Soc Choice Welfare (2003) 20: 283–306

Social Choice and Welfare

On the probabilities of correct or incorrect majority preference relations

Ilia Tsetlin[1], **Michel Regenwetter**[2]

[1] Decision Sciences, Fuqua School of Business, Duke University, Box 90120, Durham, NC 27708-0120, USA (e-mail: imt@duke.edu)
[2] Quantitative Division, Department of Psychology, University of Illinois at Urbana-Champaign, 603, East Daniel Street, Champaign, IL 61820, USA (e-mail: regenwet@uiuc.edu)

Received: 13 November 2000/Accepted: 19 March 2002

Abstract. While majority cycles may pose a threat to democratic decision making, actual decisions based inadvertently upon an incorrect majority preference relation may be far more expensive to society. We study majority rule both in a statistical sampling and a Bayesian inference framework. Based on any given paired comparison probabilities or ranking probabilities in a population (i.e., culture) of reference, we derive upper and lower bounds on the probability of a correct or incorrect majority social welfare relation in a random sample (with replacement). We also present upper and lower bounds on the probabilities of majority preference relations in the population given a sample, using Bayesian updating. These bounds permit to map quite precisely the entire picture of possible majority preference relations as well as their probabilities. We illustrate our results using survey data.

1 Introduction

Arrow's *Impossibility Theorem* [1] proves that no universal preference aggregation method satisfies a certain set of simple axioms of rationality. Thus, every preference aggregation rule requires paying a price in terms of violating one or more of Arrow's axioms. The price to pay for using majority rule is to violate the axiom of transitivity: majority rule preferences need not be transi-

This collaborative work was carried out while Regenwetter was a faculty member at the Fuqua School of Business, Duke University. We thank Fuqua for sponsoring our collaboration and the National Science Foundation for grant SBR-97-30076 to Michel Regenwetter. We are indebted to the editor and the referees, as well as to Jim Adams, Bob Clemen, Bernie Grofman, Bob Nau, Saša Pekeč, Jim Smith and Bob Winkler for helpful comments and suggestions.

tive. In order to be able to weigh this cost by its probability, researchers have investigated the theoretical probability and the empirical frequency of majority cycles.

In the theoretical arena, one influential strand of literature focuses on the probability of majority rule cycles in samples drawn from a uniform distribution and investigates the role of sample size and number of candidates [6, 11, 20]. Empirical research suggests that majority cycles are extremely rare and, moreover, that when they occur they tend to occur among the minor (lower ranked) candidates [9, 26, 31].

In practice, an occurrence of a cycle means that a social choice may be prevented or delayed. To us there exist potentially more dramatic costs to society than a social choice being prevented or delayed (postponed). The potential for high cost to society arises when there is a possibility that decisions are based on incorrect assessments of the social choice function and when this incorrect assessment remains unnoticed and uncorrected. For example, during the 2000 U.S. presidential election, the world was transfixed for several weeks with the question whether the correct presidential candidate would receive a majority of the electoral college vote. While we do not mean to take sides, it is fair to say that in the minds of some voters the election outcome was both incorrect and extremely costly for society. We do not know of any empirical work on the frequency of incorrect election outcomes. In an effort to make such studies feasible we provide tools to evaluate the likelihood of correct and incorrect assessments of majority rule preferences.

A second important theoretical strand of literature, more similar to our approach, is the work on the Condorcet efficiency of various social choice methods [12, 16, 18, 17, 21] and the work on the impact of voter turnout on the outcome of majority rule [14, 15]. The Condorcet efficiency of a voting method is the likelihood that this method will elect a majority winner (in a sample from some culture), given a majority winner exists (in that sample). As a consequence, the Condorcet efficiency of a social choice method is "in effect a conditional probability that two sample statistics coincide, given certain side conditions" [29]. For instance, Gehrlein and Valognes [17] study the Condorcet efficiency of weighted scoring rules. Our approach is conceptually different in that we study the probability that the majority preference in a random sample and the majority preference in the underlying population do or do not coincide.[1] In comparison to Condorcet efficiency, the work on voter abstention and turnout [14, 15] is more similar to our work here, in that it studies how likely the population Condorcet winner of all eligible voters coincides with the Condorcet winner of the nonabstaining voters, under certain assumptions about voter turnout and for infinite sample sizes.

Whenever we draw a random sample from a population, we refer to the population majority preference relation as the *correct majority preference relation*. (This relation may or may not be known.) Whenever the majority preference relation in a sample does not match the majority preference relation

[1] See [29] for a detailed discussion of these conceptual differences.

in the population that the sample was drawn from, then we call the sample majority preference relation an *incorrect majority preference relation*. We provide the conceptual framework and the mathematical tools to evaluate the probability of a correct or incorrect majority preference relation in a random sample from a given population. We also discuss correct and incorrect inferences about population majority preferences extrapolated from sample data. This allows the collective decision maker to assess the probability of basing a decision on a correct or incorrect majority preference relation regardless of whether Arrow's axiom of transitivity is violated. Even if a cycle effectively prevents a social choice, it is nevertheless highly informative and therefore valuable as a snapshot of society's distribution of preferences, if correct.

We rely on a general concept of majority rule introduced by Regenwetter, Marley and Grofman [33] which, they show, is applicable to virtually any kind of preference, choice or rating data. They have shown that there is no loss of generality by projecting all binary relations down to their asymmetric part because this transformation does not affect the majority rule outcomes.[2] We therefore assume that the preferences of a population are accurately captured by a probability distribution over asymmetric binary relations. Obvious special cases are when individual preferences are linear orders or strict weak orders. Using a trivial property of probabilities of joint events, we derive upper and lower bounds on the probability of any possible majority ordering. It turns out that this innocuous approach yields surprisingly strong results: Whenever the population majorities are not tied, then the upper and lower bounds on the sample majority preference probabilities are very tight, even for relatively small sample sizes. The analogue holds for population majority preferences, given sample data. The commonly studied case of drawing random samples from the impartial culture can be thought of as an example of the sampling problem where the population has a uniform distribution over preference relations.[3]

Our tools put virtually no constraint on the nature or the distribution of individual preferences, in contrast to the methods in the recent literature, which assume, e.g., the maximal culture condition over linear orders [16], the impartial weak order condition [17], the existence of cardinal utilities [35], or the impartial culture over weak orders [38]. Following Van Deemen's [38] recommendation in this journal, our illustrations are based on realistic distributions. We construct such probability distributions from observed relative frequencies in real world survey data.[4] Our method yields general theoreti-

---

[2] Also, Regenwetter et al.'s [33] equivalence between such a probability distribution and tallies on utility functions or random utilities applies, i.e., everything we say in terms of preference relations can be restated as equivalent statements about utility functions.

[3] We use the term *population* in the statistical sense. We think of a *culture* as a special case of a population.

[4] As one referee correctly pointed out, the study of preference distributions and the comparison of voting procedures under a particular culture may, however, be justified for other reasons than realism.

cal results because it does not depend on constraining theoretical assumptions about voter preferences, and it generates strong empirical results because some illustrations allow us to pin down all possible scenarios with surprising accuracy.

The paper is organized as follows: We first assume that the distribution of preferences in the population is given and we derive the sampling distribution of pairwise majority preferences in samples randomly drawn (with replacement) from that population (Sect. 2). We provide upper and lower bounds on the probability of any possible sample majority preference order, as well as upper and lower bounds on the probability that a random sample will correctly reflect the majority preference order of the population that we are sampling from (Sect. 3). Then we move from sampling distributions to statistical inference by assuming that the given information is only a random sample, and that the goal is to make an inference about the majority preferences of the population which the sample is drawn from (Sect. 4). Along the way, we illustrate our results with examples from survey research. Section 5 provides conclusions and open questions, whereas the Appendix collects some mathematical derivations for the inference problem.

