Choosing subsets: a size-independent probabilistic model and the quest for a social welfare ordering

Michel Regenwetter[1], Bernard Grofman[2]

[1] Decision Sciences Department, The Fuqua School of Business,
Duke University, Durham, NC 27708-0120, USA
(e-mail: regenwet@mail.duke.edu;http://www.fuqua.duke.edu/faculty/descs.htn)
[2] Department of Politics and Society, University of California at Irvine, Irvine,
CA 92697-5100, USA

Abstract. "Subset voting" denotes a choice situation where one fixed set of choice alternatives (candidates, products) is offered to a group of decision makers, each of whom is requested to pick a subset containing any number of alternatives. In the context of subset voting we merge three choice paradigms, "approval voting" from political science, the "weak utility model" from mathematical psychology, and "social welfare orderings" from social choice theory. We use a probabilistic choice model proposed by Falmagne and Regenwetter (1996) built upon the notion that each voter has a personal ranking of the alternatives and chooses a subset at the top of the ranking. Using an extension of Sen's (1966) theorem about value restriction, we provide necessary and sufficient conditions for this empirically testable choice model to yield a social welfare ordering. Furthermore, we develop a method to compute Borda scores and Condorcet winners from subset choice probabilities. The technique is illustrated on an election of the Mathematical Association of America (Brams, 1988).

1. Introduction

In an election there are many ways of tallying choices so as to determine one or more election winners (Levin and Nalebuff, 1995; Lijphart and Grofman, 1984). Among them, approval voting (AV) is particularly interesting, especially in political science. Under AV each candidate that a voter approves of scores a point[1]. If there are k alternatives to be elected then the k candidates with the highest AV scores are the winners.

---

[1] We use the words "alternative" and "candidate" interchangeably, and the same holds for the terms "subject", "agent" and "voter".

Approval voting is a particular example of a choice paradigm that we call *subset voting* (SV) or sometimes *subset choices*. By subset choices we mean a voting, polling, or other choice situation, in which the participants are presented with a finite set $\mathscr{C}$ of alternatives, candidates, products, or brands, and they are asked to choose a subset of any size containing those alternatives that they "approve" of, i.e. regard as satisfactory. Although variants of AV have been independently invented by a number of authors this method is most closely associated with S. Brams and P. Fishburn, who have extensively explored its properties.

Approval voting has been used e.g. by the American Statistical Association, the Institute of Electrical and Electronics Engineers, the Mathematical Association of America, the National Academy of Sciences, the Institute of Management Science (now the Institute for Operations Research and Management Science), the Society for Social Choice and Welfare, and the United Nations. There is an extensive literature about AV (e.g. Baigent and Xu, 1991; Brams, 1988, 1990; Brams and Fishburn, 1983, 1985, 1988, 1992; Brams, Fishburn, and Merrill, 1988a, 1988b; Brams and Nagel, 1991; Carter, 1990; Cox, 1985; De Maio and Muzzio, 1986; Felsenthal, Maoz, and Rapoport, 1986, 1990; Fishburn and Little, 1988; Merrill, 1988; Merrill and Nagel, 1987; Niemi, 1984; Nurmi, 1987; Rapoport and Felsenthal, 1990; Saari, 1994; Saari and Van Newenhizen, 1988a, 1988b; Weber, 1995).

The main purpose of the present paper is to 1) use a simple empirically testable descriptive model of subset choice behavior, 2) formulate social welfare concepts in a model based framework, and 3) show how one can estimate the optimal social choice (according to the Borda and/or Condorcet criteria) from the SV data by using the psychological model as a measurement device. Put differently, within the subset voting context, we integrate axiomatic concepts from social choice theory in the Arrowian tradition (Arrow, 1951; Heiner and Pattanaik, 1983; Murakami, 1968; Pattanaik, 1971; Sen, 1966, 1969, 1970) with descriptive probabilistic models of judgment and decision making as they are encountered in mathematical psychology (e.g. Luce and Suppes, 1965). While there exists a vast literature on computer simulations and analytical results in social choice theory, the present paper emphasizes how the social choice theoretic performance of SV can be determined *empirically* for *real* elections.

The paper proceeds as follows. First, we introduce a simple (probabilistic) psychological model of subset choice behavior, originally proposed by Falmagne and Regenwetter (1996). This *size-independent model* assumes that each agent has a linear preference order on the alternatives and approves of a subset at the top of that ranking. After summarizing the relevant known properties of the model (c.f. Doignon and Regenwetter, 1977, Regenwetter, Marley, and Joe, 1996) we translate some key concepts of social choice theory, such as *transitive majorities*, *Condorcet* (*majority*) *winner*, *Borda scores* and *Borda winner* as well as Sen's (1966) *value restriction* into probabilistic terms appropriate for the choice model.

For instance, Sen's (1966) value restriction is a sufficient condition for social choice to be transitive and is restated here in probabilistic terms. In the case of three candidates, the choice model yields a unique probability distribution on the latent preference rankings underlying the choices if and only if our probabilistic version of Sen's value restriction holds.

However, we may have transitivity also when more than one solution exists for the distribution of preferences underlying the size-independent model, even though Sen's value restriction would not hold. In order to account for possible transitivities in settings where value restriction is violated, we provide new results that generalize Sen's theorem in terms of *net preference probabilities* (for related concepts, see Feld and Grofman, 1986; Gaertner and Heinecke, 1978). This generalization involves two observable properties that are together necessary and sufficient for the existence of a social welfare ordering for a given probability distribution on rankings. The two conditions are called *net value restriction* and *net preference majority*, respectively, and are weaker than Sen's value restriction condition.

When the size-independent model is satisfied by a set of subset choice data, then we can use the latent probability distributions on rankings that are compatible with those data to check whether a transitive social welfare order exists. The latter exists if either of the two conditions above is satisfied. Moreover, even when no social welfare ordering exists, the size-independent model allows us to check for a *Condorcet* winner. Furthermore, we offer a model based analytical solution to compute *Borda scores* directly from the subset choice probabilities. The estimated social welfare order and the estimated ordering according to Borda scores can both be compared with the ordering induced by the AV scores. This allows us to evaluate the performance of AV as a social choice mechanism.

We illustrate our technique on a data set[2] from an election of the Mathematical Association of America (see Brams, 1988; Brams and Fishburn, 1992). On this data set we show that the conditions for the size-independent model hold, we check for the existence of a social welfare order and/ or Condorcet candidates, and we compute the Borda scores. The estimated social welfare order is not unique, while the ordering according to Borda scores turns out to match the ordering derived from the AV scores. An empirical companion paper (Regenwetter and Grofman, in press) systematically and successfully applies the tools of the present paper to several election data sets from various sources.

2. A size-independent model of subset voting

Throughout the paper, a *weak order* is a transitive, connected binary relation, a *strict weak order* is a negatively transitive, asymmetric binary relation and a

---

[2] This was kindly provided by Prof. S. J. Brams.

