ELSEVIER

Mathematical Social Sciences 43 (2002) 451–466

mathematical social sciences

www.elsevier.com/locate/econbase

On the model dependence of majority preference relations reconstructed from ballot or survey data

Michel Regenwetter[a,*], Bernard Grofman[b], A.A.J. Marley[c]

[a] *Department of Psychology, University of Illinois at Urbana-Champaign, IL, USA*
[b] *School of Social Sciences, University of California at Irvine, Irvine, CA, USA*
[c] *Department of Psychology, McGill University, Montreal, PQ, Canada*

Abstract

The Condorcet criterion requires a majority winner to be elected whenever it exists. In practice, virtually no social choice procedure or survey research study reports complete pairwise comparisons or a complete ranking of the choice alternatives for each voter, as required by standard majority rule. Thus, majority rule can almost never be computed directly from ballots or survey data. We show that it is typically impossible to unambiguously check a given set of ballot or survey data against the Condorcet criterion because any reconstruction of majority preferences is sensitive to the underlying implicit or explicit model of decision making.

Keywords: Condorcet criterion; Majority rule; Semiorders; Survey data; Weak orders; Voting

1. Introduction

Majority rule for a set of candidates $\mathscr{C}$ is usually defined either in terms of comparisons of every pair of candidates in $\mathscr{C}$, or in terms of complete rankings of the candidates in $\mathscr{C}$. In particular, for $A, B \in \mathscr{C}$, A is (strictly) majority preferred to B if, say, more than 50% of the population prefers A to B. A majority rule winner is a candidate $A \in \mathscr{C}$ which is majority preferred to each $B \in \mathscr{C} - \{A\}$. The *Condorcet criterion*, arguably the most universally accepted normative benchmark for rational social choice, says that a majority winner should be elected whenever it exists.

Quite interestingly, though, virtually no social choice procedures that are actually used in practice, explicitly require or record all pairwise comparisons or complete rankings

*Corresponding author. Tel.: +1-217-333-0763; fax: +1-217-244-5876.
E-mail address: regenwet@uiuc.edu (M. Regenwetter).

0165-4896/02/$ – see front matter
PII: S0165-4896(02)00022-7

and few social choice procedures record any. For instance, the *single nontransferable vote* and *limited voting* require the voter to choose k out of $N \geq k$ candidates, *approval voting* requires the voter to choose any subset of the candidates, the *single transferable vote* typically requires the voters to rank order the top k choices out of $N \geq k$ candidates, *cumulative voting* requires that the voters distribute m points among k many alternatives, etc. Polling and experimental data often take the form of Likert scale ratings, feeling thermometer ratings, elicitations of real valued utility ratings, buying or selling prices, or partial paired comparisons.

To put it sharply, majority rule is undefined in many empirical circumstances because the standard definition is stated in terms of hypothetical but unobserved pairwise comparisons. A companion paper (Regenwetter et al., 2002b) that focuses on theoretical aspects places virtually all empirical paradigms on an equal footing by defining majority rule in terms of theoretical primitives that virtually all choice, rating and ranking paradigms share. To phrase it loosely, our general concept of majority rule states that A is majority preferred to B if the total probability (or frequency) of all preference relations in which A is preferred to B exceeds the total probability (or frequency) of all preference relations in which B is preferred to A. In utility terms, the general concept of majority rule states that A is majority preferred to B if the total probability (or frequency) of all utility functions in which A has higher utility than B exceeds the total probability (or frequency) of all utility functions in which B has higher utility than A. We refer the reader to the theory paper for more precise formulations, references, definitions, theorems and other theoretical results that are used here.

When an argument is made for or against a social choice procedure, it is common practice to challenge the procedure's Condorcet efficiency, i.e., the likelihood that the procedure satisfies the Condorcet criterion. In other words, the opponents of a given voting procedure may bring forth the proposition that this voting procedure does not have a very high likelihood of electing a Condorcet winner when one exists. At the same time, explicitly collecting all paired comparisons among candidates to directly calculate majority preferences is rather tedious and expensive. Thus, it is generally desirable to find a simpler, cost-effective voting procedure, which, at the same time, carries the benefit of satisfying the Condorcet criterion with high probability. In this vein, a significant effort has been made in the social choice literature to discuss the Condorcet efficiency of various voting methods. There is a substantial literature on the Condorcet efficiency of various aggregation methods (Adams, 1997; Felsenthal and Machover, 1995; Felsenthal et al., 1990, 1993; Gehrlein, 1987, 1992, 1998a,b; Gehrlein and Berg, 1992; Gehrlein and Lepelley, 1998, 1999, 2001; Gehrlein and Valognes, 2001; Lepelley, 1993; Lepelley and Gehrlein, 2000; Lepelley et al., 2000; Lepelley and Valognes, 1999; Tataru and Merlin, 1997). In this literature, very little effort has been made to empirically check whether a given aggregation procedure satisfies the Condorcet criterion on a given set of ballot or survey data. Rather, almost all of the literature is of a theoretical and hypothetical nature. Furthermore, much of the theoretical work is based on implausible or unrealistic assumptions about the nature or distribution of individual preferences. In this paper, we show that the virtual absence of empirical work is not a surprise. We argue and illustrate that majority rule can not in practice be calculated without an explicit or implicit model that explains voting or polling behavior as a function of underlying preferences or utilities. We also show that the reconstructed majority preferences can completely depend on the chosen model. Therefore, we call on the researchers interested in Condorcet efficiency to check how robust their findings are to model variations or violations, and to combine their theoretical work with empirical analyses aimed at testing the validity of their modeling assumptions.

