Logics for Social Choice Theory

Eric Pacuit, University of Maryland

Lecture 5

ESSLLI 2022

Social choice correspondence

A **voting method** is a function F on the domain of all profiles such that for any profile P, $\varnothing \neq F(\mathrm{P}) \subseteq X(\mathrm{P})$ (also called a **variable social choice correspondence** VSCC).

- A (V, X)-SCC is a social choice correspondence defined on (V, X)-profiles.
- A voting method F is **resolute** if for all P, $|F(\mathrm{P})| = 1$. Resolute SCCs are called **social choice functions**.

Social choice correspondence

A **voting method** is a function F on the domain of all profiles such that for any profile P, $\varnothing \neq F(\mathrm{P}) \subseteq X(\mathrm{P})$ (also called a **variable social choice correspondence** VSCC).

- A (V, X)-SCC is a social choice correspondence defined on (V, X)-profiles.
- A voting method F is **resolute** if for all P, $|F(\mathrm{P})| = 1$. Resolute SCCs are called **social choice functions**.

There are many examples of voting methods.

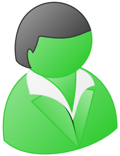
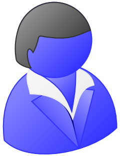
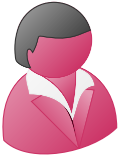
See `https://pref_voting.readthedocs.io` for a Python package that provides computational tools to study different voting methods.

We are interested in voting methods that:

1. respond in a reasonable way to **new voters** joining the election;
2. respond in a reasonable way to **new candidates** joining the election.

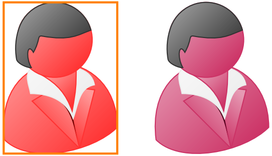
Stability for Winners

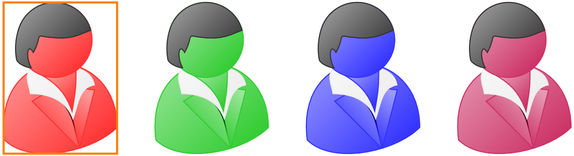
Stability for Winners

If  wins   and  wins 

Stability for Winners

If  wins

and  wins

then  wins .

Stability for Winners

Definition

A voting method F satisfies **Stability for Winners** if for any profile P and $a, b \in X(\mathrm{P})$, if $a \in F(\mathrm{P}_{-b})$ and $Margin_{\mathrm{P}}(a, b) > 0$, then $a \in F(\mathrm{P})$.

Stability for Winners

Definition

A voting method F satisfies **Stability for Winners** if for any profile P and $a, b \in X(\mathrm{P})$, if $a \in F(\mathrm{P}_{-b})$ and $Margin_{\mathrm{P}}(a, b) > 0$, then $a \in F(\mathrm{P})$.

Spoilers, Stealers

Definition

Let F be a voting method, $\mathsf{P} \in \mathrm{dom}(F)$, and $a, b \in X(\mathsf{P})$. Then we say that:

1. *b spoils the election for a in* P
   if $a \in F(\mathsf{P}_{-b})$, $Margin_{\mathsf{P}}(a, b) > 0$, $a \notin F(\mathsf{P})$, and $b \notin F(\mathsf{P})$;
2. *b steals the election from a in* P
   if $a \in F(\mathsf{P}_{-b})$, $Margin_{\mathsf{P}}(a, b) > 0$, $a \notin F(\mathsf{P})$, and $b \in F(\mathsf{P})$.

Spoilers, Stealers

Definition

Let F be a voting method.

1. F satisfies *immunity to spoilers* if for $\mathrm{P} \in \mathrm{dom}(F)$ and $a, b \in X(\mathrm{P})$, b does not spoil the election for a.
2. F satisfies *immunity to stealers* if for $\mathrm{P} \in \mathrm{dom}(F)$ and $a, b \in X(\mathrm{P})$, b does not steal the election from a.
3. F satisfies *stability for winners* if for $\mathrm{P} \in \mathrm{dom}(F)$ and $a, b \in X(\mathrm{P})$, if $a \in F(\mathrm{P}_{-b})$ and $Margin_{\mathrm{P}}(a, b) > 0$, then $a \in F(\mathrm{P})$.

	Split Cycle	Ranked Pairs	Beat Path	Mini-max	Copeland	GETCHA /GOCHA	Uncov. Set	Ranked Choice	Plurality
Immunity to Spoilers	✓	–	–	✓	✓	✓	✓	–	–
Immunity to Stealers	✓	✓*	–	–	–	✓	✓	–	–
Stability for Winners	✓	–	–	–	–	✓	✓	–	–
Expansion Consistency	✓	–	–	–	–	✓/–	✓†	–	–

W. Holliday and EP. *Split Cycle: A New Condorcet Consistent Voting Method Independent of Clones and Immune to Spoilers*. https://arxiv.org/abs/2004.02350, 2021.

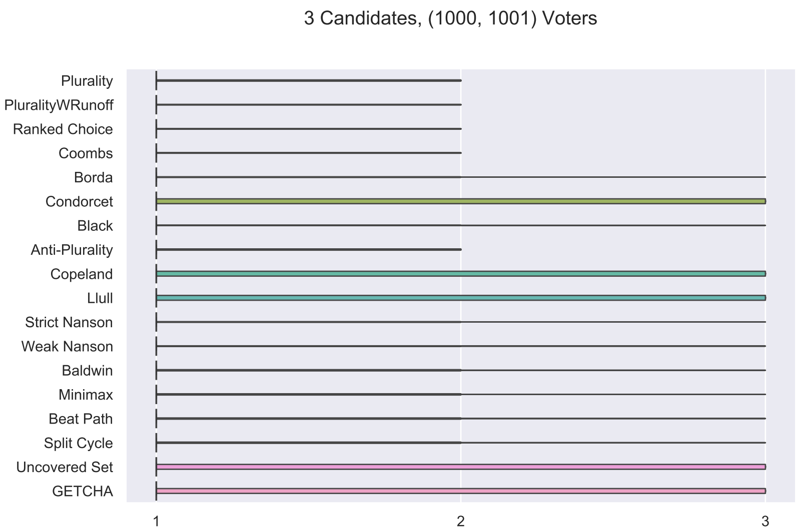
Violations of Quasi-Resoluteness

The known methods that satisfy Binary Expansion violate Asymptotic Resolvability/Quasi-Resoluteness.

Voting Method	3	4	5	6	7	8	9	10	20	30
Split Cycle	1	1.01	1.03	1.06	1.08	1.11	1.14	1.16	1.42	1.62
Uncovered Set	1.17	1.35	1.53	1.71	1.9	2.09	2.26	2.46	4.56	6.82
Top Cycle	1.17	1.44	1.8	2.21	2.72	3.31	3.94	4.68	13.55	22.94

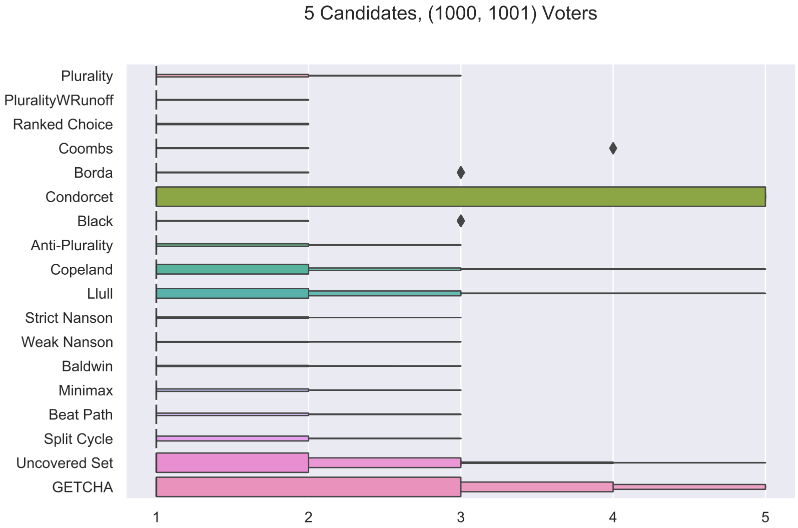
Figure: Estimated **average sizes of winning sets** for profiles with a given number of candidates (top row) in the limit as the number of voters goes to infinity, obtained using the Monte Carlo simulation technique in M. Harrison Trainor, "An Analysis of Random Elections with Large Numbers of Voters," arXiv:2009.02979.

