Sophisticated Approval Voting, Ignorance Priors, and Plurality Heuristics: A Behavioral Social Choice Analysis in a Thurstonian Framework

Michel Regenwetter
University of Illinois at Urbana–Champaign

Moon-Ho R. Ho
Nanyang Technological University

Ilia Tsetlin
INSEAD

This project reconciles historically distinct paradigms at the interface between individual and social choice theory, as well as between rational and behavioral decision theory. The authors combine a utility-maximizing prescriptive rule for *sophisticated approval voting* with the *ignorance prior* heuristic from behavioral decision research and two types of *plurality heuristics* to model approval voting behavior. When using a *sincere plurality heuristic*, voters simplify their decision process by voting for their single favorite candidate. When using a *strategic plurality heuristic*, voters strategically focus their attention on the 2 front-runners and vote for their preferred candidate among these 2. Using a hierarchy of Thurstonian random utility models, the authors implemented these different decision rules and tested them statistically on 7 real world approval voting elections. They cross-validated their key findings via a psychological Internet experiment. Although a substantial number of voters used the plurality heuristic in the real elections, they did so sincerely, not strategically. Moreover, even though Thurstonian models do not force such agreement, the results show, in contrast to common wisdom about social choice rules, that the sincere social orders by Condorcet, Borda, plurality, and approval voting are identical in all 7 elections and in the Internet experiment.

Keywords: approval voting, behavioral social choice, expected utility, decision heuristics, sophisticated voting

Supplemental materials: http://dx.doi.org/10.1037/0033-295X.114.4.994.supp

Behavioral decision research has invested decades of intense study into the interface of normative theory and descriptive data in individual decision making, with particular emphasis on heuristics and biases. Behavioral decision research has profoundly influenced economics, including, most prominently, game theory (e.g., Camerer, 2003) and finance (e.g., Thaler, 1993). The importance of behavioral decision research was most unambiguously acknowledged through D. Kahneman's 2002 shared Nobel prize in economics, in particular for his seminal collaborative work with A. Tversky (e.g., Kahneman & Tversky, 1979; Tversky & Kahneman, 1974, 1981). In contrast, social choice theory has yet to fully incorporate behavioral and descriptive methods. Few researchers have attempted to systematically develop and test formal descriptive models of collective choice behavior or to integrate insights from behavioral decision research into social choice theory. Our project links the individual, social, normative, and behavioral decision sciences (see Figure 1).

The first facet (Quadrant 1 in Figure 1) concerns a prescriptive–normative rule on how to cast an optimal vote in approval voting (AV) elections. In AV, each voter approves of any subset of the candidates. Each candidate in that set (and no other) scores a point from that voter and the winner(s) is (are) the candidate(s) with the highest point total(s), summed over all voters. According to Merrill (1981) and Brams and Fishburn (1983), a *sophisticated*, that is, utility-maximizing, voter ought to cast votes for those and only
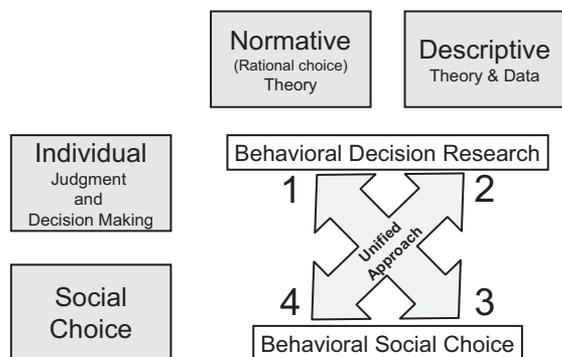
Figure 1. Two major conceptual distinctions and a unified approach to synergize the decision sciences. The four major facets of this project approximately correspond to the four quadrants: Expected utility theory (Quadrant 1), heuristics and biases in individual judgment and decision making (Quadrant 2), heuristics in voting behavior (Quadrant 3), and normative social choice theory (Quadrant 4).

those candidates who, if they are elected, will give the voter at least as much utility as the voter can expect to receive from the election if she or he abstains.[1] This rule is a straightforward implementation of normative expected utility theory (Savage, 1954; von Neumann & Morgenstern, 1947).

To compute his or her expected utility of an election, a voter must assess the candidates' probabilities of being elected. This leads to the second facet (Quadrant 2 in Figure 1), which involves the research program in heuristics and biases of individual behavioral decision theory. A well-known heuristic in probability judgment is the *ignorance prior* (e.g., Fox & Rottenstreich, 2003), in which each event in a partitioning of a probability space has equal probability of occurring. In our context, we interpret an ignorance prior to mean that the voter assumes that some or all candidates have equal (subjective) probability of being elected.

The third facet (Quadrant 3 in Figure 1) also involves heuristics, namely, the frequent concern by opponents of AV that many voters appear to treat AV as plurality by just voting for a single candidate (e.g., Saari, 2001). We refer to this type of voting behavior as the *plurality heuristic*. We model the plurality heuristic in two ways. One is to assume that voters simplify the voting process by voting for their single favorite candidate. The other is to assume that voters strategically concentrate their attention on the two front-runners and effectively ignore the remaining candidates. If many or all voters treat AV as plurality, then AV may inherit some of the normative problems that social choice theorists have identified for plurality rule. This problem would be especially severe if plurality were used in the second, strategic, fashion. One of the largest organizations that used to vote by AV, the Institute of Electrical and Electronics Engineers (IEEE), gave up AV in 2001 because only about 15% of voters voted for more than one candidate.[2] It is thus important to find ways to evaluate the extent to which the plurality heuristic appears to be at play in real elections. If it is, then it is equally important to evaluate whether it is used sincerely or strategically.

This leads to the fourth facet (Quadrant 4 in Figure 1), namely, a range of classical normative concepts in social choice theory (e.g., Mueller, 2003; Riker, 1982). These include the Condorcet

criterion, according to which a majority winner should be elected whenever it exists, and its main normative competitor, the Borda criterion, according to which the legitimate winner is the candidate with the highest Borda score. To expand on the fourth facet, we also evaluate some preference aggregation rules that play a role in the literature on group decision making in psychology. Specifically, our approach allows us to statistically infer, from AV data, the population outcomes under what Hastie and Kameda (2005) called the average, median, and random member rules of group decision making.

The engine behind our theoretical, statistical, and empirical results, linking the above facets, is a family of Thurstonian random utility models in which the utilities of the candidates (or, more generally, of the available choice options) follow a multivariate normal distribution.

The empirical analysis of the ballot data requires walking on a "statistical tightrope": On the one hand, our real world ballot data only have 5 degrees of freedom, thereby dramatically limiting the complexity of any model we can test. On the other hand, five of the seven election data sets have sample sizes in the thousands, thus giving standard statistical techniques extraordinary power to reject models on the basis of even slight deviations between predicted and observed frequencies. We rely on various statistical tools, including nested likelihood-ratio tests, Agresti's dissimilarity index for categorical data, and the Akaike (AIC) and Bayesian (BIC) information criteria for model selection as well as on checking for the conceptual coherence of estimated model parameters. Our substantive conclusions are the product of the converging and redundant multimethod evidence. Furthermore, the design of our experimental cross-validation circumvents the statistical challenges we encounter in ballot data by yielding 46 degrees of freedom (at the cost of the much smaller sample size of 190 participants).

In the next section, we review prior work and introduce our notion of a plurality heuristic. In the Random Utility Models of Sophisticated AV section, we define the general class of models that interconnect our various substantive questions. In The Thurstonian Ignorance Prior Random Utility Model (IP-RUM), we introduce the first fully testable model and test it on ballot data. In Two Thurstonian Random Utility Models of Plurality Voting (SINCE-RUM and STRAT-RUM), we introduce alternative models. In the sincere plurality heuristic random utility model (SINCE-RUM), voters vote for their single most preferred candidate, whereas according to the strategic plurality heuristic random utility model (STRAT-RUM), voters vote for their favorite among the restricted set containing the two front-runners. In the Thurstonian Mixture Random Utility Models section, we move to hybrid models, which combine IP-RUM and the plurality models. We study the sincere plurality heuristic mixture model (SINCE-MIX), the strategic plurality heuristic mixture model (STRAT-MIX), and an overarching mixture model (ALL-MIX) that combines all component models. We apply the mixture models as well as various

---

[1] Among the other early publications that discuss appropriate voting strategies under AV are Fishburn and Brams (1981), Merrill (1979), and Weber (1977).

[2] Personal communication by A. Wyckoff, IEEE Corporate Activities (November 14, 2003).

nested submodels to the empirical AV data. In particular, we estimate the proportion of voters who appear to have used the plurality heuristic in each AV election under each model. In the Behavioral Social Choice Analysis section, we lay out the social choice theoretical implications of these models. We use the parameter values that we estimated earlier from empirical data to derive model-based estimated ranking distributions and then inspect these preference distributions with regard to their social choice and preference aggregation properties. In the Internet Experiment section, we report on a psychological Web experiment, in which we cross-validate our main findings. In the Conclusion section, we discuss key findings and open questions. The supporting online supplemental materials provide Lemmas, proofs, and additional details on the Internet experiment.

Normative Versus Behavioral Social Choice Theory

In the 18th century, the Marquis de Condorcet (see Condorcet, 1785; McLean & Urken, 1995) and his compatriot and intellectual rival, Borda (1770, cited in McLean & Urken, 1995), engaged in a heated debate on how to hold fair elections when there are more than two candidates running for office. Condorcet advocated the *Condorcet criterion,* by which a candidate should be elected as winner if and only if this candidate would have won in a pairwise contest against any other candidate. The Condorcet criterion can also be stated as requiring that a *majority winner* should be elected whenever it exists. To avoid any ambiguity, note that we use the term *majority rule* exclusively in Condorcet's sense and that we formally state its general definition in Formula 10 of the Behavioral Social Choice Analysis section. Borda pointed out that the Condorcet criterion need not generate a winner and advocated an alternative procedure based on rank positions of candidates in individual voters' preference rankings. This method relies on the so-called *Borda score.* Here, the last ranked candidate in a voter's preference ranking scores nothing, the voter's second least favorite scores one point, the third last ranked receives two points, and so on. The individual voters' scores are added up for each candidate across all voters and the candidate with the highest total Borda score is declared the winner. The latter is also called the *Borda criterion.* Condorcet countered with his strong belief that only his rule followed "logically" from principles of probability theory. A frequently made argument against Borda is that the winner under Borda may not be the most preferred candidate of even a single voter. The debate about the relative merits of Condorcet and Borda is ongoing. For translations of original texts, see McLean and Urken (1995); for nontechnical reviews, see Mueller (2003) or Riker (1982); and for mathematical analyses of the distinction between competing social choice criteria, see Saari (1995). There also exists a very small body of experimental work on social choice procedures (e.g., Forsythe, Rietz, Myerson, & Weber, 1996) and extensive related work on group decision making (e.g., Davis, 1973; Feddersen & Pesendorfer, 1998; Guarnaschelli, McKelvey, & Palfrey, 2000; Hastie & Kameda, 2005; Hinsz, Tindale, & Vollrath, 1997; Kameda & Sugimori, 1995; Kameda, Tindale, & Davis, 2003; Kerr, MacCoun, & Kramer, 1996; Kerr, Stasser, & Davis, 1979; Kerr & Tindale, 2004; Laughlin, 1999; Sorkin, Hays, & West, 2001; Sorkin, West, & Robinson, 1998; Stasser, 1999; Stasser & Titus, 2003; Turner & Pratkanis, 1998), as well as a related engineering literature on distributed detection

(e.g., Han & Kim, 2001; Hoballah & Varshney, 1989; Pete, Pattipati, & Kleinman, 1993; Viswanathan & Varshney, 1997; Willett, Swaszek, & Blum, 2000).

The Borda method is a special case of a *scoring rule* or *positional voting method* that takes full or partial rankings of the candidates on a ballot as input and then assigns a numerical score to each candidate on that ballot as a decreasing function of his or her rank. The aggregate score of each candidate is the sum of his or her scores over all ballots. The most commonly used voting procedure, *plurality rule,* is also a special case of a scoring rule. Here, the voter's top ranked candidate receives a single point and the remaining candidates receive nothing. Among the standard criticisms of plurality rule are hypothetical scenarios in which the plurality winner is the worst choice by majority rule (see, e.g., Riker, 1982), that is, a *Condorcet loser* (Condorcet, 1785).

Classical social choice research has primarily provided normative (rational choice) insights into preference aggregation through the discussion of paradoxes, possibility, and impossibility theorems. Condorcet's *paradox of majority cycles* holds whenever for each candidate there exists another candidate such that some majority of voters prefers the second candidate to the first. Arrow's (1951) *impossibility theorem* shows that no deterministic aggregation method can logically combine a certain basic set of desirable features of rational social choice. Recent work grounded in psychology, in the emerging field of *behavioral social choice,* has begun to compare classical notions of social choice against empirical data (Regenwetter, Grofman, Marley, & Tsetlin, 2006, Regenwetter, Kim, Kantor, & Ho, 2007). Our behavioral social choice analysis here tests the conjecture that the classical Condorcet and Arrow paradoxes of social choice theory, although pivotal in their theoretical and normative contributions, are of limited descriptive and empirical applicability.

Prior Research on (Sophisticated) AV, the Ignorance Prior, and the Plurality Heuristic

Although AV is rarely used for political mass elections, it has been a popular voting method in professional and scientific organizations, such as the American Statistical Association (ASA) and Institute for Operations Research and the Management Sciences (INFORMS), the IEEE, the Mathematical Association of America (MAA), the National Academy of Sciences of the United States of America (NAS), the Society for Judgment and Decision Making (SJDM), and the Social Choice and Welfare Society (SCWS). We analyze ballots from two TIMS (the predecessor organization of INFORMS) elections, two MAA elections, one IEEE election, one SJDM election, and one SCWS election. An awesome feature of these data, which sets them apart from laboratory studies, is that some involve over 1,000 decision makers. One involves over 45,000 individuals.