One of our illustrating paradigms is approval voting, whose relationship to the Condorcet criterion has been extensively discussed in the literature (Brams and Fishburn, 2001; Felsenthal et al., 1990; Gehrlein and Lepelley, 1998; Lepelley, 1993; Regenwetter and Grofman, 1998a; Saari, 2001; Tabarrok, 2001; Wiseman, 2000). Under approval voting each voter is asked to cast a vote in the form of a subset of candidates that the voter 'approves of'. Each candidate in the approved set receives a point. The approval voting score of a candidate is the total number of points received from the collection of all voters. The approval voting winner(s) is (are) the candidate(s) with the highest approval voting score(s).

As a motivating example, consider an approval vote over the set $\{A, B, C\}$ with the following ballot tallies: the empty set (10 votes), candidate A alone (20 votes), B alone (10 votes), C alone (10 votes), A and B (1 vote), A and C (4 votes), B and C (5 votes), all three candidates (5 votes). The approval voting score of A is $20+1+4+5=30$ points, the approval voting score of B is $10+1+5+5=21$ points, whereas the approval voting score of C is $10+4+5+5=24$ points. Accordingly, the approval voting winner is candidate A. Is there a majority rule (Condorcet) winner and can the three candidates be rank ordered by majority rule (i.e., does there exist a transitive social welfare order)?

As we show, two different but natural ways of generating a majority rule social welfare relation for this set of approval voting ballots yield two drastically different results: According to one method, the majority rule social welfare order is the semiorder where C is majority preferred to B, but A is majority tied with both B and C. According to the other method, A is the single best candidate (i.e., the Condorcet winner) and B is the single worst candidate (i.e., the Condorcet loser), with C majority ranked in between the other two. An opponent could use one method to argue that approval voting, with winner A, fails the Condorcet benchmark on these ballots. At the same time, a proponent of approval voting could use the other method to argue that approval voting, which is extremely easy to use, would have satisfied the Condorcet criterion and would have been more cost effective than collecting the data that are required to generate majority outcomes directly. In line with this example, and as we explain in more detail in the body of the paper, the outcome of majority rule is not theory-free, rather, it may depend on the particular set of theoretical primitives (implicitly or explicitly) used when tallying a particular set of empirical data.

The focus of this paper is on model dependence of majority rule social welfare functions. Much empirical work has attempted to 'estimate' preference profiles from ballot counts (Brams and Fishburn, 1992; Brams et al., 1988a,b; Brams and Nagel, 1991; Saari and Van Newenhizen, 1988a,b; Young, 1986). We question whether even aggregate preferences such as majority preferences can be estimated unambiguously (i.e., in a model free way) from voting ballots or survey data. We adopt a framework that incorporates issues of statistical inference, such as identifiability and testability, into the analysis of actual data.

We investigate, in a random utility framework, four sets of National Election Study feeling thermometer ratings of presidential candidates in certain years in which there were three viable candidates (1968, 1980, 1992, 1996). The feeling thermometer ratings are treated as outcomes of integer valued random utilities, which we translate into probabilistic strict weak order and semiorder preference representations. These various translations account for the basic psychological intuition that feeling thermometer ratings cannot be taken as literal reflections of the respondents' preferences. For instance, there are many response biases that can play a role in the data generating process, such as the use of rating values that are multiples of 10 (or 5) and the excessive use of the value 50 on a 100 point scale. Another concern is the fact that respondents' rating values tend to be somewhat unreliable: For example, a respondent who rates candidate A as 50 and candidate B as 55 might, a short time later, rate candidate A as 60 and candidate B as 50. It is psychologically plausible to expect that the more similar the recorded ratings of two candidates are, the less sure we can be that the implied ordering of the choice alternatives is the correct one (or a stable one). This is why we consider also a variable threshold of utility discrimination in our analysis of the thermometer data.[1]

We also analyze 12 sets of approval voting data from various professional organizations using two alternative probabilistic choice models for subset choice data. In all cases where the two models apply, we (re)construct and compare the majority preference relation of the electorate.

We find a mixed bag of results as a function of the underlying model of preferences or utilities. For some data sets there is no theory-free majority winner, while in other data sets the modeling assumptions do not seem to affect the outcome of the social welfare order. These results suggest that we need to study in more detail when and why, in practice, the computation of majority or other social welfare functions tends to be robust under model variations or model violations.

In Section 2 we report our findings of the 4 national election survey data sets, and in Section 3 we report the findings for the 12 approval voting data sets. The last section is a summary of main findings and conclusions.

2. National election study feeling thermometers

We look at 1968, 1980, 1992, and 1996 National Election Study feeling thermometer data (Sapiro et al., 1998) for the major presidential candidates. Each of these presidential elections had three major candidates (Humphrey, Nixon, Wallace in 1968; Anderson, Carter, Reagan in 1980; Bush, Clinton, Perot in 1992; and Clinton, Dole, Perot in 1996).

Feeling thermometer ratings assign integer values between 0 and 100 to each of the candidates, where a value of 0 denotes a 'very cold feeling towards' a given candidate, a value of 50 denotes a 'neutral feeling towards' the candidate, and a value of 100 denotes

[1] A more appropriate semiorder representation of thermometer data would even allow different threshold values for different respondents. Such elaborations would only further complicate the picture of results, not simplify it. In that sense, we can make our point that different models matter, since they matter even with the simplest possible representations.

a 'very warm feeling towards' the candidate. Feeling thermometer ratings can be viewed as elicitations of natural valued utility functions that are constrained between 0 and 100.[2] More correctly, since the data form a random sample, we can view the sampled joint values as realizations of natural valued jointly distributed random variables each taking values between 0 and 100. In turn we can recode the observed data (relative frequencies) as a probability distribution over strict weak orders. The resulting net probabilities (and probabilities) are reported in Fig. 1 for the 1968, 1980, 1992, and 1996 data.