Sizes of Winning Sets



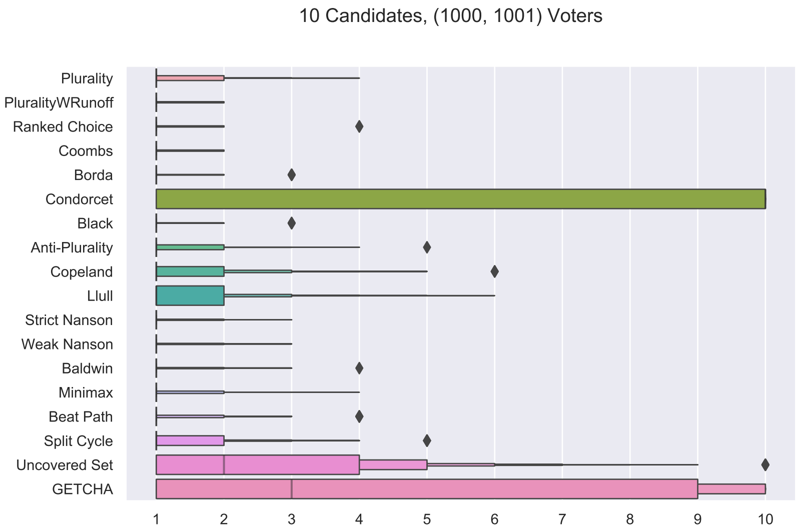
Sizes of Winning Sets

Sizes of Winning Sets

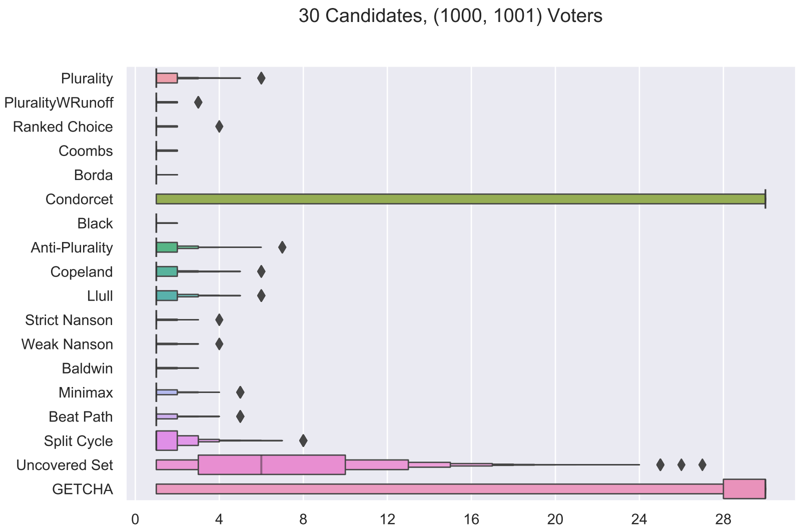


Sizes of Winning Sets

Sizes of Winning Sets



Sizes of Winning Sets



The Cost of Quasi-Resoluteness

Theorem (W. Holliday, EP, and S. Zahedian)

There is no Anonymous and Neutral voting method that satisfies Binary Expansion and Quasi-Resoluteness.

Moral: Making room for tiebreaking (runoff, lottery, etc.) is necessary and sufficient to find voting methods that satisfy Binary Expansion.

High-level idea of proof:

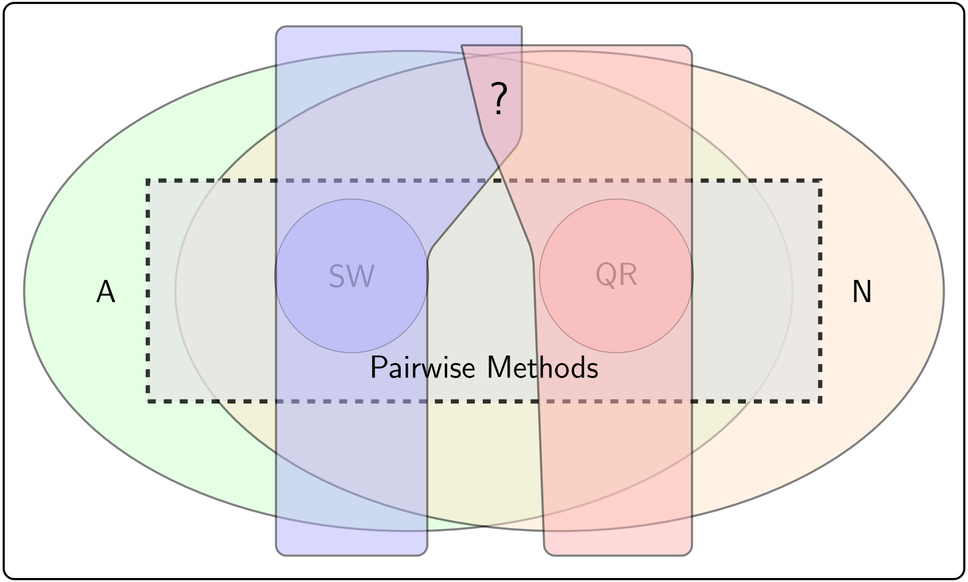
High-level idea of proof:

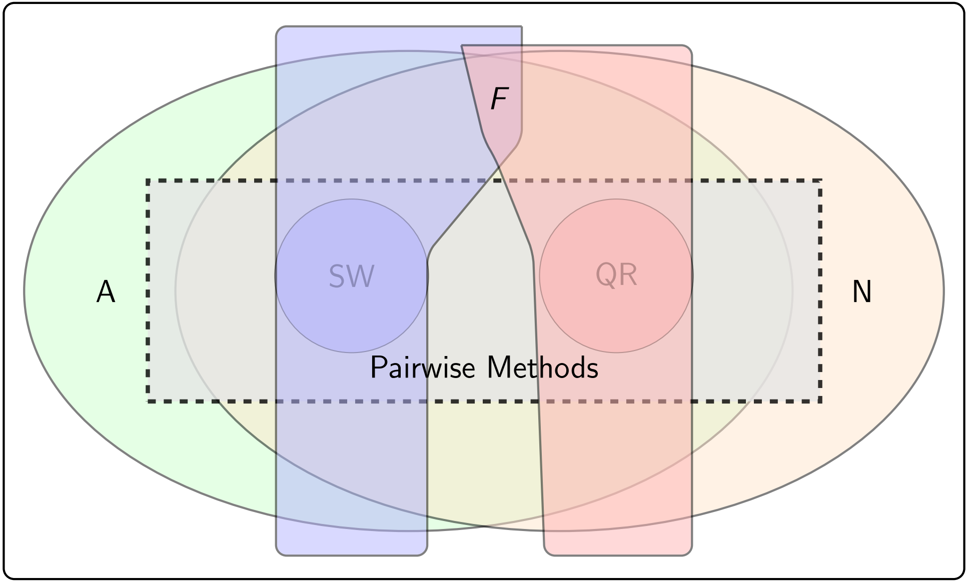
We first use a SAT solver to prove that there is no pairwise voting method, i.e., voting method that outputs the same set of winners for any two profiles whose margin graphs are the same, satisfying the stated axioms.

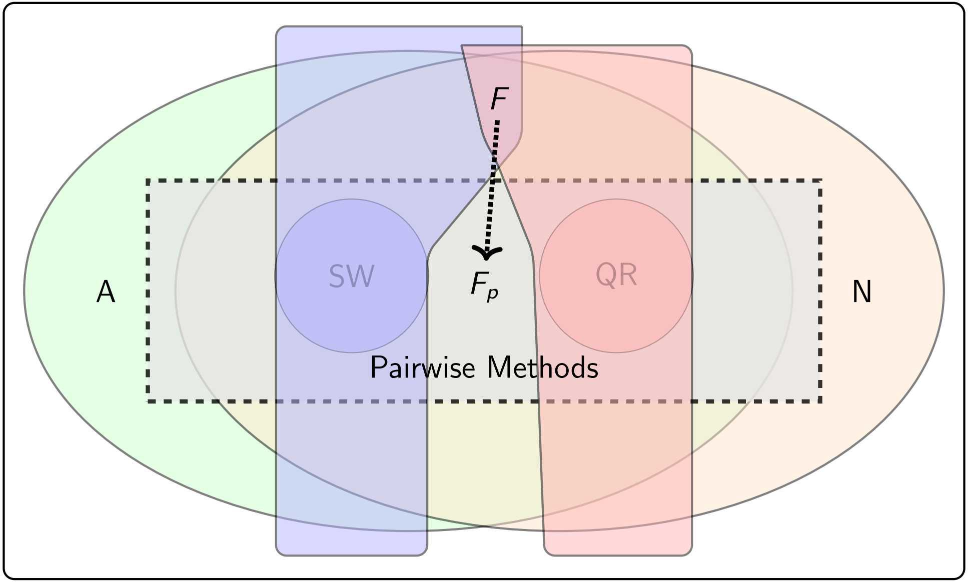
High-level idea of proof:

We first use a SAT solver to prove that there is no pairwise voting method, i.e., voting method that outputs the same set of winners for any two profiles whose margin graphs are the same, satisfying the stated axioms.

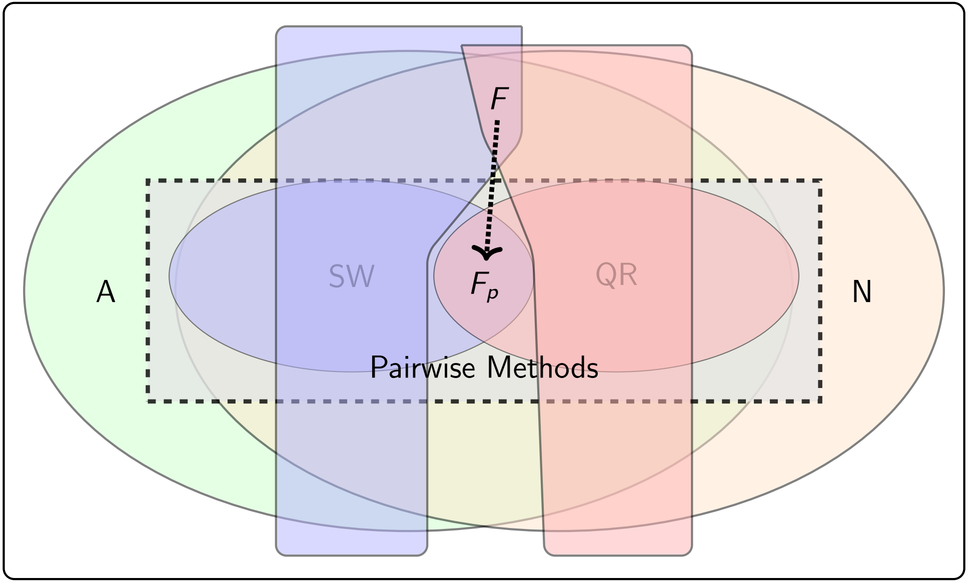
We then apply a theory that allows us to transfer impossibility theorems involving certain kinds of axioms from pairwise voting methods to all voting methods.


