American Economic Review 2026, 116(5): 1648–1681
https://doi.org/10.1257/aer.20211491

Social Preferences over Ordinal Outcomes[†]

By Sandro Ambuehl and B. Douglas Bernheim*

We study social preferences in settings where someone who chooses on behalf of others knows how those individuals rank the available options but may lack cardinal information concerning those comparisons. Contrary to majoritarian principles, most people place more weight on preventing least preferred outcomes for others than on enabling most preferred outcomes. Ranks matter both intrinsically and because they provide a basis for inferring cardinal utility. Ordinal aggregation principles are stable across domains and countries with divergent political traditions. Designing attractive social choice mechanisms is challenging in practice partly because aggregation principles that make manipulation difficult yield outcomes people consider normatively unappealing. (*JEL* C91, D71, D72)

Envision a benevolent party (a "Planner") who must make a decision that impacts the members of some group ("Stakeholders"). The Planner possesses reliable information about the Stakeholders' ordinal preferences over the available options ("within-menu rankings") but not about their cardinal preferences (intensity). How would the typical individual in the role of Planner determine the "best" choice for the group? In other words, what is the nature of social preferences over ordinal outcomes?[1]

This issue is of broad importance for at least three reasons. First, cardinal information is often unreliable or unavailable in practice. For example, a supervisor charged with allocating a diverse set of assignments may learn which tasks individual team members prefer without being able to distinguish strong tastes from a tendency to complain. More fundamentally, in the spirit of Arrow (1951), people may intuitively question the validity of cardinal interpersonal utility comparisons. For example, they may distrust their ability to gauge differences between the pleasure different people derive from incremental income. The greater reliability of ordinal information may help explain why procedures for making group decisions typically

*Ambuehl: Department of Economics and UBS Center for Economics in Society, University of Zurich (email: sandro.ambuehl@econ.uzh.ch); Bernheim: Department of Economics, Stanford University (email: bernheim@stanford.edu). Pietro Ortoleva was the coeditor for this article. We are grateful to Eva Küpper for truly outstanding research assistance. This project was funded by the Department of Economics at Stanford University and by the Department of Economics at the University of Zurich. All experiments were approved by the Stanford IRB (protocol 53339) and by the Ethics Review Board of the Department of Economics at the University of Zurich (OEC IRB 2020–063).

[†] Go to https://doi.org/10.1257/aer.20211491 to visit the article page for additional materials and author disclosure statement(s).

[1] As we use the term, an *ordinal outcome* is the vector of ranks Stakeholders assign to a given social action for a fixed opportunity set.

focus on rankings; examples include matching markets, committee deliberations, and elections. Second, people may care about rankings intrinsically. For example, they may wish to avoid selecting an option that any group member ranks last, or they may feel ethically obliged to follow the will of the majority. For these reasons, inquiries into social preferences that impose the cardinal Bergson-Samuelson structure—the predominant approach in the pertinent literature (e.g., Engelmann and Strobel 2004; Bolton and Ockenfels 2006; Jackson and Yariv 2014)—may improperly identify underlying objectives. Third, it is arguable that, in representative democracies, policymakers ought to defer to citizens' judgments about the importance of ordinal outcomes and the appropriate criteria for ordinal preference aggregation. Under that view, our positive analysis of social choice has important normative implications.

We conduct experiments designed to address four main questions. First, what aggregation principles best capture people's social preferences over ordinal outcomes? For instance, do they seek compromise solutions or insist that the majority should prevail? Second, are these principles stable, or do they vary from context to context? Third, do people intrinsically care about ordinal outcomes, or do they merely use them to draw inferences about cardinal outcomes and then aggregate using standard Bergson-Samuelson social welfare functions (in the spirit of Apesteguia, Ballester, and Ferrer 2011)? Fourth, do aggregation preferences vary across cultures with divergent political and social traditions?

In our experiments, subjects acting as Planners make potentially consequential decisions for groups of Stakeholders. These decisions are of two types: distributing work tasks among five Stakeholders (the *work domain*) and assigning a contribution on behalf of five Stakeholders to a single political entity (the *political domain*). Because we are interested in identifying structural aggregation preferences, we design the experiment to remove considerations arising from self-interest, paternalism (i.e., the tendency to ignore or discount Stakeholder judgments with which the Planner disagrees, as in Ambuehl, Bernheim, and Ockenfels 2021), and the potential incentive incompatibility of truthful preference revelation by Stakeholders.

Social preferences over ordinal rank profiles manifest themselves through desired *social choice rules*—that is, functions mapping each ordinal preference profile over a given set of alternatives to a selection from that set. The choices Planners make on behalf of others reveal the social choice rules that embody the normative principles they favor. Distilling those principles is, nevertheless, conceptually challenging. Even in simple social choice problems (e.g., with five people and three options), the set of possible mappings from group members' preference profiles to choices is astronomically large, and merely cataloging the observed rules provides little insight. We therefore proceed in three steps. First, drawing on the theoretical literature, we identify a reasonably large set of plausible aggregation rules. Each rule implies a distinctive *fingerprint* of implied best choices over the set of conceivable five-person three-option preference profiles. Second, we assign each subject to one of various prespecified rules using a Bayes classifier. Third, we use a clustering algorithm to determine whether our prespecified rules omit empirically important possibilities.

In answer to the first question, we find that the overwhelming majority of subjects behave as if they assign a cardinal utility level (a *score*) to each rank and select the option that maximizes the sum of these utilities (*scoring rules*). Many well-known

social choice criteria fall within this category. For example, plurality rule assigns a score of 1 to each person's favorite alternative and 0 to all others. The curvature of the mapping from ranks to scores determines whether improvements from low or high ranks are more important. Concave scoring rules promote compromise by placing more weight on gains from low ranks. All else equal, they favor an option ranked n-th by two people over an alternative ranked $(n-1)$-th by one and $(n+1)$-th by another. Convex scoring rules do the opposite because they place more weight on gains from high ranks. All else equal, they favor an option ranked $(n-1)$-th by one person and $(n+1)$-th by another over an alternative ranked n-th by both. To illustrate, suppose there are three options and two Stakeholders. Consider a rule that assigns scores of 1, s, and 0 to each Stakeholder's favorite, second-favorite, and least favorite option, respectively. In that case, an option that both Stakeholders rank second yields a total score of $2s$, while an option that one ranks first and the other ranks third yields a total score of 1. If $s > 0.5$, the rule is concave, and it selects the first option (the compromise) over the second. If $s < 0.5$, the rule is convex, and it rejects the compromise option. According to our findings, the two most prevalent as-if aggregation criteria are the linear scoring rule (*Borda*), which places equal weight on gains from high and low ranks, and a rule with nearly maximal concavity resembling *antiplurality*, which minimizes the number of last-place ranks. Most subjects ($> 60\%$) employ strictly concave scoring rules. Choices conform with plurality rule (which is the most convex scoring rule), the Condorcet rule (also known as pairwise majority rule), supermajority rule, and the unanimity criteria for relatively few subjects. The classification's fit is excellent, and clustering analysis identifies no important omitted patterns.

The prevalence of strictly concave scoring rules calls to mind the familiar cardinal concept of an additive, strictly concave Bergson-Samuelson social welfare function, which aggregates by applying a concave transformation w to each individual's cardinal utility U_n before summing them; that is, $W = \sum_n w(U_n)$. While related, these concepts are distinct. Even if social preferences over ordinal outcomes derive from cardinal inferences, the curvature of the scoring function would depend, in part, on the curvature of the mapping from ranks to inferred cardinal values. Similarly, the prevalence of highly concave scoring rules calls to mind the maximin criterion (Rawls 1971), a familiar cardinal concept that also focuses on the less fortunate. Concave scoring rules, however, judge misfortune according to an individual's own rankings within a menu, whereas maximin judges misfortune according to interpersonal comparisons.

In answer to the second question, we find that social preferences over ordinal outcomes entail stable structural elements. Specifically, our classifications are highly predictive of choices out of sample, even across domains. However, there is also a meaningful degree of domain dependence: Subjects prefer slightly more concave scoring rules when distributing work tasks than when aggregating political preferences.

In answer to the third question, we find that people use ordinal ranks to draw inferences about cardinal utilities but also care about them intrinsically. Two findings support the importance of cardinal inferences. First, best-fit scoring parameters are correlated with direct measures of Planners' beliefs about cardinal outcomes. Second, the cardinal inference hypothesis correctly predicts the ways in which

Planner's choices respond to information about the rankings of unavailable options. We establish the intrinsic importance of ordinal outcomes by studying treatments for which we disable the cardinal inference mechanism. In these treatments, distributions of monetary payoffs, which we reveal to the Planner, serve as the social alternatives. Planners' choices are sensitive to changes in menus that alter the profile of ranks associated with each alternative without changing the profile of cardinal payoffs. Moreover, Planners are willing to choose alternatives that yield strictly inferior cardinal payoff distributions (in the sense of first-order dominance) to achieve more favorable profiles of Stakeholder ranks.

To answer the fourth question, we run supplemental experiments with general population samples, wherein social choices determine the allocation of a contribution over well-known charities. We find that the distributions of aggregation preferences in the United States and Sweden, countries with divergent political and social traditions, are remarkably similar, and both resemble the distribution for the student sample used in our main experiment.[2] Moreover, we find suggestive evidence that the use of more concave scoring rules in experimental decisions correlates with a preference for electing compromise candidates.

Our paper contributes primarily to the literature on social preferences (reviewed in Fehr and Charness 2025) by focusing on judgments involving the aggregation of ordinal outcomes. The prior literature contains hints about the nature of those judgments, for example, regarding *last place aversion* (Kuziemko et al. 2014; Camerer et al. 2016; Martinangeli and Windsteiger 2021), but little more. One exception is Kara and Sertel (2005), which uses hypothetical choices over abstract options for four preference profiles to determine the best fit among three ordinal aggregation rules. Unlike the current paper, they did not examine the relation between ordinal and cardinal aggregation. Furthermore, our analysis uses far more exhaustive lists of rules and preference profiles, studies real choices, checks for the presence of omitted rules, tests out-of-sample predictive accuracy, and examines stability across domains and cultures. Another exception is Herreiner and Puppe (2010), who study bargaining in settings where parties only learn the ordinal rankings associated with each alternative. The study finds that some people forgo Pareto improvements to reduce differences between the ranks parties assign to the selected outcome, possibly because they use ranks as proxies for monetary value. These results are related to our finding concerning the popularity of concave scoring rules. However, Herreiner and Puppe's (2010) bargaining setting implicates selfishness and strategic concerns; unlike our study, they do not focus on disinterested judgments concerning aggregation. Indeed, even in the cardinal domain, most of the literature studies self–other trade-offs rather than other–other trade-offs; see, for example, Jackson and Yariv (2014); Engelmann and Strobel (2004); Bolton and Ockenfels (2006).

Our paper also contributes to a large literature on social choice. Instead of asking which criteria people *ought* to use when aggregating ordinal preferences (in the tradition of Arrow 1950), we investigate the criteria they *actually* use. Our results imply that most people violate Arrow's Independence of Irrelevant Alternatives

[2] Ambuehl et al. (2021) provide further evidence on robustness by comparing distributions of best-fit social choice rules for elected representatives in German federal and state parliaments and the German general population.

(IIA) when making choices for groups.[3] By interpreting a social choice rule as an expression of social preferences rather than as a procedure for group decision-making, we depart from an important branch of this literature, beginning with Gibbard (1973) and Satterthwaite (1975), that studies implementability. While issues of manipulability inevitably arise in practice, they pertain to the constraints that appear in mechanism design problems, rather than to the objective functions decision-makers seek to maximize subject to those constraints. It is the latter we seek to illuminate. In contrast, studies that elicit subjects' preferences over voting procedures (such as Engelmann and Grüner 2017; Hoffmann and Renes 2017; and Engelmann et al. 2020) seek to illuminate judgments that account for all practical considerations impacting a procedure's performance. Our findings are nevertheless relevant for this branch of social choice theory because the rules that best describe our subjects' ordinal aggregation preferences are known to be easily manipulable. Thus, the pursuit of procedures that are more difficult to manipulate may render outcomes less normatively appealing to most of the affected parties.[4]

Finally, our paper contributes to the literature on *positive welfare economics*, which uses empirical methods to determine how people evaluate the well-being of other individuals and groups (see, e.g., Konow 2003; Gaertner 2009; Gaertner and Schokkaert 2012; Andreoni et al. 2020; Almås, Cappelen, and Tungodden 2020; Ambuehl, Bernheim, and Ockenfels 2021; Arrieta and Bolte 2023).

