u> propositionwise majority voting)

**Premise-based procedure.**

- The group applies majority voting on propositions $P$ and $Q$, but not on $R$;
- the connection rule, $(R \leftrightarrow (P \wedge Q))$ dictates the collective judgment on $R$.

**Conclusion-based procedure.** The group applies majority voting directly and only on $R$.

**Note.** The premise-based and conclusion-based procedures may produce divergent outcomes.

**Question.** What is the likelihood that the premise- and conclusion-based procedures reach the correct decision on $R$?

## Tracking the truth for the right reasons versus tracking the truth regardless of reasons (Bovens and Rabinowicz 2001)

**Tracking the truth for the right reasons:** deducing the correct decision on the conclusion (*R*) from correct decisions on each of the premises (*P* and *Q*).

**Tracking the truth regardless of reasons:** includes the possibility of reaching the correct decision on the conclusion (*R*) accidentally, while making a wrong decision on at least one of the premises (*P* or *Q*).



**Proposition (List 2001b).** *Let the connection rule be* ($R \leftrightarrow (P \wedge Q)$). *The probabilities, as n tends to infinity, that the premise- and conclusion-based procedures reach a correct decision on R (i) regardless of reasons and (ii) for the right reasons, under various scenarios, are as shown in table* 6.

See also Bovens and Rabinowicz (2001) for a related set of results.

**Table 6: Probability, as $n$ tends to infinity, of a correct decision on $R$ (given ($R \leftrightarrow (P \wedge Q)$)) under the premise- and conclusion based procedures (i) regardless of reasons and (ii) for the right reasons, under various scenarios**

<table>
<tr><th rowspan="2"></th><th colspan="2">Premise-based procedure: Probability, as n tends to infinity, of …</th><th colspan="2">Conclusion-based procedure: Probability, as n tends to infinity, of …</th></tr>
<tr><th>a correct decision on R regardless of reasons</th><th>a correct decision on R for the right reasons</th><th>a correct decision on R regardless of reasons</th><th>a correct decision on R for the right reasons</th></tr>
<tr><td>0.5 < p, q < √(0.5)<br>P and Q both true</td><td colspan="2" rowspan="3">1</td><td colspan="2">0</td></tr>
<tr><td>0.5 < p, q < √(0.5)<br>not both<br>P and Q true</td><td>1</td><td>0</td></tr>
<tr><td>p, q > √(0.5)</td><td colspan="2">1</td></tr>
</table>

## Conclusions

- We have identified conditions under which the probability of (profiles leading to) collective inconsistencies under propositionwise majority voting converges to 1 and conditions under which it converges to 0.

- Convergence of the probability of the paradox to 1 occurs when all premises are true and individual competence is not particularly high.

- Convergence of the probability of the paradox to 0 occurs when either at least one of the premises is false or individual competence is very high.

- Since decision problems with medium individual competence seem empirically plausible, the occurrence of (profiles leading to) the doctrinal paradox may be quite likely.
  ➔ underlines the importance of identifying escape-routes from the paradox and associated impossibility result.

- The results suggest that in conjunctive decision tasks, the epistemic quality of a collective decision might be improved by disentangling the decision into one on multiple conjuncts -- so long as this ensures that individual competence on each conjunct is greater than individual competence on the conjunction as a whole. (See also Grofman 1985.)



## Appendix: The Probability of Inconsistent Collective Sets of Judgments Compared with the Probability of Cycles

**Condorcet's paradox**

1: $x > y > z$
2: $y > z > x$
3: $z > x > y$

**Impartial culture.** All logically possible individual preference rankings are equally likely to be held by an individual.

**Recent set of results**
(Tangian 2000; Tsetlin, Regenwetter and Grofman 2000; List and Goodin 2001)

- In an impartial culture, the probability of a cycle increases as the number of individuals increases (Gehrlein 1983).
- Given suitable systematic, however slight, deviations from an impartial culture, the probability of a cycle under pairwise majority voting will converge to either 0 or 1 as the number of individuals increases.