2 Pairwise majority in a sample drawn from a given (population) probability distribution over a set of binary relations

Throughout the paper, we refer to a fixed finite set $\mathscr{C}$ of choice alternatives, candidates, parties or consumption bundles. A binary relation on $\mathscr{C}$ is a set $B \subseteq \mathscr{C} \times \mathscr{C}$ of ordered pairs of elements in $\mathscr{C}$. If B is a binary relation describing a person's state of preference, and $(a, b) \in B$, then this person prefers a to b. In other words, this person finds that a is *better than* b, and we also write that as aBb. For any given binary relation B, we write B^{-1} for the inverse relation, that is, $B^{-1} = \{(b, a) \text{ such that } (a, b) \in B\}$. A binary relation B is asymmetric if $B \cap B^{-1} = \varnothing$. In words, a binary preference is asymmetric if there is no pair a, b of choice alternatives such that a person in this state of preference (strictly) prefers a to b and at the same time (strictly) prefers b to a.

We assume throughout that the possible states of preference form a collection $\mathscr{B}$ of asymmetric binary relations over $\mathscr{C}$. We will be interested in situations where the distribution of individual preferences over a population can be conceptualized as a probability distribution over $\mathscr{B}$. There are two important circumstances in which such a probabilistic framework may be called for:

1. Relative frequencies are a special case of a probability measure, and thus the proportion of people out of a given group that have a particular state of preference can be quantified with a probability measure on the set of possible states of preference.
2. It is often the case that people experience uncertainty as to their own preferences, and thus any preference statements they provide may be generated through an internal sampling process. Famous examples of proba-

bilistic models of individual decision making are Luce's *choice axiom* [22, 23], Tversky's *elimination by aspects model* [37], or various types of random utility models [2, 7, 10, 25, 36]. The same models are often alternatively interpreted as capturing a researcher's uncertainty about (possibly deterministic) preferences of a group of respondents.

All these scenarios have in common that we may be able to describe the overall distribution of preferences in a population by a single probability distribution over preference relations. The need for probabilistic models to describe voting behavior has been discussed for some time in the literature [5]. Recent work has also linked probabilistic choice models from psychology and econometrics to social choice and the analysis of voting or survey data [8, 30–32]. In this section we study properties of random samples drawn from the given population preference distribution. Let us denote the probability of each relation $B \in \mathscr{B}$ in the population by p_B. For example, suppose for a moment that $B = abc$, a complete linear ranking where a is single best, and c is single worst. Then $p_B = p_{abc}$, and in a sampling framework, this denotes the probability that a person drawn at random from the population has the linear preference order abc. If $\mathscr{R} \subseteq \mathscr{B}$, then we write $p_{\mathscr{R}}$ for the sum of all p_B, where $B \in \mathscr{R}$. To make the notation more readable, we write p_{aBb} for $p_{\mathscr{R}}$ when $\mathscr{R} = \{B \in \mathscr{B}$ such that $aBb\}$. So, for example, in the case of complete linear orders of 3 candidates (without indifference) $p_{aBb} = p_{abc} + p_{acb} + p_{cab}$. We use the analogous notation for the observed frequency of each binary relation in a sample, e.g. $N_{aBb} = N_{abc} + N_{acb} + N_{cab}$ (for complete rankings over 3 candidates without indifference). We use boldface letters to denote random variables and regular font to denote numbers.

We will assume throughout the paper that sampling is done independently (and, in particular, with replacement). Therefore, the sampling distribution of preference relations has a multinomial distribution. Other sampling schemes that have been studied before include *anonymous preference profiles* [3, 4, 13], that lead to distributions other than the multinomial. For instance, Berg and collaborators [3, 4] suggest a sampling scheme, where, for large assemblies, the sampling distribution can be taken to be a Dirichlet distribution.[5]

We first derive the probability that a is preferred to b by a majority in a sample, with replacement, of size N (i.e. $\boldsymbol{N}_{aBb} > \boldsymbol{N}_{bBa}$) given the probabilities of all binary relations (i.e., p_B, $B \in \mathscr{B}$) in the population. If there are no indifferences in the population, i.e. $B \cup B^{-1} = \mathscr{C} \times \mathscr{C}$, for all $B \in \mathscr{B}$, then $\boldsymbol{N}_{aBb} + \boldsymbol{N}_{bBa} = N$, and $\boldsymbol{N}_{aBb}$ has a binomial distribution with number of trials N and probability of success p_{aBb}. Let us denote by $f_{Bin}(X, N, p)$ the binomial distribution and by $F_{Bin}(X, N, p)$ the cumulative binomial distribution of the binomial random variable $\boldsymbol{X}$ with number of trials N and probability of success p, i.e.

[5] Although the Dirichlet distribution appears in the Bayesian inference part of the paper (for other purposes), we assume throughout that the sampling distribution is multinomial.

$$F_{Bin}(X, N, p) = \sum_{i=0}^{X} \frac{N!}{i!(N-i)!} p^i (1-p)^{N-i} = \sum_{i=0}^{X} f_{Bin}(i, N, p).$$

Also, to deal with even and odd N, we will need to use the *floor* of $\frac{N}{2}$ defined standardly as

$$\left\lfloor \frac{N}{2} \right\rfloor = \begin{cases} \frac{N}{2} & \text{if } N \text{ is even,} \\ \frac{N-1}{2} & \text{if } N \text{ is odd.} \end{cases} \tag{1}$$

We write aEb if and only if $(a, b) \notin B$, $(b, a) \notin B$. In other words, aEb denotes the situation where a and b are *equivalent* (hence the notation). A person in this state of preference has no preference either way, i.e., s/he is indifferent between the two options. If indifference (i.e., aEb) is allowed, then $p_{aBb} + p_{aEb} + p_{bBa} = 1$ and $\boldsymbol{N}_{aBb}|\boldsymbol{N}_{aEb}$ (i.e., $\boldsymbol{N}_{aBb}$ given $\boldsymbol{N}_{aEb}$) has a binomial distribution with number of trials $N - \boldsymbol{N}_{aEb}$ and probability of success $\frac{p_{aBb}}{p_{aBb}+p_{bBa}}$.

Proposition 1. *The probability that a is preferred to b by a strict majority in a sample of size N is given by:*

$$P(\boldsymbol{N}_{aBb} > \boldsymbol{N}_{bBa}) = \sum_{N_{aEb}=0}^{N} \left(f_{Bin}(N_{aEb}, N, p_{aEb}) \times \left(1 - F_{Bin}\left(\left\lfloor \frac{N - N_{aEb}}{2} \right\rfloor, N - N_{aEb}, \frac{p_{aBb}}{p_{aBb}+p_{bBa}} \right) \right) \right). \tag{2}$$

Furthermore, for preferences without indifference, i.e. $p_{aEb} = 0$, $a \neq b$, *the probability that a is preferred to b by a strict majority in a sample of size N is given by:*

$$P(\boldsymbol{N}_{aBb} > \boldsymbol{N}_{bBa}) = 1 - F_{Bin}\left(\left\lfloor \frac{N}{2} \right\rfloor, N, p_{aBb} \right). \tag{3}$$

Let us emphasize that Proposition 1 is valid for any number of candidates and regardless of the exact nature of the preference relations in $\mathscr{B}$.

We now illustrate (2) using attitudinal ratings in the 1996 American National Election Study (1996 ANES) [34]. Consistent with Van Deemen's [38] recommendation to use realistic preference distributions we treat the relative frequencies in the survey as a population distribution. The reason we use survey data to serve as a population is so that we use realistic probabilities rather than arbitrary ones (like the impartial culture).[6] More precisely, our population distribution is a probability distribution over strict weak orders that matches the relative frequencies of the 1996 ANES feeling thermometer ratings for c (Clinton), d (Dole), and p (Perot). Figure 1a displays these

[6] We do not make any claims about how similar other realistic distributions are or should be to the ones we use in the current illustration.