The remainder of the paper is organized as follows. Section I lays out our main strategy for understanding the criteria people implicitly use to make decisions on behalf of groups when they only have ordinal information about others' preferences. Section II details our experimental design. Section III provides our classification results and conducts robustness analyses. Section IV documents the structural stability and context dependence of social preferences over ordinal outcomes. Section V examines whether subjects draw cardinal inferences from ordinal information and whether they care about ordinal outcomes intrinsically. Section VI examines the robustness of our results using Swedish and US general population samples. Section VII concludes.

I. Conceptual Framework

We are concerned with settings in which a decision-maker, the *Planner*, must make a selection from a set of K social options, $\mathcal{A}$, on behalf of N *Stakeholders*. Each option has direct consequences for the Stakeholders but not for the Planner. We assume throughout that each Stakeholder i has a strict linear preference ordering

[3] Davies and Shah (2003) also document violations of IIA when subjects aggregate others' preferences.

[4] Knowing that people favor the normative values underlying highly manipulable social choice rules helps to explain why Planners sometimes defect from announced incentive-compatible decision criteria after group members have submitted their preferences (see Featherstone 2019 for evidence). Many of the rules that mirror our subjects' social preferences are used in practice despite their susceptibility to manipulation. The Econometric Society uses the Borda rule to elect Officers and Council (Econometric Society 2022). The Borda rule also determines Slovenian parliamentary elections (Toplak 2006) and elections within the Irish Green Party (Baker 2008). Swiss parliamentary elections incorporate disapproval votes, which follow the spirit of the antiplurality rule (Portmann and Stojanović 2019). Several nations use multistage methods for presidential elections (see, e.g., Richie 2004). The Debian Project uses a Condorcet extension for all decision-making by committees and plebiscites (Debian Project 2016).

$\succ_i$ over $\mathcal{A}$. Before making a decision, the Planner learns the group's ordinal *preference profile*, $P = (\succ_1, \ldots, \succ_N)$. We interpret *social choice rules* as expressions of Planners' social preferences. Any such rule provides a complete account of the Planner's choices for every preference profile; in other words, it is a mapping R from preference profiles into nonempty subsets of $\mathcal{A}$ (*best choices*). A *resolute* rule admits no ties, in the sense that it maps every profile to a single option. An *irresolute* rule maps at least one preference profile to a nonsingleton set.

It bears emphasis that, in this article, we do *not* interpret social choice rules as *procedures* for group decision-making. While any given social choice rule, such as a voting rule, may have a procedural flavor, we do not contend that, in making decisions for their assigned groups, our subjects consciously apply such procedures. Nor does our experiment speak to questions concerning attributes of these procedures, such as procedural fairness, other than the outcomes they produce. Our approach parallels consumer theory, which uses utility functions to summarize preference relations, but does not assert that consumers choose by computing the maxima of those functions.

The task of identifying the rule that best describes a Planner's aggregation preferences is challenging because the set of possibilities is astronomically large.[5] Therefore, our analysis proceeds in three steps. First, based on the literature, we identify a reasonably large collection of social choice rules encompassing the most plausible alternatives. Second, we classify Planners according to the prespecified rules that most closely match their actual choices. Third, we deploy clustering analysis to determine whether our prespecified rules omit empirically important alternatives.

The literature sorts social choice rules into three broad classes. Members of the first class, *Scoring rules*, assign points to ranks and select all point-maximizing alternatives. The *Borda rule*, for instance, assigns $f(k_i) = N - k_i$ points to an alternative if Stakeholder i ranks it k_i-th out of N alternatives, and selects the alternative for which the sum of points, $\sum_{i=1}^{N} (N - k_i)$, is highest. We call a scoring rule *convex* or *concave* depending on the curvature of the mapping f from ranks to scores. Convex rules place emphasis on ensuring Stakeholders obtain highly ranked alternatives, which gives them a "winner-take-all" flavor. In contrast, concave rules place emphasis on preventing Stakeholders from receiving alternatives with low ranks, which means they encode a preference for compromise. The most convex rule assigns points only if Stakeholders obtain their most preferred alternatives, which means it maximizes the number of first-place ranks (*plurality rule*). The most concave rule assigns a point to all alternatives except for the least preferred one, which means it minimizes the number of last-place ranks (*antiplurality rule*). The Borda rule is linear.

The second class consists of rules that generalize the well-known *Condorcet* criterion. Each such rule involves a binary relation that deems alternative A to be better than alternative B if the fraction of Stakeholders who prefer A over B is at least q for some $q \geq \frac{1}{2}$. Setting $q = \frac{1}{2}$ yields the Condorcet relation; setting $q = 1$ yields

[5] To illustrate, consider a simple setting with five Stakeholders and three alternatives. In that case, there are $7^{42} = 3.1 \times 10^{35}$ possible social choice rules that satisfy *anonymity* (all Stakeholders treated symmetrically) and *neutrality* (all social options treated symmetrically), of which 3.5×10^{28} exhibit no Pareto violations.

the unanimity relation. The rule then selects maximal elements according to that relation (for example, a Condorcet winner in the case of $q = \frac{1}{2}$). A well-known limitation of rules in this class is that maximal elements may not exist (the Condorcet paradox). *Condorcet extensions* select the maximal elements when they exist and differ according to the selections they make when there is no such element (see, e.g., Brandt et al. 2016).

The third class consists of *multistage rules*, which winnow down the set of alternatives in a series of steps. They introduce almost limitless possibilities, inasmuch as they can, in principle, apply different criteria in each step. We will focus on *scoring runoffs*, which repeatedly eliminate the alternative with the lowest score until only one option remains (see Freeman, Brill, and Conitzer 2014).

When classifying Planners according to the prespecified rules that most closely match their actual choices, we focus primarily on choice problems involving five Stakeholders and three options. This focus offers three advantages.

First, five-Stakeholder three-option problems are the simplest and hence most cognitively manageable settings that provide adequate scope for differentiation among a broad collection of rules.

Second, with three options, we can associate each scoring rule with a single parameter, $s \in [0, 1]$, representing the score assigned to the middle alternative (after normalizing so that highest- and lowest-ranked alternatives receive scores of 1 and 0, respectively). This property facilitates the interpretation of our results. Specifically, the function relating ranks to weights is concave when $s > 1/2$ and convex when $s < 1/2$. In this context, concavity (convexity) implies that replacing a third-place rank with a second-place rank is more (less) valuable than replacing a second-place rank with a first-place rank. In other words, concave scoring rules codify an aversion to giving Stakeholders their least favorite of the three alternatives, while convex scoring rules codify an attraction to giving Stakeholders their favorite alternatives.

Third, with 5 Stakeholders and 3 options, we can in principle investigate choices exhaustively on the entire preference domain (42 distinct strict preference profiles) by eliciting a choice for every conceivable profile. In practice, many of the 42 preference profiles provide little or no discernment among rules, so we omit them from the experiment. For example, if all Stakeholders have the same ranking, the best social choice is obvious.

Our analysis is based on the 17 discerning profiles listed in Table 1,[6] which also shows how these profiles differentiate among familiar social choice rules. For example, if a subject uses a scoring rule, we can determine the parameter s from the choices they make for profiles 1 through 11. To understand this point, first consider profile 11. Because all Stakeholders rank option B second, its score is $S_s(B) = 5s$. Four Stakeholders rank option A first and one ranks it third, so its score is $S_s(A) = 4$. Option C is rank-dominated by A, so no scoring rule will select it. Therefore, the subject will choose option B if $S_s(B) > S_s(A)$, or equivalently, $s > 0.8$, and will choose option A if $s < 0.8$. If $s = 0.8$, the subject is indifferent between options A and B. Now consider profile 6, which differs from profile 11 only in that Stakeholder 2's preferences between A and C are reversed. Reasoning

[6] Supplemental Appendix A.1 lists the omitted profiles and exhibits their limited ability to distinguish among rules.

TABLE 1—THREE-ALTERNATIVE PROFILES AND CHOICES OF SELECTED AGGREGATION RULES

		Rule predictions										
		Scoring cutoff	Scoring rules, $s \in$					Runoff cutoff	Runoff rules, $s \in$			
Index	Profile	$\bar{s}$	$\{0\}$	$(0,\bar{s})$	$\{\bar{s}\}$	$(\bar{s},1)$	$\{1\}$	$\hat{s}$	$[0,\hat{s})$	$\{\hat{s}\}$	$(\hat{s},1]$	Condorcet
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)
1*	B B A A A C C B B B A A C C C	1/3	A	A	{A,B}	B	B	–	A	A	A	A
2	B C C A A A B B B B C A A C C	1/3	{A,C}	A	{A,B}	B	B	1/4	A	A	B	B
3*	A A A B B C B B C C B C C A A	1/2	A	A	{A,B}	B	B	1	A	B	–	A
4*	B B A A A A C B B B C A C C C	1/2	A	A	{A,B}	B	B	–	A	A	A	A
5	C A A A B B B B B C A C C C A	1/2	A	A	{A,B}	B	B	–	A	A	A	A
6*	C C A A A B B B B B A A C C C	3/5	A	A	{A,B}	B	B	–	A	A	A	A
7	C A A A B B C B B C A B C C A	2/3	A	A	{A,B}	B	B	1	A	B	–	A
8*	C A A A B B B B B A A C C C C	2/3	A	A	{A,B}	B	B	–	A	A	A	A
9*	A A A C C C B B B B B C C A A	3/4	A	A	{A,B}	B	B	1	A	B	–	A
10	B A A A A C B B B B A C C C C	3/4	A	A	{A,B}	B	B	–	A	A	A	A
11*	C A A A A B B B B B A C C C C	4/5	A	A	{A,B}	B	B	–	A	A	A	A
12	C B B A A A C C B B B A A C C	0	-	-	{A,B}	B	B	1	A	B	–	{A,B,C}
13	B C C A A A B B C B C A A B C	1/2	{A,C}	{A,C}	{A,B,C}	B	B	1/2	A	{A,B,C}	B	{A,B,C}
14	C B B A A A C A B B B A C C C	–	{A,B}	{A,B}	{A,B}	{A,B}	{A,B}	–	A	A	A	A
15	B C C A A A B A B B C A B C C	0	–	–	{A,C}	A	{A,B}	–	A	A	A	A
16	C B B A A A C A B C B A C C B	0	–	–	{A,B}	A	A	–	A	A	A	A
17	B C C A A A A A B B C B B C C	0	–	–	{A,C}	A	A	–	A	A	A	A

Notes: Each profile is displayed as a 3×5 matrix. Columns correspond to Stakeholders and rows to preference ranks. A Stakeholder's first-, second-, and third-ranked alternatives are listed in the first, second, and third rows, respectively. For Condorcet-cyclical profiles, we indicate the set of options in the top cycle. For decisions in the political domain, we only use the profiles indicated with an asterisk.

as before, we see that scoring rules select B over A when $s > 0.6$. Indeed, for each of these 11 profiles, there is a threshold $\overline{s}$ at which the optimal choice for a scoring rule with parameter s switches from the alternative listed in column 7 to the alternative listed in column 5. These thresholds divide the interval $[0,1]$ into a sequence of subintervals with boundaries in the set $\mathcal{C} = \left\{0, \frac{1}{3}, \frac{1}{2}, \frac{3}{5}, \frac{2}{3}, \frac{3}{4}, \frac{4}{5}, 1\right\}$. The subject's decisions place their scoring parameter in one of these intervals. In some cases, we can also distinguish choices corresponding to scoring parameters at the boundaries of these intervals.[7]

The remaining profiles provide further scope for differentiating among social choice rules. For example, because profiles 12 and 13 exhibit Condorcet cycles, all 3 options are best choices according to the top-cycle Condorcet extension but not according to many other rules.

Each social choice rule generates a "fingerprint" of selections across the 17 preference profiles; see Figure 1. Each column in the figure represents a five-Stakeholder three-option preference profile, which we identify in the first panel. For example, the first column corresponds to the profile wherein three Stakeholders prefer A to B to C, while two prefer B to C to A. Each subsequent panel corresponds to a specified social choice rule; it shows the options that rule selects for each preference profile. Each subject's sequence of choices also generates a fingerprint. We will assign each subject to the rule with the fingerprint that matches her own most closely.