Fig. 1 reports the net probabilities of the strict weak orders, as well as the probabilities of the strict weak orders (given in parentheses). States with positive net probabilities are more darkly shaded. The majority rule orderings are also indicated.

The majority preference relations can be obtained either by applying Definition 2 of the theory paper to the strict weak order net probabilities, or Definition 2′ of the theory paper directly to the feeling thermometer ratings where the feeling thermometer ratings are interpreted as the (integer valued) joint outcomes of the utility random variables ***U***.

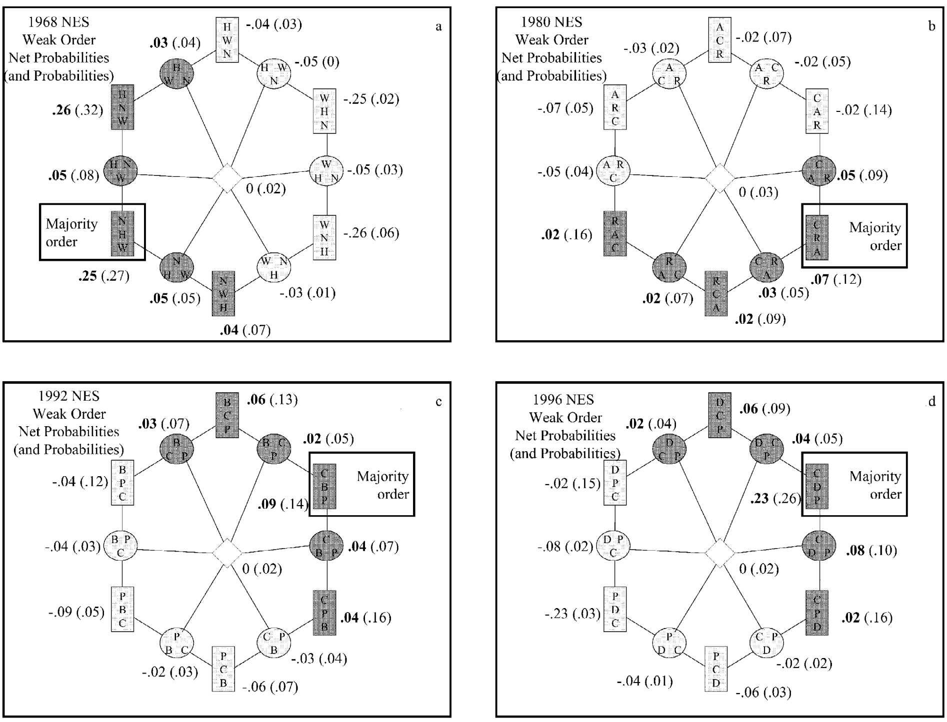

Fig. 1. Weak order probabilities and net probabilities for the presidential candidates in the 1968, 1980, 1992 and 1996 National Election Studies.

[2]As mentioned in an earlier footnote, one could complicate matters further by not viewing the thermometer scores as necessarily directly reflecting the utilities in the decision makers' heads. This would only further emphasize our main point that the underlying model of preferences and utilities is important.

Recoding feeling thermometer data as strict weak orders discards the strength of preference that may have been expressed by the reported thermometer scores and treats the data in a rather literal fashion. One may therefore instead count preference relationships between two candidates as relevant (reliable) only when the utility values indicated by the thermometer scores differ by a 'large enough amount.'

We now use the thermometer scores to illustrate our main point that the method by which we derive majority relationships from utilities (or ratings) makes an implicit or explicit choice of preference representations. By specifying 'large enough amount' as a particular threshold value $\epsilon > 0$ we specify a specific semiorder representation for thermometer data. Specifically, by choosing a threshold of discrimination, we can derive a probability distribution over semiorders. By varying the magnitude of the threshold, we can establish how sensitive majority rule analyses are to the possible lack of reliability in respondents' thermometer ratings. Very large threshold values discard thermometer ratings that are very close to each other, and possibly unreliable, whereas small threshold values treat the comparisons of rating values as highly reliable.

For example, a threshold of 10 means that we consider *a* as being preferred to *b* if and only if the feeling thermometer score for *a* exceeds that for *b* by more than 10 points. Note also that a threshold of zero would lead us back to the strict weak order case that we have already discussed.

We begin the semiorder based analysis with the 1992 data set, because our findings for this case are the most straightforward. We use the graph of all strict partial orders over three elements to illustrate our results about finding majority rule relations.[3] Fig. 2 illustrates what we find when analyzing the 1992 election data by varying the possible values of ϵ. We display the results for threshold values of 10, 29, 50. The semiorder probabilities and the corresponding net probabilities for these three threshold values are shown in Fig. 2. Semiorders with positive net probabilities are more darkly shaded.

The majority rule social welfare order is Clinton$\succ$Bush$\succ$Perot, no matter which threshold value we use for ϵ in the half-open interval [0, 100). For a threshold of 100 all net probabilities obviously vanish, i.e., the majority social welfare order trivially becomes a three-way tie. Notice that, even though the majority order does not change with ϵ, the probability distributions and net probability distributions on the partial order graph change dramatically across the three threshold values that we report.

Similarly, in the 1968 election data, the majority rule social welfare order is Nixon$\succ$Humphrey$\succ$Wallace, no matter which threshold value we use for ϵ in the half-open interval [0, 97). (All thermometer scores of 97 and above were coded as 97 in this data set.)