?
A
SW
QR
N
Pairwise Methods

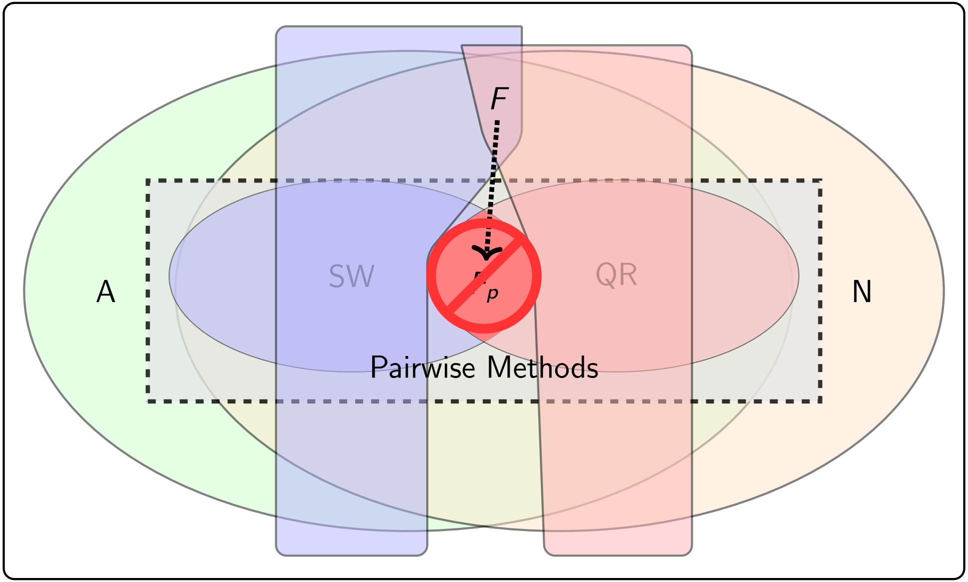

F
A
SW
QR
N
Pairwise Methods



F
SW
F_p
QR
A
N
Pairwise Methods



F
SW
F_p
QR
A
N
Pairwise Methods

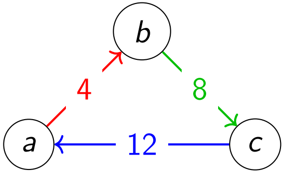


F
SW
QR
A
N
F_p
Pairwise Methods

Leveraging pairwise impossibilities for general impossibilities

Theorem (McGarvey 1953, Debord 1987)

Every weak tournament is the majority graph of a profile, and every weighted weak tournament in which all weights have the same parity is the margin graph of a profile.



$$\Big\Downarrow \ \mathrm{Deb}(\cdot)$$

2	2	4	4	6	6
a	c	b	a	c	b
b	a	c	b	a	c
c	b	a	c	b	a

$$
\begin{tikzcd}
\mathsf{P} \arrow[r, mapsto] \arrow[d, Rightarrow, "\mathcal{M}(\cdot)"'] & G(\mathcal{M}(\mathsf{P})) \\
\mathcal{M}(\mathsf{P}) \arrow[r, mapsto, "G"] & G(\mathcal{M}(\mathsf{P})) \arrow[u, Rightarrow]
\end{tikzcd}
\qquad
\begin{tikzcd}
\mathsf{Deb}(\mathbb{T}) \arrow[r, mapsto, "F"] & F(\mathsf{Deb}(\mathbb{T})) \arrow[d, Rightarrow] \\
\mathbb{T} \arrow[u, Rightarrow, "\mathsf{Deb}(\cdot)"] \arrow[r, mapsto] & F(\mathsf{Deb}(\mathbb{T}))
\end{tikzcd}
$$

From weighted tournaments to pairwise methods

Corollary

1. *There is a neutral and anonymous voting method that is* **pairwise** *and satisfies quasi-resoluteness and stability for winners on the domain* $\{\mathrm{P} \mid |X(\mathrm{P})| \leq 3\}$.
2. *For any* $Y \subset \mathcal{X}$ *with* $|Y| \geq 4$, *there is no neutral and anonymous voting method that is* **pairwise** *and satisfies quasi-resoluteness and stability for winners on the domain*

$$\{\mathrm{P} \mid X(\mathrm{P}) \subseteq Y, \mathcal{M}(\mathrm{P}) \textit{ is uniquely weighted, and all positive weights belong to } \{2,4,6,8,10,12\}\}.$$

There is a problem going beyond pairwise methods: The Debord/McGarvey constructions do not commute with transposition/restriction...

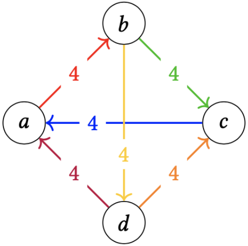
If F is Neutral, then G where $G(\mathbb{T}) = F(\mathrm{Deb}(\mathbb{T}))$ is Neutral?

If F is Neutral, then G where $G(\mathbb{T}) = F(\mathrm{Deb}(\mathbb{T}))$ is Neutral?

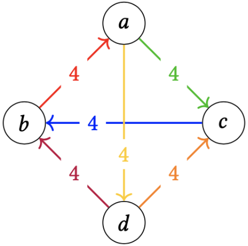
$$
\begin{tikzcd}
\mathrm{Deb}(\mathbb{T}) \arrow[r, mapsto, "F"] & F(\mathrm{Deb}(\mathbb{T})) \arrow[d, Rightarrow] \\
\mathbb{T} \arrow[u, Rightarrow, "\mathrm{Deb}(\cdot)"] \arrow[r, mapsto] \arrow[d, "(\cdot)_{a \leftrightarrows b}"'] & F(\mathrm{Deb}(\mathbb{T})) \arrow[d, "(\cdot)_{a \leftrightarrows b}\ ??"'] \\
\mathbb{T}_{a \leftrightarrows b} \arrow[r, mapsto] \arrow[d, Rightarrow] & F(\mathrm{Deb}(\mathbb{T}_{a \leftrightarrows b})) \\
\mathrm{Deb}(\mathbb{T}_{a \leftrightarrows b}) \arrow[r, mapsto, "F"] & F(\mathrm{Deb}(\mathbb{T}_{a \leftrightarrows b})) \arrow[u, Rightarrow]
\end{tikzcd}
$$

If F is Neutral, then G where $G(\mathbb{T}) = F(\mathrm{Deb}(\mathbb{T}))$ is Neutral?

$$
\begin{tikzcd}
\mathrm{Deb}(\mathbb{T}) \arrow[r, mapsto, "F"] \arrow[ddd, bend right=60, "??(\cdot)_{a \leftrightarrows b}??"'] & F(\mathrm{Deb}(\mathbb{T})) \arrow[d, Rightarrow] \arrow[ddd, bend left=60, "??(\cdot)_{a \leftrightarrows b}??"] \\
\mathbb{T} \arrow[r, mapsto] \arrow[u, Rightarrow, "\mathrm{Deb}(\cdot)"] \arrow[d, "(\cdot)_{a \leftrightarrows b}"'] & F(\mathrm{Deb}(\mathbb{T})) \arrow[d, "(\cdot)_{a \leftrightarrows b}"', "??"] \\
\mathbb{T}_{a \leftrightarrows b} \arrow[r, mapsto] \arrow[d, Rightarrow] & F(\mathrm{Deb}(\mathbb{T}_{a \leftrightarrows b})) \\
\mathrm{Deb}(\mathbb{T}_{a \leftrightarrows b}) \arrow[r, mapsto, "F"] & F(\mathrm{Deb}(\mathbb{T}_{a \leftrightarrows b})) \arrow[u, Rightarrow]
\end{tikzcd}
$$



$\overset{a \leftrightarrows b}{\Longrightarrow}$



$\Downarrow_{\mathrm{Deb}(\cdot)}$

$\Downarrow_{\mathrm{Deb}(\cdot)}$

2	2	2	2	2	2	2	2	2	2	2	2
b	d	a	d	c	d	d	c	d	b	a	c
a	c	c	b	b	a	b	a	c	a	d	b
c	b	b	a	a	c	a	d	a	d	b	a
d	a	d	c	d	b	c	b	b	c	c	d