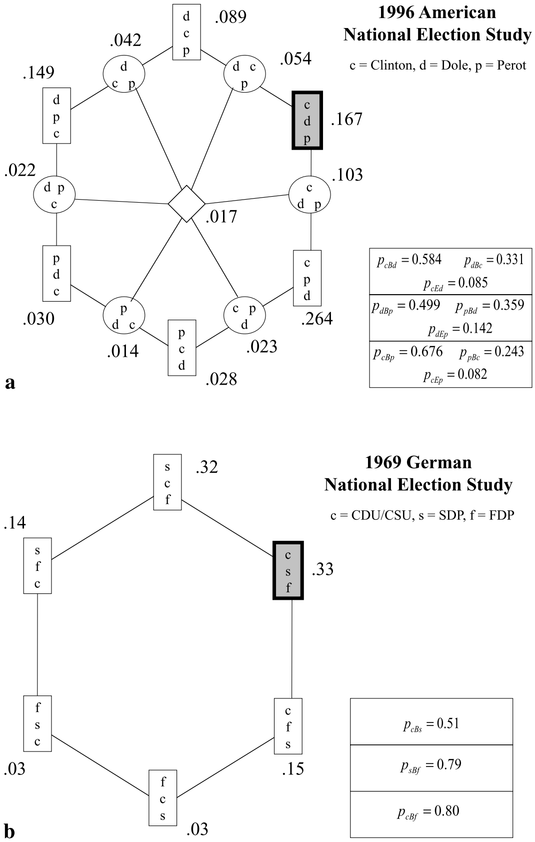


Fig. 1a,b. Probabilities of all possible strict weak orders in the population (i.e., the survey) for ANES 1996 (**a**). Probabilities of rankings in the population for GNES 1969 (**b**)

strict weak order probabilities and Table 1 shows the corresponding majority preference probabilities in samples of size $n = 50, 101, 500$, respectively. The information in Table 1 allows us, for instance, to conclude that even a random sample of 50 voters is virtually certain to correctly identify the majority preference of Clinton over Perot, whereas 50 randomly sampled voters have an 84% chance of correctly ranking Dole over Perot (by a majority in the sample).

3 Upper and lower bounds on the probabilities of majority preference relations based on probabilities of pairwise majorities

We now show how to evaluate the probability that any given complete, asymmetric binary relation $\succ \subseteq \mathscr{C} \times \mathscr{C}$ is the majority preference relation, if

Table 1. Probabilities that one candidate is preferred to another by a majority in the sample as a function of sample size for the ANES 1996 survey (Fig. 1a). For any two candidates a, b we use the following abbreviations in the table: $P(a \succ b) = P(\boldsymbol{N}_{aBb} > \boldsymbol{N}_{bBa})$, $P(a \sim b) = P(\boldsymbol{N}_{aBb} = \boldsymbol{N}_{bBa})$

N	$P(c \succ d)$	$P(c \sim d)$	$P(d \succ c)$	$P(p \succ d)$	$P(p \sim d)$
50	0.967	0.010	0.024	0.125	0.034
101	0.996	0.001	0.003	0.057	0.013
500	1.000	<1E −9	<1E −9	<1E −3	<1E −4

N	$P(d \succ p)$	$P(p \succ c)$	$P(p \sim c)$	$P(c \succ p)$
50	0.841	<1E −3	<1E −3	1.000
101	0.930	<1E −6	<1E −6	1.000
500	1.000	0.000	<1E −26	1.000

we know for each pair a, b the probability that a is majority preferred to b. In this section we illustrate the approach in a sampling framework. (Sect. 4 discusses the same method in an inference framework.)

Suppose we know the following quantities of the population distribution: For each pair (a, b) we know p_{aBb} and p_{aEb}, and therefore Eq. (2) provides the probability $P(\boldsymbol{N}_{aBb} > \boldsymbol{N}_{bBa})$ that a is majority preferred to b in a sample (of size N) drawn from the population. When it is important to distinguish between majority preferences in a population and majority preferences in a sample, we write $\succ_p$ for the majority preference relation in the population and $\succ_s$ for the majority preference relation in the sample when these preferences are known or given, and we write $\succ_{\boldsymbol{p}}$ and $\succ_{\boldsymbol{s}}$ for the corresponding majority preference relations when they are uncertain (i.e., the result of a random process). Throughout this section we focus on the case where $\succ_p$ is a complete asymmetric relation.

We first point out (without proof) a basic property of the probability of a joint event.

Proposition 2. *For any collection $A_1, A_2, \ldots, A_K$ of events, the probability of the joint event $A = A_1 \cap A_2 \cap \cdots \cap A_K$ has the following upper and lower bounds:*

$$\max\left(0, 1 - \sum_{i=1}^{K}(1 - P(A_i))\right) \leq P(A) \leq \min_i P(A_i). \tag{4}$$

We will exploit the simple fact that if the probabilities of some events A_i are close to 0 or 1, then these bounds may become very close to each other, and therefore they may become very good approximations for the probability of the joint event.

We assume for now that $a \succ_p b$, i.e., a is majority preferred to b in the population. Writing $\delta_{aBb} = p_{aBb} - p_{bBa}$, the assumption $a \succ_p b$ is equivalent to the assumption that $\delta_{aBb} > 0$. We also refer to δ_{aBb} as the *pairwise margin* (for pairwise comparisons) of a over b in the population. Let us denote by

Table 2. Sample size, as a function of the pairwise margin ($\delta = \delta_{aBb}$) and probability of indifference ($p = p_{aEb}$), sufficient for $Err(N, \delta, p)$ to be less than 0.1%

δ \ p	0.0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
0.9	7	3								
0.8	9	8	5							
0.7	15	13	10	6						
0.6	21	20	17	13	8					
0.5	33	31	27	22	17	10				
0.4	55	51	45	38	31	24	14			
0.3	101	94	83	72	60	49	36	20		
0.2	235	215	191	166	142	117	92	65	31	
0.1	951	865	769	673	577	480	384	286	186	66

$Err(N, \delta_{aBb}, p_{aEb}) = 1 - P(\mathbf{N}_{aBb} > \mathbf{N}_{bBa})$ the probability that b is erroneously majority preferred to a in a sample of size N contrary to the majority preference relation between alternatives a and b in the population.

Proposition 3. *Equation (2) implies that*

$$Err(N, \delta_{aBb}, p_{aEb}) = \sum_{i=0}^{N} f_{Bin}(i, N, p_{aEb}) F_{Bin}\left(\left\lfloor \frac{N-i}{2} \right\rfloor, N-i, \frac{\delta_{aBb} + 1 - p_{aEb}}{2(1 - p_{aEb})}\right). \tag{5}$$

If no preferences with indifference are allowed in the population (i.e., if $p_{aEb} = 0$), then according to (3) we have

$$Err(N, \delta_{aBb}, 0) = Err(N, \delta_{aBb}) = F_{Bin}\left(\left\lfloor \frac{N}{2} \right\rfloor, N, \frac{1 + \delta_{aBb}}{2}\right). \tag{6}$$

Obviously $Err(N, \delta)$ decreases in δ and N (separately for odd and even N) and approaches 0 for $N \to \infty$ (by definition δ is positive). The risk $Err(N, \delta_{aBb}, p_{aEb})$ of an error regarding a, b decreases with respect to p_{aEb}. This means, in particular, that $Err(N, \delta_{aBb}, p_{aEb}) < Err(N, \delta_{aBb}, 0) = Err(N, \delta_{aBb})$.

In Table 2 we tabulate the sufficient sample size for $Err(N, \delta_{aBb}, p_{aBb})$ to be less than .1% for various values of δ_{aBb} and p_{aBb}.

For instance, it can be seen from Table 2 that if indifference is ruled out, and $p_{aBb} = 0.55 = 1 - 0.45 = 1 - p_{bBa}$, and thus $\delta_{aBb} = 0.10$, that is, a has a margin of ten percentage points over b in the population, then we need to sample at least 951 observations (with replacement) in order to be at least 99.9% sure that the sample will not accidentally contain a (reversed) majority preference of b over a.

On the other hand, if 30% of the population are indifferent between a and b and the remaining 70% are split into 40% of the population preferring a over b and 30% preferring b over a, i.e., a still has a margin of 10 percentage points over b, then a sample size of 673 observations will be sufficient to be 99.9% certain that a random sample will yield the correct majority preference (for the pair a, b), namely that a is majority preferred to b.