The most attractive feature of AV is that the voters may cast a vote for any subset of the alternatives. There exists a substantial literature on the strengths and weaknesses of AV (e.g., Brams & Fishburn, 1983, 2001; Felsenthal, Maoz, & Rapoport, 1986; Merrill, 1981; Regenwetter et al., 2006; Saari, 2001; Weber, 1995; Wiseman, 2000). Although that literature has analyzed AV empir-

ically, its methodology is best characterized as exploratory in nature.[3]

For a rare example of an experimental comparison of AV against other voting methods in the laboratory, see Forsythe et al. (1996). We attribute the lack of an extensive experimental literature to the massive challenges involved in emulating a realistic election situation in the laboratory. For example, before the advent of Web experimentation, it was practically impossible to run the thousands of participants required to (a) move from group decision making to mass electoral decision making and (b) empirically tap into the potential diversity of opinions and preferences that an election taps into. On the other hand, some real AV ballot data from professional organizations are publicly available. Thus, we place a major emphasis on developing and applying state-of-the-art techniques to analyze real world ballots. We complement that analysis with a psychological Web experiment.

Empirical evidence in the general voting literature suggests that voters may take into account not just the desirability of candidates but also their likelihood of being elected (e.g., Stone & Abramowitz, 1983). These voters are usually referred to as *sophisticated voters,* who may weigh their preferences/utilities by the candidates' probabilities of getting elected. Merrill (1981), Brams and Fishburn (1983), and others have argued from a normative point of view that in AV, a sophisticated voter should first assess each candidate's probability of winning and also his or her own utility of each candidate if that candidate is, in fact, elected. The expected utility of a set of candidates (to a given voter) is the utility of each candidate to the voter, weighted by the candidate's probability of getting elected if the voter abstains, summed over all candidates. Sophisticated voters should vote for those candidates whose utilities exceed the expected utility of the set. When we formulate a statistical model for Merrill's rule in Formula 1, we see that a statistical test in its full generality is currently computationally intractable. We consider simplifying constraints. Some of the constraints are consistent with the hypothesis (Brandstätter, Gigerenzer, & Hertwig, 2006) that decision makers might process probabilities and outcomes of gambles sequentially, without trade-off.

Two other assumptions are frequently made in the theoretical literature on social choice. One is the assumption that a voter behaves as though the candidates' probabilities of getting elected are all equal. In the context of Merrill's prescriptive rule, this assumption could be interpreted as saying that the voters use a heuristic to simplify the decision making process. A second common assumption in social choice theory is that each voter's preferences can be adequately and fully described by a complete linear ordering of the choice alternatives, that is, that utility is ordinal. This could also be interpreted as a heuristic, according to which the voter relies only on ordinal information about the candidates' utilities as a basis for his or her decision. In fact, most of the famous social choice and preference aggregation functions, such as majority rule, the Borda score, and plurality rule, rely on ordinal input only. The combination of these two assumptions leads to the widely known and advertised prescriptive rule that one should vote for approximately half of the candidates under AV, regardless of one's preference (see, e.g., Brams & Fishburn, 1983, pp. 77, 82). We know of hardly any AV data in which many, let alone all, voters approve of approximately half of the candidates. Therefore we do not model or estimate that simple rule here. Although we

have extensively relied on the assumption of ordinal preferences in all our past work on AV (see Regenwetter et al., 2006, for an overview and extensive references), we assume throughout this article that the utilities satisfy an interval scale. On the other hand, we do investigate the common assumption of equal probabilities in this article.

The use of an equal probability assumption is well established, not just as a working hypothesis among social choice theorists. Rather it is a well-documented phenomenon of individual judgment and decision making and has been investigated heavily in the literature on heuristics and biases in behavioral decision research. That literature, and particularly the work on support theory (e.g., Fox & Rottenstreich, 2003; Tversky & Kahneman, 1974; Tversky & Koehler, 1994; Rottenstreich & Tversky, 1997), has provided a wealth of evidence that individual judges fail to follow the laws of probability theory in assessing subjective probabilities. In particular, there is experimental evidence for a heuristic called *ignorance prior,* by which individuals may assign equal probability weight to each event in a partition of probability space. Even sophisticated decision makers, who know that assigning equal probability to each event can be correct only for very specific partitions, are systematically biased toward equal probabilities. Some of our models integrate an ignorance prior by assuming that voters assign equal probability of winning to all candidates.

Although we refer to this behavior as a heuristic, there are rationales for assuming an equal chance of winning, such as the *principle of insufficient reason,* commonly attributed to Bernoulli in the 17th and Laplace in the 19th century (but see, e.g., Baron, 2001; Keynes, 1921; Von Mises, 1981, for critiques of this principle). The principle of insufficient reason may be particularly adequate in situations in which many voters know little or nothing about the candidates. We do not distinguish between an ignorance prior and the principle of insufficient reason here. We use the more general term *ignorance prior.* A voter who uses an ignorance prior across all candidates no longer needs to calculate expected utilities but can, instead, make a decision based solely on the utilities, in effect negating the suggestion made by Stone and Abramowitz (1983) that voters may be sophisticated. Merrill's rule in combination with an ignorance prior is thus to approve of those and only those candidates whose utility exceeds the arithmetic average of all candidates' utilities.

We also consider a completely different heuristic that simplifies Merrill's rule. One of the most outspoken critics of AV, Donald Saari, has repeatedly made the case for the frequent occurrence of what we label the *plurality heuristic,* that is, the phenomenon whereby voters fail to fully use the flexibility offered by AV and, instead, tend to treat AV like a plurality election (see, e.g., the heated debate in Brams, Fishburn, & Merrill, 1988; Saari & Van Newenhizen, 1988, and their follow-up articles). Plurality is generally viewed by social choice theorists as a questionable voting procedure with problematic social choice theoretical properties. For instance, it may yield an election winner who, by majority rule, is the worst candidate. As a consequence, according to Saari, AV

---

[3] Only few studies on AV emphasize the need to construct an empirically testable formal psychological model of subset choice behavior and/or to incorporate statistical considerations in the analysis of AV ballot data (e.g., Laslier, 2003, 2006; Regenwetter et al., 2006).

inherits the normative problems of plurality voting, while not living up to its own normative promise (see, most recently, Saari, 2001). Although AV with only two serious candidates may encourage plurality voting behavior, there also exist theoretical arguments for the notion that "majoritarian" systems without "runoff" and plurality, in turn, lead to a two-party system. This is called *Duverger's law* (Duverger, 1967).

Whereas the ignorance prior offers one way to simplify the expected utility calculus in Merrill's rule, the plurality heuristic may offer an alternative way. More specifically, voters can also avoid this calculus if they restrict their attention to their two favorite candidates. In light of Merrill's recommendation, if there are only two candidates, then voters need to choose only the single candidate with the highest utility. We call the latter behavior the *sincere plurality heuristic,* regardless of the process that leads to it. Although sincere plurality is consistent with Merrill's rule by giving only the voter's two favorite candidates positive subjective probability of being elected, it could also be based on different rationales. For instance, sincere plurality resembles the *take the best* heuristic, which is characterized by some researchers as *fast and frugal* (Gigerenzer & Goldstein, 1999; Hoffrage & Reimer, 2004). Although not implying either of these, sincere plurality is consistent with the idea that voters misperceive AV as plurality and then vote sincerely for their top choice, or that they process outcomes and uncertainty sequentially (Brandstätter et al., 2006).

Finally, we consider yet another scenario that also involves a type of plurality heuristic. This heuristic is completely at odds with Merrill's rule and we call it *strategic plurality voting.* In this scenario there is an objective set of two front-runners who form the consideration set and voters vote for their preferred candidate among those two front-runners. The reason why this rule is completely at odds with Merrill's rule and why we call it *strategic* is because some voters will fail to vote for their favorite candidate, namely, those voters who assign their highest utility to a minor candidate, that is, a candidate who is not a front-runner.

Figure 2 summarizes the different decision rules (in italics on banners) and the models they imply (in boldface). For instance, Merrill's rule implies the general models stated in Formulas 1 and 2 on the next page. Combined with an ignorance prior, Merrill's rule leads to a model labeled IP-RUM. Similarly, the sincere plurality heuristic, which either can be viewed as a simplification of Merrill's rule or could be motivated by other rationales, leads to a model labeled SINCE-RUM. The strategic plurality, which is incompatible with Merrill's rule (as indicated in the figure by a lightning bolt), leads to a model labeled STRAT-RUM.

One of the best-known probabilistic choice model families, the Thurstonian family, was introduced to psychology by Thurstone (1927). The key feature of Thurstone's model is the representation of stimuli as distributions (or random variables) to reflect the fluctuations in the psychological evaluations of objects. Thurstonian models continue to be heavily used (but, recently, also criticized) in psychometrics, psychophysics, and statistics (e.g., Böckenholt, 2002; Busemeyer & Townsend, 1993; Critchlow, Fligner, & Verducci, 1991; Dzhafarov, 2003; Roe, Busemeyer, & Townsend, 2001; Yao & Böckenholt, 1999). In a utility context, a Thurstonian model can be interpreted as follows: Whenever a stimulus is presented to a respondent, it elicits a utility (in Thurstone's terminology, a "discriminal process"), which is normally distributed. Normally distributed random utility models are also
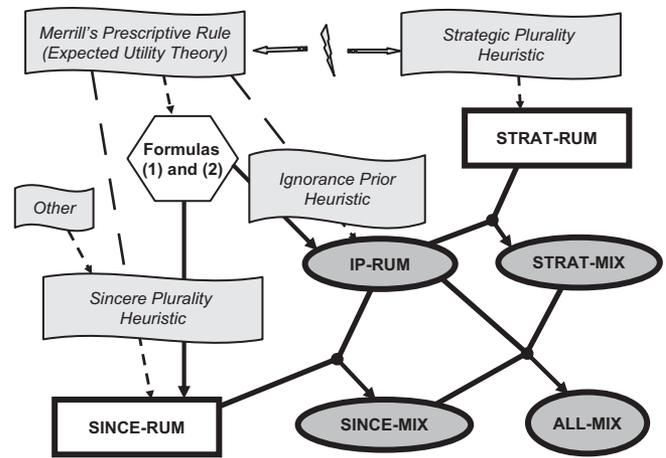

Figure 2. Overview of decision-making mechanisms and their implied models. Decision mechanisms are shown in italics in shaded banners. The lightning bolt indicates that the two processes are incompatible. All models are shown in boldface. Dashed arrows indicate model implementations of decision methods. Solid arrows link related model specifications. The models in white hexagons are too general to be testable at present. The models in white rectangles are too restrictive to warrant testing. The models in shaded ovals are tested against empirical election data. STRAT-RUM = strategic plurality heuristic random utility model; IP-RUM = ignorance prior random utility model; STRAT-MIX = strategic plurality heuristic mixture model; SINCE-MIX = sincere plurality heuristic mixture model; ALL-MIX = overarching mixture model.

commonly known as *Probit* models in econometrics, where the normal distribution is typically used to model measurement error (or exogenous shocks; see, e.g., McFadden, 2001; Train, 1986).

For examples of models with a Thurstonian flavor in a social choice context in psychology, see Hastie and Kameda (2005) and Sorkin et al. (2001). Both of these articles treat group decision making as a group signal detection problem, in which group members must process noisy information about one or more choice options with an objectively known correct or best choice. In both articles, the noise has a (multivariate) normal distribution. We do not use the normal distribution to model noise. We pursue an alternative interpretation of the random utilities in the context of social choice theory. We capitalize on the fact the a random utility model of a population can capture the distribution and heterogeneity of preferences among population members in a concise quantitative fashion. Although the normal distribution is primarily a technical assumption for purposes of model fitting and hypothesis testing, it is also very convenient in computing social welfare functions.

Random Utility Models of Sophisticated AV

We now translate the prescriptive rule of Merrill, Brams, and Fishburn into a probabilistic choice model. We denote the set of all candidates by C, the number of candidates by $|C| = n$, and we use boldfaced symbols for random variables, to differentiate them from numbers. A general random utility model for the original prescriptive rule requires, for each candidate s, random variables $\mathbf{Pr}_s$ and $\mathbf{U}_s$ that denote the voter's random (subjective) probability that candidate s gets elected and the voter's random utility for

candidate *s,* respectively. Thus, let $(\mathbf{Pr}_r)_{r \in C}$ and $(\mathbf{U}_r)_{r \in C}$ be $2n$ jointly distributed real-valued random variables.[4]

Writing P_A for the probability that a randomly chosen voter approves of the nonempty subset $A \subset C$, the *general (distribution-free) random expected utility model of sophisticated AV* states that

$$P_A = P\left[\min_{r \in A}(\mathbf{U}_r) \geq \sum_{s \in C} \mathbf{Pr}_s \mathbf{U}_s > \max_{t \notin A}(\mathbf{U}_t) \right]. \quad (1)$$

As in the original prescriptive rule, each voter first generates their utilities of the candidates and their subjective assessments of the probability of any given candidate getting elected (should the voter abstain) via a realization of the jointly distributed random variables. Then each voter computes the (random) expected utility of the full set of candidates. The model formulation in Formula 1 states this process, taking into account that the utilities and probabilities of being elected are (outcomes of) random variables.

Because this model involves products of random variables, it generates overwhelming methodological challenges. Deriving the distribution of a product of random variables and estimating parameters of the distributions of these random variables by fitting a function of the product to empirical data is, at present, prohibitively expensive computationally. A natural simplification is to consider the case in which each candidate's probability of getting elected is constant across voters (i.e., a number rather than a random variable). Writing Pr_s for the probability that candidate $s \in C$ is elected, this model states that, $\forall A \subseteq C$,

$$P_A = P\left[\min_{r \in A}(\mathbf{U}_r) \geq \sum_{s \in C} Pr_s \mathbf{U}_s > \max_{t \notin A}(\mathbf{U}_t) \right]. \quad (2)$$

General random expected utility theory is not a new idea. The version of random expected utility in which the utilities are random variables and the probabilities are numbers, which we use in Formula 2, has been introduced by Becker, DeGroot, and Marschak (1963) and discussed briefly by Luce and Suppes (1965) for ranking models. Gul and Pesendorfer (2006) study a closely related model theoretically. The case in which the probability terms are random variables and the utilities are numbers is discussed in Busemeyer and Townsend (1993) and labeled *random subjective expected utility theory.* To our knowledge, general random expected utility theory as stated in the central term of Formula 1, or even with constant probability terms (Pr_s) as in Formula 2, has never been systematically investigated on empirical data.