(*strict*) *linear order* is a transitive, asymmetric and weakly connected binary relation (c.f.Roberts, 1979).

Notation. Let Π denote the set of all linear orders on $\mathscr{C} = \{1, 2, \ldots, N\}$ with $N \geq 3$. In some cases, N is explicitly assumed to be equal to three. We denote by $X \subseteq \mathscr{C}$ any subset of $\mathscr{C}$, and for any subset X of $\mathscr{C}$, we write Π_X for the collection of all linear orders in Π which rank the members of X ahead of all other members of $\mathscr{C}$. A glossary at the end of the paper gives an overview of the notation.

The *size-independent model of subset voting* of Falmagne and Regenwetter (1996)[3] assumes that the choice behavior of a voter in the population takes the form of the joint realization of three random variables: **V** (as in vote), **R** (as in ranking) and **S** (as in set size). Accordingly we write $\mathbf{V} = X$; $\mathbf{R} = \pi$; and $\mathbf{S} = s$, respectively, for the events that a randomly drawn voter approves of subset $X \subseteq \mathscr{C}$; has latent linear preference order $\pi \in \Pi$; or approves of $s \in \{0, 1, \ldots, N\}$ many elements, respectively.

Because the subset chosen by the voter and the number of candidates contained in their vote are observable, the realizations of the random variables **V** and **S** are observable. Since the realizations of **R** are, however, unobservable, this is a latent variable, the distribution of which we would like to at least partially characterize using the observed votes. The **size-independent model** assumes that

$$\mathbb{P}(\mathbf{V} = X) = \mathbb{P}(\mathbf{S} = |X|) \times \mathbb{P}(\mathbf{R} \in \Pi_X).$$

The probability that a randomly picked voter chooses the set X is the product of the probability that s/he votes for as many alternatives as X contains times the probability that s/he likes all candidates in X better than all others.[4]

Although the product form suggests a certain degree of independence between the latent preferences and the number of objects chosen, Regenwetter et al. (1996) show that the model can take into account some directionality biases, say, in a case with a single candidate from one party, and two candidates from an opposing party. They also compare the model with competing probabilistic models of subset choice.

In statistical terms this model has more parameters (namely $N! + N - 1$) than the choice probabilities have degrees of freedom (namely $2^N - 1$). Although this might suggest that the model is empirically vacuous, Falmagne and Regenwetter (1996) give examples of empirical constraints implied by the model. Doignon and Regenwetter (1997) show that a simple transformation of the model-compatible choice probabilities produces a $2^N - N - 1$ di-

---

[3] In Falmagne and Regenwetter (1996) subset voting is called approval voting. Here we have chosen the more general term subset voting to abstract from the particular tallying procedure.

[4] Whenever we view the vote participants as being the entire sample space (endowed with the trivial relative frequency probability measure), the above statements are equivalent to relative frequency statements.

mensional convex polytope in $\mathbb{R}^{2^N}$. They have characterized the resulting *approval-voting polytope for linear orders* for $N \leq 5$, whereas the problem remains unsolved for larger N. The so-called facet-defining inequalities that they derive (for $N \leq 4$) all have interpretations as constraints on the probabilities over the observable subset choices.

For the case of three alternatives, Doignon and Regenwetter (1997) also derive closed form solutions for the possible distributions of $\mathbf{R}$ underlying $\mathbf{V}$ when the size-independent model holds. Writing $\mathbf{C}(X) = \frac{\mathbb{P}(\mathbf{V}=X)}{\mathbb{P}(\mathbf{S}=|X|)}$, when it is well defined, $C = \mathbf{C}(\{c\}), DE = \mathbf{C}(\{d,e\})$ etc., and *cde* for the ranking (over $\{c,d,e\}$) in which c is best and e is worst, the following theorem summarizes some crucial results from Doignon and Regenwetter (1997). Note that the ratios $\mathbf{C}(X)$ can be directly estimated from the data.

Theorem 1. (Doignon and Regenwetter, 1997) *Suppose that $\mathscr{C} = \{c,d,e\}$.*

1) *Then the size-independent model is satisfied if and only if for each element $x \in \mathscr{C}$ and for each $k \in \{1,2,3\}$ the quantity*

$$\mathbb{P}_k(x) = \sum_{\substack{Y \subseteq c-\{x\} \\ |Y|=k-1}} \mathbf{C}(\{x\} \cup Y) - \sum_{\substack{Z \subseteq c-\{x\} \\ |Z|=k-2}} \mathbf{C}(\{x\} \cup Z) \tag{1}$$

   is nonnegative (where the second sum vanishes when $k=1$).
2) *The quantity $\mathbb{P}_k(x)$ is the marginal ranking probability that $\mathbf{R}$ ranks alternative x at position k (this holds also when $|\mathscr{C}| > 3$).*
3) *Given that $\mathbb{P}_k(x) \geq 0$ for all x and k, the possible probability distributions on the latent rankings can be computed in closed form: Writing $\mathbb{P}_\pi$ for $\mathbb{P}(\mathbf{R} = \pi)$,*

$$\begin{pmatrix} \mathbb{P}_{cde} \\ \mathbb{P}_{dec} \\ \mathbb{P}_{ecd} \\ \mathbb{P}_{ced} \\ \mathbb{P}_{dce} \\ \mathbb{P}_{edc} \end{pmatrix} = \begin{pmatrix} C + CD + DE \\ C + D + DE \\ 1 \\ -CD - DE \\ -C - DE \\ -C - D \end{pmatrix} + \lambda \begin{pmatrix} -1 \\ -1 \\ -1 \\ +1 \\ +1 \\ +1 \end{pmatrix} \tag{2}$$

   where $\lambda \in \left[\max(C+D, CD+DE, C+DE), \min(C+CD+DE, C+D+DE, 1) \right]$.
4) *The distribution of $\mathbf{R}$ is unique if and only if at least one $\mathbb{P}_k(x)$ is equal to zero.*

Note that the range of λ is not a statistical confidence interval but rather a continuum of possible closed form solutions. As Falmagne and Regenwetter (1996) show, their model naturally belongs to the class of (nonparametric) random utility models of discrete choice (Anderson, de Palma, and Thisse, 1992; Barberá and Pattanaik, 1986; Falmagne, 1978; Fishburn, 1992; Heyer and Niederée, 1992; Koppen, 1995; Marley, 1990; Marschak, 1960; McFadden and Richter, 1970; Suck, 1992). Before applying Theorem 1, we introduce further concepts.

3. Linking social choice concepts to subset voting through a probabilistic model of preferences

3.1. Stochastic transitivity

The social choice literature says that a vote is transitive if the following property holds. If candidate c has a majority over d and d has a majority over e, then c has a majority over e. Unless we explicitly ask the voters to perform paired comparisons, it is not quite clear what this should mean in general: No binary choices are being observed in subset voting. As we shall see, the framework of a probabilistic choice model can bridge the gap between the observable choices and the theoretical concept of transitive majorities through the intermediary of (latent) preference rankings underlying the choices.