For large enough N the multinomial distribution is approximated by a multivariate normal distribution, and one may derive a normal approximation of (5). Define *adjusted pairwise margins* δ^*_{aBb} as

$$\delta^*_{aBb} = \frac{\delta_{aBb}}{\sqrt{1 - p_{aEb} - \delta^2_{aBb}}} \tag{7}$$

and let $F_N(x)$ be the standard normal cumulative distribution function. Then, for large enough N, $Err(N, \delta_{aBb}, p_{aEb})$ is approximated by the quantity $Err_N(N, \delta^*_{aBb})$, given by

$$Err_N(N, \delta^*_{aBb}) = F_N(-\sqrt{N}\delta^*_{aBb}) = F_N\left(-\sqrt{N}\frac{\delta_{aBb}}{\sqrt{1 - p_{aEb} - \delta^2_{aBb}}}\right). \tag{8}$$

We will now apply the method of bounds of Proposition 2 to the probabilities of possible majority relations in a sample. Let us denote by M the total number of pairwise margins: for m many alternatives $M = m(m-1)/2$. It is straightforward to see that as N goes to infinity, the probability of obtaining the correct majority relation in the sample goes to 1, provided the pairwise margins are nonzero.[7]

Let us index all adjusted pairwise margins (given by (7)) in such a way that $\delta^*_1 \leq \delta^*_2 \cdots \leq \delta^*_M$. Consider the case when there is a unique minimal adjusted pairwise margin, i.e., $\delta^*_1 < \delta^*_2$. Suppose that $a \succ_p b$ and that $\delta^*_1 = \delta^*_{aBb}$. Let $\succ^* = (\succ_p \cup \{(b,a)\}) \backslash \{(a,b)\}$, obtained by replacing $a \succ_p b$ by $b \succ^* a$. Using (8) we then obtain the following theorem.

Theorem 1. *If a unique minimal adjusted pairwise margin exists, then the following holds, with $\succ^*$ defined as above.*

$$\lim_{N \to \infty} P(\succ_s = \succ^* \mid \succ_s \neq \succ_p) = 1.$$

In words, for large enough sample size the only "possible" incorrect (sample) majority relation is the majority relation $\succ^* = (\succ_p \cup \{(b,a)\}) \backslash \{(a,b)\}$ in which for all pairs, except the pair (a, b) with the smallest adjusted pairwise margin, the majority preference relation is the same as in the population.

Proof. $P(\succ_s = \succ^*)$ has the following lower bound:

$$Err_N(N, \delta^*_1) - (M-1)Err_N(N, \delta^*_2) \leq P(\succ_s = \succ^*).$$

[7] From this perspective, the impartial culture assumption has a rather odd feature in that it is *impossible* for a random sample of any odd size to match the majority preference relation of the population.

The sum of the probabilities of all incorrect majority relations $P(\succ_s \neq \succ_p)$ has the following upper bound:

$$Err_N(N, \delta_1^*) + (M-1)Err_N(N, \delta_2^*) \geq P(\succ_s \neq \succ_p).$$

Therefore

$$P(\succ_s = \succ^* \mid \succ_s \neq \succ_p) = \frac{P(\succ_s = \succ^*)}{P(\succ_s \neq \succ_p)} \geq \frac{Err_N(N, \delta_1^*) - (M-1)Err_N(N, \delta_2^*)}{Err_N(N, \delta_1^*) + (M-1)Err_N(N, \delta_2^*)}.$$

Because

$$\lim_{N\to\infty} \frac{Err_N(N, \delta_2^*)}{Err_N(N, \delta_1^*)} = 0,$$

therefore

$$\lim_{N\to\infty} P(\succ_s = \succ^* \mid \succ_s \neq \succ_p) = 1. \quad \square$$

The relevance of this theorem can be seen in Fig. 2 and Table 3 and will be discussed below.

A case of particular interest involves the property

$$\left.\begin{array}{l}\delta_{aBb}^* > 0\\ \delta_{bBc}^* > 0\end{array}\right\} \Rightarrow \delta_{aBc}^* > \min(\delta_{aBb}^*, \delta_{bBc}^*) \quad (\forall a, b, c \in \mathscr{C}).$$

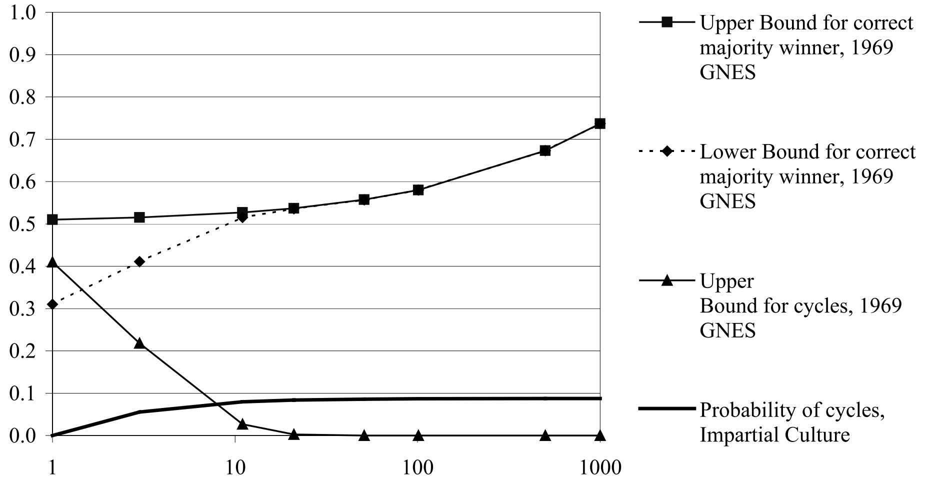
We call this property *moderate stochastic transitivity with strict inequalities* because of its similarity to *moderate stochastic transitivity* [24], which is the same implication, with strict inequality signs, applied to the adjusted pairwise margins.

Theorem 2. *If moderate stochastic transitivity with strict inequalities holds in the population then for any sufficiently large sample drawn from this population the next most probable majority preference relation after the population majority preference relation is a transitive one.*

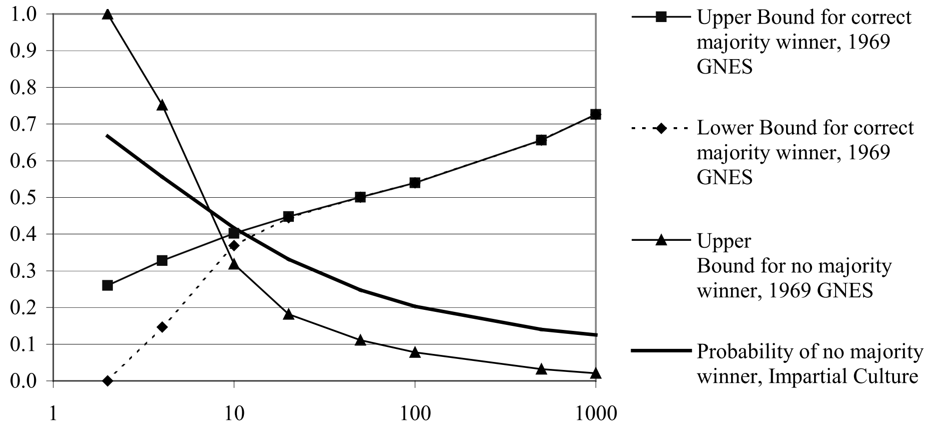
Proof. If a unique minimal adjusted pairwise margin exists, then Theorem 1 trivially implies the result because a reversal of the smallest adjusted pairwise margin, under moderate stochastic transitivity with strict inequalities, leads again to a transitive majority relation.