Even the model in Formula 2 still raises major technical challenges. First, unless we have a large number of candidates, the number of degrees of freedom in empirical data is much too small to allow estimation of the number of parameters built into the model. On the other hand, for a large number of candidates, we face the computational complexity and expense of fitting a very elaborate model. For the remainder, we therefore consider more restrictive models.

In Figure 2, the general models are shown in a white hexagon to highlight that we do not empirically test these models. We further comment on the general models in the conclusion.

As mentioned earlier, other random utility models for AV have been proposed (see Regenwetter et al., 2006, for an overview and references). These models use only ordinal information about utility. Here, we rely on the cardinal (interval scale) nature of

utility in the computation of expected utilities.[5] The model in Formula 2 can be thought of as a special case of Regenwetter et al.'s (2006) *random utility threshold model,* in which the decision threshold is a weighted average of the random utilities (either everywhere or almost everywhere).

The Thurstonian Ignorance Prior Random Utility Model (IP-RUM)

We first consider Merrill's rule in combination with an ignorance prior over all candidates, that is, the heuristic that all candidates have an equal chance of winning. In other words, we have the special case of Formula 2 where each of the n candidates is believed to have constant probability $Pr_s = \dfrac{1}{n}$ of being elected. We call this the *ignorance prior random utility model* or *ignorance prior model* (IP-RUM), for short. We later consider alternative models. The nested relationship between IP-RUM and Formula 2 is indicated by an arrow in Figure 2. IP-RUM dramatically simplifies the general models in Formulas 1 and 2 as follows. Writing $P_A{}^{IP}$ for the probability, under IP-RUM, that the set A is chosen, IP-RUM states that for $\varnothing \neq A \subseteq C$,

$$P_A^{IP} = P\left(\min_{r \in A}\mathbf{U}_r \geq \frac{\sum_{s \in C}\mathbf{U}_s}{n} > \max_{t \notin A}\mathbf{U}_t \right). \quad (3)$$

Even in the case of three candidates, IP-RUM is not testable without making additional assumptions, such as distributional assumptions about the random variables involved. This is why we focus on the classical and manageable case in which the utility random variables (**U**) are Thurstonian. (We comment on alternative distributional families in the *Summary and Discussion of Key Findings* section.)

With these three assumptions, Merrill's rule, the ignorance prior, and the Thurstonian, we now discuss maximum likelihood estimation. We focus on the computationally tractable three-candidate case. This is the most common scenario in empirical data. Throughout, we use *r, s,* and *t* as generic (interchangeable, yet distinct) labels for candidates and *a, b,* and *c* as specific (literal) labels for the candidates in a given data set (given in Tables 1 and 8).

Subset Choice Probabilities Under IP-RUM

We assume that there are random utilities $\mathbf{U}_a$, $\mathbf{U}_b$, and $\mathbf{U}_c$ associated with the three candidates, *a, b,* and *c,* with these random variables following a joint multivariate normal distribution. For any relabeling $\{r, s, t\}$ of $\{a, b, c\}$ we can write that according to the ignorance prior model in Formula 3, the single candidate set $\{r\}$ is approved of (i.e., chosen) if and only if one of the following two conditions holds: $\mathbf{U}_r - \mathbf{U}_s > \mathbf{U}_s - \mathbf{U}_t > 0$ or $\mathbf{U}_r - \mathbf{U}_t > \mathbf{U}_t - \mathbf{U}_s > 0$. Similarly, the set of candidates $\{r, s\}$ is chosen if and only

---

[4] We make no distributional assumptions about any of these random variables yet.

[5] The computation of weighted averages requires at least interval scale level input.

Table 1
Approval Voting Data From Real Elections

Subset	TIMS E1	TIMS E2	MAA 1	MAA 2	IEEE	SJDM	SCWS
{a}	66	199	1,257	413	14,365	22	26
{b}	350	248	626	1,798	12,254	24	11
{c}	290	273	1,434	1,019	11,478	9	26
{a,b}	175	84	68	29	2,934	3	2
{b,c}	442	220	141	199	1,811	4	0
{a,c}	105	114	260	52	2,746	4	2

Note. TIMS E1 and TIMS E2 are elections of the Institute of Management Science; MAA 1 and MAA 2 are elections of the Mathematical Association of America; IEEE is an election by the Institute of Electrical and Electronics Engineers; SJDM is an election of the Society of Judgment and Decision Making; and SCWS is an election of the Social Choice and Welfare Society.

if one of the following two conditions holds: $\mathbf{U}_s - \mathbf{U}_t > \mathbf{U}_r - \mathbf{U}_s > 0$ or $\mathbf{U}_r - \mathbf{U}_t > \mathbf{U}_s - \mathbf{U}_r > 0$.

Because the random utilities are unobserved (latent) variables, we can represent them by their relative positions, which can, in turn, be expressed by another set of two variables, namely, the random utility differences $\mathbf{D}_{rs} = \mathbf{U}_r - \mathbf{U}_s$ and $\mathbf{D}_{st} = \mathbf{U}_s - \mathbf{U}_t$. As the joint distribution of $\mathbf{U}_r$, $\mathbf{U}_s$, and $\mathbf{U}_t$ is multivariate normal, the joint distribution of $\mathbf{D}_{rs}$ and $\mathbf{D}_{st}$ is bivariate normal. Notice that the random utility difference between $\mathbf{U}_r$ and $\mathbf{U}_t$ can be expressed as the sum of $\mathbf{D}_{rs}$ and $\mathbf{D}_{st}$, that is, $\mathbf{U}_r - \mathbf{U}_t = \mathbf{D}_{rs} + \mathbf{D}_{st}$. In general, the bivariate normal distribution of $\mathbf{D}_{rs}$ and $\mathbf{D}_{st}$ has five parameters, $\Theta = \{\mu_{rs}; \mu_{st}; \sigma_{rs}; \sigma_{st}; \rho_{rs,st}\}$ (i.e., the means, variances, and correlation of $\mathbf{D}_{rs}$ and $\mathbf{D}_{st}$). We obtain the parameter estimates by likelihood maximization. These five parameters are not uniquely identifiable, because $\mathbf{D}_{rs}$, $\mathbf{D}_{st}$, and $\mathbf{D}_{rt}$ are not directly observable. For any choice of relabeling $\{r, s, t\} = \{a, b, c\}$, we can uniquely identify four parameters. These are functions of the original five parameters, namely $\alpha_{rs} = \frac{\mu_{rs}}{\sigma_{rs}}, \alpha_{st} = \frac{\mu_{st}}{\sigma_{st}}, \gamma_{st,rs} = \frac{\sigma_{st}}{\sigma_{rs}}$, and $\rho_{rs,st}$. Alternatively, without loss of generality (for a fixed labeling r, s, t), we may set $\sigma_{rs} = 1$, in which case we have $\mu_{rs} = \alpha_{rs}$, $\sigma_{st} = \gamma_{st,rs}$, and $\mu_{st} = \alpha_{st}\gamma_{st,rs}$. Lemma 1 in the supplemental online materials provides the exact mathematical formulae to compute the subset choice probabilities under IP-RUM using these parameters.

We can now apply IP-RUM to empirical data. Figure 2 shows IP-RUM in a shaded oval to highlight that we tested this model on empirical data. Table 1 provides an overview of seven real world AV data sets. Under AV, a vote for either all or none of the candidates is a wasted vote (equivalent to abstaining). Because our models in this article do not account for full or empty sets, we leave out the data on such ballots.

Empirical Data Analysis Using IP-RUM

We fit IP-RUM to seven data sets (see Table 1), all of which have previously been studied (most of them multiple times).[6] Data sets TIMS E1 and TIMS E2 were originally reported in Fishburn and Little (1988). Data sets MAA 1 and MAA 2 were first discussed in Brams (1988). Data set IEEE was first discussed in Brams and Nagel (1991). Data set SJDM was previously discussed, for example, in Regenwetter et al. (2006). Data set SCWS was first discussed in Brams and Fishburn (2001). Each election had three candidates running for a single seat.

Table 2 shows a summary of results when fitting IP-RUM, using the candidate labels from Table 1. We use the multinomial log-likelihood as a benchmark against which to compare the model log-likelihood. For TIMS E1, there is no significant difference between the two models in the likelihood-ratio test, $\chi^2(1) = 1.66$, $p = .20$. For illustrative purposes, we also report the parameter estimates and the standard errors for this data set. The ignorance prior model is rejected on all other data sets, with very significant chi-squares.

Some data sets have such a large sample size that it may be hopeless to obtain a good fit by maximum likelihood criteria. This is because for very large sample size, the maximum likelihood method has extremely large power to reject models on the basis of even minute deviations between predicted and observed frequencies. Thus, in order to better evaluate the goodness of fit of the model for very large sample data, we then also rely on a standard alternative fit index, namely Agresti's (1996) dissimilarity index, defined as follows:

$$D = \frac{\sum_i |n_i - \hat{n}_i|}{2n}, \tag{4}$$

where n_i is the observed frequency in cell i, $\hat{n}_i$ is the predicted frequency in cell i, and n is the total sample size.

The likelihood-ratio test examines whether the model perfectly fits the data. In many situations, including ours, this is unrealistic, because we know from the onset that the model is a simplification of a much more complex reality. Agresti's dissimilarity index describes the lack of fit of the model to the data. It is the proportion of sample cases that must move to different cells in order for the model to achieve a perfect fit. The statistic D ranges from 0 to 1. The smaller the value of D, the better the model fits the data. A rule of thumb says that if the value D is smaller than .03, then the model accounts for the data quite closely, though not perfectly (Agresti, 1996). A close look at Table 2 shows that the misfit to the six data sets is confirmed by the dissimilarity indices in all cases.

For the TIMS E1 data set, a close inspection of the parameter estimates reveals parameter values that are all significantly different from zero. The fit of the IP-RUM model to the TIMS E1 data is

---

[6] Unfortunately, INFORMS and IEEE turned down our requests for more ballot data.

Table 2
Summary of Results for Maximum Likelihood Estimation for IP-RUM

TIMS E1

Model log-likelihood $= -17.61$; multinomial log-likelihood $= -16.78$; $\chi^2(1) = 1.66$; $p = .20$; Agresti $D = .01$

θ	Parameter estimate ($\hat{\theta}$)	SE ($\hat{\theta}$)	$\dfrac{\hat{\theta}}{SE\,(\hat{\theta})}$
α_{ab}	$-.70$	.04	-20
α_{bc}	.12	.03	4
$\gamma_{bc,ab}$	1.18	.05	24
$\rho_{ab,bc}$	$-.47$	.04	-13

	Data set					
Statistic	TIMS E2	MAA 1	MAA 2	IEEE	SJDM	SCWS
Model log-likelihood	-65.56	$-1,059$	-988	$-11,209$	-20.72	-33.73
Multinomial log-likelihood	-16.57	-18.26	-17.05	-25.10	-8.75	-6.82
χ^2	97.98	2,082	1,942	22,369	23.94	53.82
df	1	1	1	1	1	1
p	.00	.00	.00	.00	.00	.00
Sample size	1,138	3,786	3,510	45,588	66	67
Agresti D	.14	.30	.24	.33	.24	.33

Note. IP-RUM = ignorance prior random utility model.

consistent with the idea that the voters used an ignorance prior over the full set of candidates and simply voted for those candidates whose utilities were above average. Note that given the large sample size, the fit to TIMS E1 is quite impressive. The strong rejection of IP-RUM on all other data sets suggests that the simple heuristic underlying the IP-RUM model, as a single process shared by the entire electorate, is descriptively inadequate for each of those elections, as confirmed by the Agresti index (unless one attributes the misfit to violations of technical assumptions or to excessive power of the likelihood-ratio test). The next section shows that if we augment IP-RUM by a second process to incorporate the plurality heuristic, then we are able to account for the other data sets as well.

Two Thurstonian Random Utility Models of Plurality Voting (SINCE-RUM and STRAT-RUM)

We now restore the original notion that voters may indeed incorporate the probability of a candidate getting elected differently for different candidates. We take two alternative approaches. Both incorporate Saari's critique of AV and assume that voters treat the election as a plurality election by voting for only one candidate.

In one model, voters are sincere and simply vote for their single top choice (and no other candidates). We call this the *sincere plurality heuristic.* One possible, but by no means necessary, explanation for this behavior would be that voters use Merrill's rule as formulated in Formulas 1 and 2 but assign zero probability of getting elected to all but their two favorite candidates. In this case, the probabilities of getting elected, among the two candidates under consideration, no longer matter and can be ignored in the decision calculus. The sincere plurality heuristic thus offers an elegant, yet dramatically different, alternative to the ignorance prior in simplifying Merrill's rule. This is indicated accordingly in Figure 2.

In the second model, voters are able to distinguish between major candidates and minor candidates. Because we are studying only three-candidate elections, we focus on the case in which these

voters strategically reduce their attention to a consideration set of the two front-runners and vote for their preferred candidate among them. We call this the *strategic plurality heuristic,* because, according to this model, some voters will vote for a candidate who is not their favorite choice (if their favorite candidate is not a front-runner). In particular, some of the voters who follow this procedure will violate Merrill's rule because the latter always includes the voter's favorite candidate in his or her approval set. This is indicated by the lightning bolt in Figure 2. In all relevant analyses, we treat a as a minor candidate in TIMS E1, TIMS E2, and MAA 2, b as a minor candidate in MAA 1 and SCWS, and c as a minor candidate in IEEE and SJDM.[7] If the strategic model were to hold, this would be a particularly devastating blow to AV as a procedure that was designed to promote sincere voting.