The finding that for the 1992 and 1968 data sets the majority rule social welfare order is the same for all applicable threshold values can be interpreted as an indication that in these two cases the concept of majority rule social welfare is robust under variations of our implicit or explicit model of preferences and utilities. This is not an artifact of our method, but rather an empirical finding for these two particular elections, as the

[3]Note that all partial orders over three alternatives are also semiorders. If we drop those semiorders from the graph that are not strict weak orders, then the resulting strict weak-order-hedron is closely related to Saari's (1994, 1995) triangle representation when his lines in the triangle represent strict weak orders.

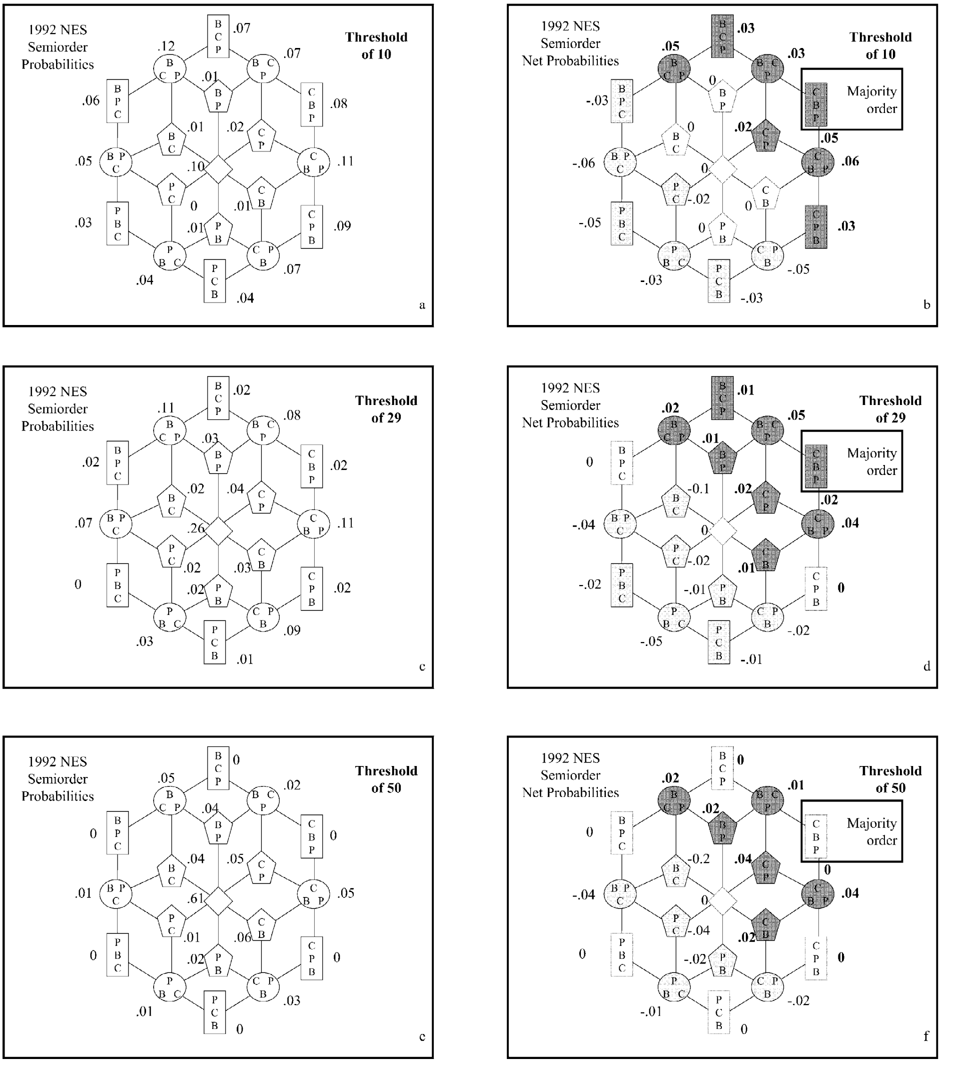

Fig. 2. Semiorder probabilities and net probabilities for the 1992 National Electional Study.

remaining analyses demonstrate. As we will see when we consider the 1980 and 1996 cases, the majority rule ordering can, however, be sensitive to variations in ϵ.

For the 1980 election data, by varying the possible values of the threshold we obtain two different majority rule orderings. For threshold values up to 30, the majority ordering is Carter$\succ$Reagan$\succ$Anderson, whereas for threshold values above 30, the majority ordering is Reagan$\succ$Carter$\succ$Anderson. Similarly, for the 1996 election data we see two different aggregate preference orderings. As we vary the threshold from 1 to

100, the majority preference begins and ends with Clinton≻Dole≻Perot. For threshold values from 50 to 84, we find Dole≻Clinton≻Perot.

The latter two data sets demonstrate that the majority ordering can change (even back and forth) as we vary the value of the threshold of utility discrimination, and thus, the majority ordering depends on the modeling assumptions that implicitly or explicitly enter the data analysis. Next, we turn from thermometer data, which are survey data, to approval voting data, which are ballot data.

3. Approval voting

We consider two methods for computing the majority social welfare relation from approval voting data, each derived from a different model of subset choice behavior. These two models, the topset model (the basis for method 1) and the size-independent model (the basis for method 2) are written out formally in Appendix A.