If a unique minimal adjusted pairwise margin does not exist, i.e., $\delta_1^* = \delta_2^*$, then the second most probable majority preference relation is still transitive. Indeed, if there is no triple a, b, c involving the pairs with margins $\delta_1^* = \delta_{aBb}^*$ and $\delta_2^* = \delta_{bBc}^*$, then we can apply Theorem 1 directly to each triple. If the pairs with adjusted pairwise margins δ_1^* and δ_2^* are in one triple, it may only be the case that $a \succ_p b \succ_p c$, $\delta_{aBb}^* = \delta_{bBc}^* = \delta_1^* = \delta_2^*$ for some appropriate choice of distinct a, b, c. From the normal approximation of the multinomial distribution it follows that

$$\lim_{N\to\infty} \frac{P((\boldsymbol{N}_{bBa} > \boldsymbol{N}_{aBb}) \cap (\boldsymbol{N}_{cBb} > \boldsymbol{N}_{bBc}))}{P(\boldsymbol{N}_{bBa} > \boldsymbol{N}_{aBb}) + P(\boldsymbol{N}_{cBb} > \boldsymbol{N}_{bBc})} = 0$$



a



b

Fig. 2a,b. Bounds on probabilities of the correct majority winner and of no majority winner, in samples from the GNES 1969 survey; probability of no majority winner in samples from the impartial culture. The horizontal axis displays the sample size, the vertical axis displays the probability. **a** is for odd N; **b** is for even N

unless the values $\boldsymbol{N}_{bBa} - \boldsymbol{N}_{aBb}$ and $\boldsymbol{N}_{cBb} - \boldsymbol{N}_{bBc}$ are perfectly correlated. But these values are perfectly correlated if and only if $p_{aBbBc} + p_{cBbBa} = 1$. In this case, the probability of a cycle in the triple $\{a, b, c\}$ is zero for any N. Applying bounds similar to the proof of Theorem 1 we get the result:

$$\lim_{N \to \infty} \frac{P(a \succ_s b, b \succ_s c, c \succ_s a) + P(b \succ_s a, c \succ_s b, a \succ_s c)}{P(b \succ_s a, b \succ_s c, a \succ_s c) + P(a \succ_s b, c \succ_s b, a \succ_s c)} = 1. \quad \square$$

Two illustrations. In Fig. 2 we present various properties for samples drawn

Table 3. Upper and lower bounds and Monte-Carlo simulation results based on the ranking probabilities in Fig. 1a (ANES 1996). We use the same abbreviations as in Table 1

	Social welfare order	$d \succ c \succ p$	$d \sim c \succ p$	$c \succ d \succ p$	$c \succ d \sim p$	$c \succ p \succ d$
Formulae for upper and lower bounds	Upper bound	$P(d \succ c)$	$P(c \sim d)$	$P(d \succ p)$	$P(p \sim d)$	$P(p \succ d)$
	Lower bound	0	0	$P(d \succ p) - (1 - P(c \succ p)) - (1 - P(c \succ d))$	$P(d \sim p) - (1 - P(c \succ p)) - (1 - P(c \succ d))$	$P(p \succ d) - (1 - P(c \succ p)) - (1 - P(c \succ d))$
Sample Size						
50	Upper Bound	0.024	0.010	0.841	0.034	0.125
	Monte-Carlo	0.03	0.01	0.80	0.03	0.13
	Lower Bound	0	0	0.807	0.000	0.091
101	Upper Bound	0.003	0.001	0.930	0.013	0.057
	Monte-Carlo	0	0	0.93	0.01	0.06
	Lower Bound	0	0	0.926	0.009	0.053
500	Upper Bound	<1E –9	<1E –9	1.000	6.08E –05	3.15E –04
	Monte-Carlo	0	0	1	0	0
	Lower Bound	0	0	1.000	6.08E –05	3.15E –04

from a realistic distribution based on the 1969 German National Election Survey (GNES) [27]. The 1969 GNES data are presented in Fig. 1b. We treat the survey as a population from which we can draw random samples (with replacement). Figure 2a treats the case of odd N, whereas Figure 2b covers the case of even N. We plot 1) upper and 2) lower bounds on the probability of the correct majority winner, 3) an upper bound on the probability of no majority winner (for odd N this translates into an upper bound on the probability of cycles) and 4) for comparison, the probability of cycles when drawing a sample of corresponding size from an impartial culture (instead of GNES). There, our best lower bound on the probability of cycles is zero, and therefore not displayed in the figure.

The three choice alternatives were c (CDU/CSU), f (FDP) and s (SDP), the major German parties. The true majority ranking in the survey is $c \succ_p s \succ_p f$, and the pairwise margins are: $\delta_{cBs} = 0.02$, $\delta_{cBf} = 0.6$, $\delta_{sBf} = 0.58$. According to Theorems 1 and 2, if we now draw random samples (with replacement) from the survey, the probability of the correct majority winner in such a sample approaches 1 much more slowly (with increasing sample size) than does the probability of transitive majority preferences in the sample. Figure 2 also shows that the probability of a majority cycle for samples drawn from an impartial culture are systematically higher than the corresponding probabilities derived from our realistic distribution, the 1969 GNES.

We now look once more at the 1996 ANES survey to illustrate how we can map out the complete picture of all possible (sample) majority preference outcomes from pairwise majority preference probabilities. The 1996 ANES survey also differs from the 1969 GNES survey in that indifference was allowed at the level of individual preferences. As we saw before, Table 1 contains the derived probabilities for pairwise sample majority preferences as a function of sample size derived from Eq. (2). Table 3 summarizes our analytical upper and lower bounds on the probability of the correct majority preference relation (of Proposition 3) and compares them with the Monte-Carlo simulations of Regenwetter et al. [29].

From Table 3 we can see that a sample of size 50 has a 96% chance of correctly ranking Clinton ahead of Dole by majority rule, and a 84% chance of correctly ranking Dole ahead of Perot. A group of 50 is virtually guaranteed to correctly rank Clinton ahead of Perot by majority rule.

For a random sample of 50, our upper bound on the chance of their correctly recovering the true majority order is 0.841, the lower bound is 0.807. Regenwetter et al. [29] determined the probability of correctly recovering the population majority preference in a sample of size 50 to be 0.8 by simulating 10,000 repeated samples (of size 50) from the 1996 ANES survey. For $N \geq 100$ the upper and lower bounds become extremely close and correspond to the results of Regenwetter et al.'s [29] Monte-Carlo simulations.

These illustrations show that our upper and lower bounds provide a surprisingly accurate assessment of the sampling probabilities for several of the most probable majority relations in the sample (as long as the pairwise margins in the population are nonzero). For N of several hundred the correct majority order is virtually guaranteed. For even rather small N the sample majority order $c \succ_s p \succ_s d$ has much higher probability than any other incorrect majority order, or any cycle, in particular. This illustrates Theorem 1, with $\succ^* = \{(c,p),(p,d),(c,d)\}$.

4 Inferring population majority preferences from sample data

We now turn to the inverse problem of drawing inferences from observations made in a sample of data. Suppose that we have a random sample of asymmetric binary preference relations, and we want to draw an inference about the probability of majority winners and majority social welfare relations in the population from which the sample (of size N) was drawn.

The natural approach is a Bayesian updating framework. Here, any possible population probability distribution over $\mathscr{B}$ is conceptualized as a set of parameters $p = (p_B)_{B \in \mathscr{B}}$ with the restriction that $\sum_{B \in \mathscr{B}} p_B = 1$. In order to take into account the uncertainty in the values of these parameters, we consider a family of jointly distributed random variables $\boldsymbol{p} = (\boldsymbol{p}_B)_{B \in \mathscr{B}}$ satisfying the constraint that $\sum_{B \in \mathscr{B}} \boldsymbol{p}_B = 1$ (everywhere). The joint distribution of $\boldsymbol{p}$ will capture the probability that $\boldsymbol{p}$ takes any particular parameter values. In particular, we will be interested in the probability that the population exhibits any particular majority preference relation $\succ$ for a given set of data. This translates into finding the probability

$$P(\succ_{\boldsymbol{p}} = \succ) = P\left(\begin{array}{c} \bigcap_{a \succ b} \boldsymbol{p}_{aBb} > \boldsymbol{p}_{bBa} \\ \cap \\ \bigcap_{c \not\succ d} \boldsymbol{p}_{cBd} \leq \boldsymbol{p}_{dBc} \end{array} \right). \tag{9}$$

Clearly, $\succ$ has to be asymmetric, since otherwise $P(\succ_{\boldsymbol{p}} = \succ) = 0$. If, furthermore, $\succ$ is complete, then (9) reduces to

$$P(\succ_{\boldsymbol{p}} = \succ) = P\left(\bigcap_{a \succ b} \boldsymbol{p}_{aBb} > \boldsymbol{p}_{bBa} \right). \tag{10}$$

The first task towards finding this probability is to estimate the distribution of $\boldsymbol{p}$ given the sample $D = (N_B)_{B \in \mathscr{B}}$. According to the general Bayesian approach (e.g., [19]) we can use Bayes' formula:

$$P(D \cap p) = P(D|p)P(p) = P(p|D)P(D).$$

Here $P(p)$ is the prior probability distribution of $\boldsymbol{p}$ (before observing the data, i.e., an assumption about this distribution or prior information) and $P(p|D)$ is the posterior distribution of $\boldsymbol{p}$ (after observing the sample of data, i.e., it is the distribution that we wish to estimate).