$\not\Updownarrow$

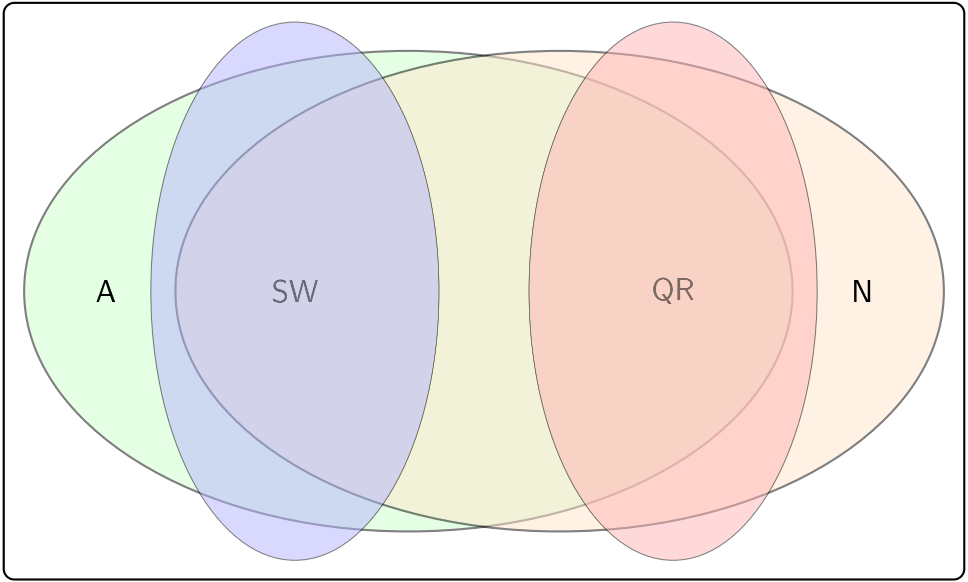
2	2	2	2	2	2	2	2	2	2	2	2
a	d	b	d	c	d	d	c	d	b	b	c
b	c	c	a	a	b	a	b	c	a	d	a
c	a	a	b	b	c	b	d	a	d	a	b
d	b	d	c	d	a	c	a	b	c	c	d

$\overset{a \leftrightarrows b}{\Longrightarrow}$

2	2	2	2	2	2	2	2	2	2	2	2
b	d	a	d	c	d	d	c	d	a	a	c
a	c	c	b	b	a	b	a	c	b	d	b
c	b	b	a	a	c	a	d	b	d	b	a
d	a	d	c	d	b	c	b	a	c	c	d

$$\mathcal{M}(\mathrm{Deb}(\mathbb{T})_{a\leftrightarrows b}) = \mathcal{M}(\mathrm{Deb}(\mathbb{T}_{a\leftrightarrows b}))$$

$$\begin{tikzcd}
\mathrm{Deb}(\mathbb{T}) \arrow[r, mapsto, "F"] \arrow[ddd, bend right] & F(\mathrm{Deb}(\mathbb{T})) \arrow[d, Rightarrow] \arrow[ddd, bend left, "(\cdot)_{a\leftrightarrows b}"] \\
\mathbb{T} \arrow[u, Rightarrow, "\mathrm{Deb}(\cdot)"] \arrow[r, mapsto] \arrow[d, "(\cdot)_{a\leftrightarrows b}"'] & F(\mathrm{Deb}(\mathbb{T})) \arrow[d, "(\cdot)_{a\leftrightarrows b}"'] \\
\mathbb{T}_{a\leftrightarrows b} \arrow[r, mapsto] \arrow[d, Rightarrow] & F(\mathrm{Deb}(\mathbb{T}_{a\leftrightarrows b})) \\
\mathrm{Deb}(\mathbb{T}_{a\leftrightarrows b}) \arrow[r, mapsto, "F"] & F(\mathrm{Deb}(\mathbb{T}_{a\leftrightarrows b})) \arrow[u, Rightarrow]
\end{tikzcd}$$



All Voting Methods

Multiple claims based on stability

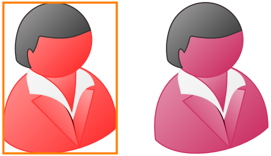
The basic problem is that inevitably there are profiles with *multiple* candidates who have the same kind of claim to winning based on stability for winners:

Multiple claims based on stability

The basic problem is that inevitably there are profiles with *multiple* candidates who have the same kind of claim to winning based on stability for winners:

wins

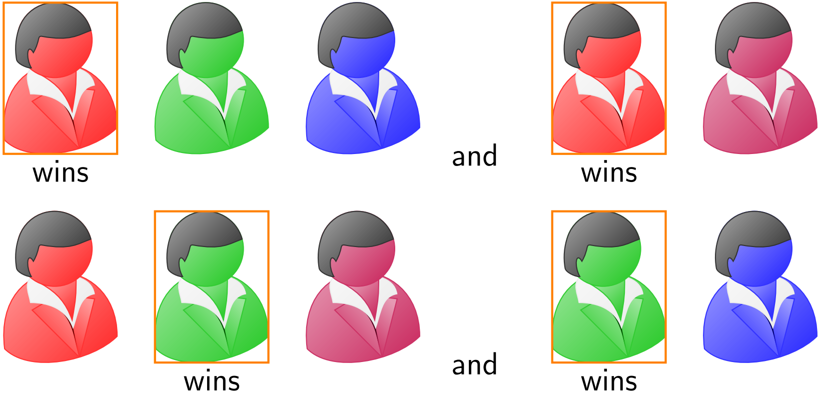
and



wins

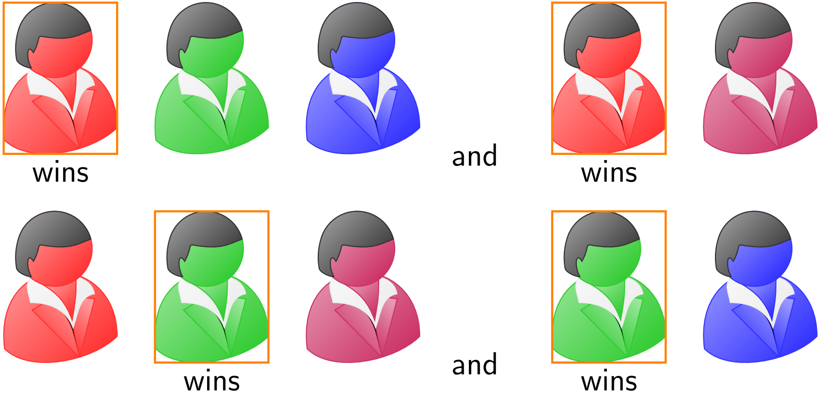
Multiple claims based on stability

The basic problem is that inevitably there are profiles with *multiple* candidates who have the same kind of claim to winning based on stability for winners:



Multiple claims based on stability

The basic problem is that inevitably there are profiles with *multiple* candidates who have the same kind of claim to winning based on stability for winners:

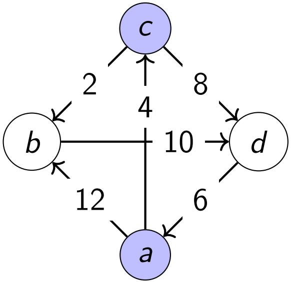


In such a situation—and only such a situation—it is legitimate to violate stability for winners for one of red or green in the name of tiebreaking between them.

Condorcetian candidates

Definition

Given a voting method F, profile P, and $a \in X(\mathrm{P})$, we say that *a is Condorcetian for F in* P if there is some $b \in X(\mathrm{P})$ such that $a \in F(\mathrm{P}_{-b})$ and $Margin_{\mathrm{P}}(a, b) > 0$.



- There are two Condorcetian candidates a and c
- Beat Path elects c
- Ranked Pairs elects a

Stability for Winners with Tiebreaking

Definition

A voting method satisfies **Stability for Winners with Tiebreaking** if for any profile $\mathbf{P}$ and $a, b \in X(\mathbf{P})$, if a wins in $\mathbf{P}_{-b}$ and $Margin_{\mathbf{P}}(a, b) > 0$,

Stability for Winners with Tiebreaking

Definition

A voting method satisfies **Stability for Winners with Tiebreaking** if for any profile $\mathbf{P}$ and $a, b \in X(\mathbf{P})$, if a wins in $\mathbf{P}_{-b}$ and $Margin_{\mathbf{P}}(a, b) > 0$, then either

Stability for Winners with Tiebreaking

Definition

A voting method satisfies **Stability for Winners with Tiebreaking** if for any profile P and $a, b \in X(\mathrm{P})$, if a wins in P_{-b} and $Margin_{\mathrm{P}}(a, b) > 0$, then either

- a wins in P or

Stability for Winners with Tiebreaking

Definition

A voting method satisfies **Stability for Winners with Tiebreaking** if for any profile $\mathbf{P}$ and $a, b \in X(\mathbf{P})$, if a wins in $\mathbf{P}_{-b}$ and $Margin_{\mathbf{P}}(a, b) > 0$, then either

- a wins in $\mathbf{P}$ or
- there are $a', b' \in X(\mathbf{P})$ such that a' wins in $\mathbf{P}_{-b'}$, $Margin_{\mathbf{P}}(a', b') > 0$, and a' wins in $\mathbf{P}'$.

That is, all winners are Condorcetian.

Recursion to the Rescue: Stable Voting

Our proposed voting method is Stable Voting, defined *recursively* as follows:

Recursion to the Rescue: Stable Voting

Our proposed voting method is Stable Voting, defined *recursively* as follows:

- If only one candidate a appears on all ballots, then a wins.

Recursion to the Rescue: Stable Voting

Our proposed voting method is Stable Voting, defined *recursively* as follows:

- If only one candidate a appears on all ballots, then a wins.
- Otherwise list all head-to-head matches a vs. b, where a is undefeated according to Split Cycle, in order from the largest to the smallest margin of a vs. b.

Recursion to the Rescue: Stable Voting

Our proposed voting method is Stable Voting, defined *recursively* as follows:

- If only one candidate a appears on all ballots, then a wins.
- Otherwise list all head-to-head matches a vs. b, where a is undefeated according to Split Cycle, in order from the largest to the smallest margin of a vs. b.

  Find the first match such that a wins according to Stable Voting *after b is removed from all ballots*; this a is the winner for the original set of ballots.

