The Repeated Insertion Model for Rankings: Missing Link between Two Subset Choice Models

Article *in* Psychometrika · March 2004

DOI: 10.1007/BF02295838 · Source: RePEc

Jean-Paul Doignon
Université Libre de Bruxelles

Michel Regenwetter
University of Illinois Urbana-Champaign

psychometrika—vol. 69, no. 1, 33–54
march 2004

THE REPEATED INSERTION MODEL FOR RANKINGS: MISSING LINK BETWEEN TWO SUBSET CHOICE MODELS.

Jean-Paul Doignon

department of mathematics, université libre de bruxelles
c.p. 216, bd du triomphe, 1050 bruxelles, belgium.
e-mail: doignon@ulb.ac.be.

Aleksandar Pekeč

fuqua school of business, duke university
durham, nc 27708, u.s.a.
e-mail: pekec@duke.edu.

Michel Regenwetter

department of psychology, university of illinois at urbana-champaign
603 e. daniel street, champaign, il 61820, u.s.a.
e-mail: regenwet@uiuc.edu.

Several probabilistic models for subset choice have been proposed in the literature, e.g., to explain approval voting data. We show that Marley et al.'s latent scale model is subsumed by Falmagne and Regenwetter's size-independent model, in the sense that every choice probability distribution generated by the former can also be explained by the latter. Our proof relies on the construction of a probabilistic ranking model which we label the 're-peated insertion model'. This model is a special case of Marden's orthogonal contrast model class and, in turn, includes the classical Mallows ϕ-model as a special case. We explore its basic properties as well as its relationship to Fligner and Verducci's multistage ranking model.

Key words: Approval voting, probabilistic choice models, probabilistic ranking models, subset choice.

The authors are grateful to the National Science Foundation for grants SES98-18756 to Regenwetter and Pekeč, and SBR97-30076 to Regenwetter. This collaborative research was carried out in the context of the conference Random Utility 2000 held at Duke University and sponsored by NSF, the Fuqua School of Business and the Center for International Business Education and Research. We thank the editor and four referees for helpful suggestions and we are grateful to Prof. J. I. Marden for providing useful information on contrast models. We thank Moon-Ho Ho for programming and running the data analyses.

Introduction

In order to be of general interest, a model of judgment and decision making needs to be generalizable beyond the specific empirical paradigm for which it was originally designed. Thus, a major open challenge in probabilistic modeling of preference and choice behavior is to formulate models that are applicable to (or can easily be reformulated for) a broad range of empirical paradigms. In this vein, the present paper investigates the interrelationship among a specific set of models for ranking and/or choice data.

There is a substantial past and evolving literature on probabilistic models specifically for ranking data (Critchlow, Fligner & Verducci, 1991, 1993; Marden, 1992, 1995). Every probabilistic ranking model in turn naturally induces a series of choice models: In a binary choice task, the respondent is asked to choose one of two choice alternatives. Given a latent or overt preference ranking of all choice alternatives, the respondent can simply choose the higher ranked among the two choice alternatives. In a multiple choice task, the respondent is asked to choose one object out of k many objects. Again, given a preference ranking over the choice alternatives, the respondent can simply choose the highest ranked among the k available options. In a subset choice task, the respondent is asked to choose a subset of any size from a given available set of choice alternatives. This is a more complex situation as there is no single canonical way in which a ranking model induces a choice model. A plausible assumption, analogous to the previous two models, states that, for any choice alternative included in the chosen set, all objects that the respondent ranks better than this alternative should also be chosen. But, a ranking model *per se* does not, for instance, explain why some respondents choose the empty set, i.e., decline to choose any of the available options. (Of course, one could add an extra object in the rankings, called a 'threshold', and assume that the respondent chooses all objects ranked ahead of the threshold. See Regenwetter et al., 1998, for such models.)

For this and many other reasons, it comes to no surprise that the multi-disciplinary and multi-faceted literature on probabilistic choice has mostly focused on binary choice (i.e., two-alternative forced choice) data (e.g., Ahn, 1995; Atkinson, Wampold, Lowe, Matthews & Ahn, 1998; Baier & Gaul, 1999; Bockenholt & Dillon, 1997; Courcoux & Semenou, 1997; Fishburn, 2001; Ichimura & Thompson, 1998; Koppen, 1995; Peterson & Brown, 1998; Tsai, 2000) and multiple choice data (e.g., Baltas & Doyle, 2001; Barberá & Pattanaik, 1986; Billot & Thisse, 1999; Chen & Kuo, 2001; Falmagne, 1978; Huang & Nychka, 2000; Lewbel, 2000; McFadden & Train, 2000; Zeng, 2000). Most models in this domain were either originally formulated or can easily be reformulated as choice models based on latent ranking probabilities. For instance, most random utility models can be reformulated as ranking models (Block & Marschak, 1960). It therefore seems natural that ranking models offer a reasonable and often used common ground between alternative models of preference and/or choice. For subset choice, however, as indicated above, a ranking model by itself does not help to determine how many objects the respondent will or should choose. Nevertheless, we show here that ranking models are highly relevant for the study of subset choice models and their interrelationship.

Arguably, the most natural ranking-based model for subset choices is the 'size-independent model' of Falmagne and Regenwetter (1996). According to this model the respondent latently rank orders the choice alternatives and decides how many objects to choose, say k many. The respondent then chooses a k-element subset whose elements are all ranked better than those not chosen. In particular, the size-independent model is able to take any ranking model and transform it into a subset choice model (after adding a mechanism for determining the subset size). An alternative model for subset choice probabilities, which, at first sight, seems to have nothing in common with ranking models, is Marley's (1993) 'latent scale model' according to which the respondent considers each object in turn and independently decides (with some probability that depends only on the choice alternative) whether or not to include it in the chosen set.

While the present paper was originally motivated by the quest for a deeper understanding of how these two models for subset choice behavior are related to each other, it turns out that a particular ranking model, which we call the 'repeated insertion model' allows us to *construct* a size-independent model from any latent scale model. More specifically, every latent scale model induces a repeated insertion ranking model, which, in turn, when combined with the size-independent model, generates subset choice probabilities that are identical to the subset choice probabilities associated with the initial latent scale model.

The paper is organized as follows. We first provide some background on subset choice data and we formally introduce the two probabilistic models for subset choice, namely the size-independent model and the latent scale model. We then introduce a ranking model, which we call the repeated insertion model, and which is equivalent to a special case of an orthogonal contrast model (Marden, 1992, 1995). We are then able to compare the size-independent model and the latent scale model via the use of the repeated insertion model. Finally, we study the basic properties of the repeated insertion model and sketch its relationship to some known ranking models. In particular, we compare it to the Mallows ϕ-model (Mallows, 1957) and to Fligner and Verducci's (1988) multistage ranking model. We also illustrate the repeated insertion model and its comparison with the multistage model through a brief empirical application of both models to three important sets of political survey data. All proofs, as well as two technical Lemmas, are relegated to the Appendix.

Probabilistic models of subset choices

Consider a choice task where the decision maker is required to choose a subset (of any size) of a given set $\mathcal{C}$ of n choice alternatives. In other words, the decision maker chooses an element in the power set $\mathcal{P}(\mathcal{C})$ of $\mathcal{C}$. If $\mathcal{C} = \{a, b, c, d\}$, s/he may choose the subset $\{c, b\}$ (which is the same as $\{b, c\}$), or the empty set, or $\mathcal{C}$, or any one of thirteen other subsets.

A prominent real world example of subset choice data are the ballots collected under 'approval voting'. Under the approval voting method, each voter selects the subset of $\mathcal{C}$ consisting of all alternatives s/he 'approves' of. Each candidate in the approved set receives a point and the winner is the candidate with the highest total of points. Approval voting is used by various organizations, including the American Statistical Association, the Mathematical Association of America, the National Academy of Sciences, the Institute for Operations Research and the Management Sciences, the Society for Social Choice and Welfare and, most recently, the Society for Judgment and Decision Making.

We focus on two probabilistic models for subset choices, namely the 'latent scale model' introduced in Marley (1993), and named in Regenwetter et al. (1998), as well as the 'size-independent model' of Falmagne and Regenwetter (1996). Both models are based on rather natural assumptions, although of a completely different qualitative nature.

A remark is in order, concerning notation. We will use the notation $P(\ )$ to denote probabilities of observable quantities generated by a model. So, for a subset choice model and a subset X, we write $P(X)$ for the probability of the set X. Analogously, when we are discussing a ranking model, for a ranking π, we write $P(\pi)$ for the probability of the ranking π.

Formally, focusing first on subset choice models, let $P : \mathcal{P}(\mathcal{C}) \to [0, 1]$ be a probability distribution over all subsets of $\mathcal{C}$, i.e. $0 \leq P(X)$ and $\sum_{X \in \mathcal{P}(\mathcal{C})} P(X) = 1$.

The latent scale model assumes that each alternative x in $\mathcal{C}$ is retained or rejected, with some fixed probability l_x, independently of the other alternatives. The probabilities $P(X)$, for $X \in \mathcal{P}(\mathcal{C})$, satisfy a latent scale model if there exist parameters l_x, for $x \in \mathcal{C}$

with $0 \leq l_x \leq 1$, such that

$$P(X) \;=\; \prod_{x \in X} l_x \cdot \prod_{y \in \mathcal{C} \setminus X} \Big(1 - l_y\Big). \tag{1}$$

The resulting probability distribution P on the power set $\mathcal{P}(\mathcal{C})$ is said to be generated by the *latent scale model for subset choice*, with parameters l_x. We use the abbreviation LSMS for "latent scale model of subset choice" (the last letter in the abbreviation, S, indicates a model for subset choice; later, a different ending letter, R, will indicate a ranking model).