The *sincere plurality heuristic random utility model,* SINCE-RUM, states that the voter votes only for his or her single favorite candidate. Formally, writing the probability of choosing the subset A, under this model, as P_A^{SINCE}, SINCE-RUM states that voters only choose single element subsets and $P_{\{r\}}^{SINCE} = P(\mathbf{U}_r = \max_{s \in C} \mathbf{U}_s)$. Lemma 2 in the supplemental online materials specifies all resulting subset choice probabilities. These choice probabilities are special cases of the probabilities under the general models in Formulas 1 and 2.

---

[7] In principle, the set containing the two front-runners could be estimated as a free parameter. However, in the three-candidate scenario, we do not have enough degrees of freedom to estimate that extra parameter. In our data analyses, we assume that the minor candidates are objectively given: These choices of minor candidates are in line with comments that have previously been made about some of these elections in the existing literature using additional information besides the ballot data. In a separate analysis, we also consider every conceivable candidate as potentially being minor. All best fitting models yield conceptually inconsistent parameter estimates.

In contrast to the sincere plurality heuristic random utility model, the *strategic plurality heuristic random utility model,* STRAT-RUM, states that the voter considers only the two strongest candidates, without loss of generality and for notational purpose denoted here by *r* and *s,* and votes for the candidate with the highest utility among these two. Formally, writing the probability of choosing the subset *A* under this model as P_A^{STRAT}, STRAT-RUM states that voters choose only either {*r*} or {*s*}, and for $v \in \{r, s\}$, $P_{\{v\}}^{STRAT} = P(\mathbf{U}_v = \max_{w \in \{r,s\}} \mathbf{U}_w)$. The resulting subset choice probabilities are given in Lemma 3 in the supplemental online materials. These choice probabilities are not consistent with Formulas 1 or 2.

Because we are unaware of AV data in which everyone voted for a single choice, we omit an empirical study of these models. (To highlight that we could test these models on our data sets, but that there is no point in carrying out the test, we indicated them in rectangular white boxes in Figure 2.) We proceed to hybrid models that allow for a portion of the electorate to satisfy IP-RUM and the rest to satisfy either SINCE-RUM or STRAT-RUM.

Thurstonian Mixture Random Utility Models

We now combine SINCE-RUM and STRAT-RUM with IP-RUM. More specifically, we consider mixtures in which some proportion q_{since} (respectively, q_{strat}) of the population vote according to the plurality heuristic SINCE-RUM (respectively, STRAT-RUM), and the remaining voters follow the ignorance prior, IP-RUM. We call these *mixture models* SINCE-MIX (respectively, STRAT-MIX). We do not consider a mixture of just the two plurality models because the model needs to account for ballots that approve of more than a single candidate. These models and their relationships are indicated in Figure 2.

Subset Choice Probabilities Under SINCE-MIX and STRAT-MIX

SINCE-MIX can be written as follows, where $P_A^{SINCE\text{-}MIX}$ denotes the probability of *A* to be chosen, according to the mixture model involving IP-RUM and SINCE-RUM:

$$P_A^{SINCE-MIX} = q_{since}P_A^{SINCE} + (1 - q_{since})P_A^{IP}. \quad (5)$$

Similarly, STRAT-MIX can be written as follows, where $P_A^{STRAT\text{-}MIX}$ denotes the probability of *A* to be chosen, according to the mixture model involving IP-RUM and STRAT-RUM:

$$P_A^{STRAT-MIX} = q_{strat}P_A^{STRAT} + (1 - q_{strat})P_A^{IP}. \quad (6)$$

As before, we use maximum likelihood estimation for the model parameters. Statistically, the difference between IP-RUM and the mixture models is that the latter add an extra parameter, namely, q_{since} or q_{strat}, which is the proportion of voters who use either version of the plurality heuristic. As a consequence, the mixture models have five parameters (i.e., four in the random utility component, plus a mixture proportion parameter). Because there are only five degrees of freedom in the data, there are no degrees of freedom left in the likelihood-ratio test. Nonetheless, our data analyses show that these models do not span the entire sample space.[8] In other words, although the likelihood-ratio test has no degrees of freedom, these models, in fact, do place constraints on

the predicted probabilities of the possible responses. (To save space, we omit the details of this statistical fact.)

Empirical Data Analysis Using SINCE-MIX

Figure 2 shows SINCE-MIX and STRAT-MIX on a gray background to highlight that we test both models against the empirical data. For the remainder of the article, we leave out any results on the SCWS data set because the existence of an empty cell for {*b, c*} creates unsurmountable parameter estimation and identifiability problems for both mixture models on this data set. For TIMS E1, we replicate the results found by IP-RUM: The proportion of voters who use the plurality heuristic is estimated to be zero and the remaining parameters are estimated to be the same values as reported in Table 2 for IP-RUM. In other words, the fit of the model, by allowing a mixture between SINCE-RUM and IP-RUM, does not improve for the TIMS E1 data, and the parameter estimates do not change. For the data sets in which IP-RUM does not fit, we find throughout that the model log-likelihood of the mixture model SINCE-MIX is exactly the same as the multinomial log-likelihood. In other words, SINCE-MIX fits these data sets perfectly, even though a perfect fit is not automatic.

Given that SINCE-MIX has five parameters for five degrees of freedom in the data, the standard next step is to introduce parameter constraints. This creates one or more degrees of freedom in the likelihood-ratio test. Rather than reporting on the details of the fit for the five parameter version, we immediately move on to nested submodels.

Empirical Fit of Nested SINCE-MIX Submodels

We consider nested submodels with equal variances, such as $\sigma_{rs} = \sigma_{st}$, that is, $\gamma_{st,rs} = 1$; independence, such as $\rho_{rs,st} = 0$; and zero means, such as $\alpha_{rs} = 0$. These are the natural and standard special cases of Thurstonian models. Depending on the combination of constraints, we need to estimate only one to four parameters and we can carry out likelihood-ratio tests among nested models. Furthermore, as we discuss in the social choice analysis, a constraint of the type $\alpha_{rs} = 0$ in a Thurstonian model translates into the hypothesis that *r* and *s* are majority tied (for any given choice of *r, s*). Table 3 reports the results under the most constrained nested submodel that fits each given data set, by both maximum likelihood and Agresti standards. We use the same candidate labels as in Table 1. Note that in the context of these nested models, the choice of labels matters. For instance, the hypothesis that $\rho_{ab,bc} = 0$ and the hypothesis that $\rho_{bc,ca} = 0$ differ from each other. The first can best be tested by setting *r* = *a*, *s* = *b*, and *t* = *c*, whereas the second can best be implemented via *r* = *b*, *s* = *c*, and *t* = *a*. We fit each mixture model with every possible choice of relabelings and report as final nested model the most restrictive model that still fits the data well.

For each data set, we report the nested model constraints, the goodness of fit by maximum likelihood standards and by Agresti's *D*, as well as AIC and BIC model selection criteria (see, e.g.,

---

[8] Because we cannot completely rule out that our results come from local optima, we cannot completely rule out that the model could span the entire space. However, because we have taken (various) precautionary measures, we very much doubt that this is the case.

Table 3

Summary of Results for Maximum Likelihood Estimation for Nested Submodels Models of SINCE-MIX

	Data set				
Statistic	TIMS E2	MAA 1	MAA 2	IEEE	SJDM
Nested model constraints	$\gamma_{ab,ca} = 1$	$\gamma_{ca,bc} = 1$	$\gamma_{ca,bc} = 1$ $\rho_{bc,ca} = 0$	$\gamma_{ab,ca} = 1$	$\gamma_{ca,bc} = 1$ $\rho_{bc,ca} = 0$
Submodel log-likelihood	−17.28	−18.98	−18.29	−25.12	−8.79
Multinomial log-likelihood	−16.57	−18.26	−17.05	−25.1	−8.75
χ^2_c	1.42	1.44	2.48	.04	.08
df	1	1	2	1	2
p	.23	.23	.29	.84	.96
Agresti D	.01	.01	.01	<.01	.01
AIC	42.56	45.96	42.58	58.24	23.58
BIC	62.71	70.92	61.07	93.15	30.15

Note. TIMS E2 is an election of the Institute of Management Science; MAA 1 and MAA 2 are elections of the Mathematical Association of America; IEEE is an election by the Institute of Electrical and Electronics Engineers; SJDM is an election of the Society of Judgment and Decision Making. SINCE-MIX = sincere plurality heuristic mixture model; AIC = Akaike information criterion; BIC = Bayesian information criterion.

Akaike, 1981; Schwartz, 1978). The latter are useful for comparing the fit of models that are not nested within each other. For reasons of brevity we omit the final parameter estimates. We mention here only that all parameter estimates are very significant and that we have not been able to fit any more restrictive nested models.

Table 3 shows for TIMS E2 that the most constrained nested model that fits the data (with a p value of .23) is the one with the labeling $r = c$, $s = a$, and $t = b$ (thus, we write the model parameters as α_{ca}, α_{ab}, $\gamma_{ab,ca}$, and $\rho_{ca,ab}$) and with the constraint $\gamma_{ab,ca} = 1$. That is, the most constrained well-fitting model is one in which $\mathbf{D}_{ab}$ and $\mathbf{D}_{ca}$ have equal variances. The rest of Table 3 is organized in the same way and shows that equal variance works for SINCE-MIX in all data sets, combined, in some cases, with independence. All in all, each model has one to two degrees of freedom in the likelihood-ratio test. Despite the very large sample size for most of the data sets, SINCE-MIX performs impressively well.

Empirical Data Analysis Using STRAT-MIX

With the exception of TIMS E1, the picture of results for STRAT-MIX is very different from that of SINCE-MIX. In TIMS E1, we obtain an estimated value of p_{strat} of zero, and all remaining parameter estimates match those we obtained for IP-RUM. In other words, for TIMS E1, STRAT-MIX reverts back to IP-RUM, just like SINCE-MIX does. Thus, there is no point in conjecturing that either sincere or strategic plurality played a role in TIMS E1.

Besides TIMS E1, STRAT-MIX is unable to coherently account for any other data sets. All remaining data sets yield conceptually inconsistent or borderline inconsistent parameter estimates: In all but two data sets (TIMS E2 and MAA 2), the minor candidate has the highest estimated standardized mean utility. For TIMS E2 the point estimates for α_{ab} and for MAA 2 the point estimates for α_{ca} are not significant. This means that the latter two cases are not distinguishable from a situation in which the minor candidate has equal mean utility as one of the major candidates. The inability of STRAT-MIX to fit the data in an internally coherent fashion suggests that this model is not valid; that is, it suggests that voters

who used the plurality heuristic did not do so in a strategic fashion. We omit the details on this model's fit or parameter estimates. Because it is so important to evaluate whether voters voted strategically, we give STRAT-MIX the benefit of the doubt and include it for TIMS E2 and MAA 2 in some of our remaining analyses. We now consider a hybrid of IP-RUM, SINCE-RUM, and STRAT-RUM.

The ALL-MIX Model

In the analysis of STRAT-MIX, we dismissed STRAT-MIX for MAA 1, IEEE, and SJDM, and we strongly questioned its validity for TIMS E2 and MAA 2. To get a better grip on STRAT-MIX and to better distinguish SINCE-MIX and STRAT-MIX, we consider one more model. We label this model ALL-MIX because it is a mixture model that combines all three component models, IP-RUM, SINCE-RUM, and STRAT-RUM. Writing $P_A^{ALL\text{-}MIX}$ for the probability of A to be chosen under ALL-MIX, this model states that

$$P_A^{ALL-MIX} = q_{since}P_A^{SINCE} + q_{strat}P_A^{STRAT} + (1 - q_{since} - q_{strat})P_A^{IP}. \quad (7)$$

This model contains six free parameters for five degrees of freedom in the data. Therefore, to make the parameters identifiable on the ballot data, we focus on nested submodels. We use the constraints that previously fit for SINCE-MIX. The results are provided in Table 4. For TIMS E2, MAA 1, and IEEE, we have zero degrees of freedom in the likelihood-ratio, and we obtain a perfect fit (this is, again, not automatic). For TIMS E2, MAA 1, and IEEE, we have also tried a host of more restrictive nested submodels, without success. For MAA 2 and SJDM, we have one degree of freedom and an excellent fit. Especially for the large sample data, the fit is an indication that the model is performing very well.

The most important information we can extract from the analysis of the ALL-MIX model is the parameter estimates for q_{strat} (see Table 5). In all cases, these are nonsignificant. In other words, the analysis suggests that none or virtually none of the voters used the strategic plurality heuristic. This further validates our earlier

Table 4
Summary of Results for Maximum Likelihood Estimation for ALL-MIX

	Data set				
Statistic	TIMS E2	MAA 1	MAA 2	IEEE	SJDM
Nested model constraints	$\gamma_{ab,ca} = 1$	$\gamma_{ca,bc} = 1$	$\gamma_{ca,bc} = 1$	$\gamma_{ab,ca} = 1$	$\gamma_{ca,bc} = 1$
			$\rho_{bc,ca} = 0$		$\rho_{bc,ca} = 0$
Minor candidate	a	b	a	c	c
Submodel log-likelihood	-16.57	-18.26	-17.09	-25.10	-8.76
Multinomial log-likelihood	-16.57	-18.26	-17.05	-25.10	-8.75
χ^2	0	0	.08	0	.02
df	0	0	1	0	1
p	—	—	.77	—	.91
Agresti D	.00	.00	<.01	.00	<.01
AIC	43.14	46.52	42.18	60.20	25.52
BIC	68.32	77.72	66.84	103.84	34.28

Note. TIMS E2 is an election of the Institute of Management Science; MAA 1 and MAA 2 are elections of the Mathematical Association of America; IEEE is an election by the Institute of Electrical and Electronics Engineers; SJDM is an election of the Society of Judgment and Decision Making. ALL-MIX = overarching mixture model; AIC = Akaike information criterion; BIC = Bayesian information criterion.

conclusion that STRAT-MIX cannot account for the data in a coherent fashion, that is, that voters do not appear to have voted strategically for one of the front-runners in any of these elections.