Method 1 makes two implicit assumptions, namely that 1) a voter is indifferent between any two choice alternatives that are either both included in or both excluded from their approval set and 2) that the voter prefers any chosen alternative to any nonchosen alternative.[4]

Thus A is counted as preferred to B if and only if A is in the chosen set and B is not. This computation is consistent with a majority rule winner as defined in the theory paper, computed via a version of Regenwetter's (1997) *topset voting model* for approval voting (see also Niederée and Heyer, 1997) with the additional assumption that each respondent's preference relation consists of at most two equivalence classes. Whenever there is a unique approval voting winner, method 1 also provides a unique solution.

Method 2 uses the *size-independent model* of approval voting of Falmagne and Regenwetter (1996), which essentially assumes that the respondents completely rank order the candidates mentally, decide how many candidates—say, k many—to approve of and then provide a subset containing the k best elements in the ranking (without rank order information) when asked. Even when there is a unique approval voting winner, method 2 does not always find a unique solution or even apply to a given set of data. (Remember that method 1 does.)

Recall that in our motivating example of the Introduction, we have considered an approval vote over a set {A, B, C} where 20 voters approve of candidate A alone, 10 approve of B alone, 10 approve of C alone, one voter approves of A and B, 4 voters approve of A and C, and 5 approve of B and C. Five voters vote for all candidates, and 10 for none of the three candidates. We posed the question whether there is a majority rule winner (Condorcet winner) here and whether the three candidates can be rank ordered by majority rule.

We now provide the majority rule social welfare relations implied by the two methods when applied to this example.

[4]Mathematically speaking, this means that each preference relation consists of at most two equivalence classes. If one allows individual voters to be nonindifferent between two candidates that they either both approved or both disapproved of, then one cannot use method 1 and claim that it generates the majority winner.

Method 1 (Majority relation derived from the topset model): We tally the paired comparison A versus B as $20+4$ (out of 65) in favor of A because 24 people chose sets that contained A but not B and $10+5$ (out of 65) in favor of B because 15 people chose sets that contained B but not A. With this computation A has a strict majority over B. Similarly, this method of tallying supports a strict majority of C over B and a strict majority of A over C, i.e., A is the Condorcet winner and the majority rule social welfare order is the linear order $A \succ C \succ B$. Formally, according to the topset model (with each preference relation having at most two equivalence classes), writing $P(X)$ for the probability of the subset X, we have

$$\mathcal{NP}_{ab} = \sum_{\substack{X \subseteq \mathscr{C} \\ a \in X \\ b \notin X}} P(X) - \sum_{\substack{Y \subseteq \mathscr{C} \\ a \notin X \\ b \in X}} P(Y).$$

Method 2 (Majority relation derived from the size-independent model): We omit the details of this computation as such detail would require stating various results published earlier (Doignon and Regenwetter, 1997; Regenwetter and Grofman, 1998a,b). Calculations based on the size-independent model lead to the conclusion that C is majority preferred to B, and that A is majority tied with each of B and C. This majority order is a semiorder.

This simple example illustrates that the majority social welfare order can in principle change dramatically as we change our implicit or explicit model of preferences, that is, of the mechanisms generating the observable data. Furthermore, proponents and opponents of approval voting could, in principle, selectively use a model that would support their point of view regarding the Condorcet efficiency of approval voting.

Another important difference between the two methods is the following: The majority rule social welfare order generated by method 1 automatically coincides with the rank order given by the approval voting scores, irrespective of the distribution of subset choices because, by method 1, A is majority preferred to B if and only if A and not B is chosen more often than B and not A, which in turn is determined by who has more approval votes. Previous work has shown that method 2 does not force the approval voting orderings and the majority ordering to match.

It is worth emphasizing once more our main point, namely that the majority orderings can differ depending on the method used to calculate them. One has to expect that this pattern holds beyond the models used here, i.e., other possible models of subset choice behavior are likely to only further complicate the picture, not to simplify it.

Now we turn to the analysis of actual empirical data. Of the 12 data sets we analyze for approval voting elections, 10 were previously considered in Regenwetter and Grofman (1998a). Their results are included in our summary. However, we extend their analyses by using both method 1 and method 2 (they only studied method 2) and by adding two new data sets. Their study used data labeled TIMS E1, TIMS E2, MAA1, MAA2, IEEE, A16, A25, A29, A30 and A72 drawn from professional associations. For example, TIMS E1 is a presidential election of the Institute of Management Science. We are adding to that list recent elections of the Society for Social Choice and Welfare (labeled SSCW) and the Society for Judgment and Decision Making (labeled SJDM).

Table 1
Strict majority social welfare orders for approval voting data sets using the topset voting and size-independent models

	Tims E1	Tims E2	MAA 1	MAA 2	A16	A25	A29	A30	A72	IEEE	SSCW	SJDM
empty set	68	78	224	510	0	0	0	0	0	6001	2	0
$\{a\}$	66	199	1257	413	48	3	13	7	11	14365	26	22
$\{b\}$	350	248	626	1798	4	32	0	6	4	12254	11	24
$\{c\}$	290	273	1434	1019	2	9	10	1	17	11478	26	9
$\{a, b\}$	175	84	68	29	0	2	0	0	6	2934	2	3
$\{a, c\}$	105	114	260	52	5	1	0	1	68	2746	2	4
$\{b, c\}$	442	220	141	199	0	2	1	0	13	1811	0	4
$\{a, b, c\}$	71	161	30	20	69	134	44	43	153	4380	2	1
Total number of votes cast	1567	1377	4040	4040	128	183	68	58	272	55969	71	67
Approval voting	*bca*	*cba*	*cab*	*bca*	*acb*	*bca*	*acb*	*abc*	*cab*	*abc*	*acb*	*bac*
Majority topset	*bca*	*cba*	*cab*	*bca*	*acb*	*bca*	*acb*	*abc*	*cab*	*abc*	*acb*	*bac*
Majority SIM	*cba* *bca*	*bca* *cba*	*acb* *cab*	*bca*		*bca*			*cab*	one cycle *acb* *abc*	$a(b \sim c)$	both cycles *abc* $(a \sim b)c$ *bac*