Why do we want transitive votes? Typically we like to construct an aggregate preference relation of the voters, ideally a strict linear order or a weak order: a *social welfare ordering* (Arrow, 1951; Heiner and Pattanaik, 1983; Pattanaik, 1971; Sen, 1966, 1969, 1970). Transitivity of votes is closely related to what the psychological choice literature calls "weak stochastic transitivity" and to the well known "weak utility model" (Luce and Suppes, 1965). This model assumes that in a binary choice paradigm each paired comparison is a Bernoulli trial with a well defined probability of a choice for either alternative. The following definition is from Luce and Suppes (1965).

Definition 1. A weak utility model *is a set of binary choice probabilities for which there exists a real-valued function w over $\mathscr{C}$ such that*

$$p_{cd} \geq \tfrac{1}{2} \Leftrightarrow w(c) \geq w(d),$$

where we write p_{cd} for the probability of a choice of c when c versus d are being offered.

The weak utility model is equivalent[5] to weak stochastic transitivity of the binary choice probabilities, which we define next.

Definition 2. Weak stochastic transitivity of binary choice probabilities *means that*

$$p_{cd} \geq \tfrac{1}{2} \quad \& \quad p_{de} \geq \tfrac{1}{2} \quad \implies \quad p_{ce} \geq \tfrac{1}{2}.$$

For a given probability distribution $\pi \mapsto \mathbb{P}(\pi)$ over π, we write $\mathbb{P}_{cd} = \sum_{(c,d)\in\pi} \mathbb{P}(\pi)$ for the marginal pairwise ranking probability of c to be ranked ahead of d. Note that there exists a substantial literature trying to explain probabilities of (observable) binary choice by probabilities of (latent and unobserved) rankings through

$$p_{cd} = \mathbb{P}_{cd}. \tag{3}$$

---

[5] This uses the assumption that $\mathscr{C}$ is finite.

Given a set of binary choice probabilities, the existence of probabilities on rankings satisfying (3) is not at all a trivial matter and is still the object of intense research (Block and Marschak, 1960; Campello de Souza, 1983; Cohen and Falmagne, 1990; Dridi, 1980; Fishburn, 1990, 1992; Fishburn and Falmagne, 1989; Gilboa, 1990; Heyer and Niederée, 1989, 1992; Koppen, 1995; Marley, 1990; Marschak, 1960; McFadden and Richter, 1970; Suck, 1992).

In the probabilistic framework it is appealing and straightforward to define an aggregate preference relation through: "c is aggregately preferred to d if and only if the choice probability $p_{cd} \geq \frac{1}{2}$ in the Bernoulli trial". Such a preference relation is transitive if and only if weak stochastic transitivity holds. Thus, for probabilistic binary choice, the existence of a transitive social welfare order, weak stochastic transitivity and the weak utility model are equivalent. The function w in Definition 1 can thus be labeled a *social welfare function*.

Probabilistic subset choice models offer an elegant way of defining transitivity also for subset voting. We call a subset vote (induced by linear orders) transitive if the underlying (theoretical) probability distribution on the strict linear orders satisfies weak stochastic transitivity (for rankings), defined as follows.

Definition 3. *A probability distribution $\mathbb{P}$ on Π satisfies* **weak stochastic transitivity** *(for rankings) iff the induced marginal (pairwise) ranking probabilities satisfy*

$$\mathbb{P}_{cd} \geq \tfrac{1}{2} \quad \& \quad \mathbb{P}_{de} \geq \tfrac{1}{2} \quad \implies \quad \mathbb{P}_{ce} \geq \tfrac{1}{2}.$$

Suppose for a moment that, for a given vote, the size-independent model holds, and that we have reconstructed a probability distribution on the linear orders from the data using the model. If this probability distribution is stochastically transitive, then we can derive a social welfare ordering from it, and call that the *social welfare order of the subset vote*.

Remark. It should be emphasized that the weak utility model, i.e. the concept of weak stochastic transitivity is only ordinal, and thus that the social welfare function w is unique only up to arbitrary monotonic transformations (Luce and Suppes, 1965). We might like to think that the higher the margin of c over d, the more c and d are different. To be meaningful, this statement requires a stronger model such as the "strong utility" model (or "Fechnerian model" in psychophysics, e.g. Falmagne, 1985; Luce and Suppes, 1965). Here the probability of choosing c over d is a monotonic function of the difference in utility between c and d (alternatives with equal utility value are discriminated at random, with equal probability). Such a representation implies strong stochastic transitivity, i.e.

$$\mathbb{P}_{cd} \geq \tfrac{1}{2} \quad \& \quad \mathbb{P}_{de} \geq \tfrac{1}{2} \quad \implies \quad \mathbb{P}_{ce} \geq \max(\mathbb{P}_{cd}, \mathbb{P}_{de}),$$

which is however not sufficient. The Fechnerian model is unique only up to linear transformations, thus ratios of utilities have no meaningful interpr-

etation. We do not consider models as strong as the Fechnerian model or the strict utility model (also known as the Bradley-Terry-Luce model) here. The interested reader is referred to Critchlow, Fligner, and Verducci (1991), Falmagne (1985) or Luce and Suppes (1965).

3.2. Condorcet winners

A second important normative concept in the social choice literature is that of the Condorcet winner, especially when there exists no social welfare order. A Condorcet candidate (also known as a majority winner) is usually defined as the candidate(s) (if they exist) who would get a majority against all other candidates if they were to compete pairwise (Black, 1958; Condorcet, 1785; Felsenthal et al., 1990; Young, 1986, 1988). The Condorcet winner is the most commonly accepted normative criterion for a legitimate winner of a social choice procedure to select a single alternative.

   Again, absent data on binary comparisons, we define a Condorcet winner in terms of (latent and unobserved) probabilistic rankings.

Definition 4. *Given a probability* $\mathbb{P}$ *on* Π, *candidate* $c \in \mathscr{C}$ *is* **a Condorcet winner** *iff*

$$\mathbb{P}_{cd} \geq \tfrac{1}{2} \quad \forall \, d \in \mathscr{C} - \{c\}.$$

This concept of a Condorcet candidate is compatible with the idea that, if the voters were indeed asked to do a paired comparison instead of a subset choice, they would choose the alternative that is ranked ahead of the other in the corresponding sampled preference order.

3.3. Value restriction

Sen's *Value Restriction* is a sufficient condition for the existence of a transitive social welfare ordering. It actually consists of 3 conditions on triples of alternatives, often referred to as NW ("never worst"), NM ("never middle"), and NB ("never best"). Assuming that each voter has a strict linear preference order, a triple of alternatives satisfies NW iff there is one alternative among the three that is never ranked worst. NM and NB are defined analogously. Sen's value restriction states that all triples satisfy either NW, NM, or NB. The underlying intuition is that in each triple of candidates there should be at least one about which all voters agree that s/he is not the worst (middle, best)[6].

   We translate these conditions into probabilistic terms.