We also compare SINCE-MIX and ALL-MIX on TIMS E2, MAA 1, MAA 2, IEEE, and SJDM, as well as STRAT-MIX on TIMS E2 and MAA 2, by AIC and BIC model selection criteria. We find that, by AIC, SINCE-MIX is the best model, except for MAA 2, where it is slightly outperformed by ALL-MIX. According to BIC, SINCE-MIX is the best model throughout. By both criteria, STRAT-MIX is unambiguously the worst model.

The Prevalence of Plurality Heuristics

As we show in Table 5, the estimated proportion of voters who use the plurality heuristic varies dramatically between data sets, ranging from none to 86%. In some cases, the estimated number of voters using a plurality heuristic also varies between models. The parameter estimates support Saari's claim that substantial numbers of voters treat AV as if it were plurality. They also show that the

choice of a model can have a substantial impact on the parameter estimates. However, the ALL-MIX analyses also heavily lean toward SINCE-MIX by suggesting in all cases that the proportion of voters following a strategic plurality heuristic is zero or non-significantly different from zero. Overall, the combined evidence suggests that at most a negligible number of voters voted by plurality in a strategic fashion. Given the converging evidence against it, we treat STRAT-MIX as rejected on all ballot data (except its nested submodel IP-RUM fitting TIMS E1) for the rest of the article. The natural next question is whether the high incidence of (sincere) plurality voters has a detrimental effect on the social choice theoretical outcomes in these elections.

Behavioral Social Choice Analysis

Order Statistics Models, Domain Restrictions, and Transitivity of Majority Outcomes

Whereas the previous sections discussed the interface between Thurstonian random utilities and AV ballots, this section lays out

Table 5
Estimated Proportion of Plurality Heuristic Users

		Data set				
Model	TIMS E1	TIMS E2	MAA 1	MAA 2	IEEE	SJDM
SINCE-MIX						
Unconstrained	0	.30**	.76**	.85**	.67**	.64**
Final nested	0	.30**	.77**	.86**	.67**	.64**
STRAT-MIX	0	—	—	—	—	—
ALL-MIX		.27* / .03	.68** / .09	.71** / .15	.67** / .003	.55# / .10

Note. Cells where a given model is rejected are marked with a dash. For ALL-MIX, the first entry is the point estimate q_{since}, and the second entry is the point estimate for q_{strat}. Entries with asterisks are significantly different from zero. TIMS E1 and TIMS E2 are elections of the Institute of Management Science; MAA 1 and MAA 2 are elections of the Mathematical Association of America; IEEE is an election by the Institute of Electrical and Electronics Engineers; SJDM is an election of the Society of Judgment and Decision Making. SINCE-MIX = sincere plurality heuristic mixture model; STRAT-MIX = strategic plurality heuristic mixture model; ALL-MIX = overarching mixture model.
$p = .054$. * $p < .05$. ** $p < .001$.

the interface between Thurstonian models and social choice functions and, thus, the relationship between AV ballots and social choice functions through the intermediary of Thurstonian models. First, in order to investigate the social choice theoretical properties of our models, we translate the multivariate normal distribution into probabilities on linear orders, using a standard *order statistics model* approach.

Any linear order π is a set of pairs of objects that satisfies certain properties (see, e.g., Roberts, 1979, for a standard definition). For instance, the linear order on three candidates, in which a is single best and c is single worst, is the set $\{(a, b), (a, c), (b, c)\}$. According to the *Thurstonian order statistics model* (see, e.g., Critchlow et al., 1991), the probability p_π of a linear order π is given by

$$p_\pi = P\left[\bigcap_{(r,s)\in\pi} (\mathbf{U}_r > \mathbf{U}_s) \right], \quad (8)$$

where $(\mathbf{U}_t)_{t\in C}$ is multivariate normal. Writing p_{rs} for the induced paired comparison probabilities and Φ for the standard normal cumulative distribution function, we have

$$p_{rs} = \sum_{\substack{\pi \\ (r,s)\in\pi}} p_\pi = P(\mathbf{U}_r > \mathbf{U}_s) = P(\mathbf{D}_{rs} > 0) = \Phi(\alpha_{rs}), \quad (9)$$

with $\mathbf{D}_{rs} = \mathbf{U}_r - \mathbf{U}_s$, $\mu_{rs} = \mu_r - \mu_s$, $\sigma_{rs} =$ the standard deviation of $\mathbf{D}_{rs}$, and $\alpha_{rs} = \dfrac{\mu_{rs}}{\sigma_{rs}}$, as before.

We define *majority rule preference* (i.e., the *Condorcet criterion*) as follows:

$$\text{r is (strictly) majority preferred to } s \Leftrightarrow p_{rs} > \frac{1}{2} \quad (10)$$

$$\Leftrightarrow \Phi(\alpha_{rs}) > \frac{1}{2} \Leftrightarrow \alpha_{rs} > 0 \Leftrightarrow \mu_{rs} > 0 \Leftrightarrow \mu_r > \mu_s, \quad (11)$$

where Formula 10 is the general definition of majority preference for probabilistic preference and random utility models of Regenwetter et al. (2006), and Formula 11 is what we obtain after substitution of Formula 9. Thus, the majority rule outcomes follow in an elegant and straightforward way from the expected values of the multivariate normal. In particular, the majority preferences do not depend on the variance–covariance structure of the reparameterized Thurstonian model $(\mathbf{D}_{rs})_{\substack{r \neq s \\ r,s \in C}}$ or of the original model $(\mathbf{U}_t)_{t\in C}$.

Black (1958) and Sen (1966) provided sufficient conditions on a distribution of linear order preferences for majority rule preferences to be transitive, that is, to rule out the Condorcet paradox. The more general condition, Sen's *value restriction* condition, says that (among any triple of candidates) if there exists a candidate and a rank such that no single voter ranks this candidate at that rank, then majority rule outcomes are transitive. We can now summarize two important properties of majority preferences under the Thurstonian model.

Proposition 1: The Thurstonian order statistics model, Formula 8, implies the following:

1. The majority preferences are transitive (at the population distribution level).

2. Except for the degenerate cases where $\sigma_{rs} = 0$, or where $\rho_{rs,st} = \pm 1$, for some distinct labels r, s, and t, all conceivable domain restrictions conditions are violated, including Sen's (1966) value restriction, and thus also Black's (1958) single peakedness condition.

Because the degenerate cases are uninteresting in empirical settings, the second property means that we cannot rely on domain restriction conditions like Sen's value restriction to explain why majority cycles are ruled out by the Thurstonian order statistics model. As Regenwetter et al. (2006) have shown, value restriction is much too strong to hold up as a descriptive condition on real voter preferences, unless the electorate is extremely homogeneous in their preferences. They have provided a different, yet closely related, set of conditions. For the sake of brevity we provide only a partial list of *net value restriction* conditions (Regenwetter et al., 2006), which are sufficient (but not necessary) to rule out the Condorcet paradox. Writing p_{rst} for the probability of the ranking in which r is single best and t is single worst,

Net never worst of r, denoted by *NetNW(r)*, holds iff

$$p_{rst} \geq p_{tsr} \text{ and } p_{rts} \geq p_{str}, \quad (12)$$

Net never best of r, denoted by *NetNB(r)*, holds iff

$$p_{tsr} \geq p_{rst} \text{ and } p_{str} \geq p_{rts}. \quad (13)$$

Intuitively, Formulas 12 and 13 mean that after cancellation of opposite linear orders (by subtracting one side of each inequality from the other), the remaining net distribution satisfies a case of Sen's value restriction. We report on net value restriction in the *Empirical Social Choice Analysis and Model Dependence* section.

In Lemma 4, reported in the supplemental online materials, we consider the relationship between the Thurstonian order statistics model and other well-known models in the psychological literature, such as the weak, strong, and strict utility models. (See Luce & Suppes, 1965, for definitions and many mathematical results on these models.)

Congruence Among Aggregation Procedures for the Thurstonian Order Statistics Model

Consistent with a previous general definition given by Young (1974) for deterministic preferences, we define the *Borda score* of candidate r at the population distribution level as

$$\text{Borda score of } r = \sum_{s \neq r} p_{rs} \quad (14)$$

$$= \sum_{s \neq r} \Phi(\alpha_{rs}), \quad (15)$$

where Formula 14 is the general definition. Formula 15 follows by substituting Formula 9, that is, the information that we have a Thurstonian order statistics model. The *Borda ordering* is the ordering induced on the candidates according to their Borda scores. Note that the literature uses a range of different definitions for the Borda score, all of which are order preserving affine transformations of Formulas 14 and/or the corresponding formula when the quantities p_{rs} are relative frequencies. All these defini-

tions induce the same Borda order for given values of p_{rs}. Like Condorcet, Borda scores and the Borda order are directly computable from the model parameters. However, in contrast to Condorcet, Borda depends on both the expected values and the variance–covariance structure of the Thurstonian model $(\mathbf{D}_{rs})_{\substack{r \notin s \\ r,s \in C}}$, and thus of the original model $(\mathbf{U}_t)_{t \in C}$. We thus have the following proposition.

Proposition 2: In the Thurstonian order statistics model the Condorcet and Borda criteria need not coincide.

Finally, we define the *(sincere) plurality score* of candidate r as the following quantity:

$$\text{Plurality score of } r = P[\mathbf{U}_r = \max(\mathbf{U}_r, \mathbf{U}_s, \mathbf{U}_t)] \quad (16)$$

$$= P^{SINCE}_{\{r\}}, \quad (17)$$

where Formula 16 is the general definition, and Formula 17 is what we obtain when substituting the Thurstonian order statistics model. The plurality order is simply the order of the candidates according to their plurality scores. The following proposition states the relationship between plurality and Condorcet or Borda in Thurstonian models.

Proposition 3: In the Thurstonian order statistics model the plurality outcomes need not match either the Condorcet or the Borda outcomes.

We now consider other prominent rules of group decision making. According to the *average* rule, we socially rank the candidates by their average utility (here, μ_r for each candidate r). According to the *median* rule, we rank the candidates according to their median utility (because the utilities are normally distributed, the median value of the utility is equal to the expected value μ_r, for each candidate r).

Proposition 4: In the Thurstonian order statistics model, the Condorcet order, the social order by the average rule, and the social order by the median rule are identical.

Because we do not need to report separate results for these three rules, we refer only to the Condorcet outcomes in the subsequent empirical analysis. Notice that the automatic agreement among these three rules contrasts the findings of Hastie and Kameda (2005). Although the Condorcet, average, and median criteria will automatically coincide at the population level in a Thurstonian model (i.e., at the level of the multivariate normal distribution), they may well disagree in random samples drawn from such a distribution, like the ones studied in Hastie and Kameda (2005). The latter authors study the sampling distributions of, for example, Condorcet, Borda, plurality, the median, and the mean rules as sample statistics in computer simulated random samples drawn from known distributions. In contrast, we deal with statistical inference and hypothesis testing on the basis of empirical human data.

Empirical Social Choice Analysis and Model Dependence

We use the various Thurstonian models to evaluate the sincere population AV, Condorcet, Borda, plurality, and random member rule outcomes for each data set. Obviously, we are not in a position

to infer the ballot frequencies that voters would have generated in the hypothetical situation that an alternative voting method had been used. Rather, we infer the latent distribution of preferences under each model for each data set and then compute the sincere outcomes under these social choice rules, thus bypassing the counterfactual question of how voters would have voted under an alternative procedure.

The results of these analyses (except for the random member rule) are collected in Table 6. We compute each social choice result using the parameter constraints and best fitting parameter estimates of Tables 2–5 for any given data set and model.

For TIMS E1 (where IP-RUM fits), we find that the AV order computed from the ballots; the AV order computed from the best IP-RUM parameter point estimates; and the Condorcet, Borda, and plurality orders under IP-RUM are all identical, namely, *bca*. This is also one of the three possible Condorcet orders and the Borda order that Regenwetter and Grofman (1998) reported for TIMS E1 using the (ordinal) *size-independent model*. (Note that the Condorcet order is often not uniquely identifiable under the size-independent model.) Regenwetter et al. (2006) and its predecessor articles identify *bca* as the Condorcet order via the (ordinal) *topset model* of AV. Furthermore, Regenwetter and Tsetlin (2004) find that *bca* has the highest posterior probability of being the plurality and Borda order in a Bayesian analysis, again using the size-independent model. We thus have maximal congruence among social choice procedures and maximal robustness across modeling assumptions and statistical methodologies for TIMS E1.

For TIMS E2, all adequate methods and models yield *c* as winner and *a* as loser, and we have *cba* as social order under all aggregation methods. This is also one of three majority orders and the Borda order via the size-independent model identified by Regenwetter and Grofman (1998), and the majority outcome reported by Regenwetter et al. (2006) using the topset model. It furthermore matches the plurality and Borda outcomes with highest (Bayesian) posterior probability in Regenwetter and Tsetlin (2004).

For MAA 1, IEEE, and SJDM, we find complete congruence among all social choice methods and according to both final analysis methods. These outcomes also match the results found under the size-independent model and the topset model.

For MAA 2, the social order *bca* under both models again matches the Condorcet order and Borda order under the size-independent model as well as the Condorcet order under the topset voting model. It also matches the plurality and Borda order with highest posterior probability via the size-independent model in Regenwetter and Tsetlin (2004).

Overall, the perfect agreement between social choice outcomes derived from competing models almost completely alleviates the differences in parameter constraints and parameter estimates that we reported in Tables 2–5. Furthermore, the perfect agreement among social choice rules also contrasts the common wisdom in axiomatic social choice theory, according to which Condorcet, Borda, and plurality are conceptually mutually incompatible. Because Thurstonian models do not force these social choice rules to coincide, the amazing agreement among procedures still warrants an explanation in the future.