Example 1. Take $\mathcal{C} = \{1, 2, 3\}$, let $l_1 = l_2 = \frac{1}{2}$, and let $l_3 = \frac{1}{3}$. Then

$$\begin{aligned} P(\emptyset) &= \left(1 - \frac{1}{2}\right)\left(1 - \frac{1}{2}\right)\left(1 - \frac{1}{3}\right), & P(\{3\}) &= \left(1 - \frac{1}{2}\right)\left(1 - \frac{1}{2}\right)\frac{1}{3}, \\ P(\{1\}) &= \frac{1}{2}\left(1 - \frac{1}{2}\right)\left(1 - \frac{1}{3}\right), & P(\{1,3\}) &= \frac{1}{2}\left(1 - \frac{1}{2}\right)\frac{1}{3}, \\ P(\{2\}) &= \left(1 - \frac{1}{2}\right)\frac{1}{2}\left(1 - \frac{1}{3}\right), & P(\{2,3\}) &= \left(1 - \frac{1}{2}\right)\frac{1}{2} \cdot \frac{1}{3}, \\ P(\{1,2\}) &= \frac{1}{2} \cdot \frac{1}{2}\left(1 - \frac{1}{3}\right), & P(\{1,2,3\}) &= \frac{1}{2} \cdot \frac{1}{2} \cdot \frac{1}{3}. \end{aligned}$$

We will later use the following property of LSMS. For x, X, Y such that $x \in X \subset \mathcal{C}$ and $x \notin Y \subset \mathcal{C}$,

$$P(X) \cdot P(Y) = P(X \setminus \{x\}) \cdot P(Y \cup \{x\}). \tag{2}$$

This condition was labeled *constant ratio condition* by Regenwetter et al. (1998) because it translates into

$$\frac{P(X)}{P(X \setminus \{x\})} = \frac{P(Y \cup \{x\})}{P(Y)}$$

whenever the probabilities in the denominators are not zero. Regenwetter et al. (1998) have shown that the constant ratio condition characterizes LSMS.

According to the second model, the voter decomposes the decision into two components. S/he rank orders the alternatives and decides how many alternatives to choose. A voter with ranking π and subset size s will then select the first s alternatives in the ranking. The resulting model involves a probability distribution p on possible subset sizes and a probability distribution q on all rankings of $\mathcal{C}$. (We write q_π rather than $P(\pi)$, to emphasize the fact that the ranking probabilities are latent parameters of the model.) Let Π be the collection of all rankings of $\mathcal{C}$. The size-independent model has parameters p_s for $0 \leq s \leq n$, and q_π for $\pi \in \Pi$, that satisfy $0 \leq p_s \leq 1$, $\sum_{s=0}^{n} p_s = 1$, $0 \leq q_\pi \leq 1$, and $\sum_{\pi \in \Pi} q_\pi = 1$. The probabilities $P(X)$, for $X \in \mathcal{P}(\mathcal{C})$, satisfy a *size-independent model of subset choices* if there exist such parameters, so that

$$P(X) \;=\; p_{|X|} \cdot \sum_{\pi \in \Pi_X} q_\pi, \tag{3}$$

where Π_X stands for the set of all rankings of $\mathcal{C}$ in which the elements of X are ranked before all the elements of $\mathcal{C} \setminus X$. We use the abbreviation SIMS for "size-independent model of subset choice".

The question of comparing the two models appears in Regenwetter et al. (1998). It is easily checked that LSMS does not generate all of the probability distributions that

SIMS generates (cf. Example 3 and Lemma 7 in the Appendix). On the other hand, Regenwetter and Doignon (1998) have shown nonconstructively that any probability distribution generated by LSMS is recoverable in SIMS at least for up to five alternatives. It appears possible, in principle, to subsume LSMS by SIMS for all values of n in view of the numbers of parameters associated with the two models: LSMS has only n independent real-valued parameters, whereas SIMS has $2^n - 1$ degrees of freedom (for the latter, see the proof of Theorem 1 in Doignon and Regenwetter, 1997). On the other hand, the first model is highly nonlinear, while the second is essentially bilinear. (See also Doignon and Fiorini, in press; Doignon and Fiorini, 2002; Doignon and Regenwetter, 2002, for further relevant work on the related 'approval voting polytope'.)

Despite this apparent mathematical discrepancy, we will prove that LSMS is always subsumed by SIMS. Even more, for any instance of LSMS, we will explicitly provide an instance of SIMS so that the two instances generate exactly the same probability distribution on $\mathcal{P}(\mathcal{C})$. Our construction in the general case involves an interesting class of probability distributions on the collection Π of all rankings of $\mathcal{C}$. We call this family of distributions the 'repeated insertion model'.

The Repeated Insertion Model

We are now ready to introduce the 'repeated insertion model' or RIMR, which will later be key in establishing the relationship between LSMS and SIMS.

As before, let $\mathcal{C}$ be a set of n alternatives. A *ranking* π of $\mathcal{C}$ is a bijective mapping $\pi : \{1, 2, \ldots, n\} \to \mathcal{C}$, that we specify by writing

$$\pi \quad = \quad \Big(\pi(1), \pi(2), \ldots, \pi(n)\Big) \quad = \quad \langle\, \pi_1, \pi_2, \ldots, \pi_n \,\rangle\,. \tag{4}$$

Thus, $\pi(i)$ or π_i denotes the alternative at rank i in π, and the rank in π of an alternative a equals $\pi^{-1}(a)$. We sometimes also think of the ranking π as a binary relation on $\mathcal{C}$, in which case we write $a\pi b$ to mean $\pi^{-1}(a) < \pi^{-1}(b)$, where a, $b \in \mathcal{C}$. For simplicity, we assume from now on that $\mathcal{C} = \{1, 2, \ldots, n\}$.

One can think of RIMR as a process model according to which the respondent constructs a ranking of the alternatives inductively, relative to some 'reference ordering'. More specifically, when a respondent rank orders the alternatives, s/he considers each alternative in turn and in the order provided by the reference ordering. S/he then inserts it among the ones already placed in her/his ranking. For instance, when the respondent considers the second object in the reference order, s/he inserts this object either at the first rank, with some probability w_1^2, or at the last rank, with probability w_2^2 (we will formally introduce the parameters below). The next object will be inserted at the first rank with probability w_1^3, at the middle rank, with probability w_2^3, or at the last rank, with probability w_3^3. The process continues until all objects have been inserted. At the last step, the last object in the reference order has n many different ranks at which it can be inserted, each rank i having a fixed probability w_i^n.

For a given *reference ordering* ρ of $\mathcal{C}$ and a given ranking π of $\mathcal{C}$, we denote by $f_{\rho,\pi}(k)$ the relative rank of alternative ρ_k in π among alternatives $\{\rho_1, \ldots, \rho_k\}$. (In particular, if ρ is the canonical ordering $\langle 1, 2, \ldots, n\rangle$ of $\mathcal{C}$, $f_{\rho,\pi}(k)$ is the relative rank of alternative k in π among alternatives $\{1, \ldots, k\}$.) In other words, $f_{\rho,\pi}(k)$ equals one plus the number of alternatives ranked before alternative ρ_k in both ρ and π:

$$\begin{aligned} f_{\rho,\pi}(k) &= 1 + |\,\{\, a \in \mathcal{C} \;:\; a\,\rho\,\rho_k \text{ and } a\,\pi\,\rho_k \,\}\,| \\ &= 1 + |\,\{\, a \in \mathcal{C} \;:\; \rho^{-1}(a) < k \text{ and } \pi^{-1}(a) < \pi^{-1}(\rho_k) \,\}\,| \qquad (5) \\ &= |\{t \in \{1, 2, \ldots, k-1\} \;:\; \pi^{-1}(\rho_t) < \pi^{-1}(\rho_k) \,\}|. \qquad (6) \end{aligned}$$

It can be checked that the reference ordering ρ and the function $f_{\rho,\pi}$ completely specify the ranking π (indeed, if ρ is the canonical ordering, π can be rebuilt from $f_{\rho,\pi}$ by first writing $\langle 2,1\rangle$ if $f_{\rho,\pi}(2)=1$, or $\langle 1,2\rangle$ if $f_{\rho,\pi}(2)=2$, then inserting 3 according to the value of $f_{\rho,\pi}(3)$, and successively inserting 4, 5, ..., n). In passing, we also mention that, when ρ is the canonical ordering, the function $f_{\rho,\pi}$ is close to the so-called 'inversion table' of π, whose k-th component equals $n - f_{\rho,\pi}(k) - 1$. Knuth (1997) explains how inversion tables provide a useful tool for the analysis of sorting methods.

In RIMR we parametrize the ranking probabilities using real-valued parameters w_i^k, with $1 \leq i \leq k \leq n$, and where $0 \leq w_i^k$ and $\sum_{i=1}^k w_i^k = 1$ (thus, $w_1^1 = 1$). Notice that in addition to these $n(n-1)/2$ real-valued parameters, the reference ordering is also a (discrete) parameter of the model. For any ranking π of $\mathcal{C}$, RIMR assigns ranking probabilities P as follows:

$$P(\pi) \quad = \quad w^1_{f_{\rho,\pi}(1)} \cdot w^2_{f_{\rho,\pi}(2)} \cdot \ldots \cdot w^n_{f_{\rho,\pi}(n)}, \tag{7}$$

where ρ is the reference ordering. The equation encapsulates the mutual independence of the values of $f_{\rho,\pi}$: the insertion of the jth alternative, ρ_j, at some position in the ranking π is independent of the positioning of all previously inserted alternatives ρ_1, ρ_2, ..., ρ_{j-1}.

Note that Equation (7), when the product is evaluated from right to left, indicates an alternative interpretation of RIMR: assign $\pi(n)$ to rank $f_{\rho,\pi}(n)$ with probability $w^n_{f_{\rho,\pi}(n)}$, assign $\pi(n-1)$ to relative rank $f_{\rho,\pi}(n-1)$ among the remaining $n-1$ ranks with probability $w^{n-1}_{f_{\rho,\pi}(n-1)}$, etc.

The following proposition (whose proof is in the Appendix) shows formally that this construction yields a well-defined probability distribution on the collection Π of all rankings of $\mathcal{C}$; we say that this distribution is obtained according to the *repeated insertion model for rankings*, or RIMR.