---

[6] Feld and Grofman (1988, 1990, 1992) discuss some evidence of the degree to which candidate preference rankings in actual elections satisfy value restrictedness.

Definition 5. *Consider a probability* $\mathbb{P}$ *on* π. *For any given triple of alternatives, we say that the marginal ranking probabilities induced by* $\mathbb{P}$ *on that triple satisfy NW(c) iff the (marginal) probability for c to be ranked worst (in the triple) is zero. When NW(c) holds, c is (almost surely)* **never worst**. *NM(c) and NB(c) are defined analogously.* $\mathbb{P}$ *is (a.s.)* **value restricted** *iff in each triple* $\{x, y, z\} \subseteq \mathscr{C}$ *there exists c with either NW(c), NM(c) or NB(c).*

3.4. Borda winner

A fourth social choice concept is that of Borda scores (Black, 1958; Young, 1974, 1986, 1995). The Borda winner, i.e. the candidate with the highest Borda score, is arguably the second most accepted normative criterion for the "best" outcome of a social choice procedure to select a single alternative. We redefine the Borda score in probabilistic terms as follows.

Definition 6. *Given a probability* $\mathbb{P}$ *on* Π, *the* **Borda score B(c)** *of candidate c is*

$$B(c) = \sum_{k=1}^{N-1}(N - k) \times \mathbb{P}_k(c), \tag{4}$$

where $\mathbb{P}_k(c)$ *is the (marginal) probability that alternative c is ranked at position k. The reader may remember from Theorem 1 that* $\mathbb{P}_k(c)$ *can be computed from the subset choices when the size-independent model holds. We derive a more direct computation of* $B(c)$ *later.*

Alternatively, we may calculate Borda scores from binary preference probabilities through

$$B(c) = \sum_{d \neq c} \mathbb{P}_{cd}.$$

The two formulas are probabilistic versions of the usual definition, and their equivalence is analogous to an equivalence result proven by Young (1974) linking pairwise preferences to the more familiar point counting definition of Borda scores.

Definition 7. *The* **Borda winners** *are the alternatives with the highest Borda score.*

In sum, social welfare orders and Condorcet candidates are traditionally defined in terms of pairwise choices, which are observable only if the voters perform paired comparisons. Similarly, Borda winners are defined in terms of probabilistic preference rankings or pairwise preference probabilities. Subset voting does not provide either data on paired comparisons or preference rankings. However, we can use the subset choice data to reconstruct voter preferences through a testable choice model such as the size-independent model. We have therefore redefined social welfare orders and Condorcet

winners in terms of preference rankings (rather than choices). This reconciles axiomatic social choice theory with empirical subset choice[7].

4. A probabilistic version of Sen's theorem on value restriction

The following theorem is a variation of Sen's theorem on value restriction (Sen, 1966, 1969, 1970), generalized to probabilistic terms.

Theorem 2. *Given a probability $\mathbb{P}$ on Π consider the relations R and P by*

$$cRd \quad \Leftrightarrow \quad \mathbb{P}_{cd} \geq \mathbb{P}_{dc} \Leftrightarrow \mathbb{P}_{cd} \geq \tfrac{1}{2}, \tag{5}$$

$$cPd \quad \Leftrightarrow \quad \mathbb{P}_{cd} > \mathbb{P}_{dc} \Leftrightarrow \mathbb{P}_{cd} > \tfrac{1}{2}. \tag{6}$$

If $\mathbb{P}$ is (a.s.) value restricted, then 1) R is a weak order, 2) P is a strict weak order, and 3) if $\mathbb{P}_{cd} \neq \mathbb{P}_{dc}, \forall c \neq d$ then P is a strict linear order. Thus, (a.s.) value restriction implies transitivity.

This and Theorem 1 imply the following result for the special case of three candidates.

Theorem 3. *Suppose that $\mathscr{C} = \{c, d, e\}$ and that the size-independent model of subset voting holds. Then the following are equivalent (using the notation of Theorem 1).*

- *The distribution of **R** is unique.*
- *At least one of the marginal probabilities $\mathbb{P}_k(x)$ in (1) is zero.*
- *The probability distribution induced by **R** on the linear orders is (a.s.) value restricted.*
- $\max(C + D, CD + DE, C + DE) = \min(C + CD + DE, C + D + DE, 1).$

This gives us an observable property on the subset choices to check whether the latent preferences are (a.s.) value restricted.

Corollary 1. *Consider a subset vote with 3 candidates. If the size-independent model holds and the probability distribution is unique, then R as defined in (5) is a weak order, P in (6) is a strict weak order and, provided that $\mathbb{P}_{cd} \neq \mathbb{P}_{dc}, \forall c \neq d$, P is a strict linear order.*

It is quite unlikely to observe data with exactly one solution for the distribution of **R**. As we show in Regenwetter and Grofman (in press), none of the elections analyzed there yields (a.s) value restriction. However, the latter can be readily replaced by a set of two conditions which, together, are necessary and sufficient for transitivity. This is shown next.

---

[7] Sen's value restrictedness is already defined originally in terms preferences and not choices.

5. Net preferences and the social welfare ordering

The following two definitions are critical throughout the rest of the paper.

Definition 8. *Given a probability* $\mathbb{P}$ *on* Π, *and denoting by* π^{-1} *the reverse order of* π, *the* **net ranking probability (net preference probability)** **NP** *(induced by* $\mathbb{P}$*) is defined as*

$$NP(\pi) = \mathbb{P}(\pi) - \mathbb{P}(\pi^{-1}).$$

The **net margins (net pairwise preference probabilities)** *(see also Feld εGrofman, 1988) are defined as* $NP_{cd} = \mathbb{P}_{cd} - \mathbb{P}_{dc}$. *Net marginal ranking probabilities of triples are defined analogously. We also write* NP_{π} *for* $NP(\pi)$; *and* NP_{cde} *to denote the net marginal ranking probability that c is ranked before both d and e, and that d is ranked before e.*

Definition 9. *Given NP on* Π *as above, define* **welfare orderings** *R and P on* $\mathscr{C}$ *through*

$$cRd \Leftrightarrow NP_{cd} \geq 0, \quad cPd \Leftrightarrow NP_{cd} > 0,$$

i.e. the relations already used in Theorem 2.