Our main classification results encompass 23 *benevolent* social choice rules (meaning that they satisfy the Pareto criterion). We start with the 15 distinguishable rules that emerge from the scoring method. We can differentiate between $s = \frac{1}{2}$ (linearity), 5 ranges of scoring parameters in the strictly convex region and 10 in the strictly concave range. Our data can also differentiate among 5 scoring runoff rules corresponding, respectively, to values of s in the following ranges: $\left[0, \frac{1}{4}\right]$, $\left(\frac{1}{4}, \frac{1}{2}\right)$, $s = \frac{1}{2}$, $\left(\frac{1}{2}, 1\right)$, and $s = 1$.[8] Finally, we can distinguish three social choice rules that emerge from the p-supermajority top-cycle method. These rules correspond to values of p in the following ranges: $\left[\frac{1}{2}, \frac{3}{5}\right]$, $\left(\frac{3}{5}, \frac{4}{5}\right]$, and $\left(\frac{4}{5}, 1\right]$. Henceforth, we adopt a slight abuse of terminology and call these rules Condorcet, Supermajority, and Unanimity, respectively. While our basic classification does not include other Condorect extensions, we conduct robustness analyses to determine whether this omission is material.

Because we distinguish between social choice rules based on their fingerprints, larger differences between fingerprints facilitate more reliable classifications. Supplemental Appendix A.1 shows, for each pair of social choice rules, the number of preference profiles (out of the 17 we use for identification) for which their implications differ. On average, pairs of rules differ on 8.58 of the 17 profiles, and nearly 90 percent of pairs differ on 3 or more profiles.

[7] When a scoring parameter coincides with an interval boundary $s \in \mathcal{C}$, the set of options selected for a given preference profile is the union of those selected with scoring parameter $s - \epsilon$ and those selected with scoring parameter $s + \epsilon$ for $\epsilon > 0$ sufficiently small. Accordingly, scoring rules with parameters $s \in \mathcal{C}$ are less resolute (produce more ties) than scoring rules with parameter $s \notin \mathcal{C}$.

[8] We implement scoring runoff rules as follows. In the first stage, we calculate the score associated with each option and drop the option with the lowest score. We then choose the majority-preferred option from the remaining alternatives. (Because this step involves at most two options, majority rule coincides with all scoring rules.) If the lowest two options are tied, we drop both of them. If all three options are tied, the runoff rule selects all of them. Three-way ties occur only for profile 13, and only for the scoring runoff rule with $s = \frac{1}{2}$. We identify the set of distinguishable scoring runoff rules using a brute-force computer script.

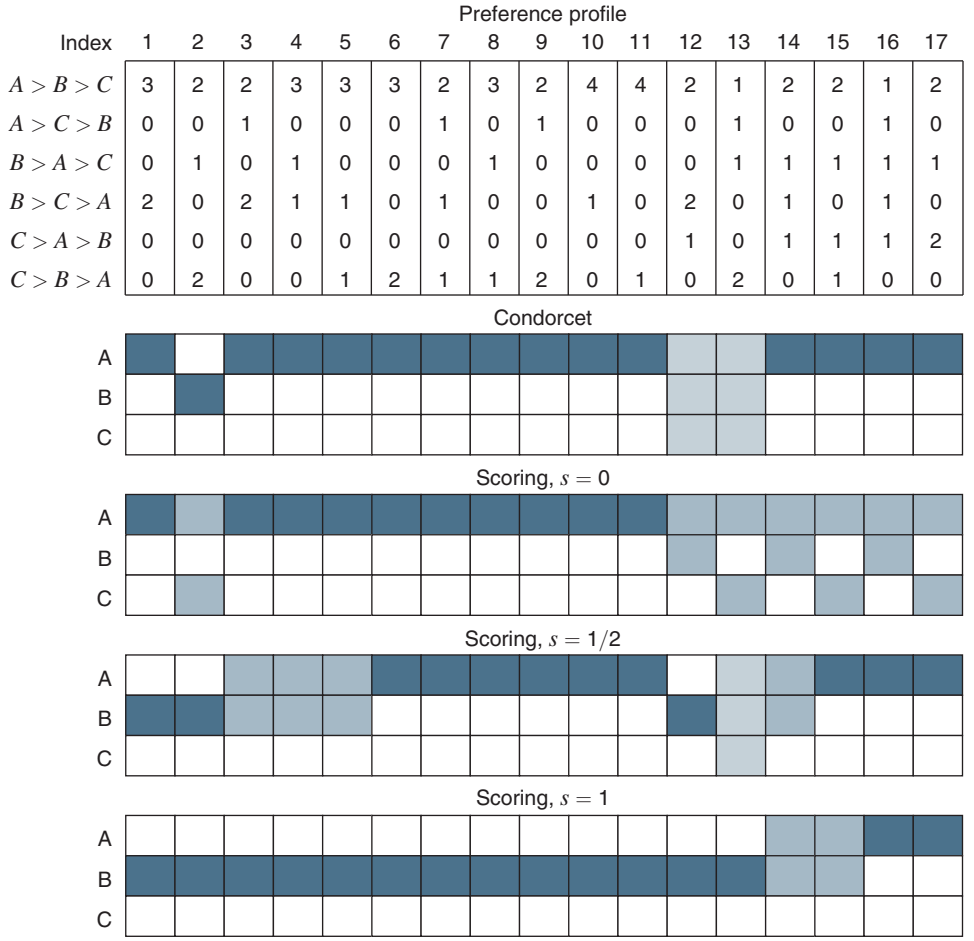

FIGURE 1. FINGERPRINTS OF AN EXAMPLE SELECTION OF ORDINAL AGGREGATION RULES

Notes: Each column corresponds to a three-alternative preference profile. The profiles appear in the top panel: Each cell shows the number of Stakeholders, with the preference ranking indicated to the left of the row. In the second through fifth panels, we use dark shading if the rule chooses the corresponding option and no shading if the rule does not choose the option. Nonresolute rules choose more than one option for some profiles. We use intermediate shades to indicate the number of tied options.

It is worth emphasizing that we selected the set of 17 profiles based on their contributions to identification rather than the frequencies with which they arise in the wild. Our approach assumes that people apply stable aggregation principles regardless of whether a profile is relatively common or uncommon. We corroborate this assumption by examining in-sample fit, out-of-sample predictive power, and out-of-context stability.

II. Experimental Design

Overview.—We assign subjects to one of two roles: Each *Planner* ("she") chooses alternatives that potentially affect groups of five *Stakeholders* ("he"). The only purpose

of including Stakeholders in the experiment is to ensure that the Planners' decisions are consequential. Choice problems fall into two domains, task assignment (*work*) problems and budget allocation (*political*) problems. For each domain, the Planner views a sequence of preference profiles and, in each instance, selects one of several alternatives. We match one out of every four Planners, selected at random, with a real group of Stakeholders. The actual preference profile for that group is among the ones that Planner considers but is not identified as such. Although we only implement decisions that pertain to actual Stakeholder groups, from each Planner's perspective any one of her decisions could turn out to be a real choice. Consequently, as long as she cares about the Stakeholders to some degree, she has an incentive to reveal her aggregation preferences truthfully for every preference profile she encounters. We focus primarily on the work domain, for which we employ the 17 profiles shown in Table 1. We examine the political domain, for which we employ a smaller set of profiles, to evaluate context dependence.

Tasks for the Work Domain.—For the work domain, Stakeholders are *workers* recruited from Amazon Mechanical Turk. Each worker receives $15 for correctly completing a single assigned task. The compensation is sufficiently high to ensure that any attrition is nonsystematic.

There are five work tasks that resemble familiar activities for online workers and that different workers might plausibly rank in different orders: image labeling, hate speech filtering, audio transcription, movie reviews classification, and a logic-based task. Supplemental Appendix A.2 provides details.

Workers reveal their preferences over tasks in a preliminary session. We ensure incentive compatibility by informing workers that their rankings determine the task they perform with 5 percent probability[9] and that some other process will determine their assigned tasks with 95 percent probability. Because we leave that process unspecified, workers cannot report preferences strategically. After the Planners make their choices, workers complete their assigned tasks in a second session.

Social Choice Problems for the Work Domain.—A social alternative is a *task assignment*: It assigns each of the five tasks to one of the five workers in a group. Planners choose from menus of task assignments. For three-option problems, menus consist of three randomly selected task assignments.

For the main portion of our experiment, Planners are students at the University of Zurich and the Swiss Federal Institute of Technology. (We examine US and Swedish general population samples in Section VI). At the beginning of the experiment, they acquaint themselves with the work tasks by reading descriptions and performing abbreviated versions (except for the matching task). To make the allocation tasks less abstract, we advise Planners that Stakeholders are mTurkers and provide information about their hourly compensation in the experiment.

A Planner proceeds through several rounds of decision-making for her group. In each round, she observes an ordinal preference profile for the task assignments. For Planners who are matched with actual workers, one of these corresponds to

[9] We preselect two tasks at random without telling the worker which they are and assign the worker to the one he ranks higher.

her workers' preferences over task assignments, which we infer from their elicited preferences over tasks. The Planner then chooses an assignment she considers best for the group. We also ask her to identify any alternative she considers just as good as the one she selects.

If the Planner knew everything about the task assignments, her decision might depend on whether she agreed or disagreed with Stakeholders' rankings. For example, she might place little weight on an expressed desire to complete the hate speech filtering task. From our perspective, those paternalistic judgments would be confounds. To ensure that Planners cannot second-guess a worker's preferences, we limit the Planner's knowledge. Specifically, we show the Planner a menu of geometric symbols, explain that each represents a task assignment, and describe each worker's ranking of those assignments. However, we do not explain which worker performs which task in any given assignment.

Each Planner makes all decisions using one of several interfaces. The interface shown in Figure 2 represents each worker as a vertical bar. Within each bar, the geometric symbols representing the assignments are ordered according to the worker's preference, with the most preferred assignment on top. By clicking buttons, Planners can highlight or hide options. Highlighting an option makes the distribution of its ranks (the data that are crucial for scoring rules) readily apparent. When an option is hidden, the display repositions the remaining assignments onto two lines. This feature makes pairwise preference counts (the data that are crucial for Condorcet extensions) readily apparent. Additionally, Planners can hide or rearrange the workers, either by dragging and dropping them, or by clicking a button to shuffle them randomly. A possible concern is that the presentation of preference profiles may influence Planners' choices by highlighting particular information. To ensure robustness, we therefore vary the interface across Planners, using *all* structurally distinct possibilities for displaying preference profiles in a two-dimensional table. Supplemental Appendix A.2 provides details.

We randomize symbols (representing options) and colors (representing workers) to ensure Planners do not conflate decisions across preference profiles. We also randomize the positions of all alternatives and of all workers in each round.

In addition to the 17 preference profiles shown in Table 1, each Planner who is matched to a real group of 5 workers views that group's actual preference profile and makes a choice, while other Planners view a randomly generated preference profile. Planners also make 10 additional decisions, randomly intermingled with the others; see Sections III and VA for details. Altogether, Planners make choices for 28 preference profiles in the work domain.

The Political Domain.—To examine the consistency of Planners' aggregation preferences across domains, we also present them with decisions involving political contributions. In this domain, Stakeholders (*citizens*) are voting-age members of the general Swiss population.[10] Planners, who are also voting-age Swiss nationals, direct a contribution of SFr30 (roughly $33.90 at the time of the experiment) to one

[10] We recruited Stakeholders mostly through the survey company pollfish.com. We supplemented this sample by placing ads on Facebook and in the laboratory at the University of Zurich. Age and citizenship are self-reported.