Clearly, $P(D|p)$ has a multinomial distribution:

$$P(D|p) = \frac{N!}{\prod_{B\in\mathscr{B}} N_B!} \prod_{B\in\mathscr{B}} p_B^{N_B}, \quad \sum_{B\in\mathscr{B}} N_B = N.$$

Proposition 4. *The posterior distribution $P(p|D)$ is given by:*

$$P(p|D) = \frac{P(D|p)P(p)}{P(D)} = \frac{P(p)}{P(D)} \frac{N!}{\prod_{B\in\mathscr{B}} N_B!} \prod_{B\in\mathscr{B}} p_B^{N_B} = const * P(p) \prod_{B\in\mathscr{B}} p_B^{N_B},$$

and const is given by the normalization condition: $\int P(p|D)\, dp = 1$.

Now it is time to make an assumption about $P(p)$. A convenient and common assumption is that $P(p)$ has a Dirichlet distribution

$$P(p) = \frac{\Gamma(\alpha)}{\prod_{B\in\mathscr{B}} \Gamma(\alpha_B)} \prod_{B\in\mathscr{B}} p_B^{\alpha_B - 1}, \quad \sum_{B\in\mathscr{B}} \alpha_B = \alpha,$$

where $\Gamma(\alpha)$ is a gamma-function (for integers, $\Gamma(N+1) = N!$). The reason we use the Dirichlet is that it provides a very flexible and general class of distributions which can approximate almost any prior distribution. The Dirichlet distribution belongs to the natural conjugate family for multinomial distributions, which means that in this case the posterior distribution is also Dirichlet.

Proposition 5. *If the prior distribution $P(p)$ is Dirichlet, then $P(p|D)$ is also Dirichlet and is given by:*

$$P(p|D) = \frac{\Gamma(\alpha + N)}{\prod_{B\in\mathscr{B}} \Gamma(\alpha_B + N_B)} \prod_{B\in\mathscr{B}} p_B^{N_B + \alpha_B - 1}. \tag{11}$$

Equation (11) provides a general framework for the estimation of majority preferences in the population given a sample. The direct calculation of the probability of any possible majority preference relation from the Dirichlet distribution requires multidimensional integration. The latter is analytically intractable and computationally extremely expensive. On the other hand, we can use upper and lower bounds if we know the probabilities with which any given candidate is preferred to any other given candidate by a majority.

Theorem 3. *Given N_{aBb} and N_{bBa} in the sample, and given the parameters α_{aBb} and α_{bBa} of the prior distribution, the posterior probability that a is preferred to b by a majority in the population is given by the following beta distribution:*

$$P((a \succ_{\mathbf{p}} b) \mid D) = F_\beta\left(\frac{1}{2}, N_{bBa} + \alpha_{bBa}, N_{aBb} + \alpha_{aBb}\right). \tag{12}$$

A proof of Theorem 3 is provided in the Appendix. Notice that the probability $P((a \succ_{\mathbf{p}} b) \mid D)$ is given by a beta distribution and depends neither upon the number of indifference relations (N_{aEb}) in the sample nor upon α_{aEb} in the prior distribution.

The appropriate choice of priors to adequately incorporate prior information plays an important role in Bayesian statistics. We do not concentrate on this choice, but illustrate our results by choosing priors that give the highest weight to the sample information. This is equivalent to the assumption that before observing a sample we have no information about $\boldsymbol{p}$, so $P(p)$ has a "flat" distribution (diffuse relative to the sample information), i.e., that each set of parameters is equally probable. Without loss of generality, a "flat" distribution can be described as a Dirichlet distribution with $\alpha_B \leq 1, \ \forall B \in \mathscr{B}$. Below we assume Eq. (12) together with $\alpha_{aBb} = \alpha_{bBa} = 1$.[8]

As in Sect. 3, we introduce a function $Err(N, \delta)$ for the probability that the results of a pairwise comparison in the population and in the sample differ.

Proposition 6. *Let* $N = N_{aBb} + N_{bBa}$, $\delta = \frac{N_{aBb} - N_{bBa}}{N}$. *Then* $Err(N, \delta)$ *is given by*

$$Err(N, \delta) = F_\beta\left(\frac{1}{2}, N\frac{1+\delta}{2} + 1, N\frac{1-\delta}{2} + 1\right). \tag{13}$$

Equation (13) follows directly from Eq. (12), after the substitution $\alpha_{aBb} = \alpha_{bBa} = 1$ and the use of the following equality:

$$1 - F_\beta\left(\frac{1}{2}, N_{bBa} + 1, N_{aBb} + 1\right) = F_\beta\left(\frac{1}{2}, N_{aBb} + 1, N_{bBa} + 1\right).$$

Having Eq. (12) for the pairwise majority probability or (13) for the probability of erroneous pairwise majority preferences, we can use the method of upper and lower bounds exactly in the same way as we did in Sect. 3. We illustrate the method of upper and lower bounds in the inference framework by analyzing survey data of the 1988 French National Election study (FNES) [28]. Contrary to our earlier analyses, we will treat these survey data as a random sample from some underlying population or culture, and we now ask inference questions about the majority preferences in that population, given the sample.

The original data consist of thermometer scores for five candidates *m* (Mitterrand), *b* (Barre), *c* (Chirac), *l* (Lajoinie) and *p* (Le Pen). After recoding the thermometer scores as strict weak orders, there are several hundred possible states of individual preference. The number of possible majority preference relations is 3^{10}. However, according to our approach, we only need to compute the population pairwise comparison probabilities. Table 4 summarizes this analysis.

As can be seen in Table 4, the probability that the pairwise majority pref-

[8] One referee pointed out the fact that the Dirichlet distribution also appears in the impartial anonymous culture assumption [3, 4]. It is important to notice that under the assumption of diffuse priors, the choice of a Dirichlet distribution is made only for analytical convenience. Any other diffuse prior distribution (e.g., a normal distribution with large enough variance) would lead to the same numerical results.

Table 4. Analysis of 1988 FNES data. For the binary preference between each pair of candidates the number of respondents preferring x to y (N_{xBy}), the number of respondents preferring y to x (N_{yBx}) and the probability of an incorrect inference of this pairwise majority preference are presented. The total number of respondents is 961

x, y	$x = m$, $y = b$	$x = m$, $y = c$	$x = m$, $y = l$	$x = m$, $y = p$	$x = b$, $y = c$
N_{xBy}	538	546	786	734	442
N_{yBx}	328	318	55	153	246
N_{xEy}	95	97	120	74	273
Probability of incorrect inference	3.8E −13	3.2E −15	3.7E −167	3.5E −92	2.8E −14

x, y	$x = b$, $y = l$	$x = b$, $y = p$	$x = c$, $y = l$	$x = c$, $y = p$	$x = l$, $y = p$
N_{xBy}	648	764	577	720	483
N_{yBx}	173	104	248	103	271
N_{xEy}	140	93	136	138	207
Probability of incorrect inference	8.2E −66	2.1E −125	1.9E −31	3.0E −115	4.0E −15

Table 5. Two most probable population majority preference relations for 1988 FNES

Ranking	$m \succ b \succ c \succ l \succ p$	$b \succ m \succ c \succ l \succ p$	Any other
Upper bound	1.0–3.8E −13	3.8E −13	2.8E −14
Lower bound	1.0–4.2E −13	3.5E −13	

erences in the population and in the sample do not coincide is very small for each pair. An upper bound on the probability of an incorrect majority preference relation (i.e., different from $m \succ b \succ c \succ l \succ p$) is given by the sum of the probabilities of erroneous pairwise majority preferences (across all pairs). This upper bound equals 4.2E −13. In other words, we are virtually certain to make the correct inference for all pairwise majority comparisons.