Definition 10. *Given NP on* Π *as before, for any triple* $\{c, d, e\} \subseteq \mathscr{C}$,

$$\begin{aligned}
NP \text{ satisfies } NW(c) &\Leftrightarrow NP_{edc} \leq 0 \ \& \ NP_{dec} \leq 0, \\
NP \text{ satisfies } NM(c) &\Leftrightarrow NP_{ecd} \leq 0 \ \& \ NP_{dce} \leq 0, \\
NP \text{ satisfies } NB(c) &\Leftrightarrow NP_{cde} \leq 0 \ \& \ NP_{ced} \leq 0.
\end{aligned}$$

Definition 11. *NP is* **marginally value restricted** *for the triple* $\{x, y, z\} \subseteq \mathscr{C}$ *iff there exists an element* $c \in \{x, y, z\}$ *such that NP satisfies NW(c) [or NB(c) or NM(c)]. If this property is satisfied, then* **marginal net value restriction** *holds on the triple* $\{x, y, z\}$. **Net value restriction** *holds on* $\mathscr{C}$ *if marginal net value restriction holds on each triple.*

Remarks. If *NP* on Π satisfies NW(c) for a triple $\{c, d, e\} \subseteq \mathscr{C}$, then

$$\begin{aligned}
NP_{ecd} \leq 0 &\Rightarrow NP \text{ satisfies } NB(e), \\
NP_{ecd} \geq 0 &\Rightarrow NP \text{ satisfies } NB(d).
\end{aligned}$$

Similarly, NB(c) implies either NW(d) or NW(e). Also, NM(c) means that $NP_{ecd} = 0$ thus that at most two rankings have strictly positive *NP* values, and that NW(d) or NW(e) hold. At most three elements in $\{cde, ced, dce, dec, ecd, edc\}$ have strictly positive net preference probabilities. Furthermore, net value restriction is weaker than value restriction:

- $\mathbb{P}$ satisfies NW(c) $\Rightarrow$ *NP* satisfies NW(c), but not conversely,
- $\mathbb{P}$ satisfies NB(c) $\Rightarrow$ *NP* satisfies NB(c), but not conversely,
- $\mathbb{P}$ satisfies NM(c) $\Rightarrow$ *NP* satisfies NW(c), but not conversely.

A further definition is required before we can state our key theorem.

Definition 12. *Given NP on* Π *as before,* $\pi \in \Pi$ *has a* **net preference majority** *iff*

$$NP(\pi) > \sum_{\substack{\pi' \in \Pi - \{\pi\}, \\ NP(\pi') > 0}} NP(\pi').$$

Similarly, for any triple $\{c, d, e\} \subseteq \mathscr{C}, cde$ *has a* **marginal net preference majority** *if*

$$NP_{cde} > \sum_{\substack{\pi' \in \{ced, dce, dec, ecd, edc\}, \\ NP_{\pi'} > 0}} NP_{\pi'}.$$

The following theorem is similar in spirit to Lemma 2 of Feld and Grofman (1986), which unfortunately omits certain knife-edge situations caused by possible ties.

Theorem 4. *R as defined in Definition 9 is transitive iff for each triple* $\{c, d, e\} \subseteq \mathscr{C}$ *at least one of the following two conditions holds:*

- *NP is marginally value restricted on* $\{c, d, e\}$ *and if at least one net preference is nonzero then it holds that* $(NP_{cde} = 0 \Rightarrow NP_{dce} \neq NP_{ced})$ *(with possible relabelings)*
- $\exists \tau \in \{cde, ced, dce, dec, ecd, edc\}$ *such that* τ *has a marginal net preference majority.*
- *Similarly, P is transitive iff for each triple* $\{c, d, e\} \subseteq \mathscr{C}$ *at least one of the following two conditions holds:*
- *NP is marginally value restricted on* $\{c, d, e\}$
- $\exists \tau \in \{cde, ced, dce, dec, ecd, edc\}$ *such that* τ *has a marginal net preference majority.*

Proof. Transitivity holds on $\mathscr{C}$ if and only if transitivity holds on each triple $\{c, d, e\}$ in $\mathscr{C}$. There is thus no loss of generality to set $\mathscr{C} = \{c, d, e\}$ and $\Pi = \{cde, ced, dce, dec, ecd, edc\}$. Recall that at most three rankings have (strictly) positive net preference probabilities.

First, suppose that none are positive, i.e. that all net ranking probabilities are zero. Then **transitivity** holds because all alternatives are tied, i.e. $P = \phi, R = \mathscr{C}^2$ and **net value restriction** holds (but there is no ranking with a net preference majority).

Second, suppose that exactly one net ranking probability NP_π is positive (i.e. four net ranking probabilities are zero). Then **transitivity** holds since $R = P = \pi$. **Net value restriction** holds, with $NP_{cde} = 0 \Rightarrow NP_{dce} \neq NP_{ced}$ (including possible relabelings), and π has a **net preference majority**.

Third, suppose that exactly two net ranking probabilities are null, w.l.o.g assume that $NP_{cde} = -NP_{edc} = 0$. Then $NM(d)$ holds, and therefore also **net value restriction**.

a) If $NP_{dce} > 0$ & $NP_{dec} > 0$ [and thus $NP_{dce} \neq NP_{ced}$] then **transitivity** follows:

$NP_{dce} > NP_{dec} \Rightarrow R = P = dce$ with a net preference majority,
$NP_{dce} = NP_{dec} \Rightarrow P = \{(d,c),(d,e)\}; R = \{(d,c),(d,e),(c,e),(e,c)\},$
$NP_{dce} < NP_{dec} \Rightarrow R = P = dec$ with a net preference majority;

b) If $NP_{dce} < 0$ & $NP_{dec} < 0$ [and thus $NP_{dce} \neq NP_{ced}$] then **transitivity** follows:

$NP_{dce} > NP_{dec} \Rightarrow R = P = ced$ with a net preference majority,
$NP_{dce} = NP_{dec} \Rightarrow P = \{(c,d),(e,d)\}; R = \{(c,d),(e,d),(c,e),(e,c)\},$
$NP_{dce} < NP_{dec} \Rightarrow R = P = ecd$ with a net preference majority;

c) If $NP_{dce} > 0$ & $NP_{dec} < 0$ then

$NP_{dce} > NP_{ced} \Rightarrow \quad R = P = dce$ with a net preference majority,
$NP_{dce} < NP_{ced} \Rightarrow \quad R = P = ced$ with a net preference majority,
$NP_{dce} = NP_{ced} \Rightarrow \begin{cases} P = \{(c,e)\}; \\ R = \{(d,c),(c,d),(d,e),(e,d),(c,e)\}; \quad (\dagger) \end{cases}$

where (†) is a violation of transitivity for R.

d) If $NP_{dce} < 0$ & $NP_{dec} > 0$ then

$NP_{ecd} > NP_{dec} \quad \Rightarrow \quad R = P = ecd$ with a net preference majority,
$NP_{ecd} < NP_{dec} \quad \Rightarrow \quad R = P = dec$ with a net preference majority,
$NP_{ecd} = NP_{dec} \quad \Rightarrow \begin{cases} P = \{(e,c)\}; \\ R = \{(d,c),(c,d),(d,e),(e,d),(e,c)\}; \quad (\dagger) \end{cases}$

where (‡) is a violation of transitivity for R.