In Table 7, we report the linear order probabilities under the order statistics model for each final model with the candidate labels of Table 1. Consistent with the broad range of parameter

Table 6
Social Welfare Orders Under Various Social Rules

	Data set					
Social rule	TIMS E1	TIMS E2	MAA 1	MAA 2	IEEE	SJDM
Data AV	*bca*	*cba*	*cab*	*bca*	*abc*	*bac*
IP-RUM AV	*bca*	—	—	—	—	—
Cond., Mn., Med.	*bca*	—	—	—	—	—
Borda	*bca*	—	—	—	—	—
Plurality	*bca*	—	—	—	—	—
SINCE-MIX AV	*bca*	*cba*	*cab*	*bca*	*abc*	*bac*
Cond., Mn., Med.	*bca*	*cba*	*cab*	*bca*	*abc*	*bac*
Borda	*bca*	*cba*	*cab*	*bca*	*abc*	*bac*
Plurality	*bca*	*cba*	*cab*	*bca*	*abc*	*bac*
STRAT-MIX AV	*bca*	—	—	—	—	—
Cond., Mn., Med.	*bca*	—	—	—	—	—
Borda	*bca*	—	—	—	—	—
Plurality	*bca*	—	—	—	—	—
ALL-MIX AV		*cba*	*cab*	*bca*	*abc*	*bac*
Cond., Mn., Med.		*cba*	*cab*	*bca*	*abc*	*bac*
Borda		*cba*	*cab*	*bca*	*abc*	*bac*
Plurality		*cba*	*cab*	*bca*	*abc*	*bac*

Note. Cells where a given model is rejected are marked with a dash. The first row, Data AV, lists the social welfare order under data-based approval voting rule (AV), and the remaining rows list the social welfare order under model-based AV; Condorcet (Cond.), mean (Mn.), median (Med.); Borda; and sincere plurality rule for the final nested submodels for each data set. TIMS E1 and TIMS E2 are elections of the Institute of Management Science; MAA 1 and MAA 2 are elections of the Mathematical Association of America; IEEE is an election by the Institute of Electrical and Electronics Engineers; SJDM is an election of the Society of Judgment and Decision Making. IP-RUM = ignorance prior random utility model; SINCE-MIX = sincere plurality heuristic mixture model; STRAT-MIX = strategic plurality heuristic mixture model; ALL-MIX = overarching mixture model.

values of Tables 2–5, the final ranking probabilities differ between alternative models. Nonetheless, just like the social orders are hardly affected by the choice of model (as we show in Table 6), we conclude that net value restriction holds in each case, regardless of our final model choice. We are thus able to provide a robust descriptive theoretical account for the absence of majority cycles in each case. With a slight abuse of notation, we write, for example, $NetNM(a, b)$ to denote that $NetNM(a)$ and $NetNM(b)$ hold simultaneously.

Finally, we discuss one more social choice rule that is important for group decision making. According to the *random member rule,* one member of a group or society is picked randomly and this member's preference is declared to be the social order. The linear order probabilities in Table 7 are our estimated (best fitting) population preference distributions for each data set (using the final parameter estimates of the final model or nested submodel in each case). They are therefore also the (estimated) outcome probabilities under the random member rule. Clearly, our findings about the random member rule are dependent on our final model choice. In Table 7, we underline the "correct social order" (as suggested by the analysis in Table 6) for each data set and we highlight in boldface the mode of the random member rule. In most cases the most likely social order under the random member rule matches the correct social ordering. Noteworthy exceptions are MAA 1 under SINCE-MIX and ALL-MIX, and SJDM under SINCE-MIX and ALL-MIX. If we focus only on the modal winner under the random member rule (given in the third column), then the most likely winner is the correct first rank social choice in all cases. Overall, the random member rule is unreliable.

Internet Experiment

We illustrate and reinforce our main findings using data from an unpublished Internet experiment (Baron, 2003) kindly provided by Jonathan Baron. Participants were told, "Suppose you are in a cooperative fixed-interest investment plan with several hundred members. The members vote on investment policies. They vote on the Internet. Almost all members vote." The participants were then asked how they, or the opposing group, would vote on various sets of proposals using different voting methods. Although the experiment collected a wealth of information on multiple methods and under many conditions, we analyze Baron's plurality, antiplurality (vote against one candidate), AV, and disapproval voting (disapprove of any number of candidates) ballots. We focus on the case in which both groups had equal size.

For each vote, the participants were given three proposals. We consider two conditions. In the high condition, which we also denote {$100, $0, $95/$0, $100, $95}, proposal *a* gave the participant's own group members $100, whereas participants of the opposing group would receive nothing. Under proposal *b,* the participant's group was going to receive nothing, and members of the other group would receive $100. Under proposal *c,* participants of each group would receive $95. In the low condition, which we also denote {$100, $0, $55/$0, $100, $55}, the outcomes for proposals *a* and *b* were the same, but proposal *c* yielded only $55 instead of $95.

Table 8 shows the data and results. The table can be read as follows: In the top half, the numbers without parentheses are the ballot counts. We refer to these as *self* because they indicate the

Table 7
Ranking Probabilities, Modal Winner (MW) by Random Member Rule, and Net Value Restriction (NetNB, NetNW)

Election and model	Ranking probabilities						MW	Net value restriction
	abc	*acb*	*bac*	*bca*	*cab*	*cba*		
TIMS E1	*abc*	*acb*	*bac*	*bca*	*cab*	*cba*	*b*	
IP-RUM[a]	.071	.050	.187	**.288**	.120	.283	*b*	*NetNB(a), NetNW(b)*
TIMS E2	*abc*	*acb*	*bac*	*bca*	*cab*	*cba*	*c*	
SINCE-MIX	.125	.138	.120	.224	.144	**.249**	*c*	*NetNB(a), NetNW(c)*
ALL-MIX	.121	.152	.114	.234	.144	**.236**	*c*	*NetNB(a), NetNW(c)*
MAA 1	*abc*	*acb*	*bac*	*bca*	*cab*	*cba*	*c*	
SINCE-MIX	.105	**.271**	.073	.119	.260	.172	*c*	*NetNB(b), NetNW(c)*
ALL-MIX	.104	**.267**	.077	.133	.246	.174	*c*	*NetNB(b), NetNW(c)*
MAA 2	*abc*	*acb*	*bac*	*bca*	*cab*	*cba*	*b*	
SINCE-MIX	.045	.086	.084	**.470**	.100	.215	*b*	*NetNB(a), NetNW(c)*
ALL-MIX	.053	.099	.095	**.452**	.102	.120	*b*	*NetNB(a), NetNW(c)*
IEEE	*abc*	*acb*	*bac*	*bca*	*cab*	*cba*	*a*	
SINCE-MIX	**.193**	.185	.188	.134	.173	.129	*a*	*NetNB(c)*
ALL-MIX	**.192**	.185	.187	.134	.174	.129	*a*	*NetNB(c)*
SJDM	*abc*	*acb*	*bac*	*bca*	*cab*	*cba*	*b*	
SINCE-MIX	.162	**.233**	.168	.256	.089	.092	*b*	*NetNW(a, b), NetNB(c)*
ALL-MIX	.151	**.236**	.157	.258	.097	.101	*b*	*NetNW(a, b), NetNB(c)*

Note. The plausible correct social order and winner are underlined and given in the top row for each data set. The modal ranking (probability) under the random member rule is boldfaced. *NetNB* = net never best; *NetNW* = net never worst. TIMS E1 and TIMS E2 are elections of the Institute of Management Science; MAA1 and MAA2 are elections of the Mathematical Association of America; IEEE is an election of the Institute of Electrical and Electronics Engineers; SJDM is an election of the Society of Judgment and Decision Making. IP-RUM = ignorance prior random utility model; SINCE-MIX = sincere plurality heuristic mixture model; ALL-MIX = overarching mixture model.
[a] Here, SINCE-MIX and STRAT-MIX give the same results as IP-RUM.

ballots that participants cast in their own group. The numbers in parentheses are the counts of respondents' expectations about the other group's modal ballot. We refer to those as *other group*, because they indicate the respondents' expectations about the other group's voting behavior.

For instance, in the high condition, out of the 190 participants, 105 cast a plurality ballot for *a*, and 7 indicated that the modal plurality vote of the other group would be for *a.* Altogether, 176 cast an antiplurality ballot against proposal *b* (which we code as an antiplurality vote in favor of {*a, c*}), and 14 responded that this would be the modal antiplurality vote in the other group. In the low condition, for instance, 78 out of the 190 respondents cast an approval vote for {*a, c*}, and 100 cast a disapproval vote against {*b*} (which we code as a disapproval vote in favor of {*a, c*}). Cells with dashes are impossible ballots. The empirical data matrix has 72 permissible cells. As the 190 voters all responded by each

method and in each condition, the data frequencies in each column add up to 190, thus yielding 56 degrees of freedom in the data matrix.

Although the addition of plurality, antiplurality, and disapproval voting ballots buys us many degrees of freedom in the data, it also requires us to extend our AV models in ways that encompass these three procedures. We only sketch the main ideas (see the online supplemental materials for more details). We model plurality ballots directly using SINCE-RUM and STRAT-RUM. We model antiplurality the same way by replacing the maximum in the model formulae for SINCE-RUM and STRAT-RUM by the minimum. Thus, a sincere antiplurality voter votes against the proposal with lowest utility, and a strategic antiplurality voter votes against the proposal with lowest utility among the two worst proposals. Similarly, we model disapproval voting ballots by taking ALL-MIX and replacing the maxima by minima or, equivalently, by substi-

Table 8
Internet Experiment Plurality (Plur), Anti-plurality (Antiplur), Approval Voting (AV), and Disapproval Voting (Dis-AV) Ballots (Modal Response Frequencies for Other Group in Parentheses), Model Fit, and Parameter Estimates (From Perspective of Self)

Subset implied	High ($100, $0, $95 / $0, $100, $95)				Low ($100, $0, $55 / $0, $100, $55)			
	Plur	Antiplur	AV	Dis-AV	Plur	Antiplur	AV	Dis-AV
{a}	105 (7)	—	41 (6)	21 (5)	143 (9)	—	85 (6)	57 (3)
{b}	3 (105)	—	2 (42)	4 (24)	5 (144)	—	4 (87)	2 (56)
{c}	82 (78)	—	33 (28)	23 (18)	42 (37)	—	18 (15)	18 (14)
{a, b}	—	5 (6)	5 (4)	2 (5)	—	5 (10)	4 (5)	3 (9)
{b, c}	—	9 (170)	0[a] (107)	9 (128)	—	16 (164)	1 (74)	10 (95)
{a, c}	—	176 (14)	109 (3)	131 (10)	—	169 (16)	78 (3)	100 (13)
Social order	acb (bca)	cab (cba)	acb (bca)	acb (bca)	acb (bca)	cab (cba)	acb (bca)	acb (bca)

Model log-likelihood $= -145.61$; multinomial log-likelihood $= -115.57$; $\chi^2 (46) = 60.07$; $p = .08$; Agresti $D = .029$

θ	$\hat{\theta}$	$SE\,(\hat{\theta})$	$\left(\dfrac{\hat{\theta}}{SE\,(\hat{\theta})}\right)$	$\hat{\theta}$	$SE\,(\hat{\theta})$	$\left(\dfrac{\hat{\theta}}{SE\,(\hat{\theta})}\right)$
α_{ab}	1.43	.04	36	Same as for high		
α_{bc}	−1.57	.06	−27	−1.46	.05	−27
$\gamma_{bc,\,ab}$	.86	.04	22	.68	.06	12
$\rho_{ab,\,bc}$	−.84	.05	−18	Same as for high		

Estimated probability of sincere plurality heuristic

θ	$\hat{\theta}$	$SE\,(\hat{\theta})$	$\left(\dfrac{\hat{\theta}}{SE\,(\hat{\theta})}\right)$	$\hat{\theta}$	$SE\,(\hat{\theta})$	$\left(\dfrac{\hat{\theta}}{SE\,(\hat{\theta})}\right)$
AV	.40	.03	15	.57	.03	21
Dis-AV	.25	.02	11	.39	.03	15

Note. Cells with dashes indicate impossible ballots.
[a] Empty cell replaced by 0.5 for the analysis.

tuting antiplurality for plurality in the component processes of ALL-MIX. Furthermore, because the experimental setting highlighted the idea of voting against a proposal, we further augment the ALL-MIX model to allow some proportion of approval voters to use sincere or strategic antiplurality as a voting heuristic (and, by symmetry, to allow some proportion of disapproval voters to use sincere or strategic plurality as a heuristic). Like in the real election voting ballot analyses, the most general model (which would allow all parameters to be different) is underdetermined. We skip the details of our extensive and systematic search through the broad range of possible nested submodels. Instead, we move directly to our final choice of best model. The lower half of Table 8 shows those analysis results. The final (most parsimonious) model uses only 10 parameters. Its likelihood-ratio test has 46 degrees of freedom, and its *p* value of .08 indicates an acceptable fit. Were it not for the one empty cell (which we replace by a count of 0.5), we would undoubtedly obtain a substantially better fit. For instance, a separate analysis of the self and the other group data yielded *p* values of .11 and .25, respectively, even though a chi-square test shows that the self and the other group frequencies (with appropriate relabeling of proposals *a* and *b* for other group) are not significantly different.

According to our final model, all voters voted sincerely and respondents' evaluation of the other group's modal ballot mirrored (in distribution) the ballots they would have cast for themselves if proposals *a* and *b* had been exchanged. Table 8 shows the remaining parameter estimates. These estimates tell a concise story that is

perfectly compatible with the natural hypotheses one might state, for example, about the differences between conditions. We are able to accommodate all four voting procedures, both high and low conditions and self and other group with a single parameter α_{ab} for the standardized mean difference of utilities among proposals *a* and *b*. On the other hand, our final model yields a lower absolute standardized mean difference $|\alpha_{bc}|$ in the low condition than in high condition. This indicates that proposal *c* was more attractive in the high condition than in the low condition. The parameter $\gamma_{bc,ab}$ also behaves in a sensible way, because, as *c* becomes more similar to *b*, the variance of their difference becomes smaller. All in all, we obtain an extremely parsimonious and interpretable Thurstonian model.