Continuing our analysis we can see that the second most probable population majority ranking (given the data) is $b \succ m \succ c \succ l \succ p$, the probability of which is bounded from below by 3.5E −13 and from above by 3.8E −13 (Table 5). Analogously to Theorem 1, this probability again dramatically exceeds the probability of any other majority relation different from the majority relation in the sample.

Notice in passing that all triples in this survey satisfy a property reminiscent of moderate stochastic transitivity with strict inequalities: for each triple $a \succ b \succ c$ we find $Err(aBc) < \max(Err(aBb), Err(bBc))$. This implies that for each triple the most probable erroneous majority preference relation is a transitive one, not a cycle. Because all pairwise margins in 1988 FNES are high enough, the probability of an incorrect inference is minuscule. If some pairwise margins are low and the number of respondents in the survey is not very high, as is the case for the GNES 1969 (Fig. 1) then the probability of incorrect inference can be quite high. The 818 respondents of the GNES 1969 may lead one to infer an incorrect population majority preference relation with probability 0.28.

5 Summary and conclusions

Van Deemen [38] in this journal concludes (p. 181, last paragraph)

> A final remark concerns the equal likelihood of the voter preferences. As is well-known, the assumption of impartial culture as employed in this paper and in most other works in the field is highly implausible. It is empirically more relevant to develop probability models that use vote frequencies, for example obtained by elections, as input.

Following Van Deemen's recommendation, we focus on the analysis of real world distributions of preferences.[9] We go beyond Van Deemen's framework by allowing individual preferences to be arbitrary asymmetric binary relations. We also go beyond just analyzing the probability of the Condorcet paradox in random samples from realistic distributions. As soon as we analyze real world data, we face the uncertainty of how reflective these data are of the true distribution of preferences in the electorate. Therefore, Van Deemen's recommendation naturally leads to statistical considerations.

We provide two frameworks to conceptualize, quantify and assess the risk of incorrect majority social welfare relations: 1) a statistical sampling framework, where we quantify the probability that a random sample exhibits the correct or an incorrect majority preference relation, relative to the population it was drawn from, 2) a Bayesian inference framework, where a random sample is used to infer or update a probability distribution over possible majority preference relations in the (unknown) population that the data were drawn from.

We provide explicit formulae for the probabilities of pairwise majorities in both frameworks. We show that the concepts of upper and lower bounds using pairwise majority probabilities is extremely useful to analyze all possible majority preference outcomes.

We prove that, for any number of candidates, if in the population (sampling problem) or in the sample (inference problem) there is an asymmetric complete majority preference relation (e.g., a strict linear ordering, or a strict majority cycle) then for a large sample that same relation will be represented in the sample (sampling problem) or inferred from the sample (inference prob-

[9] Rather than rely on ballot data, however, we start with survey data.

lem) with probability arbitrarily close to one. In particular this shows that the probability of cycles approaches zero in large samples if the majority preferences in the population (or underlying culture) form a linear order. Although a similar result for sampling has recently been published in this journal [35], that result is based on strong additional assumptions, namely that preferences are generated by cardinal utilities which are independent among voters. One particular strength and elegance of our approach is that we derive our results with virtually no constraints on the nature of individual preferences.

We define the concept of moderate stochastic transitivity with strict inequalities as a property which guarantees that for sufficiently large sample size the second most probable majority preference relation (after the correct one) is again a transitive (but incorrect) majority preference relation. The required sample size depends on the actual pairwise margins and can be calculated using the formulae for pairwise comparison probabilities.

Moderate stochastic transitivity with strict inequalities is in fact satisfied by all survey distributions that we have investigated. This supports our conjecture that incorrect transitive majority preferences are far more likely in practice than majority cycles (a fact which has to be compounded with the observation that the price to society might be higher when an incorrect decision is made than when a decision is prevented or delayed by the occurrence of a cycle).

Every social choice procedure is vulnerable to incorrect assessments based on data in the presence of uncertainty. We envision similar analyses also for other preference aggregation methods and tally procedures that generate a social ordering. This social ordering is again the conjunction of pairwise comparisons and thus the logic of upper and lower bounds on the probability of a correct or incorrect social ordering will, in principle, still apply. From an analytical point of view, possible limitations to generalizing our approach may arise from the computation of the paired comparison probabilities. From a practical point of view, the only situation where the bounds will not be tight is the knife-edge case when the pairwise tallies are tied. Another obvious limitation of our results is the fact that many real world social choice data, such as committee ballots or political election ballots can not safely be viewed as random sample data. (An interesting exception, in principle, may be juries and their group decisions.) This opens up a large array of possible refinements of our approach. For example, each stakeholder, committee (or board) member may be representative of one subpopulation, and each subpopulation may have its own distribution of preferences. The next stage of the present framework is to investigate such mixture (or hierarchical, or latent class) scenarios.

Another set of open questions is to use tools like the one presented here, to make inferences about whether or not a particular tally procedure would likely result in the election of a majority winner. This generalizes Regenwetter and Grofman [31, 32] who provided evidence that approval voting tends to elect a majority winner and Borda winner in practice.

From a decision sciences point of view, the real work lies still ahead: Once the probabilities of various erroneous majority preference assessments are understood, these probabilities have to be compounded with the costs of all possible errors. Determining the cost of using an erroneous but transitive majority preference relation in a public policy context is a task that goes far beyond the scope of this paper. For example, among those who believe that the incorrect candidate was elected in the 2000 U.S. presidential election, there undoubtedly is a broad range of opinions about the price tag of that "erroneous" outcome. We only emphasize that a majority cycle in a committee at least has the advantage of being noticed, whereas an incorrect but transitive majority preference assessment may go unnoticed, and thus, uncorrected.

Our work is similar to the standard survey and polling research in that our method can be applied to survey data. However the traditional focus of survey research has been somewhat different, such as the representativity of empirical samples and the challenge of predicting actual voting behavior from opinion surveys. Our focus is on assessing how confident we can be that the actual election outcome is correct, when the ballot casting and counting processes contain components of uncertainty.

The task of aggregating preferences raises two important questions: 1) What is a rational social choice function? 2) How accurately can we assess a given social choice function in the face of uncertainty, based on available information?

Traditional social choice theory has been preoccupied with the first question. We have addressed the second question in the case of majority rule. In majority rule, as long as majority preferences are transitive, Arrow's ideal is satisfied. Given a set of empirical rating or ranking data, our framework permits to assess the probability, not only, that Arrow's ideal is satisfied, but also that we are making a majority rule decision which, besides being rational, is also correct.