Table 1 compares the approval voting and majority outcomes for these 12 elections using both the topset voting and the size-independent model. Three of the 12 data sets, namely A16, A29, A30, violate the assumptions of the size-independent model. Accordingly, the results for these cases are left blank in Table 1. We obtain agreement between the two models in three cases out of nine, and partial agreement in the other six. For instance, in the election by the Society for Social Choice and Welfare, the size-independent model suggests that a is majority winner and that b and c are majority tied, whereas the topset voting model suggests that a is majority winner and that c is the (single) majority loser. (Recall also that the majority outcomes from the topset model and the approval voting order match automatically.) In two elections, the size-independent model allows for the possibility of a majority cycle. In five data sets where the solution space of the size-independent model includes more than one possible majority relation, this solution space could potentially be reduced to a single majority relation by adding distributional assumptions to the model and then estimating a single set of best fitting parameters for the underlying probability distribution over rankings.

A few comments are in order before we reach the Conclusion section. Method 1, which always yields a unique solution to the majority social welfare relation (matching the approval voting order), is based on a statistically untestable model of preferences and subset choice behavior (it has been shown that the topset voting model for approval voting is empirically vacuous, i.e., irrefutable by empirical data). Therefore, it gives us a positive answer in all 12 data sets. In contrast, the size-independent model is empirically testable and falsifiable, and at the same time it does not necessarily lead to a unique majority social welfare order. Furthermore, it need not yield a majority ordering that matches the ranking by approval voting scores. As already mentioned, the size-independent model is violated by three of the twelve data sets and it allows for a majority cycle in two of the remaining nine data sets.

These facts show the counterintuitive fact that more general models (from a statistical testing point of view) can be less ambiguous than more specific models, when it comes to the majority rule (or other aggregate) preference relations that they imply.

4. Conclusions

The primary motivation of this work is to show that conventional majority rule is 'counterfactual' in many empirical settings: the empirical data often do not provide the full ranking or full paired comparison information that is required for the traditional definition of majority rule to apply. In other words, majority rule can usually only be calculated by making an inference from the available data. Yet many discussions about the virtues and faults of any given voting method center around the Condorcet efficiency, namely a given voting method's ability to satisfy the Condorcet criterion according to which a majority winner ought to be elected whenever one exists. It is thus, in general, not at all straightforward to check a voting outcome against the normative benchmark imposed by the Condorcet criterion. We may address this problem by defining majority rule at an abstract level that places all rating, ranking, and choice paradigms on an equal footing. We use a general definition of majority rule for profiles of arbitrary binary relations, probability measures over arbitrary binary preference relations, real valued utility functions, and real valued utility random variables provided in the theoretical companion paper.

We emphasize that any computation of majority rule for data other than full paired comparisons or rankings implicitly or explicitly involves a model of the underlying preferences or utilities that are assumed to generate the data.[5] For instance, there is no single obvious definition of majority rule for approval voting because no paired comparisons or full rankings are provided in the ballots. However, given almost any model of approval voting in terms of binary preferences or real valued utilities, the present framework allows one to define and derive majority preferences in a natural fashion. We illustrate with 12 different empirical data sets how two different but natural ways of computing social welfare functions from approval voting ballots may yield different majority outcomes. Similarly, we conclude from the analyses of four NES data sets that using different threshold values for thermometer data to create semiorders can affect the majority ordering we find. This demonstrates that even the most basic and subtle changes in modeling approaches can affect the outcome of any analysis of voting or ballot data against the Condorcet criterion.

[5] Of course, it is conceivable that even elicited paired comparisons do not directly reflect binary preferences. Thus one may want to consider a model even for paired comparison data, if one wants to define majority preferences in terms of voter preferences, rather than elicited paired comparisons.

It is thus essential that one spell out the preference or utility representation underlying any given method of computing majority preference relations in situations where full ranking or full paired comparison data are unavailable. Similarly, it is essential, when we analyze empirical data, that we conduct a sensitivity analysis as to how robust our conclusions are under variations or violations of our modeling assumptions, in order to guard against the results being an artifact of a particular method of computation.

In addition our other findings from the NES and approval voting data are as follows: For both NES and approval voting data sets we find little evidence for majority cycles (with two possible exceptions, where the answer given by the size-independent model is ambiguous). Thus, consistent with Tangiane (1991), we conclude from our empirical applications that majority cycles are much less frequent *in practice* than normative theory suggests they could be *in principle*.

There is no mathematical feature of our analysis that would force social welfare orders to be unchanged with changing values of the preference threshold underlying a semiorder representation. Nonetheless, we find in the 1992 and 1968 data sets that the majority rule social welfare order is identical for all values of the threshold. In such a case, we can argue that the majority order is robust under variations and/or violations of our model regarding latent preferences or utilities. The main additional conclusion that we draw from the approval voting data analyses is that there is a considerable degree of agreement between the two illustrated methods for the nine data sets where both apply. This agreement between methods is not an artifact of our theory, but rather it is an empirical finding for various of these nine elections.[6]

These results, using our general concept of majority rule, open the stage for the multifaceted investigation of social choice in an inference framework:

1. Our results lead to the general statistical question—for rating, ranking or choice paradigms—of how robust majority social welfare orders (or other social welfare functions) will turn out to be in practice under violations and/or variations of the models used to explain the data. Given the available computing power, this question can easily be addressed from situation to situation by running computer simulations that allow the researcher to explore the statistical features of the estimated social welfare order given a particular set of empirical data.
2. We have treated the data in a literal fashion here: We have identified the observed thermometer scores and their relative frequencies with the sample space of the utility random variables and its probability measure. An obvious extension would be to treat the thermometer scores as the outcome of a random sample from a theoretical population and to be interested in the social welfare functions applied to that theoretical population. In some situations, it is possible to obtain analytical results about the likelihood of a voting cycle or the likelihood of a correct or incorrect Condorcet winner when drawing a random sample from a specified population of reference (DeMeyer and Plott, 1970; Regenwetter et al., 2002a; Saari and Tataru,

[6] Recall that our motivating example provides a thought experiment where the methods do not yield identical outcomes.