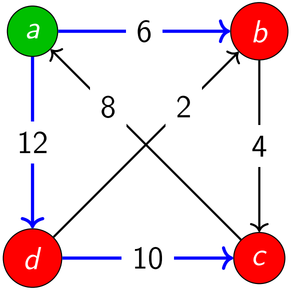
Recursion to the Rescue: Stable Voting

Our proposed voting method is Stable Voting, defined *recursively* as follows:

- If only one candidate a appears on all ballots, then a wins.
- Otherwise list all head-to-head matches a vs. b, where a is undefeated according to Split Cycle, in order from the largest to the smallest margin of a vs. b.

  Find the first match such that a wins according to Stable Voting *after b is removed from all ballots*; this a is the winner for the original set of ballots.

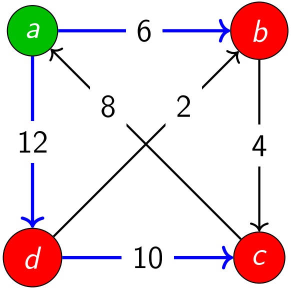
W. Holliday and EP. *Stable Voting*. arXiv:2108.00542 [econ.TH].



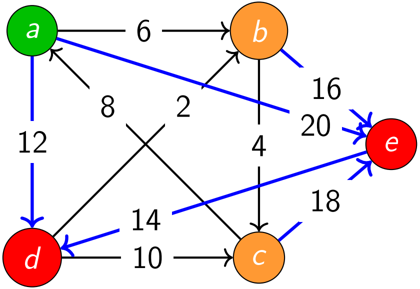
Stable Voting winner: a

Beat Path winner: a

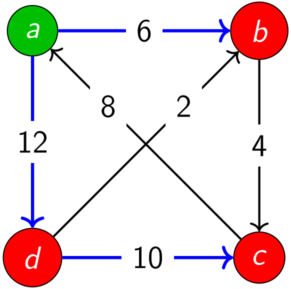
Ranked Pairs winner: a

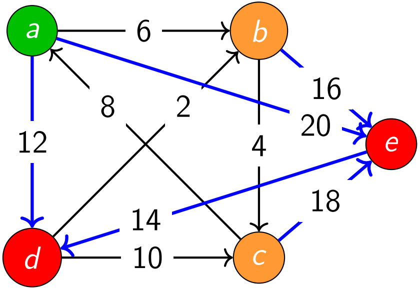


Stable Voting winner: *a*
Beat Path winner: *a*
Ranked Pairs winner: *a*



Stable Voting winner: *a*
Beat Path winner: *b*
Ranked Pairs winner: *c*

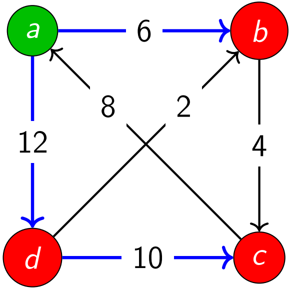




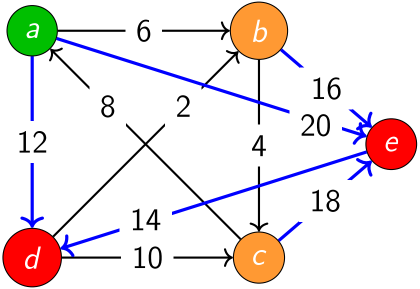
Stable Voting winner: a
Beat Path winner: a
Ranked Pairs winner: a

Stable Voting winner: a
Beat Path winner: b
Ranked Pairs winner: c

On the right, SV chooses the winner by going down the list of matches:



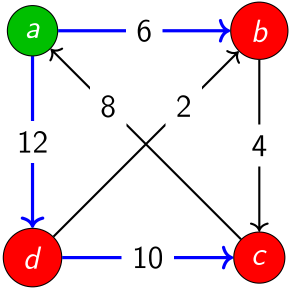
Stable Voting winner: a
Beat Path winner: a
Ranked Pairs winner: a



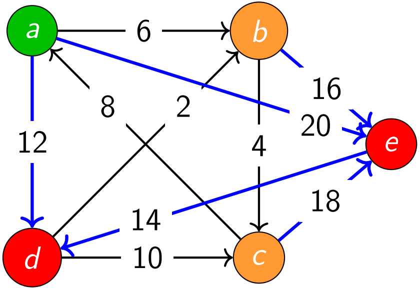
Stable Voting winner: a
Beat Path winner: b
Ranked Pairs winner: c

On the right, SV chooses the winner by going down the list of matches:

- a vs. e: margin of 20.



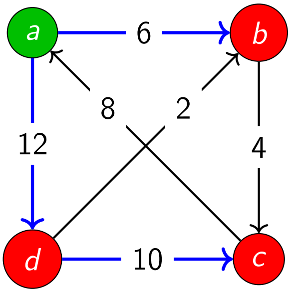
Stable Voting winner: a
Beat Path winner: a
Ranked Pairs winner: a



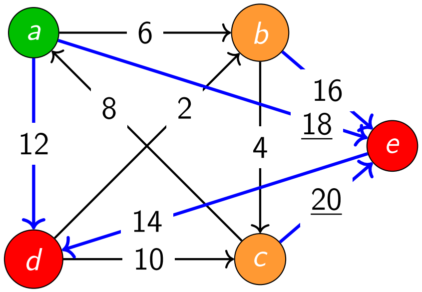
Stable Voting winner: a
Beat Path winner: b
Ranked Pairs winner: c

On the right, SV chooses the winner by going down the list of matches:

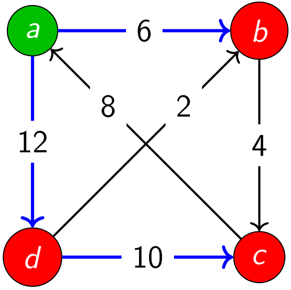
- a vs. e: margin of 20.
  a wins after removing e. Hence **a is elected**.



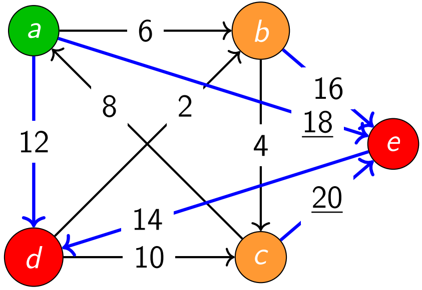
Stable Voting winner: a



Stable Voting winner: a

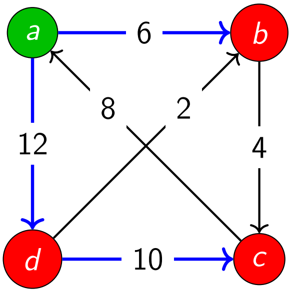


Stable Voting winner: a

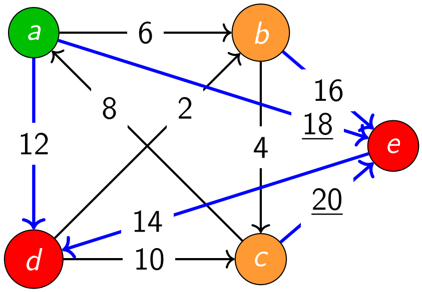


Stable Voting winner: a

On the right, SV chooses the winner by going down the list of matches:



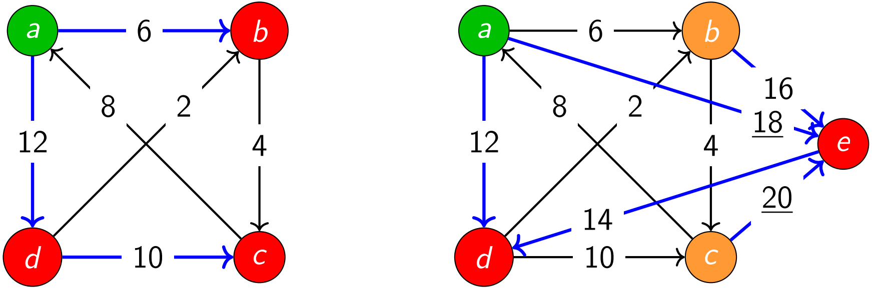
Stable Voting winner: a



Stable Voting winner: a

On the right, SV chooses the winner by going down the list of matches:

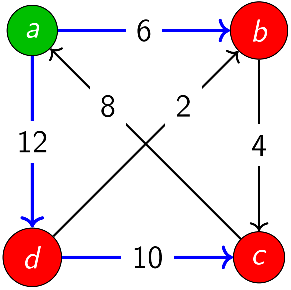
- c vs. e: margin of 20.



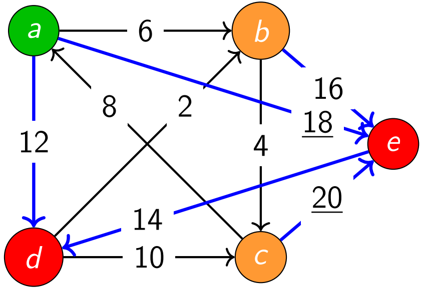
Stable Voting winner: *a* Stable Voting winner: *a*

On the right, SV chooses the winner by going down the list of matches:

- ***c*** vs. ***e***: margin of 20.
  c loses after removing ***e***. Continue to next match:



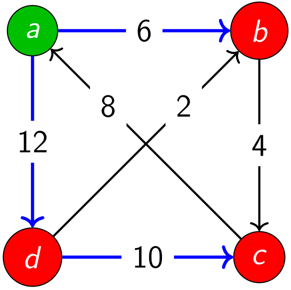
Stable Voting winner: a



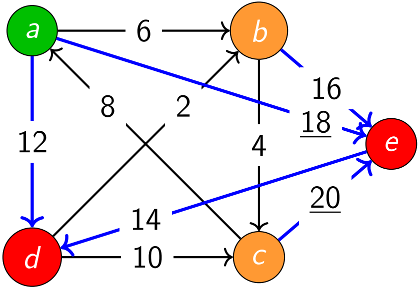
Stable Voting winner: a

On the right, SV chooses the winner by going down the list of matches:

- c vs. e: margin of 20.
  c loses after removing e. Continue to next match:
- a vs. e: margin of 18.