PROPOSITION 1. *The function P specified by the repeated insertion model in Equation (7) is a probability distribution on the collection Π of all rankings of $\mathcal{C}$.*

In fact, as pointed out by a referee, RIMR is a special case of an orthogonal contrast model (Marden, 1992, 1995). More specifically, it can be reformulated as a free orthogonal contrast model with up or down contrasts. (Writing, as before, ρ for the reference ordering, the up or down contrasts are $(\{\rho(1)\},\{\rho(2)\})$, $(\{\rho(1),\rho(2)\},\{\rho(3)\})$, ..., $(\{\rho(1),\rho(2),\ldots,\rho(n-1)\},\{\rho(n)\})$.)

Comparing Subset Models

Given any ranking model MR, we can introduce a submodel of SIMS by restricting the allowed distributions on Π to those generated by MR. In other words, given the size-independent model, any probabilistic ranking model MR naturally induces a probabilistic subset choice model SIMS(MR). The *size-independent model induced by the repeated insertion model* SIMS(RIMR) is the subset choice model obtained from the size-independent model when the ranking distribution is constrained to be consistent with a repeated insertion model. The model SIMS(RIMR) thus has parameters p_s (for the various sizes s, with $0 \leq s \leq n$), ρ (the reference ordering), and w_i^k (as in the previous section).

Our goal is to establish that LSMS is subsumed by SIMS(RIMR).

Example 2. Suppose we want to show that the subset probabilities obtained from LSMS in Example 1 can also be obtained from SIMS(RIMR). Clearly, the size probabilities can be computed from the data in Example 1. We obtain

$$p_0 = \frac{2}{12}, \quad p_1 = \frac{5}{12}, \quad p_2 = \frac{4}{12}, \quad p_3 = \frac{1}{12}.$$

As a consequence, expressing the subset probabilities according to SIMS(RIMR) leads to 6 linear equations in the 5 unknowns w_j^k:

$$P(\{1\}) = \frac{2}{12} = \frac{5}{12}(w_2^2 \cdot w_3^3 + w_2^2 \cdot w_2^3), \quad P(\{2\}) = \frac{2}{12} = \frac{5}{12}(w_1^2 \cdot w_3^3 + w_1^2 \cdot w_2^3),$$

$$P(\{3\}) = \frac{1}{12} = \frac{5}{12}(w_2^2 \cdot w_1^3 + w_1^2 \cdot w_1^3), \quad P(\{1,2\}) = \frac{2}{12} = \frac{4}{12}(w_2^2 \cdot w_3^3 + w_1^2 \cdot w_3^3),$$

$$P(\{1,3\}) = \frac{1}{12} = \frac{4}{12}(w_2^2 \cdot w_2^3 + w_2^2 \cdot w_1^3), \quad P(\{2,3\}) = \frac{1}{12} = \frac{4}{12}(w_1^2 \cdot w_2^3 + w_1^2 \cdot w_1^3).$$

Using $w_1^2 + w_2^2 = 1$, we derive $w_1^3 = \frac{1}{5}$, $w_3^3 = \frac{1}{2}$, and thus $w_2^3 = \frac{3}{10}$, and finally $w_1^2 = w_2^2 = \frac{1}{2}$.

THEOREM 2. *Any probability distribution on $\mathcal{P}(\mathcal{C})$ generated by the latent scale model LSMS on $\mathcal{C} = \{1, 2, \ldots, n\}$ can also be generated by the size-independent model induced by the repeated insertion model SIMS(RIMR). Moreover the reference ordering in RIMR can be chosen arbitrarily, e.g., it can be the canonical ordering.*

The proof, as well as two subsidiary technical Lemmas (together with their proofs), can be found in the Appendix. In particular, the subset size probabilities are computed from LSMS in Equations (12) and (27), and the parameters w_j^k are given recursively in Equations (28) and (29) for the case of a canonical reference ordering.

Notice that the converse of Theorem 2 does not hold, as the following example demonstrates.

Example 3. Let $\mathcal{C} = \{1, 2, 3\}$ and consider SIMS(RIMR) where the ranking distribution P is defined by the reference ordering $\langle 1, 2, 3\rangle$ and parameters $w_1^2 = \frac{3}{4}$ (thus $w_2^2 = \frac{1}{4}$), $w_1^3 = w_2^3 = w_3^3 = \frac{1}{3}$. In other words, the ranking distribution is given by

$$P(\langle 1,2,3\rangle) = \left(\frac{1}{4}\right)\left(\frac{1}{3}\right), \quad P(\langle 2,1,3\rangle) = \left(\frac{3}{4}\right)\left(\frac{1}{3}\right), \quad P(\langle 3,1,2\rangle) = \left(\frac{1}{4}\right)\left(\frac{1}{3}\right),$$

$$P(\langle 1,3,2\rangle) = \left(\frac{1}{4}\right)\left(\frac{1}{3}\right), \quad P(\langle 2,3,1\rangle) = \left(\frac{3}{4}\right)\left(\frac{1}{3}\right), \quad P(\langle 3,2,1\rangle) = \left(\frac{3}{4}\right)\left(\frac{1}{3}\right).$$

We thus have

$$P(\Pi_{\{1,2\}}) = \frac{1}{3}, \qquad P(\Pi_{\{1,3\}}) = \frac{1}{6},$$

$$P(\Pi_{\{2\}}) = \frac{1}{2}, \qquad P(\Pi_{\{3\}}) = \frac{1}{3}.$$

Thus, for any choice of nonzero values of the size probabilities p_1 and p_2 in SIMS(RIMR), setting $x = 1, X = \{1,2\}, Y = \{3\}$ violates the constant ratio condition in Equation (2) because

$$P(\{1,2\}) \cdot P(\{3\}) = p_2\frac{1}{3} \cdot p_1\frac{1}{3} \quad \neq \quad p_1\frac{1}{2} \cdot p_2\frac{1}{6} = P(\{2\}) \cdot P(\{1,3\}).$$

Hence the same subset distribution P cannot be generated by LSMS. Notice that we could instead have deduced the same fact from dimension considerations because the collection of probability distributions on $\mathcal{P}(\mathcal{C})$ defined by SIMS is characterized by $2^n - 1$ independent parameters (see the proof of Theorem 1 in Doignon and Regenwetter, 1997), while the similar collection defined by SIMS(RIMR) is characterized by at most $n + n(n-1)/2$ independent real-valued parameters for a fixed reference ordering.

More on the Repeated Insertion Model for Rankings

In this section we further analyze some basic properties of RIMR, such as 'label invariance', 'reversibility', 'L-decomposability' and the lack of 'unimodality' (and thus the lack of 'complete consensus'). For example, these properties were proposed by Critchlow et al. (1991) as benchmarks for classifying ranking models. We further show that RIMR encompasses the Mallows ϕ-model as a special case and compare RIMR with the 'multistage ranking model' of Fligner and Verducci (1988).

Label invariance. A ranking model MR satisfies *label invariance* when the following property holds. Given a probability distribution on rankings generated by MR, if we relabel all choice alternatives, we again obtain a probability distribution on rankings generated by MR. Formally, for any permutation σ of the choice set $\mathcal{C}$, and any probability distribution $P : \Pi \to [0,1]$ in MR, the probability distribution $P_\sigma : \Pi \to [0,1]$ given by

$$P_\sigma(\pi_1, \pi_2, \dots, \pi_n) = P(\sigma(\pi_1), \sigma(\pi_2), \dots, \sigma(\pi_n))$$

is also in MR.

Reversibility. The *reverse* of a ranking $\pi : \{1, 2, \dots, n\} \to \mathcal{C}$ is the ranking $\overleftarrow{\pi}$ defined by $\overleftarrow{\pi}_j = \pi_{n-j+1}$, that is

$$\langle \overleftarrow{\pi}_1, \overleftarrow{\pi}_2, \dots, \overleftarrow{\pi}_n \rangle \quad = \quad \langle \pi_n, \pi_{n-1}, \dots, \pi_1 \rangle.$$

A ranking model MR satisfies *reversibility* if, whenever a ranking distribution P is generated by MR, then so is also the probability distribution $\tilde{P}$ defined by

$$\tilde{P}(\pi) \quad = \quad P(\overleftarrow{\pi}).$$

L-decomposability. A ranking model satisfies *L-decomposability* if each probability distribution P on rankings, that satisfies the model, can be written in the following form:

$$P(\pi) = P_{\mathcal{C}}(\pi_1) \cdot P_{\mathcal{C} \setminus \{\pi_1\}}(\pi_2) \cdot \ldots \cdot P_{\{\pi_{n-1}, \pi_n\}}(\pi_{n-1}),$$

for some family of real numbers $P_X(y)$ defined for $X \subset \mathcal{C}$ with $2 \leq |X|$, $y \in X$, and such that $0 \leq P_X(y)$ and $\sum_{y \in X} P_X(y) = 1$. Intuitively, the meaning of the expression on the right hand side is that the probability of any ranking π can be expressed as the product of the probabilities of choosing alternative π_1 in $\mathcal{C}$, of choosing alternative π_2 in $\mathcal{C} \setminus \{\pi_1\}$, ..., and of choosing alternative π_{n-1} in $\{\pi_{n-1}, \pi_n\}$ for some alternative choice probabilities $P_X(y)$. Critchlow et al. (1991) show that a ranking model is L-decomposable if and only if according to it, the conditional probability of assigning rank $k+1$ to an alternative u, knowing which alternatives received ranks 1, 2, ..., k, is independent of the order of the latter alternatives. More precisely, for $k = 2, 3, \dots, n-1$ and for $u \in \mathcal{C}$, the function ξ_u given by the conditional probability

$$\xi_u(x_1, \dots, x_k) = P\Big(\pi(k+1) = u \mid \pi(1) = x_1, \pi(2) = x_2, \dots, \pi(k) = x_k\Big), \tag{8}$$

has to be symmetric.

Before we define strong unimodality, we need to introduce another concept. Given any two rankings π and σ of $\mathcal{C}$, an unordered pair $\{a, b\}$ of alternatives is an *inversion between* π and σ when $(\pi^{-1}(a) - \pi^{-1}(b)) \cdot (\sigma^{-1}(a) - \sigma^{-1}(b)) < 0$. The pair (a, b) is an *inversion from* π to σ when $\pi^{-1}(a) < \pi^{-1}(b)$ and $\sigma^{-1}(a) > \sigma^{-1}(b)$. Writing $\mathrm{Inv}(\pi, \sigma)$ for the set of all inversions between π and σ, the *symmetric difference distance* $d(\pi, \sigma)$ between π and σ (also known as Kendall's tau) is defined as $d(\pi, \sigma) = |\mathrm{Inv}(\pi, \sigma)|$.