Appendix

Proof of Theorem 3. Let us consider a sample in which we know N_{aBb}, N_{bBa}, and N_{aEb}; $N_{aBb} + N_{bBa} + N_{aEb} = N$. The parameters of the prior Dirichlet distribution are given by α_{aBb}, α_{bBa}, α_{aEb}, $\alpha_{aBb} + \alpha_{bBa} + \alpha_{aEb} = \alpha$. We introduce the following abbreviations: $G = N_{aBb} + \alpha_{aBb} - 1$ (a is Greater than b), $L = N_{bBa} + \alpha_{bBa} - 1$ (a is Less than b) and $S = N_{aEb} + \alpha_{aEb} - 1$ (a is the Same as b). Then for the posterior distribution of $\boldsymbol{p}_{bBa} = \boldsymbol{l}$ and $\boldsymbol{p}_{aEb} = \boldsymbol{s}$ ($\boldsymbol{p}_{aBb} = \boldsymbol{g}$ is determined by the equality $\boldsymbol{g} = 1 - \boldsymbol{s} - \boldsymbol{l}$), we have from (11):

$$P(l, s \mid G, L, S) = \frac{\Gamma(G+L+S+3)}{\Gamma(G+1)\Gamma(L+1)\Gamma(S+1)} l^L s^S (1-s-l)^G.$$

The condition that a is preferred to b by a majority in the population means that $\boldsymbol{g} > \boldsymbol{l}$ or $\boldsymbol{l} < \frac{1-\mathbf{s}}{2}$, so for the probability that a is preferred to b by a majority in the population we obtain:

$$
\begin{aligned}
&P((a \succ_{\boldsymbol{p}} b) \mid D) \\
&\quad = P\left(\boldsymbol{l} < \frac{1-\boldsymbol{s}}{2} \mid D\right) \\
&\quad = \frac{\Gamma(G+L+S+3)}{\Gamma(G+1)\Gamma(L+1)\Gamma(S+1)} \int_{s=0}^{1} s^S \left(\int_{l=0}^{(1-s)/2} l^L (1-s-l)^G \, dl \right) ds.
\end{aligned}
$$

Substituting $t = \frac{l}{(1-s)}$ yields

$$
\int_{l=0}^{(1-s)/2} l^L (1-s-l)^G \, dl = (1-s)^{L+G+1} \int_0^{1/2} t^L (1-t)^G \, dt.
$$

Denoting the cumulative beta distribution of variable x with parameters $L+1$ and $G+1$ by $F_\beta(x, L+1, G+1)$, namely

$$
F_\beta(x, L+1, G+1) = \int_0^x \frac{\Gamma(G+L+2)}{\Gamma(G+1)\Gamma(L+1)} t^L t^G \, dt,
$$

we get the following result

$$
\begin{aligned}
&P\left(\left(\boldsymbol{l} < \frac{1-\boldsymbol{s}}{2}\right) \mid D\right) \\
&\quad = \frac{\Gamma(G+L+S+3)}{\Gamma(G+1)\Gamma(L+1)\Gamma(S+1)} \frac{\Gamma(G+1)\Gamma(L+1)}{\Gamma(G+L+2)} \\
&\qquad \times F_\beta\left(\frac{1}{2}, L+1, G+1\right) \int_{s=0}^{1} s^S (1-s)^{L+G+1} \, ds \\
&\quad = \frac{\Gamma(G+L+S+3)}{\Gamma(S+1)\Gamma(G+L+2)} F_\beta\left(\frac{1}{2}, L+1, G+1\right) \frac{\Gamma(G+L+2)\Gamma(S+1)}{\Gamma(G+L+S+3)} \\
&\quad = F_\beta\left(\frac{1}{2}, L+1, G+1\right).
\end{aligned}
$$

Returning to our original notation we get formula (12) in Sect. 4. □

References

[1] Arrow KJ (1951) Social choice and individual values. Wiley, New York

[2] Barberá S, Pattanaik PK (1986) Falmagne and the rationalizability of stochastic choices in terms of random orderings. Econometrica 54: 707–715

[3] Berg S (1985) Paradox of voting under an urn model: The effect of homogeneity. Publ Choice 47: 377–387

[4] Berg S, Bjurulf B (1983) A note on the paradox of voting: Anonymous preference profiles and May's formula. Publ Choice 40: 307–316

[5] Burden B (1997) Deterministic and probabilistic voting models. Am J Polit Sci 41: 1150–1169

[6] DeMeyer F, Plott CR (1970) The probability of a cyclical majority. Econometrica 38: 345–354

[7] Falmagne J-Cl (1978) A representation theorem for finite random scale systems. J Math Psychol 18: 52–72
[8] Falmagne J-Cl, Regenwetter M (1996) Random utility models for approval voting. J Math Psychol 40: 152–159
[9] Feld SL, Grofman B (1988) Ideological consistency as a collective phenomenon. Am Polit Sci Rev 83: 773–788
[10] Fishburn PC (1992) Induced binary probabilities and the linear ordering polytope: A status report. Math Soc Sci 23: 67–80
[11] Gehrlein W (1988) Probability calculations for transitivity of the simple majority rule. Econ Lett 27: 311–315
[12] Gehrlein W (1999) Condorcet efficiency of Borda rule under the dual culture condition. Soc Sci Res 28: 36–44
[13] Gehrlein W, Fishburn P (1976) Condorcet's paradox and anonymous preference profiles. Publ Choice 26: 1–18
[14] Gehrlein W, Fishburn P (1978) Probabilities of election outcomes for large electorates. J Econ Theory 19: 38–49
[15] Gehrlein W, Fishburn P (1979) Effects of abstentions on voting procedures in three-candidate elections. Behav Sci 24: 346–354
[16] Gehrlein W, Lepelley D (1999) Condorcet efficiencies under the maximal culture condition. Soc Choice Welfare 16: 471–490
[17] Gehrlein W, Valognes F (2001) Condorcet efficiency: A preference for indifference. Soc Choice Welfare 18: 193–205
[18] Gehrlein WV, Lepelley D (2001) The Condorcet efficiency of Borda rule with anonymous voters. Math Soc Sci 41: 39–50
[19] Heckerman D (1998) A tutorial on learning with Bayesian networks. In: Jordan MI (ed) Learning in graphical models. MIT Press
[20] Jones B, Radcliff B, Taber C, Timpone R (1995) Condorcet winners and the paradox of voting: Probability calculations for weak preference orders. Am Polit Sci Rev 89(1): 137–144
[21] Lepelley D, Pierron P, Valognes F (2000) Scoring rules, Condorcet efficiency and social homogeneity. Theory Decision 49: 175–196
[22] Luce RD (1959) Individual choice behavior: A theoretical analysis. Wiley, New York
[23] Luce RD (1977) The choice axiom after twenty years. J Math Psychol 15: 215–233
[24] Luce RD, Suppes P (1965) Preference, utility and subjective probability. In: Luce RD, Bush RR, Galanter E (eds) Handbook of mathematical psychology, vol. III. Wiley, New York pp 249–410
[25] McFadden D (1998) Specification of econometric models. Econometrica (forthcoming)
[26] Niemi R, Wright J (1987) Voting cycles and the structure of individual preferences. Soc Choice Welfare 4: 73–183
[27] Norpoth H (1979) The parties come to order! Dimensions of preferential choice in the West German electorate, 1961–1976. Am Polit Sci Rev 73: 724–736
[28] Pierce R (1996) French presidential elections survey, 1988. Inter-University Consortium for Political and Social Research, Ann Arbor, MI
[29] Regenwetter M, Adams J, Grofman B (2002) On the Condorcet efficiency of majority rule: An alternative view of majority cycles and social homogeneity. Theory Decision (in press)
[30] Regenwetter M, Falmagne J-Cl, Grofman B (1999) A stochastic model of preference change and its application to 1992 presidential election panel data. Psychol Rev 106: 362–384
[31] Regenwetter M, Grofman B (1998) Approval voting, Borda winners and Condorcet winners: Evidence from seven elections. Manag Sci 44: 520–533
[32] Regenwetter M, Grofman B (1998) Choosing subsets: A size-independent proba-

bilistic model and the quest for a social welfare ordering. Soc Choice Welfare 15: 423–443
[33] Regenwetter M, Marley AAJ, Grofman B (2002) A general concept of majority rule. Math Soc Sci 43: 405–428
[34] Sapiro V, Rosenstone SJ, Miller WE (1998) American national election studies, 1948–1997. Inter-university Consortium for Political and Social Research, Ann Arbor, MI
[35] Tangian AS (2000) Unlikelihood of Condorcet's paradox in a large society. Soc Choice Welfare 17: 337–365
[36] Thurstone LL (1927) A law of comparative judgement. Psychol Rev 34: 273–286
[37] Tversky A (1972) Choice by elimination. J Math Psychol 9: 340–367
[38] Van Deemen A (1999) The probability of the paradox of voting for weak preference orderings. Soc Choice Welfare 16: 171–182