Stable Voting winner: a



Stable Voting winner: a

On the right, SV chooses the winner by going down the list of matches:

- ***c*** vs. ***e***: margin of 20.
  c loses after removing ***e***. Continue to next match:
- ***a*** vs. ***e***: margin of 18.
  a wins after removing ***e***. Hence ***a* is elected**.

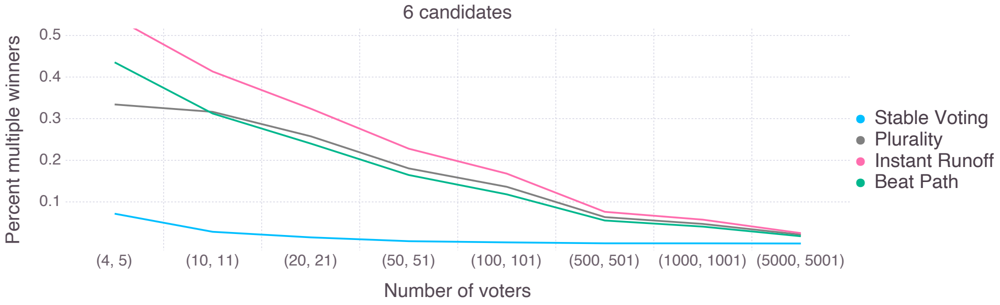
Stable Voting

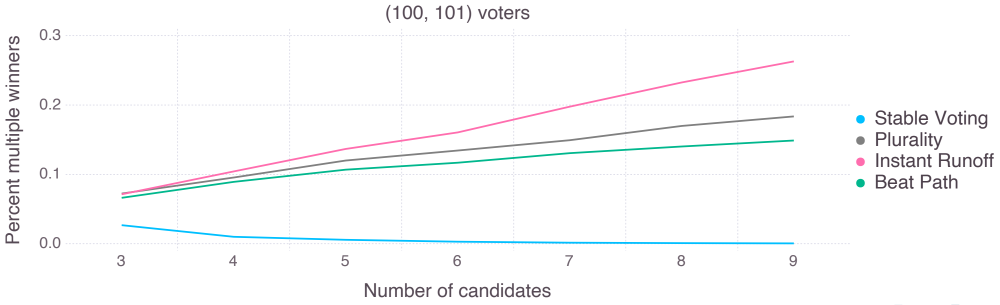
Good news: Stable Voting satisfies **Stability for Winners with Tiebreaking** and **Quasi-resoluteness**.

Stable Voting

Good news: Stable Voting satisfies **Stability for Winners with Tiebreaking** and **Quasi-resoluteness**.

In fact, SV has a remarkable ability to avoid ties even in elections with small numbers of voters that can produce tied margins.


6 candidates
Percent multiple winners
0.5
0.4
0.3
0.2
0.1
(4, 5)
(10, 11)
(20, 21)
(50, 51)
(100, 101)
(500, 501)
(1000, 1001)
(5000, 5001)
Number of voters
Stable Voting
Plurality
Instant Runoff
Beat Path


(100, 101) voters
Percent multiple winners
0.3
0.2
0.1
0.0
3
4
5
6
7
8
9
Number of candidates
Stable Voting
Plurality
Instant Runoff
Beat Path

A logic for resolute social choice correspondences

G. Cinà and U. Endriss. *Proving classical theorems of social choice theory in modal logic*. Autonomous Agents and Multi-Agent Systems, 30, pp. 963 - 989, 2016.

N. Troquard, W. van der Hoek, and M. Wooldridge. *Reasoning about social choice functions*. Journal of Philosophical Logic 40(4), 473 - 498 (2011).

T. Agotnes, W. van der Hoek, and M. Wooldridge. *On the logic of preference and judgment aggregation*. Journal of Autonomous Agents and Multiagent Systems 22(1), 4 - 30 (2011).

Language

Atomic Propositions:

- $Pref[V, X] := \{p^i_{x \succeq y} \mid i \in V, x, y \in X\}$ is the set of preference atomic propositions, where $p^i_{x \preceq y}$ means i prefers y to x.
- Each $x \in X$ is an atomic proposition.

Modality:

- $\Diamond_C \varphi$: C can *ensure* the truth of φ.

Language

Atomic Propositions:

- $Pref[V, X] := \{p^i_{x \succeq y} \mid i \in V, x, y \in X\}$ is the set of preference atomic propositions, where $p^i_{x \preceq y}$ means i prefers y to x.
- Each $x \in X$ is an atomic proposition.

Modality:

- $\Diamond_C \varphi$: C can *ensure* the truth of φ.

$$p \mid \neg\varphi \mid \varphi \wedge \psi \mid \Diamond_C \varphi$$

Model

A **model** is a triple $M = \langle N, X, F \rangle$, consisting of a finite set of agents N (with $n = |N|$), a finite set of alternatives X, and a resolute SCC $F : \mathcal{L}(X)^V \to X$.

A **world** is a profile $(\mathrm{P}_1, \ldots, \mathrm{P}_n)$

Truth

Let $w = (\mathrm{P}_1, \ldots, \mathrm{P}_n)$

- $M, w \models p^i_{x \succeq y}$ iff $x\mathrm{P}_i y$
- $M, w \models x$ if and only if $F(\mathrm{P}_1, \ldots, \mathrm{P}_n) = x$
- $M, w \models \neg\varphi$ if and only if $M, w \not\models \varphi$
- $M, w \models \varphi \wedge \psi$ if and only if $M, w \models \varphi$ and $M, w \models \psi$
- $M, w \models \Diamond_C \varphi$ if and only if $M, w' \models \varphi$ for some $w' = (\mathrm{P}'_1, \ldots, \mathrm{P}'_n)$ with $\mathrm{P}_j = \mathrm{P}'_j$ for all $j \in N - C$.

(1) $p^i_{x \succeq x}$

(2) $p^i_{x \succeq y} \leftrightarrow \neg p^i_{y \succeq x}$ for $x \neq y$

(3) $p^i_{x \succeq y} \wedge p^i_{y \succeq y} \rightarrow p^i_{x \succeq z}$

(1) $p^i_{x \succeq x}$

(2) $p^i_{x \succeq y} \leftrightarrow \neg p^i_{y \succeq x}$ for $x \neq y$

(3) $p^i_{x \succeq y} \wedge p^i_{y \succeq y} \rightarrow p^i_{x \succeq z}$

$$ballot_i(w) \;=\; p^i_{x_1 \succeq x_2} \wedge \cdots \wedge p^i_{x_{m-1} \succeq x_m}$$

$$profile(w) \;=\; ballot_1(w) \wedge \cdots \wedge ballot_n(w)$$

(4) all propositional tautologies

(5) $\Box_i(\varphi \to \psi) \to (\Box_i\varphi \to \Box_i\psi)$ (K(i))

(6) $\Box_i\varphi \to \varphi$ (T(i))

(7) $\varphi \to \Box_i\Diamond_i\varphi$ (B(i))

(8) $\Diamond_i\Box_j\varphi \leftrightarrow \Box_j\Diamond_i\varphi$ (confluence)

(9) $\Box_{C_1}\Box_{C_2}\varphi \leftrightarrow \Box_{C_1\cup C_2}\varphi$ (union)

(10) $\Box_\emptyset\varphi \leftrightarrow \varphi$ (empty coalition)

(11) $(\Diamond_i p \wedge \Diamond_i\neg p) \to (\Box_j p \vee \Box_j\neg p)$, where $i \neq j$ (exclusiveness)

(12) $\Diamond_i ballot_i(w)$ (ballot)

(13) $\Diamond_{C_1}\delta_1 \wedge \Diamond_{C_2}\delta_2 \to \Diamond_{C_1\cup C_2}(\delta_1 \wedge \delta_2)$ (cooperation)

(14) $\bigvee_{x\in X}(x \wedge \bigwedge_{y\in X\setminus\{x\}}\neg y)$ (resoluteness)

(15) $(profile(w) \wedge \varphi) \to \Box_N(profile(w) \to \varphi)$ (functionality)

Theorem (Cinà and Endriss) The logic $L[V, X]$ is sound and complete w.r.t. the class of models of resolute social choice correpsondences.