Fourth, the only remaining possibility is that three net probabilities are positive [and the others are negative, i.e. $NP_{xyz} = 0$ cannot occur]. There are eight such cases:

$$NP_{cde} > 0 \ \& \ NP_{dce} > \ 0 \ \& \ NP_{ced} > 0, \tag{7}$$

$$NP_{cde} > 0 \ \& \ NP_{dce} > \ 0 \ \& \ NP_{dec} > 0, \tag{8}$$

$$NP_{cde} > 0 \ \& \ NP_{ecd} > \ 0 \ \& \ NP_{ced} > 0, \tag{9}$$

$$NP_{cde} > 0 \ \& \ NP_{ecd} > \ 0 \ \& \ NP_{dec} > 0, \tag{10}$$

$$NP_{edc} > 0 \ \& \ NP_{dce} > \ 0 \ \& \ NP_{ced} > 0, \tag{11}$$

$$NP_{edc} > 0 \ \& \ NP_{dce} > \ 0 \ \& \ NP_{dec} > 0, \tag{12}$$

$$NP_{edc} > 0 \ \& \ NP_{ecd} > \ 0 \ \& \ NP_{ced} > 0, \tag{13}$$

$$NP_{edc} > 0 \ \& \ NP_{ecd} > \ 0 \ \& \ NP_{dec} > 0, \tag{14}$$

The cases (7)–(9) and (12)–(14) are all equivalent through relabeling of alternatives: Starting each time from (7), the relabeling $c \leftrightarrow d$ yields (8), $d \leftrightarrow e$ yields (9), $c \to d \to e \to c$ yields (12), $c \to e \to d \to c$ yields (13), and $c \leftrightarrow e$ yields (14). Similarly, (10) is equivalent to (11) through, for instance, the relabeling $c \leftrightarrow e$. We thus need to consider only (7) and (10).

If (7) holds, then **net value restriction** holds, cPe and furthermore

$$NP_{cde} < NP_{ced} + NP_{dce} \quad \Rightarrow \quad \begin{cases} dRc \Rightarrow dPe \\ eRd \Rightarrow cPd, \end{cases}$$

$$NP_{cde} \geq NP_{ced} + NP_{dce} \quad \Rightarrow \quad R = P = cde$$

each of which implies transitivity for both R and P.

If (10) holds (and thus net value restriction is violated) we obtain transitivity if and only if one of the three rankings cde, ecd and dec has a **net preference majority**: Suppose that each of the three has a net probability strictly smaller than the sum of the other two. Then $R = P = \{(c,d),(d,e),(e,c)\}$, a violation of transitivity. Also, if one of the three has a net probability equal to the sum of the other two, say $NP_{cde} = NP_{ecd} + NP_{dec}$, then $P = \{(c,d),(d,e)\}, R = \{(c,d),(d,e),(c,e),(e,c)\}$, which both violate transitivity. $\qquad\square$

The next Theorem offers a straightforward way to calculate the Borda scores directly from the subset choice probabilities when the size-independent model holds. Note that this applies also to large N, where the ranking probabilities themselves cannot easily be computed.

Theorem 5. *Whenever the size-independent model holds the Borda scores can be estimated directly from the data, as follows*:

$$B(c) = \sum_{\substack{X \subset \mathscr{C}, \\ X \ni c}} \mathbf{C}(X) \, . \tag{15}$$

Proof. The result is immediate from the definition of the Borda score and the fact that

$$\sum_{\substack{X \subset \mathscr{C}, \\ X \ni c, \\ |X|=k}} \mathbf{C}(X) \quad = \quad \sum_{i=1}^{k} \mathbb{P}_i(c). \qquad\qquad\square$$

6. An illustration

We illustrate our method with the analysis of a data set from an election of the Mathematical Association of America (see Brams, 1988; Brams and Fishburn, 1992). The raw frequencies as well as the estimates of $\mathbf{C}(X)$ (see Theorem 1) are reported in Table 1. As Table 2 shows, the necessary and sufficient condition ($\mathbb{P}_k(c) \geq 0$), for the size-independent model holds here. The probability distribution induced by **R** on the rankings must be of the form shown in Figure 1 (the net preference probabilities are reported in parentheses). Furthermore, we derive the marginal pairwise preference probabilities and net margins, reported in Table 4, where the value in row i

Table 1. Subset choice frequencies N_X, and estimated values $\widetilde{C(X)}$ of $\mathbf{C}(X)$

X	N_X	$\widetilde{C(X)}$	X	N_X	$\widetilde{C(X)}$
$\emptyset$	224	1.000	$\{c,d,e\}$	30	1.000
$\{c\}$	1257	0.379	$\{c,d\}$	68	0.145
$\{d\}$	626	0.189	$\{c,e\}$	260	0.554
$\{e\}$	1434	0.432	$\{d,e\}$	141	0.301

Table 2. Marginal ranking probabilities $\mathbb{P}_k(x)$

	c	d	e
1	0.379	0.189	0.432
2	0.320	0.257	0.423
3	0.301	0.554	0.145

Table 3. Borda scores estimated through Theorem 5

c	d	e
1.078	0.635	1.287

Table 4. Estimated pairwise preference probabilities and net margins (in parentheses) with $\lambda \in [0, 0.145]$

	c	d	e
c	–	$0.554 + \lambda\,(0.108 + 2\lambda)$	$0.524 - \lambda\,(0.048 - 2\lambda)$
d	$0.446 - \lambda\,(-0.108 - 2\lambda)$	–	$0.189 + \lambda\,(-0.622 + 2\lambda)$

$$
\begin{array}{ccc}
 & \lambda \quad (-.257 + 2\lambda) & \\
 & \text{cde} & \\
(.335 - 2\lambda) \quad .379 - \lambda & \boxed{\text{ced} \qquad \text{dce}} & .145 - \lambda \quad (-.03 - 2\lambda) \\
(.03 + 2\lambda) \quad .175 + \lambda & \boxed{\text{ecd} \qquad \text{dec}} & .044 + \lambda \quad (-.335 + 2\lambda) \\
 & \text{edc} & \\
 & .257 - \lambda \quad (.257 - 2\lambda) &
\end{array}
$$
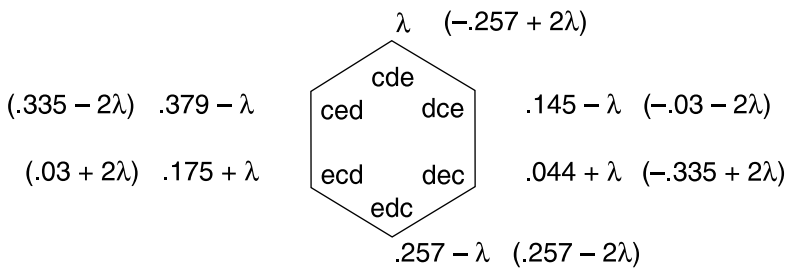
Fig. 1 Possible probability distributions on the rankings with $\lambda \in [0, 0.145]$. Net probabilities are given in parentheses.

column j is the probability of preferring alternative i to alternative j, with $\lambda \in$ [0, 0.145] (the corresponding net margins are given in parentheses).

