# Intuitive Preference Aggregation Arrow's Principles and Behaviorally-Derived Criteria 

Todd Davies Raja Shah* Renee Trochet Katerina Ling<br>Symbolic Systems Program, Stanford University

## Questions

- Does intuitive preference aggregation robustly violate Arrow's criteria of IIA and collective rationality?
- Can we characterize intuitive preference aggregation judgments in a descriptive framework?
- What is the relationship between intuitive preference aggregation and normative approaches to social choice?


## Background Theory

The problem of preference aggregation

- Construct some social welfare function (SWF) to choose or rank options for a group based on any possible set of individuals' rankings.

Arrow's criteria: The SWF should always obey...

- Collective rationality. Social ranking is complete and transitive.
- Weak Pareto. When everyone prefers (ranks one option above) another, the social ranking should also.
- Independence of irrelevant alternatives (IIA). If individuals' preferences between two options S and T remain fixed while the ranking/s of other option/s change, the social preference between S and T should not change (also known as "pairwise independence")
- Non-dictatorship. No individual should be able to always determine which of any two options is socially preferred.
Impossibility Theorem
- No SWF exists that satisfies all of Arrow's criteria (Arrow, 1951/1963).
Arrow on IIA and collective rationality:
- "The essential argument in favor of [IIA] is its direct appeal to intuition." (1952)
- IIA is "stricter than desirable" but necessary to avoid having to gather limitless information on unavailable options. (1967)
- Dead candidate example (1951): If a candidate dies after an election, the outcome should not change unless the dead candidate won. Given as intuition behind IIA but actually applies collective rationality instead.

Profile Triplets


V1 through V7 are voters, and S, T, U, V, W, X are options that the voters rank. Experiment asks survey-takers to choose or rank option/s for the group based on a given set of rankings (a "preference profile"). For each triplet, preference aggregation that chooses S in the top and T in the middle profiles violates collective rationality. Socially choosing T in the middle and S in the bottom profiles violates IIA. Triplets differ in extent of "Borda reversal" between profiles (number of rank shifts for $S$ and $T$ ).

## Experimental Variations

Between Subjects

- Each survey taker aggregates only one profile in a given triplet.

Within Subjects

- Each survey taker's aggregations are tested for adherence to IIA and/or collective rationality
Pairwise Format
- Profile is presented to survey taker as a matrix of individuals' pairwise preferences instead of in the rank format shown above


## Restricted Availability

- Independence of unavailable alternatives (IUA). Attempt to capture Arrow's intuitive argument based on dead candidate example. Survey taker is shown profile but told that only options S and T are available.


## Indifference Option

- Survey taker is told they $\mathrm{s} / \mathrm{he}$ can circle more than one option for a group of voters if both options are equally desirable given the profile.


## Presentation Order Variations

- Example: Some survey takers see one profile in a triplet before another, with other survey takers seeing them in the reverse order.
Profile Relationship Variations
- Different profiles are presented to a survey taker as representing different voters and/or different alternatives, or changes in the menu of options.


## Results

- For each of the four profile triplets, when presented in rank format, a majority of survey takers' aggregations violate both collective rationality (CR) and IIA, both between and within subjects. Examples: In the 5 -voter triplet.
-Between subjects, 34/35 pick S in 5-2, 38/39 pick T in 5-3, and 41/44 pick S in 5-3
-Within-subjects test finds $84 \%$ violate CR and $92 \%$ violate IIA. - For the 4 -voter and 7 -voter triplets, survey takers robustly violate both CR and IIA when profiles are presented in pairwise format as well, but pairwise format can induce adherence to CR and IIA when the Borda reversal is weak (e.g. the 3- and 5-voter triplets).
Examples: Between subjects
$-90 / 90$ pick T in 4-6, and 70/90 pick S in 4-6
$-100 / 100$ pick S in 5-3', and 55/100 pick S in 5-3
- Adherence to IUA is somewhat stronger than adherence to either IIA or CR, and for weak Borda reversals shows majority adherence when IIA and CR are violated
- Availability of indifference option does not significantly increase adherence to IIA or CR.
- Order of profile presentation does not affect results when Borda swings are large enough, but can mater when they are small.
- Survey takers treat each group of voters independently when the options remain the same, and treat each menu independently.
- Attendees of the Public Choice Society annual meeting show the same pattern of results as naïve survey-takers.


## Behavioral Principles

- Independence of irrelevant voters (IIV). Social choices should be based only on the preferences of individuals affected. (Strongly supported)
- Interpersonal comparisons of utility (ICU). Social choices should take into account any information about relative strengths of preference between individuals over the same options. (Strongly supported)
- Inter-menu independence (IMI). Social choices should take into account only the alternatives that are shown in a given profile. (Strongly supported)
- Independence of unavailable alternatives (IUA). Social choices should be the same as long as individual preferences do not change over the set of available options. (Weakly supported)


## Conclusions

- IIA and Collective rationality are not indispensible normatively
- Abstract intutive preference aggregation may aid rule selection.