Pareto

$$Par \quad := \bigwedge_{x \in X} \bigwedge_{y \in X - \{x\}} \left[\left(\bigwedge_{i \in N} p^i_{x \succeq y} \right) \to \neg y \right]$$

IIA

$$IIA := \bigwedge_{w \in \mathcal{L}(X)^n} \bigwedge_{x \in X} \bigwedge_{y \in X - \{x\}} [\Diamond_V(\mathit{profile}(w) \wedge x) \to (\mathit{profile}(w)(x,y) \to \neg y)]$$

- $N^w_{x \succeq y} = \bigwedge \{p^i_{x \succeq y} \mid x\mathrm{P}_i y \text{ in } w\}$
- $\mathit{profile}(w)(x,y) := N^w_{x \succeq y} \wedge N^w_{y \succeq x}$

Dictatorship

$$Dic \quad := \quad \bigvee_{i \in N} \bigwedge_{x \in X} \bigwedge_{y \in X - \{x\}} (p^i_{x \succeq y} \to \neg y)$$

Arrow's Theorem

Theorem (Ciná and Endriss) Consider a logic $L[V, X]$ with a language parameterised by X such that $|X| > 3$. Then we have:

$$\vdash Par \wedge IIA \rightarrow Dic$$

Strong Monotonicity

$$SM \;:=\; \bigwedge_{w \in \mathcal{L}(X)^n} \bigwedge_{x \in X} \left[\Diamond_V (\mathit{profile}(w) \wedge x) \wedge \left(\bigwedge_{y \in X \setminus \{x\}} N^w_{x \succeq y} \right) \rightarrow x \right]$$

Surjectivity

$$Sur := \bigwedge_{x \in X} \bigwedge_{w \in \mathcal{L}(X)^V} \Diamond_V (profile(w) \wedge x)$$

Theorem (Cinà and Endriss) Consider a logic $L[V, X]$ with a language parameterised by X such that $|X| \geq 3$. Then we have:

$$\vdash SM \wedge Sur \rightarrow Dic$$

Compare principles of group decision making in terms of the language used to express them

M. Pauly. *On the Role of Language in Social Choice Theory*. Synthese, 163, 2, pgs. 227 - 243, 2008.

M. Pauly. *Axiomatizing Collective Judgement Sets in a Minimal Logical Language*. Synthese, 158, pp. 233 - 250, 2007.

T. Daniëls. *Social choice and logic of simple games*. Journal of Logic and Computation, 21:6, pp. 883 - 906, 2011.

A Minimal Language

M. Pauly. *Axiomatizing Collective Judgement Sets in a Minimal Logical Language*. Synthese, 158, pp. 233 - 250, 2007.

Let φ_I be the set of **individual formulas** (standard propositional language)

V_I the set of individual valuations

A Minimal Language

M. Pauly. *Axiomatizing Collective Judgement Sets in a Minimal Logical Language*. Synthese, 158, pp. 233 - 250, 2007.

Let φ_I be the set of **individual formulas** (standard propositional language)

V_I the set of individual valuations

φ_C the set of **collective formulas**: $\Box\alpha \mid \varphi \wedge \psi \mid \neg\varphi$

$\Box\alpha$: *The group collectively accepts* α*.*
V_C the set of collective valuations: $v : \varphi_C \to \{0, 1\}$

A Minimal Language

Let $\mathcal{CON}_n = \{v \in V_C \mid v(\Box\alpha) = 1 \text{ iff } \forall i \leq n,\ v_i(\alpha) = 1\}$

E. $\Box\varphi \leftrightarrow \Box\psi$ provided $\varphi \leftrightarrow \psi$ is a tautology

M. $\Box(\varphi \wedge \psi) \to (\Box\varphi \wedge \Box\psi)$

C. $(\Box\varphi \wedge \Box\psi) \to (\Box\varphi \wedge \Box\psi)$

N. $\Box\top$

D. $\neg\Box\bot$

Theorem [Pauly, 2005] $V_C(\mathsf{KD}) = \mathcal{CON}_n$, provided $n \geq 2^{|\varphi_0|}$.

($\mathcal{D} = V_C$, $\mathcal{T} = \mathcal{CON}_n$, $\Delta = EMCND$, then Δ absolutely axiomatizes $\mathcal{T}$.)

A Minimal Language

Let $\mathcal{MAJ}_n = \{v \in \mathcal{V}_C \mid v([>]\alpha) = 1 \text{ iff } |\{i \mid v_i(\alpha) = 1\}| > \frac{n}{2}\}$

STEM contains all instances of the following schemes

S. $[>]\varphi \to \neg[>]\neg\varphi$

T. $([\geq]\varphi_1 \wedge \cdots \wedge [\geq]\varphi_k \wedge [\leq]\psi_1 \wedge \cdots \wedge [\leq]\psi_k) \to \bigwedge_{1\leq i\leq k}([=]\varphi_i \wedge [=]\psi_i)$
where $\forall v \in V_I : |\{i \mid v(\varphi_i) = 1\}| = |\{i \mid v(\psi_i) = 1\}|$

E. $[>]\varphi \leftrightarrow [>]\psi$ provided $\varphi \leftrightarrow \psi$ is a tautology

M. $[>](\varphi \wedge \psi) \to ([>]\varphi \wedge [>]\psi)$

Theorem [Pauly, 2005] $V_C(\text{STEM}) = \mathcal{MAJ}$.

($\mathcal{D} = V_C$, $\mathcal{T} = \mathcal{MAJ}_n$, $\Delta = STEM$, then Δ absolutely axiomatizes $\mathcal{T}$.)

SAT

Arrow's Theorem for a fixed set of alternatives (e.g., $|N| = 2, |X| = 3$) can be embedded into classical propositional logic and automatically checked as a SAT problem. (The full theorem is proved by mathematical induction).

P. Tang. *Computer-aided theorem discovery—a new adventure and its application to economic theory*. Ph.D. thesis, Hong Kong University of Science and Technology, 2010.

P. Tang and F. Lin. *Computer-aided proofs of Arrow's and other impossibility theorems*. Artificial Intelligence, 173(11), pp. 1041 - 1053, 2009.

C. Geist and U. Endriss. *Automated Search for Impossibility Theorems in Social Choice Theory: Ranking Sets of Objects*. Journal of Artificial Intelligence Research, 40, pp. 143 - 174, 2011.

SAT

F. Brandt, C. Geist, and D. Peters. *Optimal bounds for the no-show paradox via SAT solving*. Mathematical Social Science, 90, pp. 18 - 27, 2017.

F. Brandt and C. Geist. *Finding strategyproof social choice functions via SAT solving*. Journal of AI Research, 55, pp. 565 - 602, 2016.

F. Brandl, F. Brandt, and C. Geist. *Proving the incompatibility of efficiency and strategyproofness via SMT solving*. Proceedings of 25th IJCAI, pp. 116 - 122, 2016.

W. Holliday, EP, and S. Zahedian. *Extending a SAT solving approach to impossibility theorems in social choice*. under submission.

social choice theory turns out to be perfectly suitable for mechanical theorem proving."

F. Wiedijk. *Arrow's impossibility theorem*. Formalized Mathematics, 15:171–174, 2007.

T. Nipkow. *Social choice theory in HOL: Arrow and Gibbard-Satterthwaite*. Journal of Automated Reasoning, 43:289–304, 2009.

M. Eberl. *Verifying Randomised Social Choice*. International Symposium on Frontiers of Combining Systems, FroCoS 2019: Frontiers of Combining Systems pp 240-256.

F. Brandt, M. Eberl, C. Saile and C. Stricker. *The Incompatibility of Fishburn-Strategyproofness and Pareto-Efficiency*. `https://www.isa-afp.org/entries/Fishburn_Impossibility.html`.

Lean

The Lean Theorem Prover aims to bridge the gap between interactive and automated theorem proving, by situating automated tools and methods in a framework that supports user interaction and the construction of fully specified axiomatic proofs. The goal is to support both mathematical reasoning and reasoning about complex systems, and to verify claims in both domains.

`https://leanprover.github.io/`

W. Holliday, C. Norman, and EP. *Voting Theory in the Lean Theorem Prover*. Proceedings of LORI 2021, `https://arxiv.org/abs/2110.08453`.

Profiles of Preferences

Profiles

Definition

For $V \subseteq \mathcal{V}$ and $X \subseteq \mathcal{X}$, a (V, X)-*profile* is a map $\mathrm{P} : V \to \mathcal{B}(X)$.

Given a (V, X)-profile P, let $V(\mathrm{P})$ be V and $X(\mathrm{P})$ be X.

We then define a function Prof that assigns to each pair (V, X) of $V \subseteq \mathcal{V}$ and $X \subseteq \mathcal{X}$ the set $\mathrm{Prof}(V, X)$ of all (V, X)-profiles.

Profiles

Definition

For $V \subseteq \mathcal{V}$ and $X \subseteq \mathcal{X}$, a (V, X)-*profile* is a map $\mathrm{P} : V \to \mathcal{B}(X)$.

Given a (V, X)-profile P, let $V(\mathrm{P})$ be V and $X(\mathrm{P})$ be X.

We then define a function Prof that assigns to each pair (V, X) of $V \subseteq \mathcal{V}$ and $X \subseteq \mathcal{X}$ the set $\mathrm{Prof}(V, X)$ of all (V, X)-profiles.

```
def Prof : Type → Type → Type :=
λ (V X : Type), V → X → X → Prop
```

Majority Preferred

Definition

Given a profile P and $x, y \in X(\mathrm{P})$, we say that x *is majority preferred to* y *in* P if more voters rank x above y than rank y above x.

Majority Preferred

Definition

Given a profile P and $x, y \in X(\mathrm{P})$, we say that x *is majority preferred to* y *in* P if more voters rank x above y than rank y above x.

```
def majority_preferred {V X : Type} :
Prof V X → X → X → Prop := λ P x y,
cardinal.mk {v : V // P v x y} >
cardinal.mk {v : V // P v y x}
```

Margin

Definition

Given a profile P and $x, y \in X(\mathrm{P})$, the *margin of* x *over* y *in* P, denoted $Margin_{\mathrm{P}}(x, y)$, is $|\{i \in V(\mathrm{P}) \mid x\mathrm{P}_i y\}| - |\{i \in V(\mathrm{P}) \mid y\mathrm{P}_i x\}|$.