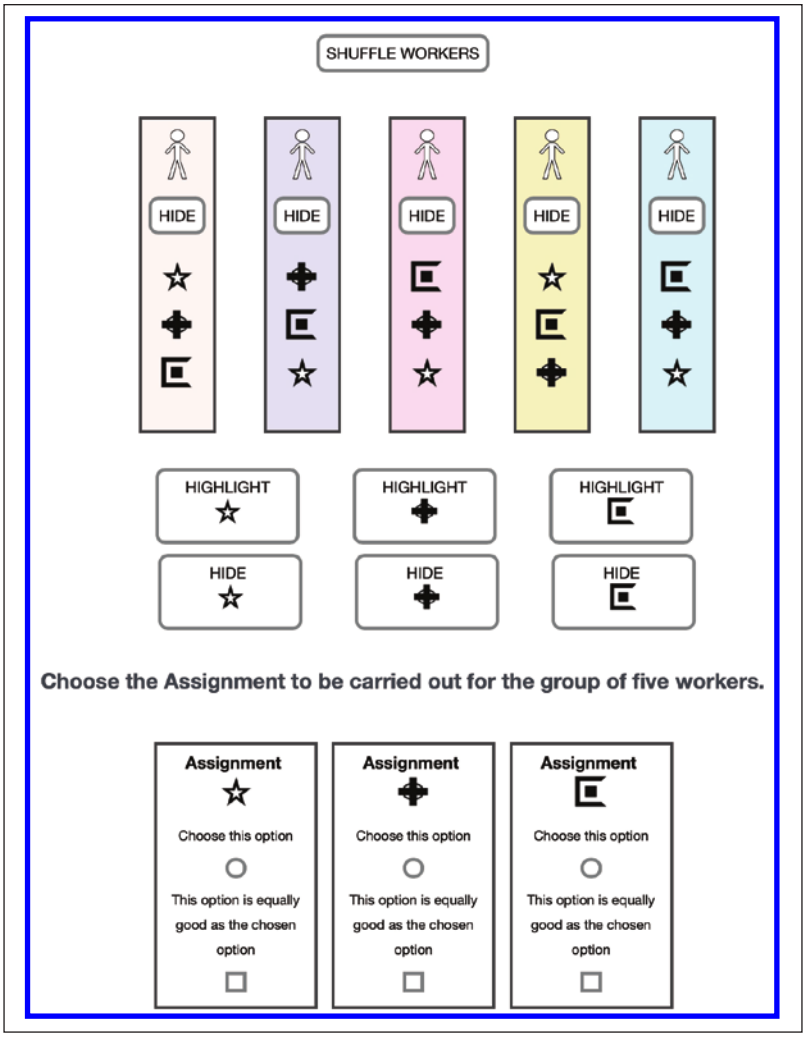

FIGURE 2. PLANNERS' DECISION INTERFACE

Notes: Subjects can drag and drop columns, click a button to shuffle columns, highlight choice options, hide choice options, and hide workers. If a subject hides an option, the remaining options are arranged on two rows regardless of their initial position. Symbols for options and colors for the bars representing each worker are randomly drawn each round. The order of workers is randomly drawn each round.

of the five largest Swiss political parties, as measured by the number of members in the Swiss National Council.

The procedures we use for political contributions are generally the same as for the work domain. Here, it is especially important to emphasize that, for any given preference profile, Planners do not know which geometric symbol corresponds to which party, so they cannot impose their own political preferences.

To keep the length of the experiment manageable for our subjects, we use a smaller set of preference profiles for the political domain. Over 12 rounds, Planners view, in random order, the starred preference profiles in column 2 of Table 1, one additional profile (either the actual profile for her assigned citizen group or a

randomly generated profile, depending on whether we assign her to a group), and a selection of four four-option profiles (see Supplemental Appendix B.2).

Additional Elicitations.—For the work domain, we ask Planners to predict the average reservation wages for the tasks workers rank first, second, and third, knowing only that rankings involve three tasks randomly selected from the five possibilities. For the political domain, we ask Planners to predict the average Swiss citizen's willingness to pay to trigger or prevent the donation the citizen ranks first, second, and third. As we discuss subsequently, we use these responses to assess the hypothesis that Planners' aggregate ordinal preferences are based on implicit inferences about cardinal utility.

We elicit risk preferences using the method developed in Holt and Laury (2002). We also elicit altruism toward workers through modified dictator games. In addition, subjects complete the four-item version of the Cognitive Reflection Test by Thomson and Oppenheimer (2016). For the last 143 subjects, we added 4 questions probing their knowledge of social choice theory. Subjects also report a variety of personal characteristics. See Supplemental Appendix A.2 for further details.

Implementation.—Subjects make all decisions within a domain (work versus political) before proceeding to the next domain. We randomize the order of the two domains, as well as the order of preference profiles within each domain, at the individual level. To avoid nudging subjects to think about cardinal utility, we elicit risk aversion and beliefs about WTA/WTP after the preference aggregation tasks are complete.

A randomly selected decision (pertaining to WTP/WTA, risk preferences, or social preferences) from part C determines the Planner's own payment. As we have explained, we incentivize social choices in the sense that each one may be consequential for other participants, and we also incentivize the elicitation of workers' preferences.

The full instructions for Planners, which we present on-screen in English, appear in Supplemental Appendix E.1. They require subjects to try out each option in the preference display (e.g., hide and highlight). Subjects must pass two multipart comprehension checks to continue with the study.

Data.—Our analysis focuses on 405 subjects in the role of Planner. Subjects participated in 11 online sessions, supervised via videoconferencing software (Zoom), in January 2021. We restricted the sample to Swiss citizens by checking each participants' government-issued ID. All data and code are available in Ambuehl and Bernheim (2026).

The median subject completed the session in 82 minutes and received SFr50 (worth approximately $56.5 at the time of the experiment). Twelve potentially eligible subjects started the study but did not complete it. Only 5 of them did so after presenting a valid ID. The median age is 23. Among the subjects who were asked, 7 percent reported having taken a class that covered social choice theory, but only 1 percent, 3 percent, and 2 percent could correctly name Arrow's theorem, the Condorcet paradox, and the Borda rule, respectively. While our subject pool

Profile	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
Condorcet scoring	A	B	A	A	A	A	A	A	A	A	A	ABC	ABC	A	A	A	A
$\bar{s}$	1/3	1/3	1/2	1/2	1/2	3/5	2/3	2/3	3/4	3/4	4/5	0	1/2	–	0	0	0
$s < \bar{s}$	A	A[a)]	A	A	A	A	A	A	A	A	A	–[b)]	AC	AB	–[a)]	–[b)]	–[a)]
$s > \bar{s}$	B	B	B	B	B	B	B	B	B	B	B	B	B	AB	A[c)]	A	A
Choice freq.																	
A	13	2	36	37	24	38	56	60	72	70	70	8	15	62	93	98	99
B	86	96	64	63	75	61	43	39	25	29	29	91	73	37	6	1	1
C	0	1	0	0	1	0	1	1	3	0	1	1	13	1	1	1	0

FIGURE 3. DISTRIBUTION OF CHOICES IN THREE-OPTION PROBLEMS

Notes: (a) For $s = 0$, $\{A, C\}$ is selected. (b) For $s = 0$, $\{A, B\}$ is selected. (c) For $s = 1$, $\{A, B\}$ is selected. Darker shades of blue indicate higher frequencies. The numerical percentage appears within each shaded cell. The profile numbering is the same as for Table 1.

includes a higher proportion of women (61 percent) than men and skews toward the political Left (70 percent, 15 percent, and 14 percent of subjects rate themselves as Left, Center, and Right, respectively), we obtain similar results in general population samples; see Section VI.

III. Classification Results

In this section, we present our main results concerning the prevalence of the various as-if social choice rules. We consider aggregate choice patterns followed by subject-level classifications.

Aggregate Choice Patterns.—Figure 3 shows choice frequencies for each of the 17 three-option profiles in the work domain. Structure is readily apparent, such as the near-unanimous decisions for profiles 2, 16, and 17. Because nearly all subjects choose the option that is both rank-dominant and a Condorcet winner in profiles 16 and 17, we can infer that they act benevolently toward their groups. Notably, disagreements are common for the score-identifying profiles 3 to 11, indicating heterogeneity in aggregation preferences.

Patterns for profiles 1 and 2 anticipate our overall conclusions. If, just for the moment, we confine attention to scoring rules and the Condorcet top-cycle extension, we would infer that 86 percent of subjects use scoring rules with $s > \frac{1}{3}$ (because they choose B in profile 1), that only 2–3 percent of subjects use scoring rules with $s < \frac{1}{3}$ (because they choose A or C in profile 2), and that 10 percent of subjects use the Condorcet rule (because the fraction of B increases from 86 percent to 96 percent between profiles 1 and 2). The third-option comparisons inherent in scoring rules also point to widespread violations of Arrow's Independence of Irrelevant Alternatives axiom.

Choices for profiles 12 and 13 corroborate the low prevalence of the Condorcet rule with the top-cycle extension, in that we see little indication of indeterminacy despite the presence of cycles. For profile 12, 91 percent of subjects agree on option B, which is rank-dominant, and for profile 13, 73 percent of subjects select option B, which minimizes the number of last-place ranks.

Choices for profiles 3, 4, and 5 corroborate the prevalence of concave scoring rules, as 64 percent, 63 percent, and 75 percent of subjects select option B, respectively. In contrast, all convex scoring rules and the Condorcet rule select option A, which is chosen by 36 percent, 37 percent, and 24 percent of subjects, respectively. Either pattern is consistent with the Borda rule, which is only partially resolute for these profiles.

Classification Methods.—Next, we assign each subject to the social choice rule that best matches their selections. We employ a Bayes classifier (see, e.g., Webb 2010), which assigns each subject the social choice rule with the greatest posterior probability conditional on the subject's observed choices. Critically, this method gives irresolute rules less "credit" (in terms of the increment to the posterior probability) than resolute rules when both turn out to be consistent with a given choice. Without this feature, a classification method would favor irresolute rules artificially. To compute posterior probabilities, we make four assumptions (analogously to Costa-Gomes, Crawford, and Broseta 2001): (i) The prior distribution over prespecified rules is uniform. (ii) For each of the 17 profiles, the subject follows her assigned rule with probability $1 - \epsilon$ and uniformly randomizes over the three options with probability ϵ. The prior distribution of ϵ is uniform over $[0, 1]$. (iii) Decision errors are independent across preference profiles. (iv) When a rule is irresolute, the subject randomizes uniformly and independently over the prescribed choices.[11] We use the popular Maximum A Posteriori (MAP) criterion, which assigns each subject the rule R^* and noise level ϵ^* that maximize the posterior probability, $P(R, \epsilon | c)$, where c is their choice vector. When more than one rule maximizes the Bayesian posterior, we assign the subject to a maximizing rule at random. Supplemental Appendix A.3 derives the classifier formally and shows that it performs well in simulations.

We assign each subject to 1 of 39 rules. To the set of 23 benevolent social choice rules discussed in Section I, we add malevolent versions that apply the benevolent versions after inverting the Stakeholder's preference rankings. For technical reasons, the Bayes classifier will never select a benevolent rule with $s \in \left\{\frac{3}{5}, \frac{4}{5}\right\}$ or a malevolent rule with $s \in \left\{\frac{1}{3}, \frac{3}{5}, \frac{2}{3}, \frac{3}{4}, \frac{4}{5}\right\}$.[12] Any subject who employs one of these seven rules is classified as using either a slightly higher or slightly lower value of s. For similar reasons, any subject with $s = 1$ tends to be classified as using $\frac{4}{5} < s < 1$, so our exposition merges the estimates of these two categories. Differentiating among all 46 rules requires us to use the subjects' expressions of

[11] Because our displays present alternatives and workers in random order, redrawn in each round, even positional criteria such as always choosing the option on the left will yield uniform distributions.

[12] Each of these rule differs from nearby scoring rules on exactly one profile. Moreover, for the single differentiating profile, each prescribes the union of the options selected when the scoring parameter is slightly larger and when it is slightly smaller. Because our classifier always favors resoluteness over irresoluteness when both possibilities are consistent with the same observation, it will never select such a rule over all nearby scoring rules.

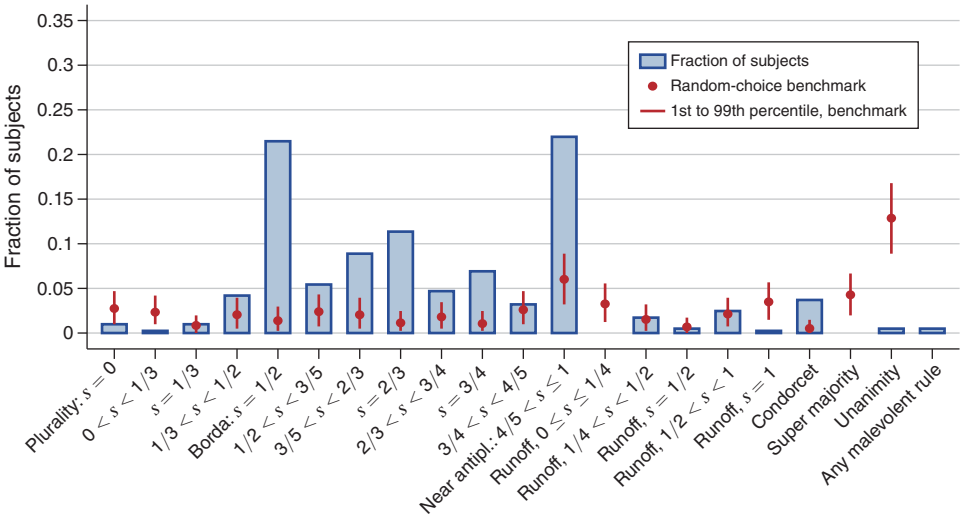

FIGURE 4. BEST-FITTING PRESPECIFIED RULES

Note: The random-choice benchmark for malevolent rules exceeds the range of the graph (44.8 percent).

indifference in addition to their choices. While those expressions do not have consequences, each concerns a consequential choice the subject has just made, and there is no reason to answer incorrectly. Supplemental Appendix B.1 shows that we obtain similar results.