1999) or making an inference about a population based on a sample of data (Tsetlin and Regenwetter, 2002).

3. While our analysis here was in the domain of 'distribution free' random utility or ranking models, a related task is to classify existing parametric families of random utility and ranking models (Fishburn, 1998; Critchlow et al., 1991) according to whether or not they (or random samples generated from them) automatically yield a majority rule social welfare order that coincides with the approval voting outcome (or the outcome of another voting method).

Our focus in this paper has been rather different from the ongoing interesting and important work on (mathematical) social welfare theory (Arrow et al., 1997; Ben-Ashar and Paroush, 2000; Pattanaik and Peleg, 1986; Saari, 1995, 1998, 1999; Sen, 1999; Tangian, 2000; Young, 1988). Here we blend traditional social choice issues with an approach that allows us to 'make sense of' actual choice data. This work also serves as a wake-up call to integrate issues of preference and utility theory as well as statistical issues into the study of social choice processes.

Acknowledgements

Most of this work was carried out while Regenwetter was an Assistant Professor in the Decision Sciences group at the Fuqua School of Business at Duke University. We thank the Fuqua School of Business for its valuable support. We thank the National Science Foundation for funding this collaborative research through NSF grants SBR 97-30076 to Regenwetter and SBR 97-30578 to Grofman and Marley. Marley was a fellow at the Hanse-Wissenschaftskolleg, Germany, during the paper's completion. We are indebted to the Interuniversity Consortium for Political and Social Research (ICPSR) for access to the 1968, 1980, 1992, and 1996 U.S. National Election Study (NES) data, to Prof. N. Tideman for access to the data sets labeled A16, A25, A29, A30, A72, and several professional organizations for access to their approval voting data. We thank Mark Berger for helping us with NES data extraction, Ilia Tsetlin for helping us tally the new approval voting data analyzed here, Saša Pekeč, the Editor and two anonymous referees for helpful comments on earlier versions.

Appendix A

Topset Model: Writing $P(X)$ for the probability that the set X is chosen, Pr for the probability measure on the set $\mathcal{B}$ of binary relations under consideration, and $\text{top}(B) = \{c \in \mathcal{C} | \forall d \in \mathcal{C}, (dBc) \Rightarrow (cBd)\}$, the topset model of approval voting states that

$$P(X) = \Pr(\{B \in \mathcal{B} | X = \text{top}(B)\}).$$

The original paper also provides a random utility formulation of the topset model.

Size-Independent Model: Writing $P(Y)$ for the probability that the set Y is chosen, $\mathbb{P}(\mathbf{S} = |Y|)$ for the probability that a set of size $|Y|$ is chosen, and letting $(\mathbf{U}_c)_{c\in\mathscr{C}}$ be a family of jointly distributed random variables, the size-independent random utility model of approval voting states that

$$P(Y) = \mathbb{P}(\mathbf{S} = |Y|)\mathbb{P}\left(\min_{c\in Y} \mathbf{U}_c > \max_{d\in\mathscr{C}-Y} \mathbf{U}_d\right).$$

The original paper also provides a formulation of the size-independent model in terms of probability distributions over linear order preferences.

References

Adams, J., 1997. Condorcet efficiency and the behavioral model of the vote. Journal of Politics 59, 1252–1263.

Arrow, K., Sen, A., Suzumura, K. (Eds.), 1997. Social Choice Reexamined, Vols. 1 & 2. St. Martin's Press, New York.

Ben-Ashar, R., Paroush, J., 2000. A nonasymptotic Condorcet jury theorem. Social Choice and Welfare 17, 189–199.

Brams, S.J., Fishburn, P.C., 1992. Approval voting in scientific and engineering societies. Group Decision and Negotiation 1, 41–55.

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.I., 1988. Rejoinder to Saari and van Newenhizen. Public Choice 59, 149.

Brams, S.J., Fishburn, P.C., Merrill, S.I., 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.

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

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

Doignon, J.-P., Regenwetter, M., 1997. An approval-voting polytope for linear orders. Journal of Mathematical Psychology 41, 171–188.

Falmagne, J.-C., Regenwetter, M., 1996. Random utility models for approval voting. Journal of Mathematical Psychology 40, 152–159.

Felsenthal, D., Machover, M., 1995. Who ought to be elected and who is actually elected—an empirical investigation of 92 elections under 3 procedures. Electoral studies 14, 143–169.

Felsenthal, D.S., Maoz, Z., Rapoport, A., 1990. The Condorcet-efficiency of sophisticated voting under the plurality and approval procedures. Behavioral Science 35, 24–33.

Felsenthal, D.S., Maoz, Z., Rapoport, A., 1993. An empirical evaluation of 6 voting procedures—do they really make any difference. British Journal of Political Science 23, 1–27.