Strong unimodality. A ranking distribution P satisfies *strong unimodality* if the following two conditions hold: 1) There exists a unique ranking π_0 that has maximal probability P, i.e. π_0 is a *modal ranking*, and 2) the ranking probabilities $P(\pi)$ are nonincreasing as we move along a path of rankings that are increasingly distant from π_0 in terms of the symmetric difference metric on rankings.

Complete Consensus. A ranking distribution P satisfies *complete consensus* if it satisfies strong unimodality with some modal ranking π_0 and, in addition, the following condition holds: For all alternatives i ranked before j in π_0 and for any ranking π in which i is ranked before j, the inequality $P(\pi) \geq P(\pi')$ holds, where π' is obtained from π be exchanging i and j.

THEOREM 3. *The repeated insertion model for rankings, RIMR, satisfies label invariance, reversibility and L-decomposability, but it violates strong unimodality and thus also complete consensus.*

The proof is provided in the Appendix.

We now proceed to show that RIMR encompasses the classical Mallows ϕ-model. Mallows' ϕ-model uses a numerical parameter ϕ with $0 \leq \phi \leq 1$, and a reference ordering ρ. It assigns to any ranking π the probability

$$P(\pi) = \frac{\phi^{d(\rho,\pi)}}{\sum_{\sigma \in \Pi} \phi^{d(\rho,\sigma)}}, \tag{9}$$

where d is the symmetric difference metric on rankings (and where, by convention, $0^0 = 1$).

PROPOSITION 4. *Mallows' ϕ-model is a submodel of the repeated insertion model for rankings. Moreover, the probability distribution defined by Equation (9) is generated by RIMR, i.e., by Equation (7), with the same reference ordering ρ and with*

$$w_i^k = \phi^{k-i} \frac{1-\phi}{1-\phi^k}$$

(notice $0 \leq w_i^k$ and $\sum_{i=1}^{k} w_i^k = 1$).

Another model that encompasses Mallows' ϕ-model is the 'multistage ranking model' (for rankings) due to Fligner and Verducci (1988), denoted here as MSMR.

Similarly to RIMR, one can think of MSMR as a process model according to which the respondent constructs a ranking π of the alternatives relative to some 'reference ordering', say, ζ. Here, when a respondent rank orders the alternatives, s/he considers each rank position of π in turn starting from the first rank position. In the first step, when assigning an alternative to rank one, s/he assigns alternative ζ_k with some probability $p(k-1, 1)$ (we will formally introduce the parameters below). At each step, s/he assigns to each rank position r an alternative a among the ones not yet placed in her/his ranking with some probability $p(m, r)$, where $m + 1$ is the relative rank of a in ζ among all alternatives that have not yet been assigned to any rank in π.

For a given *reference ordering* ζ and ranking π we denote by $g_{\zeta,\pi}(r)$ the number of alternatives ranked before π_r in ζ and with rank greater than r in π. Formally,

$$\begin{aligned} g_{\zeta,\pi}(r) &= \left|\left\{b \in \mathcal{C} \;:\; b\,\zeta\,\pi_r \text{ and } \pi_r\,\pi\,b\right\}\right| \\ &= \left|\left\{b \in \mathcal{C} \;:\; \zeta^{-1}(b) < \zeta^{-1}(\pi_r) \text{ and } r < \pi^{-1}(b)\right\}\right| \\ &= \left|\left\{u \in \{r+1, r+2, \ldots, n\} \;:\; \zeta^{-1}(\pi_u) < \zeta^{-1}(\pi_r)\right\}\right|. \end{aligned} \tag{10}$$

In MSMR, the ranking probabilities are parametrized by a discrete parameter, namely the reference ordering ρ, and by real-valued parameters $p(m, r)$, with $1 \leq r \leq n$ and $0 \leq m \leq n - r$, and where $0 \leq p(m, r)$ and $\sum_{m=0}^{n-r} p(m, r) = 1$ (thus $p(0, n) = 1$). Notice that, as in RIMR, the parameters consist again of a reference ordering plus $n(n-1)/2$ real-valued parameters. For any ranking π of $\mathcal{C}$, the *multistage model for rankings* MSMR assigns ranking probabilities P as follows:

$$P(\pi) \quad = \quad p(g_{\zeta,\pi}(1), 1) \cdot p(g_{\zeta,\pi}(2), 2) \cdot \ldots \cdot p(g_{\zeta,\pi}(n), n). \tag{11}$$

The apparent similarity between $g_{\zeta,\pi}$ here and $f_{\rho,\pi}$ in the definition of RIMR as well as the apparent similarity between Equations (7) and (11) call for an investigation of the relationship between MSMR and RIMR. We first note that the two models do not coincide, i.e., there exist probability distributions P that can be generated by one but not by both models. In particular, we first show that the nonoverlapping distributions are fundamental in our proof of LSMS being subsumed by SIMS: The proof of Theorem 2 cannot be carried out using MSMR instead of RIMR.

Proposition 5. *There exist probability distributions on subsets of $\mathcal{C}$ generated by LSMS, that cannot be generated by SIMS(MSMR).*

Corollary 6. *RIMR $\neq$ MSMR .*

The proof of the proposition as well as of the corollary is in the Appendix.

In fact, we already introduced an example of a distribution P that is generated by RIMR but that cannot be generated by MSMR. It is straightforward to check that the distribution obtained for the rankings in Example 2 belongs to RIMR but not to MSMR. To show this, we do not need to consider the 6 possible reference orderings for MSMR. Instead, we can reason indirectly, in two steps. First, we recall that with the size probabilities given in Example 2, this ranking distribution generates, according to SIMS(RIMR), the subset distribution that was obtained from the instance of LSMS in Example 1. Second, the proof of Proposition 5 implies that this particular instance of LSMS cannot be recovered from SIMS(MSMR) (since Equation (42) is violated for $a = 1$, $b = 2$ and $c = 3$; see the proof of Proposition 5 in the Appendix).

Although the two models formally differ (as stated in Corollary 6), one can think of MSMR as a 'dual' version of RIMR in which rank positions and objects are exchanged. This should be clear from the way Equations (7) and (11) were motivated. In fact, Fligner and Verducci (1988) alluded to RIMR in their discussion of MSMR by pointing out that, in the construction of MSMR, one could exchange objects and ranks to obtain a rather different model of ranking probabilities. They did, however, not name or study that alternative ranking model.

In more formal terms, the right-hand side of the first equation is transformed into the right-hand side of the second by setting $w_j^k = p(j-1, n+1-k)$ for $1 \leq j \leq k \leq n$.

As a consequence, for every probability distribution P generated by RIMR (or MSMR), there exists a probability distribution P' generated by MSMR (or RIMR) such that $\{P(\pi) : \pi \in \Pi\} = \{P'(\pi) : \pi \in \Pi\}$. In other words, there exists a permutation Ψ of the collection Π of all rankings of $\mathcal{C}$ such that for any probability distribution P generated by RIMR, the probability distribution Q defined by $Q(\pi) = P(\Psi(\pi))$ is in MSMR, and conversely. The permutation Ψ of Π can be made explicit as follows. With ρ and ζ the reference orderings of RIMR and MSMR, respectively, Ψ maps any ranking π onto the ranking γ such that $f_{\rho,\pi}$ coincides with $g_{\zeta,\gamma}$. We illustrate this 'metaisomorphism' Ψ in case $n = 3$.

Example 4. Figure 1 displays the ranking probabilities for RIMR on $\mathcal{C} = \{1, 2, 3\}$ with reference ordering $\rho = \langle 1, 2, 3\rangle$, and for MSMR on $\mathcal{C} = \{a, b, c\}$ with reference ordering $\langle a, b, c\rangle$ (notice our change of notation for the alternatives).

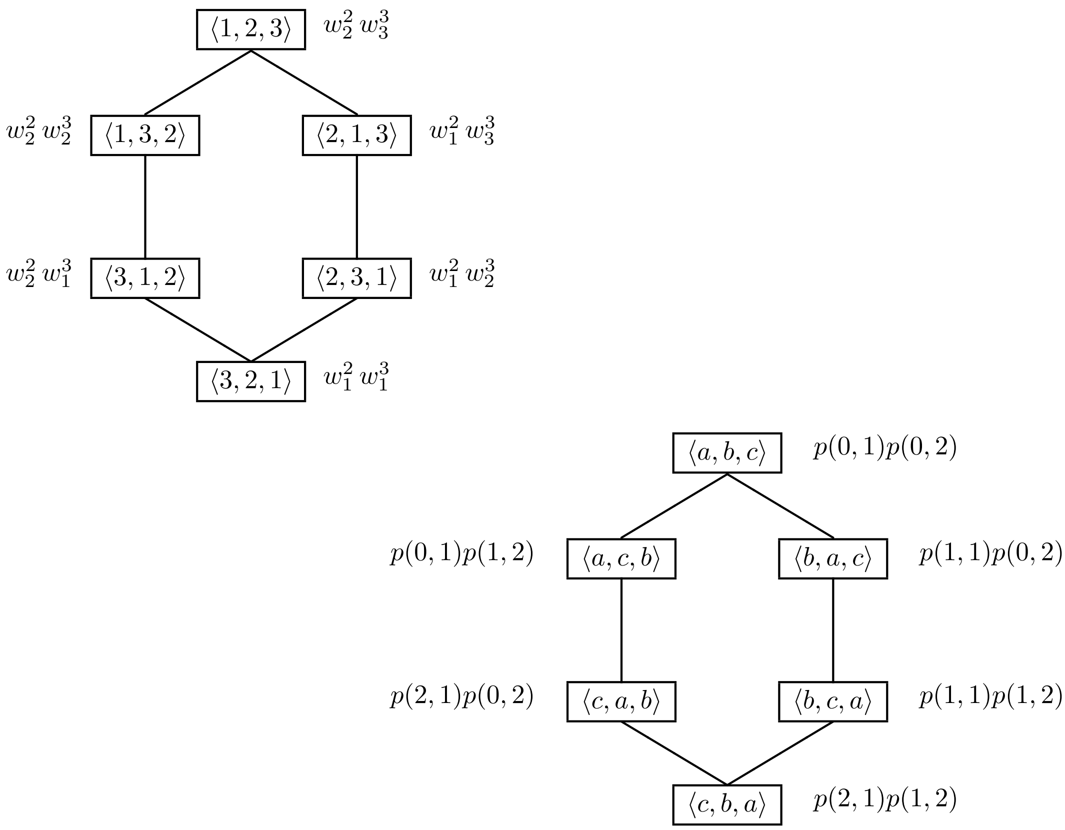

FIGURE 1.
The RIMR and MSMR probabilities in case $n = 3$ (see Example 4).