Results.—Figure 4 displays our main classification results. Four features merit emphasis.

First, choices for a clear majority of subjects are consistent with *strictly concave scoring rules* ($\frac{1}{2} < s \leq 1$; 62.5 percent). In contrast, choice patterns consistent with strictly *convex* scoring rules ($s < \frac{1}{2}$) are uncommon. The most convex rule (plurality) describes the choices of surprisingly few subjects (1 percent).

Second, we assign the largest number of subjects to the Borda rule and to near antiplurality (with $4/5 < s \leq 1$). Together, these 2 cases account for over 40 percent of all subjects.

Third, our classification procedure assigns fewer than 5 percent of subjects to the Condorcet rule and almost none to other p-supermajority rules (supermajority and unanimity). A small minority follow scoring runoff rules (4.9 percent). Because our prespecified rules contain only one Condorcet extension (top cycle), these classifications could in principle understate the prevalence of choices consistent with the Condorcet class. We therefore enlarge the set of prespecified rules to include *all conceivable* Condorcet extensions, and reclassify subjects using our first procedure.[13] The classifier assigns 8.9 percent of subjects to a Condorcet extension, a

[13] Because all but 2 of the 17 profiles admit a Condorcet Winner, there are only 9 resolute Condorcet extensions. As long as we include all of them, there is no need to add any irresolute Condorcet extensions because our classifier never assigns a subject to a less resolute rule that is equally consistent with the subject's choices.

modest increase, some of which would occur by chance. It is unlikely that the low prevalence of Condorcet reasoning is a statistical or experimental artifact because the Condorcet rule coincides on a large number of profiles with convex scoring rules, which are also unpopular. Supplemental Appendix B.2 provides further corroboration of this finding based on four-option problems.

Fourth, we see little evidence that subjects are malevolent or lazy. The fraction of subjects assigned to malevolent rules is de minimis. Lazy subjects would be inclined to select the most easily implementable rule. Plurality rule is arguably the least cognitively demanding alternative, followed by plurality runoff, yet the vast majority of subjects evidently find the distributional implications of both unappealing. With the exception of antiplurality, the concave scoring rules, which describe the majority of our subjects, embody more nuanced and complex judgments. An analysis of decision times corroborates these assessments: Decision-making is relatively rapid for subjects assigned to plurality rule and relatively slow for those assigned to scoring rules with $0.5 \leq s \leq 1$ (see Supplemental Appendix B.3).

Supplemental Appendix B.4 shows that these results are robust across our three formats for displaying preference profiles to subjects.

Random Benchmark.—We draw statistical inferences by comparing our classification to a random-choice benchmark. We generate 4,000 artificial subjects whose simulated selections for each profile are uniformly distributed across options. We use the Bayes classifier to assign each simulated subject to a prespecified social choice rule. Finally, we draw 1,000 bootstrap samples of 405 simulated subjects and compute the first and ninety-ninth percentiles for the percentages of simulated subjects assigned to each rule.

Figure 4 plots, for each rule, the mean fraction of simulated subjects assigned to that rule, as well as the first and ninety-ninth percentiles of the corresponding distribution. The average fraction of simulated subjects assigned to a malevolent rule is 44.8 percent, which exceeds the range of the figure.[14] Turning to benevolent rules, we see that the fraction of subjects assigned to every weakly concave scoring rule except $\frac{3}{4} < s < \frac{4}{5}$ exceeds the ninety-ninth percentile of the random benchmark. In contrast, the fractions of subjects assigned to the two most convex scoring rules ($s < \frac{1}{3}$) fall short of the first percentile of the random benchmark. While the fraction of subjects assigned to the Condorcet rule is small, it is larger than we would observe by chance, a conclusion for which we find corroboration in Supplemental Appendix B.2. Finally, scoring runoff rules, supermajority, and unanimity are no more prevalent (and in some cases significantly less prevalent) than for the random-choice benchmark.

Goodness of Fit.—The average value of the estimated noise parameter ϵ^* is 0.10, which signifies that the average subject chose randomly in just 1.7 of the 17 preference profiles. A whopping 42.0 percent of our subjects fit their assigned rule perfectly. For the remainder, the mean noise parameter is 0.16, equivalent to choosing randomly for 2.7 of the 17 preference profiles. In contrast, the average estimated

[14] This benchmark deviates from 50 percent because our selection of profiles is not random.

Panel A. Subjects classified as Borda ($s = 0.5$)

Empirical choices

A	8	0	66	55	40	84	94	95	95	95	97	1	28	66	99	98	100
B	92	99	34	45	60	16	5	5	5	5	3	99	48	34	1	2	0
C	0	1	0	0	0	0	1	0	0	0	0	0	24	0	0	0	0

Rule prediction

A	0	0	50	50	50	100	100	100	100	100	100	0	33	50	100	100	100
B	100	100	50	50	50	0	0	0	0	0	0	100	33	50	0	0	0
C	0	0	0	0	0	0	0	0	0	0	0	0	33	0	0	0	0

Panel B. Subjects classified as near antiplurality ($4/5 < s < 1$)

Empirical choices

A	5	0	10	14	1	1	15	11	33	14	0	4	3	63	100	100	100
B	95	100	90	86	96	99	85	89	62	86	100	96	96	37	0	0	0
C	0	0	0	0	3	0	0	0	5	0	0	0	1	0	0	0	0

Rule prediction

A	0	0	0	0	0	0	0	0	0	0	0	0	0	50	100	100	100
B	100	100	100	100	100	100	100	100	100	100	100	100	100	50	0	0	0
C	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

FIGURE 5. EMPIRICAL FINGERPRINTS OF CLASSIFIED SUBJECTS

Notes: In both panels, the top half displays the choices by subjects assigned to a certain rule. The bottom half displays the choices the rule predicts. Darker shades of gray indicate higher frequencies. The numerical percentage appears within each shaded cell. For the theoretical rule, we take the frequency of each best option to be 50 percent for a 2-way tie and 33 percent for a 3-way tie.

noise parameter for our simulated random-choice data is 0.47, and 0.58 when excluding simulated subjects classified as supermajority or unanimity (10 of 17 profiles).[15]

A more detailed picture of goodness of fit emerges from comparisons between the theoretical fingerprints associated with particular rules and the empirical fingerprints of the subjects assigned to those rules. The top half of Figure 5, panel A, plots the distribution of choices, by profile, for subjects assigned to the Borda rule. The bottom half of the same panel shows the Borda prescriptions. The fit is plainly tight. While these subjects rarely depart from the Borda rule, they do deviate somewhat from uniform resolution of ties. Figure 5, panel B provides analogous information for the near-antiplurality rule, which differs from Borda on 10 of the 17 profiles. The fit is also good, if slightly less tight.

Overall, the 21 prespecified benevolent social choice rules fit the data remarkably well given the astronomical number of possible rules in our setting. To further allay possible concerns about overfitting, we show in Section IV that our classification predicts well out of sample.

[15] The mean estimated noise parameter in the random-choice data never equals its actual value (unity) because some random-choice sequences will always match some choice rules relatively well.

Completeness of the Classification.—To determine whether our classification analysis excludes empirically important rules, we look for clusters of subjects whose choices more closely resemble each others' than any of the prespecified possibilities. Our approach is similar to that of Costa-Gomes and Crawford (2006). To identify clusters, we use the *k*-modes clustering algorithm, which is an adaptation of the well-known *k*-means method to categorical data (Huang 1998). We use our prespecified rules to create additional potential cluster modes that are always present as the algorithm iterates. Because we do not use the indifference data for this exercise, we include every resolute version of each prespecified rule. Following Costa-Gomes and Crawford (2006), if a subject is equidistant from a prespecified rule and an endogenous cluster, we assign them to the prespecified rule. We run the algorithm multiple times, gradually increasing the fixed number of endogenous clusters *k* until no further nontrivial clusters emerge.

We find only two nontrivial endogenous clusters. The larger includes just under 5 percent (19 of 405) of subjects; the smaller includes fewer than 2 percent (7 of 405) of subjects. The rule for the larger one is a one-profile deviation from near antiplurality: On profile 9, it selects option A rather than B. This profile is the only one for which near antiplurality selects an option ranked last by some Stakeholder and not ranked first by any other Stakeholder. Supplemental Appendix B.5 presents additional details.

Corroboration Based on Class-Separating Preference Profiles.—We corroborate our findings on the importance of scoring rules relative to Condorcet extensions using two additional preference profiles, each of which cleanly separates these classes. Both of these profiles involve four options rather than three. Critically, each has a unique Condorcet winner that is rank-dominated. *Proper* scoring rules (which do not assign the same score to any two ranks) never select rank-dominated options. Only 7.4 percent of choices are consistent with a preference for the Condorcet winner, which is in line with the finding in Figure 4. A related analysis corroborates the importance of scoring rules relative to the plurality runoff rule. Supplemental Appendix B.2 provides details.

IV. Structural Stability versus Context Dependence

Does our classification capture subjects' structural aggregation principles, contextual judgments, or both? To answer this question, we first compare choices across our two domains. Because subjects made decisions for fewer preference profiles in the political domain, we reestimate the rule distribution for the work domain using only the seven profiles common to both domains. Doing so reduces the number of rules we can distinguish. Specifically, we can no longer identify scoring rules involving any value of s that lies at the boundary between two intervals.[16] Moreover, the fingerprints of scoring rules with $s < \frac{1}{3}$ (including plurality rule), the Condorcet

[16] Without using subjects' expressions of indifference, we cannot distinguish any such scoring rule unless the rules for the two adjacent intervals differ on at least two preference profiles. Supplemental Appendix A.1 shows the distance between all pairs of prespecified rules in the political domain.

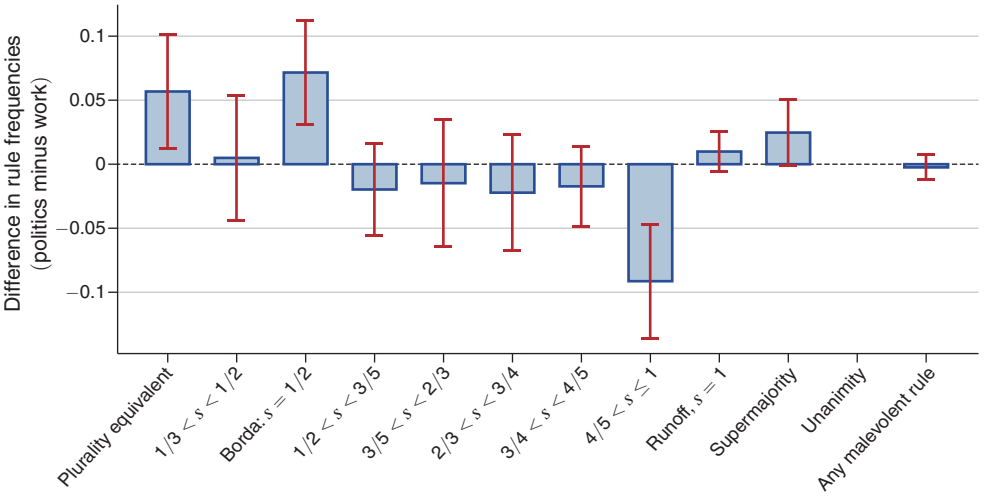

FIGURE 6. COMPARISON BETWEEN WORK AND POLITICAL DOMAINS

Notes: Classifications are based on the same set of seven preference profiles for both domains. The category "plurality equivalent" includes scoring rules with $s < 1/3$, the Condorcet rule, and all scoring runoff rules with $s < 1$. Whiskers denote 95 percent confidence intervals.

rule, and all scoring runoff rules with $s < 1$ coincide for these seven profiles. For our current purposes, we will call this group the "plurality-equivalent rules."

Figure 6 shows the differences in the resulting distributions of classifications for the work and political domains. Systematic differences are immediately apparent. Within the political domain, subjects' choices are less frequently consistent with strictly concave scoring rules, especially $\frac{4}{5} < s \leq 1$, and more frequently consistent with the Borda rule or a plurality-equivalent rule. The difference is highly statistically significant ($p < 0.01$, Wilcoxon signed-rank test applied to the 400 subjects classified as following a benevolent rule in both domains).

Crucially, these differences are of limited magnitude. The increase in the frequency of plurality-equivalent rules, for instance, is on the order of 5.7 percent, while the decrease in the frequency of all strictly concave rules is 16.5 percent. The absence of larger differences suggests a meaningful degree of structural stability.