Although the Thurstonian parameters tell an interesting and compelling story, the mixture parameters are even more interesting. In our final model, the probabilities of ignorance priors, strategic plurality, and strategic antiplurality are zero. Because the mixture parameters add up to one, we need to report only the probabilities of sincere plurality voters in the table. The probability of sincere antiplurality follows by taking one minus this number. In the high condition, 40% of approval voters voted by sincere plurality; 60% voted by sincere antiplurality; and nobody used strategic plurality, strategic antiplurality, or an ignorance prior. In the low condition, in which proposal *c* is less attractive than in the high condition, the proportion of sincere plurality users went up substantially to 57%, and still nobody used strategic plurality, strategic antiplurality, or an ignorance prior. In disapproval voting, the story is similar. Consistent with

the idea of voting against candidates, less than half of the voters used sincere plurality in the disapproval ballots. We believe that these parameters make perfect sense and reinforce the notion that we can model these elaborate data (56 degrees of freedom) with relatively simple (and sincere) heuristics (only 10 parameters).

Although the purpose of the Internet experiment was primarily to evaluate the decision model experimentally, we also briefly comment on the behavioral social choice component. The social orders computed directly from the ballots match those computed from the model on the basis of predicted frequencies. The plurality, antiplurality, AV, and disapproval voting social orders are given in Table 8. We find perfect agreement among all voting methods, including Condorcet and Borda, on the loser. Only antiplurality switches the winner and the second place. (Notice that the other group social orders match the self social orders throughout when exchanging proposals *a* and *b*.) The Thurstonian order statistics model provides a detailed view of the homogeneity of preferences in the Internet experiment. The most important findings here are that in the high condition, ranking *acb* has an estimated probability of .52, followed by ranking *cab* with an estimated probability of .39. In the low condition, ranking *acb* has an estimated probability of .74, followed by ranking *cab* with an estimated probability of .16. Given the nature of the proposals under consideration, this high homogeneity of preferences and the near perfect agreement among social choice outcomes are sensible.

Conclusion

Summary and Discussion of Key Findings

It is a complex balancing act to formulate statistically testable choice models with identifiable parameters while also formally capturing the most prominent prescriptive rule and integrating insights from behavioral decision theory. Even with today's computing resources it is a challenging task to test such models on empirical data, and it remains computationally feasible only when the number of choice alternatives is limited. We integrate several facets of AV as a collective human decision making procedure.

Figure 2 gives an overview of mechanisms and models for the real election data. The decision rules are shown in italics on shaded banners. The models are in boldface. Merrill (1981) developed a prescriptive rule, based on expected utility, for casting a rational vote under AV. The general distribution-free models in Formulas 1 and 2, which, in principle, offer empirically testable formulations of Merrill's rule, are for the time being computationally intractable (indicated by their white background in Figure 2). We have considered two tractable special cases of the general model, namely, the ones with an ignorance prior and/or sincere plurality, that both dramatically simplify the decision calculus in Merrill's rule. We are not aware of previous work aimed at statistically testing Merrill's rule or the ignorance prior and plurality heuristics on empirical voting data.

The sincere plurality heuristic links Merrill's rule to the argument (e.g., by Saari, 1995, 2001) that voters tend to treat AV as if it were a plurality election. Besides the sincere plurality heuristic, we also studied the strategic plurality heuristic, under which voters vote for the single best among the two front-runners, even if that is not their overall favorite candidate. If the strategic plurality model were to hold, this would be damaging to AV. Strategic plurality is also inconsistent with Merrill's rule. It is thus especially important that we reject STRAT-MIX throughout.

Although it is impossible to judge on the surface why a voter votes for a single candidate, models like the ones in this article allow us to disentangle various processes that can lead to the same observed outcome. With the exception of one data set, our analyses suggest that a substantial portion of the population does indeed use the plurality heuristic. At the same time, the multimodel evidence converges toward pointing away from strategic plurality in favor of sincere plurality. We do not validate the pessimistic conclusions one might draw about the social choice theoretic performance of AV from the heavy use of the plurality heuristic. Our analyses suggest that the AV outcomes are highly consistent with the sincere outcomes under Condorcet, Borda, and plurality in all data sets and across all models, and that the latter aggregation rules are highly congruent. We now review how this article fits in with the overarching goal of synergizing and reconciling the various disparate branches of the decision sciences.

Synergizing the Decision Sciences

The literature and research community in the decision sciences are roughly split along two profound conceptual distinctions (see Figure 1): (a) *individual choice* by a *single individual* versus *social choice* by *groups or societies* and (b) *normative* theories of *rational* decision making versus *descriptive* theories of *actual* decision making.

There are at least three major ways in which the research traditions in individual decision making and social choice contrast each other. We illustrate how both research traditions are interwoven and integrated in our project.

1. A trademark of (individual) behavioral decision research, and one of its greatest achievements, is the systematic comparison between normative proposals of how rational individuals should make decisions against the reality of how individuals actually make decisions. This area, called *behavioral decision research*, has produced an enormous and influential literature (for some highlights and additional references, see Gilovich, Griffin, & Kahneman, 2002; Kahneman & Tversky, 1979, 2000; Luce, 2000; Tversky & Kahneman, 1974, 1981).

In contrast, social choice theory routinely assumes individuals to be rational actors and, by and large, ignores the insights gained in individual behavioral decision research about actual decision makers. In social choice theory, the main thrust of research has been the possibility or impossibility of rational actors to make social choices under various hypothetical distributions of preference in a society. For instance, Arrow's (1951) paradox shows that there exists no voting procedure that simultaneously satisfies certain specified rationality requirements. More generally, much of social choice theory is centered around highlighting the conceptual differences and mutual incompatibilities of competing social choice criteria.

Although the vast literature on individual decision making (e.g., on heuristics and biases) has established repeatedly that individuals fail to satisfy normative benchmarks, the pessimistic conclusions that are frequently drawn from normative social choice theory appear to fall short of actual behavior: Here, as is discussed

in more detail in Regenwetter et al. (2006, 2007), the failures, paradoxes, and incompatibilities pointed out by theorists are rarely observed in practice. Thus, individual choice research finds actors to behave worse than normative theory requires, whereas the sparse empirical research on social choice appears to suggest that electorates may outperform normative expectations. However, for lack of a systematic behavioral social choice research paradigm, we are in no position yet to make general statements about how actual electorates perform compared with normative rational social choice benchmarks. We believe that psychology is the driving force behind (individual) behavioral decision theory and should also lead behavioral social choice. This article is part of our ongoing effort to establish such a research paradigm (Regenwetter et al., 2006, 2007). We go even further by integrating all four quadrants of Figure 1.

2. A second major contrast between individual and social choice research rests in the fact that social choice theory is built on the central primitive that preferences vary between people and require aggregation. In comparison, outside the extensive social psychology literature on group decision making, individual decision research pays relatively little attention to interindividual variability. The most prominent models, such as *expected utility theory* and *rank- and sign-dependent utility theory* (see Kahneman & Tversky, 1979; Luce, 2000; Savage, 1954; von Neumann & Morgenstern, 1947) fail to treat variability as a relevant theoretical primitive (but see Busemeyer & Townsend, 1993; Roe et al., 2001, for models incorporating intraindividual variability). This article shows, on the example paradigms of expected utility maximization, heuristics, and AV, that diverse research traditions can be synergized to formulate random utility versions of deterministic models in a fashion that is tailor-made for a given choice paradigm, and that allows us to bridge the gap between normative concepts, descriptive models, and empirical choice data. We also avoid the common "one heuristic fits all" assumption that is implicit in much work on individual decision making (e.g., Brandstätter et al., 2006, use observed between-subjects majority choices to test the priority heuristic of individual decision making). Rather, we allow for different decision makers to use different heuristics and we estimate the different decision making types through quantitative fit and parameter estimation.

3. Because individual decision research tests theoretical models against empirical data, it routinely and systematically relies on statistical inference and hypothesis testing. In social choice research, statistical approaches hardly ever go beyond the sampling distributions of social choice rules applied to samples drawn from hypothetical preference distributions (see DeMeyer & Plott, 1970; Gehrlein & Fishburn, 1976; Hastie & Kameda, 2005; Maassen & Bezembinder, 2002). Highly stylized artificial theoretical distributions, like the ubiquitous "impartial culture," have little descriptive value as social psychological models of the distribution of preferences in an electorate. This fact is readily acknowledged and yet routinely ignored by social choice theorists. Although our model is somewhat engineered to fit within the limited computational and statistical estimation resources that are currently available, it also grounds the study of social choice phenomena in a statistically testable model, which in turn is built on well-established theoretical, modeling, and empirical principles in several areas of psychology. Furthermore, the key social choice theoretical constructs, namely the Condorcet (and the average and median rules), the

Borda, and plurality criteria, as well as the random member rule, are elegant functions of the model parameters.

Overall, we believe that the application of Thurstonian models reported here reconciles disparate research traditions by unifying a complex and multifaceted research question under the roof of a mathematically concise and statistically testable model family. Specifically, following the tradition established by Luce and Suppes (e.g., Luce & Suppes, 1965), our approach satisfies several key prerequisites of any future unified theory of decision making, in addition to the points we already made:

1. We treat individual and group decision making in a unified fashion. Traditionally, Thurstonian models have been used primarily to model the discriminal process that takes place when a person is exposed to a stimulus. For instance, previous work on social choice in psychology has assumed that group members process cues or signals embedded in normally distributed noise (Hastie & Kameda, 2005; Sorkin et al., 2001). We do not treat political preferences as noisy reflections of a single objective truth, as do signal detection models. Here, we use the Thurstonian model class at an aggregate level.

2. A major focus here is the reconciliation of normative and descriptive approaches to individual and social choice. On the one hand, our models are based on the basic rationality requirement that individuals should make their social choice on the basis of an expected utility calculation, whereas on the other hand, they incorporate decision heuristics that simplify the expected utility calculus. We are among the first to integrate decision heuristics into a formal model of social choice. The ignorance prior heuristic is prominent in individual choice research, whereas the plurality heuristic resembles the *take the best* and the *priority* heuristics of that literature. Notably, the heuristics we study could be labeled *fast and frugal* in that they are simple and that they do not appear to have interfered with the performance (from a normative perspective) of the voting procedure.

3. We compare competing normative benchmarks, such as Condorcet and Borda, against each other via the best fitting parameters. This methodology differs substantially from traditional approaches that prove universal mathematical truths, rely on thought experiments, or use computer simulations to explore such relationships. We find striking agreement among social choice methods (see also, Regenwetter et al., 2007; Regenwetter & Rykhlevskaia, 2007, for similar findings on different data), whereas the theoretical literature highlights the conceptual differences between methods. Whether (sincere) antiplurality, which disagrees here with the other rules on the top two choices, may reliably be a notable exception to this apparent consensus is a separate open question.

Possible Extensions and Open Challenges

We focus on three-candidate elections. Theoretically, the same estimation procedures can be applied to an arbitrary number of candidates. However, for a large number of candidates this requires the numerical computation of high dimensional integrals, which is currently computationally intractable. Alternatively, one may consider other estimation procedures, for example, Gibbs sampling, to handle such difficulties (see, e.g., Yao & Böckenholt, 1999).

The choice of a multivariate normal distribution provides much mathematical convenience and is very common in psychology and

econometrics. The multivariate normal allows us to compute the Condorcet and Borda criteria, as well as the average and median rules, as soon as we know the means and variances of the random utilities. We may investigate the sensitivity of the distributional assumptions in the future. A natural alternative that is frequently used in econometrics is the generalized extreme value family of distributions, otherwise known as the *Logit* model. Investigating such models here, however, would take too much space and is, in our opinion, unlikely to generate new substantive insights. A further natural extension of either model class is to investigate other types of mixture models in which different groups or constituencies are characterized by different parameters (e.g., different means).

The greatest challenge remains in the question of whether the general model in Formula 1 can somehow be tested statistically on empirical data. Rohatgi (1976, p. 141) presented a theoretical result for determining the distribution of the product of two random variables. However, the numerical implementation of this result is very difficult. Recently, Glen, Leemis, and Drew (2004) presented an efficient algorithm for computing the probability density function of the product of two independent random variables in a computer algebra system (e.g., Maple). This new strand of work may eventually open up the possibility for estimating the general AV model (Formula 1) in the future. Once the model is expanded to more than three-choice alternatives, another difficult question arises, namely, how the model for three-choice alternatives is related to the expanded model for, say, four candidates. The simplest would be to assume *regularity*, that is, that the random utilities do not depend on the set of available candidates. However, existing work suggests that regularity is empirically invalid (Roe et al., 2001).

Another major open question for future work is how to incorporate game-theoretic notions into analyses like this one. For instance, there exist game-theoretic rationales for treating AV as plurality. We currently do not see an easy way to statistically identify the large number of free parameters that a Nash-equilibrium model would involve. Similarly, future work needs to incorporate the most prominent game-theory motivated prescriptive rule for AV, according to which one ought to vote for one's favorite among the two front-runners, as well as for all candidates whose utilities exceed the utility of that front-runner (see Brams & Fishburn, 1983; Merrill, 1988; Myerson & Weber, 1993). Myerson and Weber (1993) derived this prescriptive rule by assuming that all voters cast their vote by the same method. However, if this prescriptive rule were descriptively valid for the entire voting population, then no voter would vote for the minor candidate alone and no voter would cast a vote that includes both front-runners. Judging by the data, these constraints are empirically invalid, and the rule by itself cannot account for the data. In particular, Myerson and Weber's (1993) assumptions that all voters behave the same way, as a rationale for this rule, appears to be violated. We leave it to future work to study that prescriptive rule in a mixture model context, where it is combined with other models, such as the ignorance prior model.

Finally, prompted by a referee comment, we highlight one more important reason why psychology should more systematically study social choice procedures by mass publics. Most discussions of strategic or insincere voting behavior in the social choice literature are derived from normative social choice theoretical

(e.g., game-theoretic) considerations. In addition, however, other social or cognitive psychological mechanisms may be at play in collective decision-making environments that could lead to other kinds of insincere or biased ballots. For instance, certain cover stories might affect information processing in such a way that they could induce participants in an experiment to change or distort their ballots in systematic ways. To the extent that such hypotheses can be cast as constraints on the Thurstonian parameters, they may be testable within our modeling framework provided here.

References

Agresti, A. (1996). *An introduction to categorical data analysis.* New York: Wiley.

Akaike, H. (1981). Likelihood of a model and information criteria. *Journal of Econometrics, 16,* 3–14.