As all numbers in Table 2 are nonzero, (a.s.) value restriction is violated. Thus, Corollary 1 can not be used here to guarantee transitivity. From Theorem 4 it follows, however, that an aggregate strict linear preference order in the sense of the weak utility model exists when $\lambda \neq 0.024$. In fact, net value restriction holds regardless of the particular value of λ. The following possible solutions are derivable: *cPePd* when $\lambda \in (0,0.024)$, which has a net preference majority for *ced* in that case, or *ePcPd* when $\lambda \in (0.024,0.145]$, which has a net preference majority for *ecd* when $\lambda \in (0.094,0.145]$. The ordering according to approval voting scores is *ecd*, i.e. the same as the social welfare order in most of the solution space. A unique Condorcet winner exists here unless $\lambda = .024$, in which case e and c tie as Condorcet candidates[8]. The larger part of the solution space produces e as the Condorcet winner.

From Theorem 5 and the data in Table 1 we directly compute the Borda scores reported in Table 3. The ordering according to Borda scores is *ecd* with e the Borda winner. Recall that *ecd* is also the ordering by approval voting scores and a likely social welfare order. This is evidence that AV would have elected a candidate who is both Condorcet and Borda.

The fact that we may end up with several possible Condorcet candidates, depending on the value of λ, i.e. the particular solution of the size-independent model, may appear somewhat problematic. However, the reason for this is that the size-independent model is nonparametric, and thus very general. For a systematic analysis of several elections as well as a detailed discussion about their substantive interpretations the reader is referred to Regenwetter and Grofman (in press) where the relationships between AV score ordering, Borda score ordering and social welfare orders are studied in much detail.

7. Discussion

The work presented in this paper provides a method of crafting more realistic models of social choice by embedding social choice analysis into a psychological representation of preferences and choice behavior. By formulating social choice concepts in terms of a plausible and empirically testable cognitive model of individual choice behavior we take a model-based approach to the question of whether or not weak stochastic transitivity holds and whether or not a Condorcet winner exists in paradigms, where no paired comparisons are being observed.

---

[8] Of course the meaning of the word Condorcet candidate is given by the size-independent model as the candidate, if s/he exists, who would have a greater $\frac{1}{2}$ probability of being ranked ahead of any given other candidate.

Most social choice theory has proceeded in a purely axiomatic way (for an exception see, for instance, Chamberlin, Cohen and Coombs, 1984). In contrast, there is a strong tradition in mathematical psychology to formulate descriptive models of choice behaviour in terms of empirically observable primitives, as does the weak utility model of individual binary choice (Luce and Suppes, 1965). In combining both traditions, we believe we contribute to both literatures by showing some axiomatic underpinnings of probabilistic choice models and their empirical applications to social choice concepts.

So far we have applied the model presented here to 10 three-candidate AV elections. In 7 of these 10 instances the model satisfactorily fit, and we were able to apply the technique to test for the existence of a Condorcet winner and a social welfare order and to compute the Borda scores. These results along with further analyses are presented and discussed in detail in an empirical companion paper (Regenwetter and Grofman, in press).

It may seem problematic that some results so far are only applicable to three-candidate elections. The reason for this limitation is that the structure of the approval-voting polytope underlying the size-independent model becomes rather complex when there are more than three candidates and requires further study. Most social choice theoretic results in the present paper are formulated for arbitrary values of N. In particular, parametric submodels of the size-independent model allow for applications with N taking any value. Such parametrizations are still under development and a detailed application is left for later.

The main contribution of the present paper concerns the mathematical tools and the underlying research strategy. We demonstrate that it pays off to combine different research traditions and, in particular, to base the analysis of voting schemes on testable cognitive models of the decision making process. We also consider it useful to formulate traditional concepts encountered in social welfare in terms of such a model. By doing so, we make explicit our assumptions about what determines voting behavior. This combines the analysis of social choice theoretical properties of a voting scheme such as approval voting with an empirical test of the fundamental assumptions about the nature of voter choice. While the size-independent model may be too simple to qualify as a comprehensive model of subset choice, it serves well as a measurement instrument, the appropriateness of which we can evaluate each time we use it.

Glossary of symbols

$\exists$	there exists
$\forall$	for all
$\mid$	such that
$\mid X \mid$	cardinality of X
$\mathscr{C}$	set of choice alternatives or candidates
$\boldsymbol{C}(Y)$	conditional probability of Y given the size of Y
λ	a real valued scalar
NP	net probability
p_{ab}	binary choice probability of a over b
$\mathbb{P}_{ab}$	binary preference probability of a over b
$\mathbb{P}$	probability
P	welfare ordering
Π	collection of linear orders (rankings, permutations) over $\mathscr{C}$
Π_X	collection of rankings where the elements of X are ranked first
π	any element in Π
π^{-1}	converse of π
R	welfare ordering
$\mathbf{R}$	preference relation random variable
$\mathbf{S}$	vote size random variable
$\mathbf{V}$	subset vote random variable
X	subset of choice alternatives

References

Anderson SP, de Palma A, Thisse J.-F (1992) Discrete Choice Theory of Product Differentiation. Cambridge: MIT Press

Arrow KJ (1951) Social Choice and Individual Values. New York: Wiley

Baigent N, Xu Y (1991) Independent necessary and sufficient conditions for Approval Voting. Mathematical Social Sciences 21: 21–29

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

Black D (1958) The Theory of Committees and Elections. Cambridge: University Press

Block HD, Marschak J (1960) Random orderings and stochastic theories of responses. In: I. Olkin, S. Ghurye, H. Hoeffding, W. Madow, H. Mann (Eds.) Contributions to Probability and Statistics (pp. 97–132). Stanford: Stanford University Press

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

Brams SJ (1990) Constrained Approval Voting: A voting system to elect a governing board. Interfaces 20: 67–80

Brams SJ and Fishburn, PC (1983) Approval Voting. Boston: Birkhäuser

Brams SJ, Fishburn PC (1985) Comment on the problem of strategic voting under Approval Voting. American Political Science Review, 79: 816–818

Brams SJ, Fishburn PC (1988) Does Approval Voting elect the lowest common denominator? PS: Political Science and Politics 21: 277–284

Brams SJ, Fishburn PC (1992) Approval Voting in scientific and engineering societies. Group Decision and Negotiation 1: 41–55

Brams SJ, Fishburn PC, Merrill SI (1988a) Rejoinder to Saari and Van Newenhizen. Public Choice 59: 149

Brams SJ, Fishburn PC, Merrill SI (1988b) The responsiveness of approval voting: Comments on Saari and Van Newenhizen. Public Choice 59: 121–131

Brams SJ, Nagel JH (1991) Approval Voting in practice. Public Choice 71: 1–17

Campello de Souza FM (1983) Mixed models, random utilities, and the triangle inequality. Journal of Mathematical Psychology 27: 183–200

Carter C (1990) Admissible and sincere strategies under approval voting. Public Choice 64: 43–55

Chamberlin JR, Cohen JL, Coombs CH (1984) Social choice observed: Five presidential elections of the American Psychological Association. The Journal of Politics 46: 479–502