To evaluate the stability of aggregation preferences more comprehensively, we examine the out-of-sample and cross-domain predictive power of our classifications, using all profiles available in each domain. For each evaluation, we designate a training set of preference profiles, which we use to classify Planners, as well as a test set that we use to evaluate predictive success. We score predictions as follows. If the actual choice lies outside the predicted best-choice set, the score is 0. If it lies within the predicted best-choice set, the score is 1, $\frac{1}{2}$, or $\frac{1}{3}$ depending on whether there are 1, 2, or 3 best options. With this scoring system, if underlying choices are in fact uniformly random, the average score will be $\frac{1}{3}$ irrespective of the sizes of predicted best-choice sets. If all of our individual-level assignments were correct, we would obtain a score of 0.924 in the work domain and 0.955 in the political domain, given the estimated distributions of prespecified rules for each domain.

TABLE 2—OUT-OF-SAMPLE PREDICTIVE POWER OF THE BAYESIAN CLASSIFICATION

	Fraction of correct predictions (weighted by resoluteness)	
Test domain	Work (1)	Politics (2)
Predictions		
Training domain		
Work	0.853	0.758
Politics	0.805	0.755
Benchmarks		
1. Theoretical maximum	0.924	0.955
2. Random choice	0.333	0.333
3a. Estimated rule frequencies, work domain		
Mean	0.736	0.617
Ninety-ninth percentile	0.754	0.645
3b. Estimated rule frequencies, political domain		
Mean	0.716	0.642
Ninety-ninth percentile	0.736	0.669

Note: Within-domain predictions are based on a leave-out-one approach.

These numbers are less than 1 because some rules are irresolute, and we do not seek to predict how subjects resolve ties. We conduct both within- and cross-domain predictions. For the within-domain predictions, we use the leave-one-out approach. For cross-domain predictions, we designate all profiles in one domain as the training set and all profiles in the other domain as the test set.[17]

As shown in Table 2, the average predictive accuracy scores for our four evaluations range from 0.755 (politics to politics) to 0.853 (work to work). In all four cases, predictive performance is far superior to the random-choice benchmark of $\frac{1}{3}$. However, that benchmark is a low bar. Our predictions would outperform it even if they merely captured generally reasonable tendencies, such as avoiding dominated alternatives. We construct a more demanding benchmark by randomly assigning subjects to aggregation rules using the estimated distribution of rules for each domain. We obtain the distributions for the benchmarks by drawing 1,000 bootstrap samples of these random assignments. In each case, we find that the average score for our predictions exceeds the ninety-ninth percentile of the benchmarks' distribution by a wide margin. Considering that the theoretical maximum for the work domain is 0.924, the within-domain score of 0.853 implies that our classification achieves 62 percent of the potential gain in predictive accuracy over the work-domain benchmark (0.736). Similarly, considering that the theoretical maximum for the political domain is 0.955, the within-domain score of 0.755 achieves 36 percent of the potential gain in predictive accuracy over the political-domain benchmark (0.642). The superior performance within the work domain is driven by the inclusion of profiles for which the choices of many popular rules coincide, which makes them easier to predict.

[17] The training set does not always contain sufficient information to distinguish among all rules that have different predictions in the test set. In these cases, we treat all rules that maximize the Bayesian posterior in the training set for a given subject as equally likely and select one at random.

We conclude that the assigned rules capture important aspects of subjects' general aggregation preferences. While we find some degree of context specificity, the relative accuracy of the cross-domain predictions reassures us that ordinal aggregation entails stable structural elements.

V. Why Does Ordinal Information Matter?

What is the role of ordinal information in Planners' social preferences? Do Planners seek to aggregate cardinal utilities by inferring them from preference ranks when cardinal information is unavailable (in the spirit of Apesteguia, Ballester, and Ferrer 2011), or do they have intrinsic preferences over ordinal outcomes, such as a desire to avoid alternatives that some Stakeholders rank last? Section VA demonstrates that the cardinal inference mechanism plays an important role; Section VB shows that Planners also care about ordinal outcomes intrinsically.

A. *Do Planners Draw Cardinal Inferences?*

We use two strategies to investigate the role of cardinal inferences. First, we ask whether Planners' beliefs about Stakeholders' cardinal valuations, which we elicited, predict their ordinal aggregation principles. Second, we test an implication of the cardinal inference hypothesis concerning the effects of an informational intervention.

Relation between As-If Scoring Parameters and Beliefs about Cardinal Valuations.—A money-metric utilitarian should choose the alternative x that minimizes $\sum_n u_{R(n,x)}$, where $R(n,x)$ is the rank Stakeholder n assigns to option x and u_r is the average dollar-equivalent reservation wage the Planner associates with a Stakeholder's r-th ranked task. This objective is equivalent to applying the scoring rule that assigns, respectively, one point, $\tilde{s}$ points, and zero points for Stakeholders who receive their top-ranked, second-ranked, and third-ranked options, where $\tilde{s} = (u_3 - u_2)/(u_3 - u_1)$.

Because we elicited Planners' beliefs about Stakeholders' reservation wages, u_r, we can examine the relation between the synthetic utilitarian scoring parameter $\tilde{s}$ and the best-fitting scoring parameter s across subjects. To avoid selection effects, we assign every subject to a scoring rule even if some other rule is a better fit. We exclude nine subjects in the work domain and six in the political domain who assign lower reservation wages to lower-ranked alternatives, as well as three who are assigned to malevolent rules. We estimate OLS regressions using interval midpoints whenever best-fit scoring parameters are interval identified. We pool data across domains to increase statistical power. The specification includes fixed effects for the decision domain, preference profile presentation modes, and the order in which the subject made decisions about the work domain and the political domain. We also cluster standard errors at the subject level where appropriate.

In column 1 of Table 3, the coefficient of the synthetic scoring rule is positive and statistically significant, indicating that the as-if scoring rules are partly informed by subjects' beliefs about cardinal utilities. Because our measurements of these beliefs,

Table 3—Relation between Best-Fitting Scoring Parameters and Beliefs about Reservation Prices

Variable	Estimated scoring parameter s			
	(1)	(2)	(3)	(4)
Scoring parameter $\tilde{s}$				
Level	0.115		0.102	
	(0.056)		(0.056)	
%-rank		0.062		0.055
		(0.023)		(0.023)
Risk tolerance				
Level			−0.717	
			(0.314)	
%-rank				−0.054
				(0.028)
Altruism				
Level			0.178	
			(0.223)	
%-rank				0.048
				(0.028)
Political domain	−0.111	−0.111	−0.111	−0.112
	(0.010)	(0.010)	(0.010)	(0.010)
Observations	792	792	792	792
Subjects	405	405	405	405

Notes: Parameters estimated with OLS using midpoint values of best-fit score. Regressions include all subjects with monotonic beliefs about reservation prices. Risk tolerance is measured as the certainty equivalent of an equiprobable win 40/win 10 lottery implied by individual-level CRRA estimates. Altruism is measured as the maximum amount of Swiss francs the Planner is willing to give up to increase each Stakeholder's payoff by $2. For subjects with multiple switches in the risk preference or altruism elicitation, the corresponding variable is set to the mean value among the other subjects. We also include two indicator variables for the presence of multiple switching points, one for each characteristic. All regressions control for the type of preference profile presentation and for whether the political domain was displayed before or after the work domain. Standard errors are clustered by subject.

and hence of the synthetic scoring parameters, are potentially noisy, that coefficient may be biased. Column 2 limits the influence of outliers in $\tilde{s}$ by substituting subjects' percentile rank within the $\tilde{s}$ distribution. While this change in the predictor variable prevents a direct comparison with the previous estimate, the statistical significance of the resulting coefficient estimate increases substantially, as expected. Columns 3 and 4 replicate these regressions including controls for risk tolerance and altruism. They yield similar estimates of the coefficients on the synthetic scoring parameter and its percentile rank. The coefficient of risk tolerance is negative, as one would expect if Planners maximize the expected utility of the representative Stakeholder assessed behind a "veil of ignorance," assuming random assignment of rankings to Stakeholders. In that case, greater risk tolerance implies greater convexity of the scoring rule. We also find that more altruistic Planners tend to use more concave scoring rules. Additional analysis shows that the relation between synthetic and best-fitting scoring parameters is stronger for the political domain than the work domain, though not to a statistically significant extent (see Supplemental Appendix B.6).

Table 4—Information Treatments

	(1)	(2)	(3)
Panel A. Baseline profiles			
Profiles	B B A A A A A B B B	B A A A B A B B B A	A A B B C B B A C A C C C A B
Choice distribution	A B 0.993 0.007	A B 0.993 0.007	A B C 0.622 0.368 0.010
Panel B. Adding information about an unavailable option X			
Profiles	B B A A A A ~~X~~ B B B ~~X~~ A ~~X~~ ~~X~~ ~~X~~	~~X~~ A A A B B B B B ~~X~~ A ~~X~~ ~~X~~ ~~X~~ A	A A B B C B B A C ~~X~~ C C C ~~X~~ A ~~X~~ ~~X~~ ~~X~~ A B
Choice distribution	A B 0.427 0.573	A B 0.358 0.642	A B C 0.217 0.674 0.109
Panel C. Adding an available option X			
Profiles	B B A A A A X B B B X A X X X	X A A A B B B B B X A X X X A	A A B B C B B A C X C C C X A X X X A B
Choice distribution	A B X 0.365 0.630 0.005	A B X 0.240 0.748 0.012	A B C X 0.205 0.694 0.099 0.002
Panel D. p-values of the difference of choice distributions			
Comparison across panels			
A and B	0.000	0.000	0.000
A and C	0.000	0.000	0.000
B and C	0.531	0.021	1.000

Notes: Column 3 of panel A is profile 14 of Table 1. In panel C, column 3 is profile 18 from Supplemental Appendix Table B.1, and columns 1 and 2 are profiles 4 and 5 from Table 1, respectively. Columns 1 and 2 of panel A both display the distribution of choices that subjects made for the single two-alternative profile they encountered. All decisions concern the work domain. *p*-values are based on two-sample Kolmogorov-Smirnov tests for equality of distributions among options other than *X*.

Informational Interventions.—Next, we test an implication of the cardinal information mechanism. To appreciate this implication, consider the preference profiles depicted in column 1 of Table 4. The one in panel B is the same as the one in panel A, except that it also lists a third option, *X*, that is not available. One Stakeholder prefers *X* to *A*, while none prefer it to *B*. If Planners' choices depend on cardinal inferences, this additional information should reduce *A*'s attractiveness to the Planner. In contrast, if a Planner's choices are based solely on the ordinal ranks among *available* options, the information should have no effect. In fact, we see that the fraction selecting *A* falls sharply from 0.993 to 0.427, and the change is highly statistically significant (panel D). Likewise, under the cardinal inference hypothesis, one would expect that listing the unavailable option (*X*) shown in columns 2 and 3 of panel B would shift Planners' choices away from *A* in these columns. The data strongly support both predictions. While an aversion to last-place ranks among *listed* options would also predict the effects in columns 1 and 2, it would not predict the effect in column 3.

In all three columns, the option *X* is relatively unattractive: When we make it available, very few Planners' choose it (panel C). Consequently, if Planners use ordinal information *only* to make cardinal inferences, *X*'s availability (as opposed to its mere listing) should leave the choice frequencies for the other options unchanged.

We obtain mixed evidence concerning this implication. The choice distribution changes significantly for column 2 ($p < 0.03$) but not for columns 1 and 3. The former finding raises the possibility that, in addition to making cardinal inferences based on ordinal information, Planners may care intrinsically about the ordinal positions of available options. We explore this possibility more thoroughly in the next subsection.

The comparison of panels A and C also reveals a striking violation of the choice axiom known as Sen's α. According to that axiom, removing an unchosen option from a menu should not change the chosen option. In fact, choices change dramatically.

B. *Do People Care about Ordinal Information Intrinsically?*

To investigate the possibility that Planners may care about ranks intrinsically, we conduct an additional experiment in which social alternatives are distributions of monetary payoffs. By overtly revealing these payoffs to Planners, we ensure that they no longer use ordinal information to draw inferences about cardinal payoffs.

Design.—A laboratory subject in the role of Planner chooses among vectors of payoffs for three Stakeholders. To test whether ordinal information affects Planner's choices even when inference about cardinal payoffs can no longer play a role, we use two types of treatments, *Swap* and *Addition*.