It can be checked that no bijection from $\{1, 2, 3\}$ to $\{a, b, c\}$ (acting on alternatives) transforms the left part of Figure 1 into the right part—whatever transformation from the w's to the p's is considered. On the other hand, the 'metaisomorphism' Ψ is the bijection, mapping rankings of $\{1, 2, 3\}$ to rankings of $\{a, b, c\}$:

$\Psi(\langle 1, 2, 3\rangle) = \langle c, b, a\rangle$, $\Psi(\langle 2, 1, 3\rangle) = \langle c, a, b\rangle$, $\Psi(\langle 2, 3, 1\rangle) = \langle b, a, c\rangle$, $\Psi(\langle 3, 2, 1\rangle) = \langle a, b, c\rangle$, $\Psi(\langle 3, 1, 2\rangle) = \langle a, c, b\rangle$, $\Psi(\langle 1, 3, 2\rangle) = \langle b, c, a\rangle$.

Despite this 'duality' defined by the metaisomorphism Ψ that relates the sets of rankings (and not the sets of alternatives), there are several properties that distinguish the two models. One such distinguishing property is reversibility: as noted in Theorem 3, RIMR satisfies reversibility; on the other hand, Critchlow, Fligner & Verducci (1991) shows that MSMR does not.

Another difference between the two models arises from the different types of independence properties that are inherent in the two models. In RIMR with reference ordering ρ, the relative (marginal) rank of $\rho(k)$ among alternatives $\{\rho(1), \rho(2), \ldots, \rho(k)\}$, when seen as a random variable, is independent of the relative (marginal) ranking of the subset $\{\rho(1), \rho(2), \ldots, \rho(k-1)\}$. On the other hand, in MSMR, the object at position k is independent of the relative positions in the reference ordering ζ of $\pi(k+1)$, $\pi(k+2)$, $\ldots$, $\pi(n)$. For $n \geq 3$, neither of these two independence properties is satisfied by the other model, as can be easily seen by taking $k = 1$ and computing the required conditional probabilities.

A related feature that distinguished the two models is the way in which the reference order in each model ranges from being nearly identifiable to being maximally nonidentifiable. There are ranking probabilities generated from RIMR with some specific reference ordering which cannot be generated under the same model with any other reference ordering, except that the first two objects in the reference ordering can always be exchanged (in which case, w_1^2 and w_2^2 also need to be exchanged). This means that in some situations, the reference ordering of RIMR is recoverable from the data, up to the relative position of the first two objects.

At the other extreme, for instance, the distribution setting all the mass onto one specific ranking can be generated from RIMR with the reference ordering being any ordering whatsoever. In contrast to RIMR, the lack of identifiability of the reference ordering in MSMR is different. Here, one can without loss of generality reverse the reference ordering and exchange other parameter values accordingly (e.g., if n=3, exchange $p(0, 1)$ with $p(2, 1)$ and exchange $p(0, 2)$ with $p(1, 2)$) to obtain the same ranking distribution. Also, as with RIMR, there are ranking distributions that can be generated from MSMR with the reference ordering being any ordering whatsoever. In summary, one commonality of both RIMR and MSMR is the fact that the reference ordering, in general, is not uniquely identifiable from a given set of data. As a consequence, even if it is conceptually implausible to assume that the reference ordering is the same for all members of the population under study, it is reasonable to state this as a technical assumption: its general lack of identifiability makes the reference order into a technical parameter that has limited substantive relevance.

We now illustrate RIMR, as well as its difference from MSMR, with an empirical application to data from three German national election survey (GNES) data sets collected in 1969, 1972 and 1976. These data are taken from Norpoth (1979) and provide rankings of the three major German parties, namely, C: Christian Democrats, F: Free Democrats and S: Social Democrats. Our results are reported in Tables 1-3.

We report the predicted relative frequencies under the best fitting RIMR and MSMR as well as the observed relative frequencies in the data. For each of these three cases we also provide the (maximum) log likelihood. The goodness-of-fit of each model is reported via a nested likelihood ratio test (G^2 = -2 log likelihood ratio) of the model against a multinomial benchmark model that perfectly fits the data. We also quantify the goodness-of-fit via the p-value of the G^2 test-statistic based on a χ^2 distribution with 2 degrees of freedom.

Table 1.
Predicted relative frequencies under the best fitting RIMR and MSMR models versus observed frequencies and goodness-of-fit results for the 1969 GNES. The sample size of the survey was 818.

Ranking	Data: Observed Relative Frequency	RIMR: Predicted Relative Frequency	MSMR: Predicted Relative Frequency
SCF	.32	.319	.313
SFC	.33	.332	.327
CSF	.15	.148	.154
CFS	.03	.031	.019
FSC	.03	.029	.041
FCS	.14	.142	.147
Log Likelihood	3.986	3.9457	.3152
G^2		.0808	7.3418
p-value (2 dgf)		.96	.025

Table 2.
Predicted relative frequencies under the best fitting RIMR and MSMR models versus observed frequencies and goodness-of-fit results for the 1972 GNES. The sample size of the survey was 1595.

Ranking	Data: Observed Relative Frequency	RIMR: Predicted Relative Frequency	MSMR: Predicted Relative Frequency
SCF	.185	.167	.180
SFC	.28	.334	.273
CSF	.14	.086	.147
CFS	.025	.043	.023
FSC	.04	.076	.042
FCS	.33	.294	.335
Log Likelihood	-2.1076	-60.365	-2.9418
G^2		116.52	1.668
p-value (2 dgf)		$< .001$	.434

Table 3.
Predicted relative frequencies under the best fitting RIMR and MSMR models versus observed frequencies and goodness-of-fit results for the 1976 GNES. The sample size of the survey was 1872.

Ranking	Data: Observed Relative Frequency	RIMR: Predicted Relative Frequency	MSMR: Predicted Relative Frequency
SCF	.08	.138	.131
SFC	.19	.132	.137
CSF	.28	.319	.334
CFS	.02	.039	.023
FSC	.06	.041	.057
FCS	.37	.332	.320
Log Likelihood	-11.288	-19.206	-17.179
G^2		15.836	11.782
p-value (2 dgf)		$< .001$	.003

As Table 1 shows, the 1969 GNES data are extremely well fit by RIMR (we find a p-value of .96 for G^2) but poorly fit by MSMR (we find a p-value of .025 for G^2). This data set therefore demonstrates that there are cases where RIMR greatly outperforms MSMR in terms of its ability to fit empirical data. However, as we can see from the analysis of 1972 GNES data that is presented in Table 2, there are also data sets where the opposite is true, i.e., where MSMR fits well, while RIMR does not fit at all (p-value $< .001$). Finally, the remaining data set that we have analyzed, namely the 1976 GNES, is a case where neither of the two models is able to provide a good account for the data.

The results presented in this section are aimed at providing basic insights into RIMR, and are by no means geared toward a comprehensive analysis of RIMR. As with any new ranking model, one could attempt to address many other potentially interesting questions about RIMR. For example, it would be interesting to characterize the ranking distributions in RIMR (respectively MSMR) that do not depend on the choice of reference ordering. It would also be interesting to characterize the intersection of RIMR and MSMR, as well as their union. When both models share the same reference ordering, it can be shown that the intersection depends on $n - 1$ parameters. However, for the two models to produce the same distribution (even a 'generic one'), it is not necessary that the reference orderings be either identical or opposite. We leave these questions for further research, as well as similar questions about the relationship between RIMR and other ranking models, such as the Babington Smith (1950) model class. We note in passing that the intersection of RIMR and the Babington Smith model class is strictly larger than the Mallows model class.

Conclusion

In this paper we have demonstrated the fundamental interconnection between subset choice models and ranking models. The size-independent model for subset choice can be viewed as a link between the two classes of models, since it provides a generic and natural way to transform probability distributions on rankings into probability distributions on subsets and vice versa. The results here clearly demonstrate that useful insights for the analysis of subset choice models could come from the analysis of certain ranking models, and that the analysis of subset choice models could yield new paradigms and interesting results for the study of ranking models.

In particular, we have shown that two seemingly unrelated models of subset choice, the latent scale model and the size-independent model, are not only related but that the latent scale model is subsumed in the size-independent model (Theorem 2). In order to prove this, we have ventured from the realm of subset choice models into the realm of ranking models by considering the repeated insertion model for rankings (Proposition 1).

We have also provided a basic analysis of the repeated insertion model and have shown that it is label invariant, reversible, L-decomposable, but not necessarily strongly unimodal (Theorem 3). Furthermore, we have shown that the repeated insertion model subsumes the Mallows model (Proposition 4), and we have investigated its (dis)similarity with the multistage model. Interestingly, in our first argument that identified differences between the repeated insertion model and the multistage model (Corollary 6), we have again utilized the link between subset choice and ranking models (Proposition 5). We have also explored the relative empirical applicability of RIMR and MSMR on empirical ranking data from three important political surveys.

We believe that this work is only the first step towards a deeper understanding of the connections between subset choice models and ranking models.