Cohen M, Falmagne J.-C (1990) Random utility representation of binary choice probabilities: A new class of necessary conditions. Journal of Mathematical Psychology 34: 88–94

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

Cox GW (1985) Electoral equilibrium under Approval Voting. American Journal of Political Science 29: 112–118

Critchlow DE, Fligner MA, Verducci JS (1991) Probability models on rankings. Journal of Mathematical Psychology 35: 294–318

De Maio G, Muzzio D (1986) Mapping candidate systems via Approval Voting. Western Political Quarterly 39: 663–674

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

Dridi T (1980) Sur les distributions binaires associées à des distributions ordinales (On the binary distributions associated with ordinal distributions). Mathématiques et Sciences Humaines 69: 15–31

Falmagne J-C (1985) Elements of Psychophysical Theory. Oxford: Clarendon

Falmagne J-C (1978) A representation theorem for finite random scale systems. Journal of Mathematical Psychology 18: 52–72

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

Feld SL, Grofman B (1986) Partial single-peakedness: An extension and clarification. Public Choice 51: 71–80

Feld SL, Grofman B (1988) Ideological consistency as a collective phenomenon. American Political Science Review 83: 773–788

Feld SL, Grofman B (1990) Collectivities as actors: Consistency of collective choices. Rationality and Society 2: 429–448

Feld SL, Grofman B (1992) Who's afraid of the big bad cycle? Evidence from 36 Elections. Journal of Theoretical Politics 4: 231–237

Felsenthal DS, Maoz Z, Rapoport A (1986) Comparing voting systems in genuine elections: Approval Voting vs Selection-Plurality. Social Behavior 1: 41–53

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

Fishburn PC (1990) Binary probabilities induced by rankings. SIAM Journal of Discrete Mathematics 3: 478–488

Fishburn PC (1992) Induced binary probabilities and the linear ordering polytope: A status report. Mathematical Social Sciences 23: 67–80

Fishburn PC, Falmagne J-C (1989) Binary choice probabilities and rankings. Economic Letters 31: 113–117

Fishburn PC, Little JDC (1988). An experiment in Approval Voting. Management Science, 34: 555–568

Gärtner W, Heinecke A (1978) Cyclically mixed preferences – a necessary and sufficient condition for transitivity of the social preference relation. In: HW. Gottinger, W. Leinfellner (Eds.) Decision Theory and Social Ethics (pp. 169–185). Dordrecht: Reidel.

Gilboa I (1990). A necessary but insufficient condition for the stochastic binary choice problem. Journal of Mathematical Psychology 34: 371–392

Heiner RA, Pattanaik PK (1983) The structure of general probabilistic group decision rules. In: PK. Pattanaik and M. Salles (Eds.) Social Choice and Welfare (pp. 37–54) Amsterdam: North-Holland

Heyer D, Niederée R (1989) Elements of a model-theoretic framework for probabilistic measurement. In: EE. Roskam (Ed.) Mathematical Psychology in Progress (pp. 99–112) Berlin, Heidelberg, New York: Springer

Heyer D, Niederée R (1992) Generalizing the concept of binary choice systems induced by rankings: One way of probabilizing deterministic measurement structures. Mathematical Social Sciences 23: 31–44

Koppen M (1995) Random utility representation of binary choice probabilities: Critical graphs yielding critical necessary conditions. Journal of Mathematical Psychology 39: 21–39

Levin J, Nalebuff B (1995) An introduction to vote counting schemes. Journal of Economic Perspectives 9: 3–26

Lijphart A, Grofman B (Eds.). (1984) Choosing an electoral system. New York: Praeger

Luce RD, Suppes P (1965) Preference, utility and subjective probability. In: RD. Luce, RR. Bush, E. Galanter (Eds.) Handbook of Mathematical Psychology, Vol. III (pp. 249–410) New York: Wiley

Marley AAJ (1990) A historical and contemporary perspective on random scale representations of choice probabilities and reaction times in the context of Cohen and Falmagne's results. Journal of Mathematical Psychology 34: 81–87

Marschak J (1960) Binary-choice constraints and random utility indicators. In: KJ. Arrow, S. Karlin, P. Suppes (Eds.) Proceedings of the first Stanford symposium on mathematical methods in the social sciences, 1959 (pp. 312–329) Stanford, CA: Stanford University Press

McFadden D, Richter MK (1970) Revealed stochastic preference. Unpublished manuscript, Department of Economics, University of California at Berkeley

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

Merrill S,III, Nagel J (1987) The effect of approval balloting on strategic voting under alternative decision rules. American Political Science Review 81: 509–524

Murakami Y (1968) Logic and Social Choice. London: Routledge, K Paul

Niemi RG (1984) The problem of strategic behavior under Approval voting. American Political Science Review 78: 952–958

Nurmi H (1987) Comparing Voting Systems. Dordrecht: Kluwer Academic Publishers

Pattanaik PK (1971) Voting and collective choice: Some aspects of the theory of group decision-making. Cambridge: University Press

Rapoport A, Felsenthal DS (1990) Efficiency in small electorates under plurality and Approval Voting. Public Choice 64: 57–71

Regenwetter M, Grofman B (1997). Approval voting, Borda winners and Condorcet winners: evidence from seven elections. Management Science (in press)

Regenwetter M, Marley A, Joe H (1996) Random utility threshold models of subset choice. Submitted

Roberts FS (1979) Measurement Theory. London: Addison-Wesley

Saari DG (1994) Geometry of Voting. Berlin, Heidelberg, New York: Springer

Saari DG, Van Newenhizen J (1988a) Is Approval Voting an 'unmitigated evil'?: A response to Brams, Fishburn, and Merrill. Public Choice 59: 133–147

Saari DG, Van Newenhizen J (1988b) The problem of indeterminacy in approval, multiple, and truncated voting systems. Public Choice 59: 101–120

Sen AK (1966) A possibility theorem on majority decisions. Econometrica 34: 491–499

Sen AK (1969) Quasi-transitivity, rational choice and collective decisions. Review of Economic Studies 36: 381–393

Sen AK (1970) Collective Choice and Social Welfare. San Francisco: Holden-Day

Suck R (1992) Geometric and combinatorial properties of the polytope of binary choice probabilities. Mathematical Social Sciences 23: 81–102

Weber RJ (1995) Approval Voting. Journal of Economic Perspectives 9: 39–49

Young HP (1974) An axiomatization of Borda's rule. The Journal of Economic Theory 9: 43–52

Young HP (1986) Optimal ranking and choice from pairwise comparisons. In: B. Grofman, G. Owen (Eds.) Information Pooling and Group Decision Making: Proceedings of the Second University of California, Irvine, Conference in Political Economy (pp. 113–121) Greenwich: JAI Press

Young HP (1988) Condorcet's theory of voting. American Political Science Review 82: 1231–1243

Young HP (1995) Optimal voting rules. Journal of Economic Perspectives 9: 51–64