To understand the Swap treatments, consider the cardinal payoff profile $S1$ in column 1 of panel A of Table 5. Assuming Stakeholders only care about their own payoffs, they rank the alternatives as shown in column 2. The Swap treatment swaps participants' cardinal payoffs within alternatives, as highlighted in profile $S1$. These particular swaps yield the cardinal payoff profile $S1'$. Notice that option A delivers a payoff of six to one Stakeholder, five to a second, and three to a third in both $S1$ and $S1'$. Similar statements hold for options B and C. Therefore, the swaps do not alter the relative attractiveness of the three options according to any anonymous Bergson-Samuelson social welfare function. However, the effect on ordinal rankings is potentially consequential: While the rank-dominant alternative in profile $S1$ is B, the rank-dominant alternative in the altered profile $S1'$ is A. Hence, if the frequency with which Planners select alternative B is higher for profile $S1$ than for $S1'$, an intrinsic concern for (within-menu) rankings of alternatives provides a plausible explanation, but concern for anonymous cardinal payoff profiles does not.

For profiles $S1$ and $S1'$, any Planner who applies an anonymous Bergson-Samuelson social welfare function is indifferent between options A and B. Consequently, a comparison of choices for those profiles can only reveal whether a Planner breaks cardinal ties based on ordinal information. To determine whether Planners are also willing to sacrifice the quality of the cardinal outcome in order to improve the (within-menu) ordinal outcome, we also study Swap treatments in which the payoffs for option A are higher than in $S1$ for all three Stakeholders by a small amount $\epsilon > 0$. This change renders option A strictly better than option B according to any anonymous Bergson-Samuelson social welfare function. If the frequency

Table 5—Cardinal Payoff Profiles

(1)				(2)			
Panel A. Swap treatments							
*S*1: Original cardinal payoff profile				Implied ordinal preferences			
Option	Stakeholder			Preference	Stakeholder		
	1	2	3		1	2	3
A	6	3 swap ⟷	5	Best	A	B	B
B	3 swap ⟷	5	6	Middle	B	C	A
C	1 swap ⟷	4	2	Worst	C	A	C
*S*1′: Altered cardinal payoff profile				Implied ordinal preferences			
Option	Stakeholder			Preference	Stakeholder		
	1	2	3		1	2	3
A	6	5	3	Best	A	A	B
B	5	3	6	Middle	B	B	A
C	4	1	2	Worst	C	C	C
Panel B. Addition treatments							
*A*1: Original cardinal payoff profile				Implied ordinal preferences			
Option	Stakeholder			Pref. rank	Stakeholder		
	1	2	3		1	2	3
A	3	2	2	Best	A	B	B
B	1	3	3	Worst	B	A	A
*A*1′: Altered cardinal payoff profile				Implied ordinal preferences			
Option	Stakeholder			Preference	Stakeholder		
	1	2	3		1	2	3
A	3	2	2	Best	A	B	B
B	1	3	3	Middle	X	A	A
X	2	1	1	Worst	B	X	X

with which Planners select alternative *B* is also higher for the modified *S*1 than for the modified *S*1′, we can infer that Planners' use of ordinal information is not limited to breaking cardinal ties.

Our second type of treatment, *Addition*, resembles those studied in Section VA: We modify a profile by adding an alternative that Planners will likely avoid but that alters the rank distributions of the remaining alternatives. In contrast to those previous treatments, Planners know the cardinal payoffs associated with each option and are aware that the addition of the new alternative leaves those payoffs unchanged. Therefore, any Planner who applies a Bergson-Samuelson social welfare function will continue to select the same option. Accordingly, if choices shift toward the option that becomes more attractive from an ordinal perspective, an intrinsic concern for (within-menu) rankings provides a plausible explanation, but concern for cardinal payoffs does not. The profiles we use for this treatment include *A*1 in panel B of Table 5. For this profile, all scoring rules select the majority winner *B*. Adding the unattractive alternative *X* turns it into profile *A*1′. For this profile, scoring rules with $s > \frac{1}{2}$ will instead select option *A*. Consequently, we ask whether the addition of *X* increases the frequency with which Planners select *A* or leave it unchanged.

TABLE 6—CHOICES WITH CARDINAL PAYOFF INFORMATION

	Option *A* chosen		
Profile	*S*1	*S*1′	Difference
Panel A. Swap treatments			
$\epsilon = 0$	0.300	0.712	0.412
	(0.013)	(0.012)	(0.022)
$\epsilon > 0$	0.755	0.901	0.146
	(0.014)	(0.010)	(0.015)
Difference	0.455	0.189	−0.265
	(0.017)	(0.013)	(0.021)
	Option *B* chosen		
Profile	*A*1	*A*2	
Panel B. Addition treatments			
X unavailable	0.220	0.496	
	(0.014)	(0.019)	
X added	0.175	0.396	
	(0.013)	(0.018)	
Difference	0.045	0.099	
	(0.014)	(0.014)	

Notes: Panel A shows the frequency with which subjects choose option A. Panel B shows the frequency with which subjects choose option B. Standard errors clustered by subject. Pooled across display formats.

Implementation.—We ran the experiment at the University of Zurich's Experimental Economics Laboratory between March and May 2023 with 584 subjects in the role of Planner. To increase statistical power, we include a second profile *A*2 for the Addition treatment[18] and have Planners make decisions for multiple rescaled versions of all profiles, with scaling factors ranging from 2 to 5. (Additional minor changes to the payoff profiles *A*1 and *A*2 prevent overlapping information in visual depictions.) Supplemental Appendix C.1 presents details. Planners completed the corresponding 20 tasks in random order. Each Planner knew there was a 10 percent chance she would be assigned to an actual group of Stakeholders, in which case we would implement one of her decisions, selected at random. As in the foregoing experiment, this design ensures that, as long as the Planner is not indifferent about the Stakeholders' welfare, she has an incentive to reveal her true social preferences. The Planner's own payment does not depend on her selection. In addition, Stakeholders recruited on prolific.com received the payments associated with the option chosen by their assigned Planner in a randomly selected task.

Results.—Table 6 shows our results. Because Planners rarely choose options *C* or *X*,[19] we can focus on the distribution of choices between options *A* and *B*.

Panel A concerns the Swap treatments. The first row shows dramatic effects of changes in rankings when all ϵ_i are zero. Planners choose option *A* 30 percent of the time in profile *S*1 (where it is rank-dominated) versus more than 70 percent of

[18] The payoff vectors of profile *A*2 are $A:(2,2,2), B:(1,3,3), X:(3,1,1)$, where *X* is the added alternative.

[19] The frequencies are 1.23 percent, 1.34 percent, 1.28 percent, and 0.68 percent in profiles $S1, S1', A1'$, and $A2'$, respectively. Supplemental Appendix C presents additional results and robustness checks.

the time in profile $S1'$ (where it is rank-dominant). In both cases, we can reject a choice frequency of 50 percent, the most natural manifestation of indifference, at $p < 0.01$. The second row of panel A shows that this effect persists even when alternative *A* cardinally dominates alternatives *B* and *C*. Notably, the frequency with which Planners choose option *A* in profile *S*1 is only 76 percent, which is consistent with social preferences that value the avoidance of Stakeholders' least preferred outcomes. Furthermore, that frequency rises to more than 90 percent in $S1'$, where ordinal and cardinal considerations both favor *A*. Comparing the rows of panel A shows that Planners also care about cardinal payoffs: They select alternative *A* significantly more often when its payoff distribution dominates that of alternative *B*, even though the ordinal preference profile is the same.

The Addition treatments corroborate these findings (panel B). For both profiles *A*1 and *A*2, the addition of alternative *X* decreases the frequency with which Planners choose the majority winner, *B*, to a statistically significant degree. This violation of Sen's Alpha is consistent with the hypothesis that Planners try to avoid alternatives that Stakeholders rank lowest among the available options.

We conclude that Stakeholders' ordinal rankings over available alternatives affect Planners' choices even when cardinal information is provided.

VI. Robustness across Samples

In this section, we ask whether our main conclusions extend to general population samples. We also test whether the general public in countries with divergent social and political traditions, the United States and Sweden (Alesina and Glaeser 2004), use similar or divergent criteria when aggregating ordinal preferences, and we explore external validity by asking whether ordinal aggregation preferences among the general public are related to attitudes toward political processes.

In these supplemental experiments, each social choice entails the allocation of $20 (in the United States) or Skr170 (in Sweden, worth approximately $20 at the time of the experiment) to 1 of 4 charities: Doctors without Borders, UNICEF, Oxfam, and the International Fund for Animal Welfare. These organizations are well-known in both countries and represent diverse causes that have broad appeal across the political spectrum. We recruited 711 Swedish and 805 US voting-age citizens through Dynata and Lucid to serve as Planners. Each Planner aggregates the preferences of same-country Stakeholders, recruited through Pollfish. Planners only observe group members' ordinal preference rankings. All Planners see the same set of profiles we used for the political domain in our main experiment, as well as either the actual profile for a group of Stakeholders or a randomly generated profile. Planners know that one of the preference profiles they consider corresponds to real Stakeholders with 10 percent probability. Abridged instructions carefully explain the preference displays (see Supplemental Appendix E.3). Subjects who failed comprehension checks were excluded. To make our samples more representative of the respective general populations, we weight observations as follows. For the US sample, we use the 2018 General Social Survey (Smith et al. 2019) to generate weights for 4 population categories defined by (i) gender (male, female), and (ii) political party preference (Democrat, Republican). For the Swedish sample, we use data from Statistics Sweden (2018) to generate weights for 12 categories defined by (i) gender and (ii)

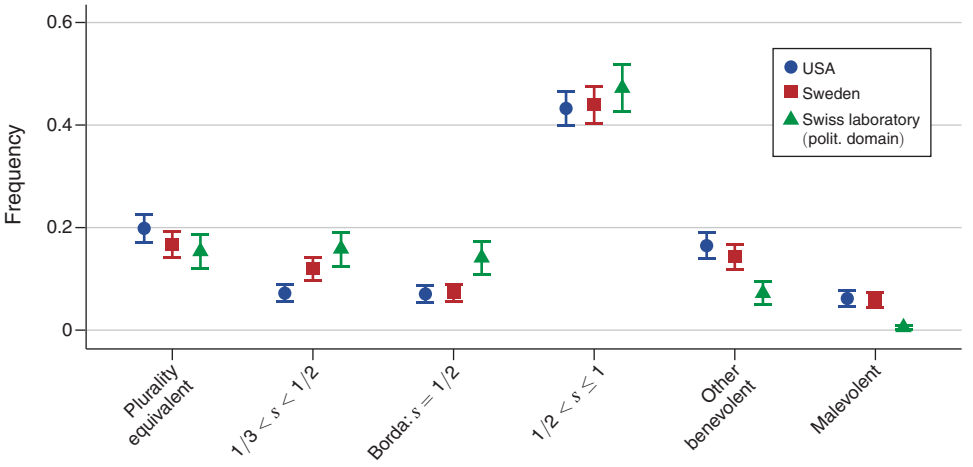

FIGURE 7. CLASSIFICATION OF GENERAL POPULATION SAMPLES TO PRESPECIFIED RULES

Notes: Bayes classifications based on consequential choices for the seven three-option preferences profiles indicated with stars in Table 1. General population observations weighted to make the samples representative with respect to gender and party preference. Whiskers denote 95 percent confidence intervals.

political party preference (Left Party, Social Democratic Party, Green Party, Centre Party, Moderate Party, Sweden Democrats). Supplemental Appendix D.1 provides additional implementation details and demographic summary statistics.

We use the Bayes classifier to assign each subject to the best-fitting rule based on their choices for the seven three-option profiles. Figure 7 shows classification frequencies for the US and Swedish samples. For comparison, we also include the frequencies for the Swiss student sample concerning the political domain, with the caveat that this domain is not identical to the one we use for the general population samples. To facilitate comparisons across samples, we consolidate several categories (strictly concave scoring rules, all benevolent alternatives aside from scoring rules, and all malevolent rules).

Differences in classification frequencies between the Swedish and US general population samples are remarkably small and statistically weak: Holm-Bonferroni corrected t-tests for each category reject the hypothesis of equal frequency at the 5 percent level only for $\frac{1}{3} < s < \frac{1}{2}$. Hence, the fundamental aggregation preferences of US and Swedish citizens are similar, which suggests that they do not contribute to policy divergences. The general population distributions also resemble the distribution for the student sample despite the difference in domains. Widespread conformance with strictly concave scoring rules proves to be a robust phenomenon. Notably, the elevated frequencies of "Other benevolent rules" are partly attributable to the antiplurality runoff rule, which is closely related to the strictly concave scoring rule alternatives. Choices for discerning four-option profiles corroborate these conclusions (see Supplemental Appendix D.2). The fact that conformance with the Borda rule in Figure 7 is lower than in Figure 4 is a mechanical consequence of classifying subjects based on 7 profiles rather than 17: With fewer profiles, it is more difficult to detect the irresoluteness that distinguishes the linear scoring rule from

nearby alternatives. The modestly higher frequencies of malevolent rules suggest that our general population samples may include slightly higher fractions of subjects who did not take the tasks seriously. These results do not substantially change if we run unweighted regressions (see Supplemental Appendix D.3).