Appendix

Proof of Proposition 1. Clearly, all values of P are nonnegative since all w_i^k are assumed to be nonnegative. In order to show that $\sum_{\pi\in\Pi} P(\pi) = 1$, we partition the rankings on $\mathcal{C}$ according to the rank they assign to alternative ρ_n (the last object in the reference ordering). For any ranking π such that $\pi^{-1}(\rho_n) = j$, we know that the expression $P(\pi)$ is a product of factors, one of which is w_j^n. By assumption, $\sum_{j=1}^n w_j^n = 1$. Therefore, we only need to prove that

$$\sum_{\pi\in\Pi,\pi^{-1}(\rho_n)=j} \frac{P(\pi)}{w_j^n} = 1,$$

for $j = 1, 2, \ldots, n$. This equality follows by induction, where the repeated insertion model for rankings on $\mathcal{C} \setminus \{\rho_n\}$ uses the ranking $\langle\rho_1, \rho_2, \ldots, \rho_{n-1}\rangle$ and the same parameters w_j^k for $1 \le j \le k \le n-1$. Q.E.D.

Notice that the above proof relied upon the fact that deleting alternative ρ_n (i.e., the last alternative in the reference order) leads again to a RIMR, now on $\mathcal{C} \setminus \{\rho_n\}$. This heredity property is repeatedly used in several proofs below.

Before proving Theorem 2 we establish two Lemmas.

Consider the LSMS, writing again l_x for its parameters, with $x \in \mathcal{C}$. Denote by P_s the probability of choosing a subset of size s according to this model. We immediately have the following expression, where we use the notation $\mathcal{P}(\mathcal{C}, s) = \{X \in \mathcal{P}(\mathcal{C}) \;:\; |X| = s\}$:

$$P_s \quad = \sum_{X\in\mathcal{P}(\mathcal{C},s)} \prod_{x\in X} l_x \cdot \prod_{y\in\mathcal{C}\setminus X} \Big(1 - l_y\Big). \tag{12}$$

Lemma 7. *The sequence P_0, P_1, ..., P_n is log-concave, that is, for $s = 1, 2, \ldots, n-1$,*

$$(P_s)^2 \quad \ge \quad (P_{s-1}) \cdot (P_{s+1}). \tag{13}$$

Proof. Substituting (12) for each term in (13) we obtain

$$\left(\sum_{X\in\mathcal{P}(\mathcal{C},s)} \prod_{x\in X} l_x \prod_{y\in\mathcal{C}\setminus X} \Big(1 - l_y\Big)\right) \left(\sum_{X^*\in\mathcal{P}(\mathcal{C},s)} \prod_{x^*\in X^*} l_{x^*} \prod_{y^*\in\mathcal{C}\setminus X^*} \Big(1 - l_{y^*}\Big)\right) \ge$$
$$\left(\sum_{Z\in\mathcal{P}(\mathcal{C},s-1)} \prod_{z\in Z} l_z \prod_{r\in\mathcal{C}\setminus Z} \Big(1 - l_r\Big)\right) \left(\sum_{T\in\mathcal{P}(\mathcal{C},s+1)} \prod_{t\in T} l_t \prod_{u\in\mathcal{C}\setminus T} \Big(1 - l_u\Big)\right). \tag{14}$$

Notice that, on each side of Equation (14), distributing the product transforms the expression into a sum of nonnegative terms which are all products of $2n$ factors. Moreover, the product obtained on the right-hand side from the pair (Z, T) of subsets gives the same value as the product obtained on the left-hand side from the pair (X, X^*) of subsets as soon as the following two equations are satisfied:

$$X \cap X^* = Z \cap T, \qquad\qquad X \cup X^* = Z \cup T. \tag{15}$$

To establish Equation (14), it thus suffices to show the existence of an injective mapping $(Z, T) \mapsto (X, X^*)$ with Z, T, X, and X^* as in Equations (14) and (15). In turn, we only need to prove for any subsets A, B of $\mathcal{C}$, with $A \subseteq B$, $0 \le |A| \le s-1$ and $|B| = 2s - |A|$, that the following holds.

$$\Big|\{(X, X^*) \in \mathcal{P}(\mathcal{C}, s) \times \mathcal{P}(\mathcal{C}, s) \;:\; X \cap X^* = A, X \cup X^* = B\}\Big| \qquad \ge$$
$$\Big|\{(Z, T) \in \mathcal{P}(\mathcal{C}, s-1) \times \mathcal{P}(\mathcal{C}, s+1) \;:\; Z \cap T = A, Z \cup T = B\}\Big|. \tag{16}$$

With $a = |A|$, the left-hand side of Inequality (16) equals $\binom{2(s-a)}{s-a}$, while the right-hand side equals $\binom{2(s-a)}{s-a-1}$. Thus, Inequality (16) holds because the inequality $\binom{2c}{c} \geq \binom{2c}{c-1}$ among binomial coefficients holds for any choice of a natural number c. Q.E.D.

A central quantity entering SIMS is the probability $P(\Pi_X)$. We now provide an explicit formula for $P(\Pi_X)$ under RIMR. For any subset X of $\mathcal{C}$, the definition of Π_X given informally in the text can be stated formally as $\Pi_X = \left\{\pi \in \Pi \,:\, X = \{\pi_1, \pi_2, \ldots, \pi_{|X|}\}\right\}$. In particular, $\Pi_\emptyset = \Pi = \Pi_{\mathcal{C}}$.

LEMMA 8. *In RIMR on $\mathcal{C} = \{1, 2, \ldots, n\}$ with the canonical reference ordering, we have*

$$P(\Pi_X) \quad = \quad \prod_{k \in X} \left(\sum_{j=1}^{|X \cap \{1,2,\ldots,k\}|} w_j^k \right) \cdot \prod_{k \in \mathcal{C} \setminus X} \left(\sum_{j=|X \cap \{1,2,\ldots,k\}|+1}^{k} w_j^k \right). \tag{17}$$

PROOF. We proceed by induction on n. Let $\mathcal{C}' = \mathcal{C} \setminus \{n\}$ and denote the collection of all rankings on $\mathcal{C}'$ by Π'. For $Y \in \mathcal{P}(\mathcal{C}')$, we denote by $P'(\Pi'_Y)$ the total probability of the collection of all rankings on $\mathcal{C}'$ that start with Y (according to the repeated insertion model on $\mathcal{C}'$ with parameters w_j^k, where $1 \leq j \leq k \leq n-1$).

If $n \in X$, the rank j of n in any ranking starting with X takes a value in $\{1, 2, \ldots, |X|\}$. Thus, writing

$$W_X^k \quad = \quad \sum_{j=1}^{|X \cap \{1,2,\ldots,k\}|} w_j^k, \qquad \textit{and} \qquad \overline{W}_X^k \quad = \quad \sum_{j=|X \cap \{1,2,\ldots,k\}|+1}^{k} w_j^k,$$

we have

$$P(\Pi_X) = \sum_{j=1}^{|X|} \left(w_j^n \cdot P'(\Pi'_{X \setminus \{n\}}) \right) \tag{18}$$

$$= \left(\sum_{j=1}^{|X|} w_j^n \right) \cdot P'\left(\Pi'_{X \setminus \{n\}}\right) \tag{19}$$

$$= W_X^n \cdot \prod_{k \in X \setminus \{n\}} W_X^k \cdot \prod_{k \in \mathcal{C} \setminus X} \overline{W}_X^k, \tag{20}$$

which establishes Equation (17) in the case when $n \in X$.

Similarly, if $n \in \mathcal{C} \setminus X$, the rank of n in any ranking starting with X equals $|X|+1$, $|X|+2$, $\ldots$, or n. Hence, with the same abbreviations as above, we have

$$P(\Pi_X) = \sum_{j=|X|+1}^{n} \left(w_j^n \cdot P'(\Pi'_X) \right) \tag{21}$$

$$= \left(\sum_{j=|X|+1}^{n} w_j^n \right) \cdot P'(\Pi'_X) \tag{22}$$

$$= \overline{W}_X^n \cdot \prod_{k \in X} W_X^k \cdot \prod_{k \in \mathcal{C} \setminus (X \cup \{n\})} \overline{W}_X^k, \tag{23}$$

which establishes Equation (17) in the case when $n \notin X$. Q.E.D.

Proof of Theorem 2. Given any instance of the latent scale model LSMS with parameters l_i on $\mathcal{C} = \{1, 2, \ldots, n\}$ we use the canonical reference ordering without loss of generality. The theorem follows as soon as we specify parameters for SIMS(RIMR) in such a way that any probability distribution P generated by LSMS is also generated by the corresponding SIMS(RIMR).

Starting from the LSMS parameters l_i (with $0 \leq l_i \leq 1$ for $i = 1, 2, \ldots, n$), we proceed by induction. The cases $n = 1, 2$ are straightforward to check. As in the proof of Lemma 8, we set $\mathcal{C}' = \mathcal{C} \setminus \{n\}$. Moreover, for $Y \in \mathcal{P}(\mathcal{C}')$, we denote by $P'(\Pi'_Y)$ the probability of the collection of all rankings on $\mathcal{C}'$ that start with Y, according to the latent scale model on $\mathcal{C}'$ with the canonical ordering and parameters l_i, where $1 \leq i \leq n-1$. Using both LSMS on $\mathcal{C}$ and its restriction to $\mathcal{C}'$, we get the following relationships for subset sizes (where P'_s is the probability according to LSMS on $\mathcal{C}'$ of a subset of size s being chosen):

$$P_0 = P'_0\,(1 - l_n), \tag{24}$$

$$P_s = P'_s\,(1 - l_n) + P'_{s-1}\,l_n, \qquad \text{for } s = 1, 2, \ldots, n-1, \tag{25}$$

$$P_n = P'_{n-1}\,l_n. \tag{26}$$

Note that this is consistent with the fact established in Equation (12) that the subset size probabilities are completely determined from LSMS.