Arrow, K. J. (1951). *Social choice and individual values.* New York: Wiley.

Baron, J. (2001). *Thinking and deciding.* Cambridge, England: Cambridge University Press.

Baron, J. (2003). *Voting methods* [Web experiment]. Retrieved July 25, 2007, from http://finzi.psych.upenn.edu/~baron/ex/vote/vote7.htm

Becker, G. M., DeGroot, M. H., & Marschak, J. (1963). Stochastic models of choice behavior. *Behavioral Science, 8,* 41–55.

Black, D. (1958). *The theory of committees and elections.* Cambridge, England: Cambridge University Press.

Böckenholt, U. (2002). A Thurstonian analysis of preference change. *Journal of Mathematical Psychology, 46,* 300–314.

Brams, S. J. (1988). More voters and more votes for MAA officers under approval voting. *Focus: The Newsletter of the Mathematical Society of America, 8,* 1–2.

Brams, S. J., & Fishburn, P. C. (1983). *Approval voting.* Boston: Birkhäuser.

Brams, S. J., & Fishburn, P. C. (2001). A nail-biting election. *Social Choice and Welfare, 18,* 409–414.

Brams, S. J., Fishburn, P. C., & Merrill, S., III. (1988). The responsiveness of approval voting: Comments on Saari and Van Newenhizen. *Public Choice, 59,* 121–131.

Brams, S. J., & Nagel, J. H. (1991). Approval voting in practice. *Public Choice, 71,* 1–17.

Brandstätter, E., Gigerenzer, G., & Hertwig, R. (2006). The priority heuristic: Making choices without trade-offs. *Psychological Review, 113,* 409–432.

Busemeyer, J. R., & Townsend, J. T. (1993). Decision field theory: A dynamic-cognitive approach to decision making in an uncertain environment. *Psychological Review, 100,* 432–459.

Camerer, C. (2003). *Behavioral game theory: Experiments on strategic interaction.* Princeton, NJ: Princeton University Press.

Condorcet, M. (1785). *Essai sur l'application de l'analyse à la probabilité des décisions rendues à la pluralité des voix* [Essay on the application of the probabilistic analysis of majority vote decisions]. Paris: Imprimerie Royale.

Critchlow, D. E., Fligner, M. A., & Verducci, J. S. (1991). Probability models on rankings. *Journal of Mathematical Psychology, 35,* 294–318.

Davis, J. H. (1973). Group decision and social interaction: A theory of social decision schemes. *Psychological Review, 80,* 97–125.

DeMeyer, F., & Plott, C. R. (1970). The probability of a cyclical majority. *Econometrica, 38,* 345–354.

Duverger, C. (1967). *Political parties: Their organization and activity in the modern state.* London: North B and North R Methuen and Company.

Dzhafarov, E. N. (2003). Thurstonian-type representations for "same-different" discriminations: Probabilistic decisions and interdependent images. *Journal of Mathematical Psychology, 47,* 229–243.

Feddersen, T., & Pesendorfer, W. (1998). Convicting the innocent: The

inferiority of unanimous jury verdicts under strategic voting. *American Political Science Review, 92,* 23–45.

Felsenthal, D. S., Maoz, Z., & Rapoport, A. (1986). Comparing voting systems in genuine elections: Approval voting vs selection-plurality. *Social Behavior, 1,* 41–53.

Fishburn, P. C., & Brams, S. J. (1981). Efficacy, power, and equity under approval voting. *Public Choice, 37,* 425–434.

Fishburn, P. C., & Little, J. D. C. (1988). An experiment in approval voting. *Management Science, 34,* 555–568.

Forsythe, R., Rietz, T., Myerson, R., & Weber, R. (1996). An experimental study of voting rules and polls in three-way elections. *International Journal of Game Theory, 25,* 355–383.

Fox, C. R., & Rottenstreich, Y. (2003). Partition priming in judgment under uncertainty. *Psychological Science, 14,* 195–200.

Gehrlein, W. V., & Fishburn, P. C. (1976). The probability of the paradox of voting: A computable solution. *Journal of Economic Theory, 13,* 14–25.

Gigerenzer, G., & Goldstein, D. G. (1999). Betting on one good reason: The take the best heuristic. In G. Gigerenzer, P. M. Todd, & the ABC Research Group (Eds.), *Simple heuristics that make us smart* (pp. 75–95). New York: Oxford University Press.

Gilovich, T., Griffin, D., & Kahneman, D. (Eds.). (2002). *Heuristics and biases: The psychology of intuitive judgment.* New York: Cambridge University Press.

Glen, A. G., Leemis, L. M., & Drew, J. H. (2004). Computing the distribution of the product of two continuous random variables. *Computational Statistics and Data Analysis, 44,* 451–464.

Guarnaschelli, S., McKelvey, R. D., & Palfrey, T. R. (2000). An experimental study of group decision rules. *American Political Science Review, 94,* 407–423.

Gul, F., & Pesendorfer, W. (2006). Random expected utility. *Econometrica, 74,* 121–146.

Han, Y. I., & Kim, T. (2001). Mutual and conditional informations for optimizing Bayes detectors. *IEEE Transactions on Aerospace and Electronic Systems, 37,* 147–157.

Hastie, R., & Kameda, T. (2005). The robust beauty of majority rules in group decisions. *Psychological Review, 112,* 494–508.

Hinsz, V. B., Tindale, R. S., & Vollrath, D. A. (1997). The emerging conceptualization of groups as information processors. *Psychological Bulletin, 121,* 43–64.

Hoballah, I. Y., & Varshney, P. K. (1989). An information theoretic approach to the distributed detection problem. *IEEE Transactions on Information Theory, 35,* 988–994.

Hoffrage, U., & Reimer, T. (2004). Models of bounded rationality: The approach of fast and frugal heuristics. *Management Revue, 15,* 437–459.

Kahneman, D., & Tversky, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica, 47,* 263–291.

Kahneman, D., & Tversky, A. (Eds.). (2000). *Choices, values, and frames.* New York: Cambridge University Press.

Kameda, T., & Sugimori, S. (1995). Procedural influences in two-step group decision making: Power of local majorities in consensus formation. *Journal of Personality and Social Psychology, 69,* 865–876.

Kameda, T., Tindale, R. S., & Davis, J. H. (2003). Cognitions, preferences, and social sharedness: Past, present and future directions in group decision-making. In S. L. Schneider & J. Shanteau (Eds.), *Emerging perspectives on judgment and decision making* (pp. 458–485). Cambridge, England: Cambridge University Press.

Kerr, N. L., MacCoun, R. J., & Kramer, G. P. (1996). Bias in judgment: Comparing individuals and groups. *Psychological Review, 103,* 687–719.

Kerr, N. L., Stasser, G., & Davis, J. H. (1979). Model-testing, model-fitting, and social decision schemes. *Organizational Behavior and Human Decision Processes, 23,* 339–410.

Kerr, N. L., & Tindale, R. S. (2004). Group performance and decision making. *Annual Review of Psychology, 55,* 623–655.

Keynes, J. (1921). *A treatise of probability.* London: MacMillan.

Laslier, J. F. (2003). Analysing a preference and approval profile. *Social Choice and Welfare, 20,* 229–242.

Laslier, J. F. (2006). Spatial approval voting. *Political Analysis, 14,* 160–185.

Laughlin, P. R. (1999). Collective induction: Twelve postulates. *Organizational Behavior and Human Decision Processes, 80,* 50–69.

Luce, R. D. (2000). *Utility of gains and losses: Measurement-theoretical and experimental approaches.* Mahwah, NJ: Erlbaum.

Luce, R. D., & Suppes, P. (1965). Preference, utility and subjective probability. In R. D. Luce, R. R. Bush, & E. Galanter (Eds.), *Handbook of mathematical psychology* (Vol. 3, pp. 249–410). New York: Wiley.

Maassen, H., & Bezembinder, T. (2002). Generating random weak orders and the probability of a Condorcet winner. *Social Choice and Welfare, 19,* 517–532.

McFadden, D. (2001). Economic choices. *American Economic Review, 91,* 351–378.

McLean, I., & Urken, A. B. (1995). *Classics of social choice.* Ann Arbor: Michigan University Press.

Merrill, S., III. (1979). Approval voting: A "best buy" method for multicandidate elections? *Mathematics Magazine, 52,* 98–102.

Merrill, S., III. (1981). Strategic decisions under one-stage multi-candidate voting systems. *Public Choice, 36,* 115–134.

Merrill, S., IIII. (1988). *Making multicandidate elections more democratic.* Princeton, NJ: Princeton University Press.

Mueller, D. C. (2003). *Public choice III.* Cambridge, England: Cambridge University Press.

Myerson, R. B., & Weber, R. J. (1993). A theory of voting equilibria. *American Political Science Review, 87,* 102–114.

Pete, A., Pattipati, K. R., & Kleinman, D. L. (1993). Optimal team and individual decision rules in uncertain dichotomous situations. *Public Choice, 75,* 205–230.

Regenwetter, M., & Grofman, B. (1998). Approval voting, Borda winners and Condorcet winners: Evidence from seven elections. *Management Science, 44,* 520–533.

Regenwetter, M., Grofman, B., Marley, A., & Tsetlin, I. (2006). *Behavioral social choice.* Cambridge, England: Cambridge University Press.

Regenwetter, M., Kim, A., Kantor, A., & Ho, R. (2007). The unexpected empirical consensus among consensus methods. *Psychological Science, 18,* 559–656.

Regenwetter, M., & Rykhlevskaia, E. (2007). A general concept of scoring rules: General definitions, statistical inference, and empirical illustrations. *Social Choice and Welfare, 29,* 211–228.

Regenwetter, M., & Tsetlin, I. (2004). Approval voting and positional voting methods: Inference, relationship, examples. *Social Choice and Welfare, 22,* 539–566.

Riker, W. H. (1982). *Liberalism against populism.* San Francisco: W. H. Freeman.

Roberts, F. S. (1979). *Measurement theory.* London: Addison-Wesley.

Roe, R. M., Busemeyer, J. R., & Townsend, J. T. (2001). Multialternative decision field theory: A dynamic connectionist model of decision making. *Psychological Review, 108,* 370–392.

Rohatgi, V. K. (1976). *An introduction to probability theory and mathematical statistics.* New York: Wiley.

Rottenstreich, Y., & Tversky, A. (1997). Unpacking, repacking, and anchoring: Advances in support theory. *Psychological Review, 104,* 406–415.

Saari, D. G. (1995). *Basic geometry of voting.* Berlin, Germany: Springer Verlag.

Saari, D. G. (2001). Analyzing a nail-biting election. *Social Choice and Welfare, 18,* 415–430.

Saari, D. G., & Van Newenhizen, J. (1988). The problem of indeterminacy

in approval, multiple, and truncated voting systems. *Public Choice, 59,* 101–120.

Savage, L. J. (1954). *The foundations of statistics.* New York: John Wiley & Sons.

Schwartz, G. (1978). Estimating the dimension of a model. *The Annals of Statistics, 5,* 461–464.

Sen, A. K. (1966). A possibility theorem on majority decisions. *Econometrica, 34,* 491–499.

Sorkin, R. D., Hays, C. J., & West, R. (2001). Signal-detection analysis of group decision making. *Psychological Review, 108,* 183–203.

Sorkin, R. D., West, R., & Robinson, D. E. (1998). Group performance depends on the majority rule. *Psychological Science, 9,* 456–463.

Stasser, G. (1999). A primer of social decision scheme theory: Models of group influence, competitive model-testing, and prospective modeling. *Organizational Behavior and Human Decision Processes, 80,* 3–20.

Stasser, G., & Titus, W. (2003). Hidden profiles: A brief history. *Psychological Inquiry, 14,* 3–4.

Stone, W. J., & Abramowitz, A. I. (1983). Winning may not be everything, but it's more than we thought: Presidential party activists in 1980. *American Political Science Review, 77,* 945–956.

Thaler, R. H. (Ed.). (1993). *Advances in behavioral finance.* New York: Russell Sage Foundation.

Thurstone, L. L. (1927). A law of comparative judgement. *Psychological Review, 34,* 273–286.

Train, K. (1986). *Qualitative choice analysis.* Cambridge, MA: MIT Press.

Turner, M., & Pratkanis, A. R. (1998). Twenty-five years of groupthink theory and research: Lessons from the evaluation of a theory. *Organizational Behavior and Human Decision Processes, 73,* 105–115.

Tversky, A., & Kahneman, D. (1974, September 27). Judgment under uncertainty: Heuristics and biases. *Science, 185,* 1124–1131.

Tversky, A., & Kahneman, D. (1981, January 30). The framing of decisions and the psychology of choices. *Science, 211,* 453–458.

Tversky, A., & Koehler, D. (1994). Support theory: A nonextensional representation of subjective probability. *Psychological Review, 101,* 547–567.

Viswanathan, R., & Varshney, P. K. (1997). Distributed detection with multiple sensors: Part I—Fundamentals. *Proceedings of the IEEE, 85,* 54–63.

Von Mises, R. (1981). *Probability, statistics, and truth.* New York: Dover.

von Neumann, J., & Morgenstern, O. (1947). *Theory of games and economic behavior.* Princeton, NJ: Princeton University Press.

Weber, R. J. (1977). *Comparison of voting systems* (Cowles Foundation Discussion Paper No. 498). New Haven, CT: Yale University.

Weber, R. J. (1995). Approval voting. *Journal of Economic Perspectives, 9,* 39–49.

Willett, P., Swaszek, P., & Blum, R. S. (2000). The good, bad and ugly: Distributed detection of a known signal in dependent Gaussian noise. *IEEE Transactions on Signal Processing, 48,* 3266–3279.

Wiseman, J. (2000). Approval voting in subset elections. *Economic Theory, 15,* 477–483.

Yao, G., & Böckenholt, U. (1999). Bayesian estimation of Thurstonian ranking models based on the Gibbs sampler. *British Journal of Mathematical and Statistical Psychology, 52,* 79–92.

Young, H. P. (1974). An axiomatization of Borda's rule. *Journal of Economic Theory, 9,* 43–52.