Fishburn, P.C., 1998. Stochastic utility. In: Barberá, S., Hammond, P.J., Seidl, C. (Eds.), Handbook of Utility Theory. Kluwer, Dordrecht, pp. 273–318.

Gehrlein, W.V., 1987. The impact of social homogeneity on the Condorcet efficiency of weighted scoring rules. Social Science Research 16, 361–369.

Gehrlein, W.V., 1992. Condorcet efficiency of simple voting rules for large electorates. Economics Letters 40, 61–66.

Gehrlein, W.V., 1998. The probability of a Condorcet winner with a small number of voters. Economics Letters 59, 317–321.

Gehrlein, W.V., 1998. The sensitivity of weight selection on the Condorcet efficiency of weighted scoring rules. Social Choice and Welfare 15, 351–358.

Gehrlein, W.V., Lepelley, D., 1999. Condorcet efficiencies under the maximal culture condition. Social Choice and Welfare 16, 471–490.

Gehrlein, W.V., Valognes, F., 2001. Condorcet efficiency: a preference for indifference. Social Choice and Welfare 18, 193–205.

Gehrlein, W.V., Berg, S., 1992. The effect of social homogeneity on coincidence probabilities for pairwise proportional lottery and simple majority rules. Social Choice and Welfare 9, 361–372.

Gehrlein, W.V., Lepelley, D., 1998. The Condorcet efficiency of approval voting and the probability of electing the Condorcet loser. Journal of Mathematical Economics 29, 271–283.

Gehrlein, W.V., Lepelley, D., 2001. The Condorcet efficiency of Borda rule with anonymous voters. Mathematical Social Sciences 41, 39–50.

Lepelley, D., 1993. Condorcet's paradox. Theory and Decision 15, 161–197.

Lepelley, D., Gehrlein, W.V., 2000. Strong Condorcet efficiency of scoring rules. Economics Letters 58, 157–164.

Lepelley, D., Pierron, P., Valognes, F., 2000. Scoring rules, Condorcet efficiency and social homogeneity. Theory and Decision 49, 175–196.

Lepelley, D., Valognes, F., 1999. On the Kim and Roush voting procedure. Group Decision and Negotiation 8, 109–123.

Niederée, R., Heyer, D., 1997. Generalized random utility models and the representational theory of measurement: a conceptual link. In: Marley, A.A.J. (Ed.), Choice, Decision and Measurement: Essays in Honor of R. Duncan Luce. Lawrence Erlbaum, Mahwah, NJ, pp. 155–189.

Pattanaik, P.K., Peleg, B., 1986. Distribution of power under stochastic social choice rules. Econometrica 54, 909–921.

Regenwetter, M., 1997. Probabilistic preferences and topset voting. Mathematical Social Sciences 34, 91–105.

Regenwetter, M., Adams, J., Grofman, B., 2002. On the (sample) Condorcet efficiency of majority rule: an alternative view of majority cycles and social homogeneity. Conditionally accepted pending minor revision (Theory and Decision).

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., 1998. Choosing subsets: a size-independent probabilistic model and the quest for a social welfare ordering. Social Choice and Welfare 15, 423–443.

Regenwetter, M., Marley, A.A.J., Grofman, B., 2002b. A general concept of majority rule. Mathematical Social Sciences. 43/3, 407–430.

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

Saari, D.G., 1994. Geometry of Voting. Springer, New York.

Saari, D.G., 1995. Basic Geometry of Voting. Springer Verlag, Berlin, New York.

Saari, D.G., 1998. Connecting and resolving Sen's and Arrow's theorems. Social Choice and Welfare 15, 239–261.

Saari, D.G., 1999. Explaining all three-alternative voting outcomes. Journal of Economic Theory 87, 313–355.

Saari, D.G., Tataru, M., 1999. The likelihood of dubious election outcomes. Economic Theory 13, 345–363.

Saari, D.G., Van Newenhizen, J., 1988. Is approval voting an 'unmitigated evil'?: A response to Brams, Fishburn, and Merrill. Public Choice 59, 133–147.

Saari, D.G., Van Newenhizen, J., 1988. The problem of indeterminacy in approval, multiple, and truncated voting systems. Public Choice 59, 101–120.

Sapiro, V., Rosenstone, S., Miller, W., 1998. American National Election Studies, 1948–1997. Inter-university Consortium for Political and Social Research, Ann Arbor, MI.

Sen, A., 1999. The possibility of social choice. American Economic Review, 349–379.

Tabarrok, A., 2001. President Perot or fundamentals of voting theory illustrated with the 1992 election. Public Choice 106, 275–297.

Tangian, A.S., 2000. Unlikelihood of Condorcet's paradox in a large society. Social Choice and Welfare 17, 337–365.

Tangiane, A.S., 1991. Aggregation and Representation of Preferences: Introduction To Mathematical Theory of Democracy. Springer-Verlag, Berlin.

Tataru, M., Merlin, V., 1997. On the relationship of the Condorcet winner and positional voting rules. Mathematical Social Sciences 34, 81–90.

Tsetlin, I., Regenwetter, M., 2002. On the probability of correct or incorrect majority preference relations. Social Choice and Welfare. Forthcoming.

Wiseman, J., 2000. Approval voting in subset elections. Economic Theory 15, 477–483.
Young, H.P., 1986. Optimal ranking and choice from pairwise comparisons. In: Grofman, B., Owen, G. (Eds.), Information Pooling and Group Decision Making: Proceedings of the Second University of California, Irvine, Conference in Political Economy. JAI Press, Greenwich, pp. 113–122.
Young, H.P., 1988. Condorcet's theory of voting. American Political Science Review 82, 1231–1243.