By the induction hypothesis, there exist parameters p'_s (for $s = 0, 1, \ldots, n-1$) and w'^k_j (for $1 \leq j \leq k \leq n-1$) such that SIMS(RIMR) (considered with the canonical ordering on $\mathcal{C}'$) generates the same probabilities for subsets of $\mathcal{C}'$ as the LSMS does with parameters l_1, $l_2, \ldots, l_{n-1}$. We now define an instance of SIMS(RIMR) on $\mathcal{C}$ that generates the same subset probabilities $P(X)$ as those generated by LSMS. Take the reference ordering to be the canonical ordering. Then, according to Equations (24) - (26), the size parameters in SIMS(RIMR) must be

$$p_s = P_s, \qquad \text{for } s = 0, 1, \ldots, n. \tag{27}$$

Now, let $w^k_j = w'^k_j$ for $1 \leq j \leq k \leq n-1$. To define the w^n_k's, we first treat the 'generic' case in which both $l_n \neq 0$ and $P_s \neq 0$ for $s = 1, 2, \ldots, n$. Let

$$w^n_1 \quad = \quad \frac{P'_0}{P_1} \cdot l_n, \tag{28}$$

which is obviously nonnegative, and, for $s = 2, 3, \ldots, n$, let

$$w^n_s \quad = \quad \left(\frac{P'_{s-1}}{P_s} - \frac{P'_{s-2}}{P_{s-1}}\right) \cdot l_n. \tag{29}$$

To check that $0 \leq w^n_s$, we only need to verify that

$$P'_{s-1} \cdot P_{s-1} \quad \geq \quad P'_{s-2} \cdot P_s, \tag{30}$$

because P_s, P_{s-1}, and l_n are all nonnegative. Using Equation (25) in the case when $s < n$, we can rewrite the latter inequality as

$$P'_{s-1} \cdot \Big(P'_{s-1}\,(1 - l_n) + P'_{s-2}\,l_n\Big) \quad \geq \quad P'_{s-2} \cdot \Big(P'_s\,(1 - l_n) + P'_{s-1}\,l_n\Big). \tag{31}$$

Therefore,

$$\left(\left(P'_{s-1}\right)^2 - P'_{s-2} \cdot P'_s\right) \cdot (1 - l_n) \quad \geq \quad 0. \tag{32}$$

Thus, the nonnegativeness of w^n_s follows from the definition of l_n and Lemma 7. The case when $s = n$ is an immediate consequence of Equation (26) and Equation (25) for $s = n-1$. Next, using Equations (26) and (28) - (29), we note that

$$w^n_1 + w^n_2 + \ldots + w^n_n = \frac{P'_{n-1}}{P_n} \cdot l_n = 1. \tag{33}$$

It remains to show that SIMS(RIMR) with the parameters just defined generates subset probabilities identical to those of the given LSMS. Take $X \in \mathcal{P}(\mathcal{C})$, and consider two cases.

First, suppose $n \notin X$.

According to SIMS(RIMR), using (22), (25), (27) with $s = |X|$ and (33), the probability of X equals

$$\begin{aligned} p_s \cdot P(\Pi_X) &= p_s \cdot \left(\sum_{j=s+1}^{n} w_j^n \right) \cdot P'(\Pi'_X) \\ &= p_s \cdot \left(1 - \sum_{j=1}^{s} w_j^n \right) \cdot P'(\Pi'_X) \\ &= p_s \cdot \left(1 - \frac{P'_{s-1}}{P_s}\, l_n \right) \cdot P'(\Pi'_X) \\ &= \left(P_s - P'_{s-1}\, l_n \right) \cdot P'(\Pi'_X) \\ &= P'_s \cdot (1 - l_n) \cdot P'(\Pi'_X) \\ &= P'(X) \cdot (1 - l_n). \end{aligned}$$

On the other hand, the probability of X according to LSMS is

$$P'(X) \cdot (1 - l_n), \tag{34}$$

since $n \notin X$. Hence, LSMS and SIMS(RIMR) both assign the same probability to the subset X.

Second, we assume $n \in X$. According to SIMS(RIMR), using (19) and (27), the probability of X is, with $s = |X|$,

$$\begin{aligned} p_s \cdot P(\Pi_X) &= p_s \cdot \left(\sum_{j=1}^{s} w_j^n \right) \cdot P'(\Pi'_{X \setminus \{n\}}) \\ &= p_s \cdot \frac{P'_{s-1}}{P_s}\, l_n \cdot P'(\Pi'_{X \setminus \{n\}}) \\ &= P'(X \setminus \{n\}) \cdot l_n. \end{aligned}$$

According to LSMS, the probability of X is

$$P'(X \setminus \{n\}) \cdot l_n, \tag{35}$$

since $n \in X$. Thus, again, the two models assign the same probability to X. Hence, they generate the same probability distribution on $\mathcal{P}(\mathcal{C})$.

This completes the proof in the 'generic' case. To treat the other cases, assume first $l_n = 0$. Here we set $w_n^n = 1$ (and thus $w_k^n = 0$ for $1 \le k \le n-1$). An induction argument readily yields that the resulting SIMS(RIMR) generates the same probability distribution P as LSMS does. The case $l_n = 1$ is analogous.

Finally, consider the case $P_s = 0$ for some $s \in \{1, 2, \ldots, n\}$ and $0 < l_n < 1$. From Equation (12) and the assumption that P is generated by LSMS, there exist sizes u, $v \in \{0, 1, \ldots, n\}$ such that

$$\begin{aligned} &P_0 = P_1 \ldots = P_{u-1} = 0, \\ &P_u \neq 0, P_{u+1} \neq 0, \ldots, P_v \neq 0, \\ &P_{v+1} = P_{v+2} = \ldots = P_n = 0. \end{aligned}$$

(In fact, $u = |\{x \in \mathcal{C} \ : \ l_x = 1\}|$ and $v = |\{x \in \mathcal{C} \ : \ L_x > 0\}|$.) From Equations (25)-(26), we see that if a denominator vanishes in Equations (28)-(29), so does the corresponding numerator. Set the resulting indefinite quotients to zero, except for

$$\frac{P'_v}{P_{v+1}} = \frac{1}{l_n}. \tag{36}$$

It is straightforward to check, along the same lines as in the 'generic case', that the above settings define an instance of SIMS(RIMR) which generates the given probability distribution P. Q.E.D.

Notice that changing the order in which we consider the choice alternatives in our inductive proof will in general lead to another instance of SIMS(RIMR).

Proof of Theorem 3. Label invariance follows immediately from the fact that any relabeling of the objects also relabels the objects in the reference ordering.

Reversibility holds because, if P is obtained for a reference ordering ρ and numerical parameters w_j^k (where $1 \leq j \leq k \leq n$), then $\tilde{P}$ is obtained with the same reference ordering ρ and parameters $\tilde{w}_j^k$ given by $\tilde{w}_j^k = w_{k-j+1}^k$.

The property of L-decomposability follows if we can prove that Expression (8) is symmetric in alternatives $x_1, x_2, \ldots, x_k$ for any choice of $x_1, x_2, \ldots, x_k$. First, notice that

$$\frac{P(\langle x_1, x_2, \ldots, x_k, u, z_{k+2}, z_{k+3}, \ldots, z_n\rangle)}{P(\langle y_1, y_2, \ldots, y_k, u, z_{k+2}, z_{k+3}, \ldots, z_n\rangle)} \tag{37}$$

$$= \frac{P(\langle x_1, x_2, \ldots, x_k, t_{k+1}, t_{k+2}, \ldots, t_n\rangle)}{P(\langle y_1, y_2, \ldots, y_k, t_{k+1}, t_{k+2}, \ldots, t_n\rangle)} \tag{38}$$

$$= C_X, \tag{39}$$

where

$$\{y_1, y_2, \ldots, y_k\} = \{x_1, x_2, \ldots, x_k\} = X, \text{ and}$$
$$\{u, z_{k+2}, z_{k+3}, \ldots, z_n\} = \{t_{k+1}, t_{k+2}, \ldots, t_n\} = \mathcal{C} \setminus X.$$

In other words, the quantity (37) depends only on the set X, hence its notation C_X. This follows immediately from the definition of RIMR, as we now show. After the numerator and denominator in Equation (37) are expanded using Equation (7), they involve exactly the same w_j^i for the superscripts i that satisfy $\rho_i \in \mathcal{C} \setminus X$. The same is true for (38). Similarly, the numerators in (37) and (38) involve identical w_j^i for the superscripts i that satisfy $\rho_i \in X$. The same holds for the denominators of (37) and (38). This completes the proof of (37)-(39).

Now, for any $x_1, \ldots, x_k$ and $y_1, \ldots, y_k$ such that $\{x_1, \ldots, x_k\} = \{y_1, \ldots, y_k\} = X$, we can expand (8) as follows.

$$\begin{aligned}
\xi_u(x_1, \ldots, x_k) &= \frac{\displaystyle\sum_{\substack{z_{k+2}, z_{k+3}, \ldots, z_n \\ \{z_{k+2}, z_{k+3}, \ldots, z_n\} = \mathcal{C}\setminus(X \cup \{u\})}} P(\langle x_1, x_2, \ldots, x_k, u, z_{k+2}, z_{k+3}, \ldots, z_n\rangle)}{\displaystyle\sum_{\substack{t_{k+1}, t_{k+2}, \ldots, t_n \\ \{t_{k+1}, t_{k+2}, \ldots, t_n\} = \mathcal{C}\setminus X}} P(\langle x_1, x_2, \ldots, x_k, t_{k+1}, t_{k+2}, \ldots, t_n\rangle)} \\
&= \frac{\displaystyle\sum_{\substack{z_{k+2}, z_{k+3}, \ldots, z_n \\ \{z_{k+2}, z_{k+3}, \ldots, z_n\} = \mathcal{C}\setminus(X \cup \{u\})}} C_X P(\langle y_1, y_2, \ldots, y_k, u, z_{k+2}, z_{k+3}, \ldots, z_n\rangle)}{\displaystyle\sum_{\substack{t_{k+1}, t_{k+2}, \ldots, t_n \\ \{t_{k+1}, t_{k+2}, \ldots, t_n\} = \mathcal{C}\setminus X}} C_X P(\langle y_1, y_2, \ldots, y_k, t_{k+1}, t_{k+2}, \ldots, t_n\rangle)} \\
&= \frac{\displaystyle\sum_{\substack{z_{k+2}, z_{k+3}, \ldots, z_n \\ \{z_{k+2}, z_{k+3}, \ldots, z_n\} = \mathcal{C}\setminus(X \cup \{u\})}} P(\langle y_1, y_2, \ldots, y_k, u, z_{k+2}, z_{k+3}, \ldots, z_n\rangle)}{\displaystyle\sum_{\substack{t_{k+1}, t_{k+2}, \ldots, t_n \\ \{t_{k+1}, t_{k+2}, \ldots, t_n\} = \mathcal{C}\setminus X}} P(\langle y_1, y_2, \ldots, y_k, t_{k+1}, t_{k+2}, \ldots, t_n\rangle)} \\
&= \xi_u(y_1, y_2, \ldots, y_k).
\end{aligned}$$

Strong unimodality is not always satisfied by RIMR. For instance, with $n = 3$, the reference ordering is the modal ranking as soon as $w_2^2 > w_1^2$ and $w_3^3 > \max(w_2^3, w_1^3)$. If, however, this is combined with $w_1^3 > w_2^3$, then strong unimodality is violated. A counter-example can be easily built for $n = 3$.