Margin

Definition

Given a profile P and $x, y \in X(\mathrm{P})$, the *margin of* x *over* y *in* P, denoted $Margin_{\mathrm{P}}(x, y)$, is $|\{i \in V(\mathrm{P}) \mid x\mathrm{P}_i y\}| - |\{i \in V(\mathrm{P}) \mid y\mathrm{P}_i x\}|$.

```
def margin {V X : Type} [fintype V] :
Prof V X → X → X → ℤ
:= λ P x y, ↑(finset.univ.filter (λ v, P v x y)).card -
↑(finset.univ.filter (λ v, P v y x)).card
```

Simple Example

Condorcet Winner and Majority Winner

Definition

Given a profile P and $x \in X(\mathrm{P})$, x is a *Condorcet winner in* P if for all $y \in X(\mathrm{P})$ with $y \neq x$, x is majority preferred to y in P.

We say that x is a *majority winner in* P if the number of voters who rank x (and only x) in first place is greater than the number of voters who do not rank x in first place.

Condorcet Winner and Majority Winner

Definition

Given a profile P and $x \in X(\mathrm{P})$, x is a *Condorcet winner in* P if for all $y \in X(\mathrm{P})$ with $y \neq x$, x is majority preferred to y in P.

We say that x is a *majority winner in* P if the number of voters who rank x (and only x) in first place is greater than the number of voters who do not rank x in first place.

```
def condorcet_winner {V X : Type} (P : Prof V X) (x :  X) :
Prop := ∀ y ≠ x, majority_preferred P x y


def majority_winner {V X : Type} (P : Prof V X) (x :  X) :
 Prop := cardinal.mk {v :  V // ∀ y ≠ x, P v x y} >
 cardinal.mk {v :  V // ∃ y ≠ x, P v y x}
```

Lemma. For any profile P, for all $x \in X(P)$, if x is a majority winner in P, then x is a Condorcet winner in P.

```
lemma condorcet_of_majority_winnner {V X : Type}
(P : Prof V X) [fintype V] (x :  X) :
majority_winner P x → condorcet_winner P x :=
```

Lean has a very well-developed library of mathematical results called mathlib. For instance, we made use of the following theorem from mathlib:

```
theorem cardinal.mk_subtype_mono {α : Type u}
{φ ψ : α → Prop} (h : ∀ x, φ x → ψ x) :
cardinal.mk {x // φ x} ≤ cardinal.mk {x // ψ x}
```

```
lemma condorcet_of_majority_winnner {V X : Type}
(P : Prof V X) [fintype V] (x :  X) :
majority_winner P x → condorcet_winner P x :=
begin
1.   intros majority z z_ne_x,
2.   have imp1 :  ∀ v, (∀ y ≠ x, P v x y) → P v x z :=
     by finish,
3.   refine lt_of_lt_of_le _ (cardinal.mk_subtype_mono imp1),
4.   have imp2 :  ∀ v, P v z x → (∃ y ≠ x, P v y x) :=
     by finish,
5.   apply lt_of_le_of_lt (cardinal.mk_subtype_mono imp2),
6.   exact majority,
end
```

Functions on Profiles

Definition

For $V \subseteq \mathcal{V}$ and $X \subseteq \mathcal{X}$, a *social choice correspondence for* (V, X), or (V, X)-SCC, is a function $F : \mathrm{Prof}(V, X) \to \wp(X)$.

Let SCC be a function that assigns to each pair (V, X) of $V \subseteq \mathcal{V}$ and $X \subseteq \mathcal{X}$ the set of all (V, X)-SCCs.

Definition

For $V \subseteq \mathcal{V}$ and $X \subseteq \mathcal{X}$, a *social choice correspondence for* (V, X), or (V, X)-SCC, is a function $F : \mathrm{Prof}(V, X) \to \wp(X)$.

Let SCC be a function that assigns to each pair (V, X) of $V \subseteq \mathcal{V}$ and $X \subseteq \mathcal{X}$ the set of all (V, X)-SCCs.

```
def SCC := λ (V X : Type), Prof V X → set X
```

Definition

For $V \subseteq \mathcal{V}$ and $X \subseteq \mathcal{X}$, a *social choice correspondence for* (V, X), or (V, X)-SCC, is a function $F : \mathrm{Prof}(V, X) \to \wp(X)$.

Let SCC be a function that assigns to each pair (V, X) of $V \subseteq \mathcal{V}$ and $X \subseteq \mathcal{X}$ the set of all (V, X)-SCCs.

```
def SCC := λ (V X : Type), Prof V X → set X
```

```
def universal_domain_SCC {V X : Type} (F : SCC V X) : Prop :=
∀ P : Prof V X, F P ≠ ∅
```

Example

The Condorcet SCC:

```
def condorcet_SCC {V X : Type} : SCC V X := λ P,
{x : X | condorcet_winner P x ∨ ¬ ∃ y, condorcet_winner P y}
```

Variable-Election Framework

Definition

A *variable-election social choice correspondence* (VSCC) is a function F that assigns to each pair (V, X) of a $V \subseteq \mathcal{V}$ and $X \subseteq \mathcal{X}$ a (V, X)-SCC.

Variable-Election Framework

Definition

A *variable-election social choice correspondence* (VSCC) is a function F that assigns to each pair (V, X) of a $V \subseteq \mathcal{V}$ and $X \subseteq \mathcal{X}$ a (V, X)-SCC.

```
def VSCC : Type 1 := Π (V X : Type), SCC V X
```

Given α `:` `Type` and β `:` α `→` `Type`, the type

$$\Pi\ \mathtt{y}\ \mathtt{z}\ :\ \alpha,\ \beta\ \mathtt{y}\ \mathtt{z}$$

is the type of functions `f` such that for each `a b :` α, we have that `f a b` is an element of β `a b`.

Variable-Election Framework

Definition

A *variable-election social choice correspondence* (VSCC) is a function F that assigns to each pair (V, X) of a $V \subseteq \mathcal{V}$ and $X \subseteq \mathcal{X}$ a (V, X)-SCC.

```
def VSCC : Type 1 := Π (V X : Type), SCC V X
```

Example: Condorcet VSCC

```
def condorcet_VSCC : VSCC := λ V X, condorcet_SCC
```

Other Functions on Profiles

```
def SCC := λ (V X : Type), Prof V X → set X

def VSCC : Type 1 := Π (V X : Type), SCC V X
```

Other Functions on Profiles

```
def SCC := λ (V X : Type), Prof V X → set X
def VSCC : Type 1 := Π (V X : Type), SCC V X

def CCR := λ (V X : Type), Prof V X → X → X → Prop
def VCCR := Π (V X : Type), CCR V X
```

Other Functions on Profiles

```
def SCC := λ (V X : Type), Prof V X → set X
def VSCC : Type 1 := Π (V X : Type), SCC V X
```

```
def CCR := λ (V X : Type), Prof V X → X → X → Prop
def VCCR := Π (V X : Type), CCR V X
```

Given an asymmetric VCCR f, we define the *maximal-element induced* VSCC f_M:

```
def max_el_VSCC : VCCR → VSCC := λ f V X P,
{x : X | ∀ y : X, ¬ f V X P y x}
```

Formalized Proofs

We verified all the results about a new voting method, *Split Cycle*, from

W. Holliday and E. Pacuit. *Split Cycle: A New Condorcet Consistent Voting Method Independent of Clones and Immune to Spoilers* . `https://arxiv.org/abs/2004.02350`.

	Split Cycle	Ranked Pairs	Beat Path	Mini-max	GETCHA/ GOCHA	Ranked Choice	Plurality
Condorcet Winner	✓	✓	✓	✓	✓	–	–
Condorcet Loser	✓	✓	✓	–	✓	✓	–
Pareto	✓	✓	✓	✓	–	✓	✓
Monotonicity	✓	✓	✓	✓	✓	–	✓
Independence of Clones	✓	✓	✓	–	✓	✓	–
Strong Stability for Winners	✓	–	–	–	✓	–	–
Reversal Symmetry	✓	✓	✓	–	✓	–	–
Positive Involvement	✓	–	–	✓	✓/–	✓	✓
Negative Involvement	✓	–	–	✓	✓/–	–	✓

Future Work

- Verify axioms of other voting methods: Not just margin-based methods (e.g., Split Cycle, Beat Path), but also scoring rules (e.g., Plurality, Borda), and *recursive* voting methods (e.g., Instant Runoff, Stable Voting).
- Formalize characterization theorems (e.g., Arrow's Theorem characterizing dictatorship, May's Theorem characterizing majority rule, Young's Theorem characterizing scoring rules, ...).
- See `https://github.com/asouther4/lean-social-choice/` by A. Souther for a verification of proofs about Stable Voting.

Justifying Voting Outcomes

`https://demo.illc.uva.nl/justify/`

A. Boixel and U. Endriss. *Automated Justification of Collective Decisions via Constraint Solving*. AAMAS-2020.

A. Boixel, U. Endriss, and R.. de Haan. *A Calculus for Computing Structured Justifications for Election Outcomes*. Proceedings of the 36th AAAI Conference on Artificial Intelligence (AAAI-2022).

Thank you!

https://pref-voting.readthedocs.io/

W. Holliday and EP, Axioms for defeat in democratic elections, Journal of Theoretical Politics, https://arxiv.org/abs/2008.08451

W. Holliday and EP, Split Cycle: A New Condorcet Consistent Voting Method Independent of Clones and Immune to Spoilers, forthcoming Public Choice, https://arxiv.org/abs/2004.02350

https://stablevoting.org