**Analogies between the probability of cycles and the probability of inconsistent collective sets of judgments:**
- An impartial culture is a special case, implying a non-zero probability of the paradox.
- Systematic deviations from an impartial culture imply convergence of that probability to either 0 or 1.

Assume (in a Condorcet jury framework) that the probability distribution over all logically possible strict preference orderings is skewed (however slightly) in favour of a preference for the "correct" option over each other option.

(Analogous to the assumption that the probability distribution over all logically possible individual sets of judgments is skewed in favour of the "correct" judgment on each premise.)

Then the probability of a cycle will converge to 0.

By contrast, turning to the aggregation over multiple propositions, suppose the premises $P$ and $Q$ are both true.

If $0.5 < p, q < \sqrt{(0.5)}$, the probability of a collective inconsistency under propositionwise majority voting converges to 1 as the number of individuals increases.



# Bibliography


Bovens, L., and W. Rabinowicz (2001) "Democratic Answers to Complex Questions - an Epistemic Perspective", *Synthese*, forthcoming.

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

Cohen, J. (1986) "An epistemic conception of democracy", *Ethics* 97(1), pp. 6-25.

Grofman, Bernard (1985) "Research Note: The Accuracy of Group Majorities for Disjunctive and Conjunctive Decision Tasks", *Organizational Behavior and Human Decision Processes* 35, pp. 119-123.

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

Kornhauser, L. A. (1992) "Modelling Collegial Courts. II. Legal Doctrine", *Journal of Law, Economics and Organization* 8, pp. 441-470.

List, C. (2001a) "An Implication of the *k*-option Condorcet jury mechanism for the probability of cycles", Appendix 3 of List and Goodin (2001); at http://www.nuff.ox.ac.uk/users/list/probcycles.pdf.

List, C. (2001b) "Complex Collective Decisions and the Probability of Collective Inconsistencies", Nuffield College Working Paper in Economics 2001-W23; at http://www.nuff.ox.ac.uk/economics/papers/2001/w23/doctrinal-wp.pdf.

List, C. (2002) "A New Proposal on Special Majority Voting", Nuffield College Working Paper in Politics 2002-W3;
at http://www.nuff.ox.ac.uk/Politics/papers/2002/w3/specialmajorityprop.pdf.

List, C., and R. E. Goodin (2001) "Epistemic Democracy: Generalizing the Condorcet Jury Theorem", *Journal of Political Philosophy* 9, pp. 277-306.

List, C., and P. Pettit (2002) "Aggregating Sets of Judgments: An Impossibility Result", *Economics and Philosophy* 18, pp. 89-110; at http://www.nuff.ox.ac.uk/users/list/AGG1.pdf.

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

Miller, N. (2000) "The Geometry of Voting Cycles: Theoretical Developments", paper presented at the Workshop on Mathematical Models of Individual and Public Choice, Institute for Mathematical Behavioral Sciences, University of California, Irvine, July 2000.

Mueller, D. C. (1989) *Public Choice II*, Cambridge (Cambridge University Press).

Pettit, P. (2001) "Deliberative Democracy and the Discursive Dilemma", *Philosophical Issues* 11.

Pettit, P., and W. Rabinowicz (2001) "The Jury Theorem and the Discursive Dilemma", Appendix to Pettit (2001).

Tangian, A. S. (2000) "Unlikelihood of Condorcet's paradox in a large society", *Social Choice and Welfare* 17, pp. 337-365.

Tsetlin, I., M. Regenwetter, and B. Grofman (2000) "The Impartial Culture maximizes the probability of majority cycles", Unpublished manuscript, Duke University, Fuqua School of Business.

Young, H. P. (1988) "Condorcet's theory of voting", *American Political Sciene Review* 82, pp. 1231-1244.