Finally, we ask whether our measures of aggregation preference correlate with self-reported attitudes toward political processes. Our survey presents subjects with two hypothetical candidates for political leadership of the nation. It describes Candidate 1 as polarizing and Candidate 2 as a compromise alternative.[20] We ask subjects which candidate better represents the will of the citizens of the nation. Pooling across the US and Swedish samples, we find evidence that Planners who deploy more concave scoring rules tend to have a stronger preference for the compromise candidate over the polarizing candidate, but statistical significance is not robust. Subjects also express their agreement with two statements regarding compromise versus majoritarian approaches in the political process. Statistical evidence on the relation between these responses and Planners' best-fitting scoring parameters is inconclusive. Supplemental Appendix D.4 presents details.

VII. Conclusion

Our objective in this paper has been to develop an empirical understanding of ordinal aggregation preferences. Because there are many settings in which groups make collective choices based primarily on ordinal rankings, this objective is of considerable practical importance. We find that choices for the overwhelming majority of people conform to *scoring rules*. The most common choice patterns correspond to the Borda and near-antiplurality rules. For a sizable majority of subjects, choices conform to strictly concave scoring rules, which indicates a pronounced preference for compromise. Subjects are generally averse to majoritarian outcomes. Thus, many familiar rules, such as plurality and the Condorcet criterion, describe the ordinal aggregation choices of relatively few individuals. The same is true of plurality runoff, supermajority, and unanimity rules. The classification's fit is excellent, and clustering analysis reveals no major omissions from our list of prespecified rules. We find systematic and significant differences in the distributions of rules between the work domain and the political domain, but these differences are of limited magnitude. Because our classifications are highly predictive of choices out of sample, including across domains, we infer that ordinal aggregation also entails stable structural elements. While we find strong indications that subjects aggregate ordinal preferences based in part on inferences about cardinal utility, a separate experiment demonstrates that subjects also care about ordinal rankings intrinsically. It follows that (within-menu) ordinal rankings may impact perceptions of fairness even when cardinal information is available. Supplemental experiments show that the distributions of aggregation preferences in the United States and Sweden, countries with divergent political and social traditions, are remarkably similar, and both

[20] Candidate 1: "Most citizens either love him or hate him. There is hardly anyone who does not have a strong opinion. If candidate 1 were elected, some citizens would be exhilarated, many others would be devastated, and nobody would be indifferent." Candidate 2: "While he is nobody's greatest favorite, most citizens would be ok with candidate 2. If he were elected, nobody would be exhilarated, nobody would be devastated." We randomize the order of the descriptions and the candidate labels.

resemble the distribution for the student sample used in our main experiment. There is suggestive evidence that conformance with more concave scoring rules in experimental decisions correlates with a preference for electing compromise candidates.

Our analysis suggests many potential directions for future work. The extent to which our as-if characterizations generalize beyond the domain we examined (small groups and small menus) is an open question. Even if they are context specific, they potentially illuminate attitudes that may apply more generally, for example, concerning the desire for compromise solutions. We have also focused exclusively on preferences over outcomes. A complementary line of inquiry would investigate the relationships between preferences over rules revealed by choices over outcomes, and preferences over rules implied by approval of axioms (in the spirit of Nielsen and Rehbeck 2022 and others). Likewise, one could also study preferences over rules revealed by choices over rules (as in Engelmann and Grüner 2017; Hoffmann and Renes 2017; Engelmann et al. 2020), but that approach would necessarily implicate the strategic issues from which we have intentionally abstracted. It would also be of interest to investigate whether an awareness of manipulability issues factors into preferences over collective choice mechanisms.

REFERENCES

Alesina, Alberto, and Edward Glaeser. 2004. *Fighting Poverty in the US and Europe: A World of Difference*. Oxford University Press.

Almås, Ingvild, Alexander W. Cappelen, and Bertil Tungodden. 2020. "Cutthroat Capitalism versus Cuddly Socialism: Are Americans More Meritocratic and Efficiency-Seeking than Scandinavians?" *Journal of Political Economy* 128 (5): 1753–88.

Ambuehl, Sandro, and B. Douglas Bernheim. 2026. *Data and Code for: "Social Preferences over Ordinal Outcomes."* American Economic Association; distributed by Inter-university Consortium for Political and Social Research. http://doi.org/10.3886/E226921V1.

Ambuehl, Sandro, B. Douglas Bernheim, and Axel Ockenfels. 2021. "What Motivates Paternalism? An Experimental Study." *American Economic Review* 111 (3): 787–830.

Ambuehl, Sandro, Sebastian Blesse, Philip Dörrenberg, Christoph Feldhaus, and Axel Ockenfels. 2021. "Politicians' Social Welfare Criteria: An Experiment with German Legislators." Unpublished.

Andreoni, James, Deniz Aydin, Blake Barton, B. Douglas Bernheim, and Jeffrey Naecker. 2020. "When Fair Isn't Fair: Understanding Choice Reversals Involving Social Preferences." *Journal of Political Economy* 128 (5): 1673–711.

Apesteguia, Jose, Miguel A. Ballester, and Rosa Ferrer. 2011. "On the Justice of Decision Rules." *Review of Economic Studies* 78 (1): 1–16.

Arrieta, Gonzalo, and Lukas Bolte. 2023. "What You Don't Know May Hurt You: A Revealed Preferences Approach." Unpublished.

Arrow, Kenneth J. 1950. "A Diffculty in the Concept of Social Welfare." *Journal of Political Economy* 58 (4): 328–46.

Arrow, Kenneth J. 1951. *Social Choice and Individual Values*. John Wiley & Sons.

Baker, John. 2008. "Election of the Green Party Cathaoirleach, 2007." *Irish Political Studies* 23 (3): 431–40.

Bolton, Gary E., and Axel Ockenfels. 2006. "Inequality Aversion, Effciency, and Maximin Preferences in Simple Distribution Experiments: Comment." *American Economic Review* 96 (5): 1906–11.

Brandt, Felix, Vincent Conitzer, Ulle Endriss, Jérôme Lang, and Ariel D. Procaccia. 2016. *Handbook of Computational Social Choice*. Cambridge University Press.

Camerer, Colin F., Anna Dreber, Eskil Forsell, Teck-Hua Ho, Jürgen Huber, Magnus Johannesson, Michael Kirchler, et al. 2016. "Evaluating Replicability of Laboratory Experiments in Economics." *Science* 351 (6280): 1433–36.

Costa-Gomes, Miguel A., and Vincent P. Crawford. 2006. "Cognition and Behavior in Two-Person Guessing Games: An Experimental Study." *American Economic Review* 96 (5): 1737–68.

Costa-Gomes, Miguel, Vincent P. Crawford, and Bruno Broseta. 2001. "Cognition and Behavior in Normal-Form Games: An Experimental Study." *Econometrica* 69 (5): 1193–235.

Davies, Todd, and Raja Shah. 2003. "Intuitive Preference Aggregation: Tests of Independence and Consistency." Paper presented at the Annual Meeting of the Public Choice Society and Economic Science Association, Nashville, TN, March 21.

Debian Project. 2016. "Constitution for the Debian Project, v1.7." https://www.debian.org/devel/constitution.

Econometric Society. 2022. "Rules and Procedures." https://www.econometricsociety.org/society/organization-and-governance/rules-and-procedures#45.

Engelmann, Dirk, and Hans Peter Grüner. 2017. "Tailored Bayesian Mechanisms: Experimental Evidence from Two-Stage Voting Games." CESifo Working Paper 6405.

Engelmann, Dirk, Hans Peter Grüner, Timo Hoffmann, and Alex Possajennikov. 2020. "Minority Protection in Voting Mechanisms: Experimental Evidence." CEPR Discussion Paper 14393.

Engelmann, Dirk, and Martin Strobel. 2004. "Inequality Aversion, Effciency, and Maximin Preferences in Simple Distribution Experiments." *American Economic Review* 94 (4): 857–69.

Featherstone, Clayton R. 2019. "Rank Effciency: Investigating a Widespread Ordinal Welfare Criterion." Unpublished.

Fehr, Ernst, and Gary Charness. 2025. "Social Preferences: Fundamental Characteristics and Economic Consequences." *Journal of Economic Literature* 63 (2): 440–514.

Freeman, Rupert, Markus Brill, and Vincent Conitzer. 2014. "On the Axiomatic Characterization of Runoff Voting Rules." *Proceedings of the AAAI Conference on Artificial Intelligence* 28 (1): 675–81.

Gaertner, Wulf. 2009. "Distributive Justice: An Overview of Experimental Evidence." In *Handbook of Rational and Social Choice*, edited by Paul Anand, Prasanta Pattanaik, and Clemens Puppe, 501–23. Oxford University Press.

Gaertner, Wulf, and Erik Schokkaert. 2012. *Empirical Social Choice: Questionnaire-Experimental Studies on Distributive Justice*. Cambridge University Press.

Gibbard, Allan. 1973. "Manipulation of Voting Schemes: A General Result." *Econometrica* 41 (4): 587–601.

Herreiner, Dorothea K., and Clemens Puppe. 2010. "Inequality Aversion and Effciency with Ordinal and Cardinal Social Preferences—An Experimental Study." *Journal of Economic Behavior & Organization* 76 (2): 238–53.

Hoffmann, Timo, and Sander Renes. 2017. "Flip a Coin or Vote: Choosing Group Decision Rules." Unpublished.

Holt, Charles A., and Susan K. Laury. 2002. "Risk Aversion and Incentive Effects." *American Economic Review* 92 (5): 1644–655.

Huang, Zhexue. 1998. "Extensions to the k-Means Algorithm for Clustering Large Data Sets with Categorical Values." *Data Mining and Knowledge* Discovery 2 (3): 283–304.

Jackson, Matthew O., and Leeat Yariv. 2014. "Present Bias and Collective Dynamic Choice in the Lab." *American Economic Review* 104 (12): 4184–204.

Kara, Ayça E. Giritligil, and Murat R. Sertel. 2005. "Does Majoritarian Approval Matter in Selecting a Social Choice Rule? An Exploratory Panel Study." *Social Choice and Welfare* 25 (1): 43–73.

Konow, James. 2003. "Which Is the Fairest One of All? A Positive Analysis of Justice Theories." *Journal of Economic Literature* 41 (4): 1188–239.

Kuziemko, Ilyana, Ryan W. Buell, Taly Reich, and Michael I. Norton. 2014. "'Last-Place Aversion:' Evidence and Redistributive Implications." *Quarterly Journal of Economics* 129 (1): 105–49.

Martinangeli, Andrea F. M., and Lisa Windsteiger. 2021. "Last Word Not Yet Spoken: A Reinvestigation of Last Place Aversion with Aversion to Rank Reversals." *Experimental Economics* 24 (3): 800–820.

Nielsen, Kirby, and John Rehbeck. 2022. "When Choices Are Mistakes." *American Economic Review* 112 (7): 2237–68.

Portmann, Lea, and Nenad Stojanovic. 2019. "Electoral Discrimination against Immigrant-Origin Candidates." *Political Behavior* 41 (1): 105–34.

Rawls, John A. 1971. *Theory of Justice*. Harvard University Press.

Richie, Robert. 2004. "Instant Runoff Voting: What Mexico (and Others) Could Learn." *Election Law Journal* 3 (3): 501–12.

Satterthwaite, Mark Allen. 1975. "Strategy-Proofness and Arrow's Conditions: Existence and Correspondence Theorems for Voting Procedures and Social Welfare Functions." *Journal of Economic Theory* 10 (2): 187–217.

Smith, Tom W., Michael Davern, Jeremy Freese, and Stephen L. Morgan. 2019. *General Social Surveys, 1972–2018*. NORC. https://gss.norc.org/content/dam/gss/get-documentation/pdf/codebook/GSS_Codebook.pdf.

Statistics Sweden. 2018. *The Party Preference Survey in May 2018*. Statistics Sweden. https://www.scb.se/publication/34426.

Thomson, Keela S., and Daniel M. Oppenheimer. 2016. "Investigating an Alternate Form of the Cognitive Reflection Test." *Judgment and Decision Making* 11 (1): 99–113.

Toplak, Jurij. 2006. "The Parliamentary Election in Slovenia, October 2004." *Electoral Studies* 25 (4): 825–31.

Webb, Geoffrey I. 2010. "Naïve Bayes." In *Encyclopedia of Machine Learning*, edited by Claude Sammut and Geoffrey I. Webb, 713–14. Springer.