Since complete consensus is a special case of strong unimodality, and since the latter fails, the former does too. Q.E.D.

On the other hand, some instances of RIMR, such as the Mallows ϕ-model, discussed below, are well known to satisfy complete consensus and thus strong unimodality (Critchlow et al., 1991).

Let us also note the straightforward (but somewhat tedious to prove) fact that a probability distribution generated by RIMR is strongly unimodal with ρ being the modal ranking if and only if $w_{i+1}^k \geq w_i^k$ for $1 \leq i < k \leq n$ with a strict inequality when $i = k - 1$.

Proof of Proposition 4. First note that for a fixed k, the number of inversions (j, k) from ρ to π equals $k - f_{\rho,\pi}(k)$. The probability of any ranking π in Π according to RIMR is

$$\begin{aligned}
P(\pi) &= w^1_{f_{\rho,\pi}(1)} \cdot w^2_{f_{\rho,\pi}(2)} \cdot \ldots \cdot w^n_{f_{\rho,\pi}(n)} \\
&= \phi^{1-f_{\rho,\pi}(1)} \cdot \phi^{2-f_{\rho,\pi}(2)} \cdot \ldots \cdot \phi^{n-f_{\rho,\pi}(n)} \cdot \frac{(1-\phi)^n}{\prod_{k=1}^n (1-\phi^k)} \\
&= \phi^{d(\rho,\pi)} \cdot \frac{(1-\phi)^n}{\prod_{k=1}^n (1-\phi^k)} \\
&= \phi^{d(\rho,\pi)} \cdot \frac{1}{\sum_{\sigma\in\Pi} \phi^{d(\rho,\sigma)}},
\end{aligned}$$

which is the probability of the same ranking according to the Mallows model. Hence, with this choice of parameters, we get the Mallows ϕ-model. Q.E.D.

Proof of Proposition 5. Suppose some probability distribution P on $\mathcal{P}(\mathcal{C})$ is generated both by SIMS(MSMR) and by LSMS with parameters l_i, $0 < l_i < 1$, where $i \in \mathcal{C}$. Let a, b, c be the first three elements of the reference ordering ζ of SIMS(MSMR), listed in that order. By the constant ratio condition (2) of LSMS we know that

$$P(\{a,b\}) \cdot P(\{c\}) \quad = \quad P(\{b\}) \cdot P(\{a,c\}).$$

Considering that P also comes fom an instance of SIMS(MSMR) with parameters $p(m,r)$, we obtain

$$(p(0,1)\,p(0,2) + p(1,1)\,p(0,2)) \cdot p(2,1) \quad = \quad p(1,1) \cdot (p(0,1)\,p(1,2) + p(2,1)\,p(0,2)),$$

from which we derive

$$p(0,2) \cdot p(2,1) \quad = \quad p(1,1) \cdot p(1,2), \tag{40}$$

because $p(0,1) \neq 0$ follows from $P(\{a\}) = l_a \cdot \left(\prod_{i\in\mathcal{C}\setminus\{a\}} (1-l_i)\right) \neq 0$. Similarly,

$$P(\{a,c\}) \cdot P(\{b\}) \quad = \quad P(\{a\}) \cdot P(\{b,c\})$$

gives

$$p(0,2) \cdot p(1,1) \quad = \quad p(0,1) \cdot p(1,2). \tag{41}$$

Now, dividing Equation (40) by Equation (41), we get

$$(p(1,1))^2 \quad = \quad p(0,1) \cdot p(2,1).$$

Since P is generated by SIMS(MSMR), it follows that

$$(P(\{b\}))^2 \quad = \quad P(\{a\}) \cdot P(\{c\}),$$

and, because P is generated by LSMS,

$$\begin{aligned}
((1-l_a)l_b(1-l_c))^2 &= (l_a(1-l_b)(1-l_c)) \cdot ((1-l_a)(1-l_b)l_c) \\
\left(\frac{l_b}{1-l_b}\right)^2 &= \frac{l_a}{1-l_a} \cdot \frac{l_c}{1-l_c}.
\end{aligned} \tag{42}$$

Thus, whenever $0 < l_i < 1$ for all $i \in \mathcal{C}$ with (42) being violated, the probability distribution generated by LSMS cannot be generated by SIMS(MSMR). For example, take $\{l_a, l_b, l_c\} = \{\frac{1}{2}, \frac{1}{2}, \frac{1}{3}\}$. Q.E.D.

Proof of Corollary 6. By Theorem 2, we have LSMS $\subseteq$ SIMS(RIMR). By Proposition 5, we have LSMS $\not\subseteq$ SIMS(MSMR). Hence, SIMS(RIMR) $\neq$ SIMS(MSMR) and, thus, RIMR $\neq$ MSMR. Q.E.D.

References

Ahn, H. (1995). Nonparametric 2-stage estimation of conditional choice probabilities in a binary choice model under uncertainty. *Journal of Econometrics*, 67:337–378.

Atkinson, D., Wampold, B., Lowe, S., Matthews, L., and Ahn, H. (1998). Asian american preferences for counselor characteristics: Application of the Bradley-Terry-Luce model to paired comparison data. *Counseling Psychologist*, 26:101–123.

Babington Smith, B. (1950). Discussion on Professor Ross's paper. *Journal of the Royal Statistical Society, Series B*, 12:153–162.

Baier, D. and Gaul, W. (1999). Optimal product positioning based on paired comparison data. *Journal of Econometrics*, 89:365–392.

Baltas, G. and Doyle, P. (2001). Random utility models in marketing research: A survey. *Journal of Business Research*, 51:115–125.

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

Billot, A. and Thisse, J. (1999). A discrete choice model when context matters. *Journal of Mathematical Psychology*, 43:518–538.

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

Bockenholt, U. and Dillon, W. (1997). Modeling within-subject dependencies in ordinal paired comparison data. *Psychometrika*, 62:411–434.

Chen, Z. and Kuo, L. (2001). A note on the estimation of the multinomial logit model with random effects. *American Statistician*, 55:89–95.

Courcoux, P. and Semenou, M. (1997). Preference data analysis using a paired comparison model. *Food Quality and Preference*, 8:353–358.

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

Critchlow, D.E., Fligner, M.A., and Verducci, J.S., editors (1993). *Probability Models and Statistical Analyses for Ranking Data*. Springer, New York.

Doignon, J.-P. and Fiorini, S. (in press). The approval-voting polytope: Combinatorial interpretation of the facets. *Mathématiques, Informatique et Sciences Humaines*.

Doignon, J.-P. and Fiorini, S. (2002). The facets and the symmetries of the approval-voting polytope. Manuscript under Review.

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

Doignon, J.-P. and Regenwetter, M. (2002). On the cominatorial structure of the approval voting polytope. *Journal of Mathematical Psychology*, 46:554-563.

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

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

Fishburn, P. (2001). Signed orders, choice probabilities, and linear polytopes. *Journal of Mathematical Psychology*, 45:53–80.

Fligner, M.A., and Verducci, J.S. (1988). Multi-stage ranking models. *Journal of the American Statistical Association*, 83:892–901.

Huang, J. and Nychka, D. (2000). A nonparametric multiple choice method within the random utility framework. *Journal of Econometrics*, 97:207–225.

Ichimura, H. and Thompson, T.S. (1998). Maximum likelihood estimation of a binary choice model with random coefficients of unknown distribution. *Journal of Econometrics*, 86:269–295.

Knuth, D.E. (1973). *The Art of Computing Programming* (vol. 3: Sorting and Searching, 2nd printing). Addison-Wesley, Menlo Park, CA.

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

Lewbel, A. (2000). Semiparametric qualitative response model estimation with unknown heteroscedasticity or instrumental variables. *Journal of Econometrics*, 97:145–177.

Mallows, C.L. (1957). Non-null ranking models I. *Biometrika*, 44:114–130.

Marden, J. (1992). Use of nested orthogonal contrasts in analyzing rank data. *Journal of the American Statistical Association*, 87:307 – 318.

Marden, J.I. (1995). *Analyzing and Modeling Rank Data*. Chapman & Hall, London.

Marley, A.A.J. (1993). Aggregation theorems and the combination of probabilistic rank orders. In Critchlow, D. E., Fligner, M. A., and Verducci, J. S., editors, *Probability Models and Statistical Analyses for Ranking Data*, pages 216–240. Springer, New York.

McFadden, D. and Train, K. (2000). Mixed MNL models for discrete response. *Journal of Applied Econometrics*, 15:447–470.

Norpoth, H. (1979). The parties come to order! Dimensions of preferential choice in the West German Electorate, 1961-1976. *The American Political Science Review*, 73:724–736.

Peterson, G. and Brown, T. (1998). Economic valuation by the method of paired comparison, with emphasis on evaluation of the transitivity axiom. *Land Economics*, 74:240–261.

Regenwetter, M. and Doignon, J.-P. (1998). The choice probabilities of the latent-scale model satisfy the size-independent model when n is small. *Journal of Mathematical Psychology*, 42:102–106.

Regenwetter, M., Marley, A., and Joe, H. (1998). Random utility threshold models of subset choice. *Australian Journal of Psychology: Special issue on mathematical psychology*, 50:175–185.

Tsai, R. (2000). Remarks on the identifiability of Thurstonian ranking models: Case v, case III, or neither? *Psychometrika*, 65:233–240.

Zeng, L. (2000). A heteroscedastic generalized extreme value discrete choice model. *Sociological Methods & Research*, 29:118–144.