Journal of Mathematical Psychology 104 (2021) 102596

Contents lists available at ScienceDirect

Journal of Mathematical Psychology

journal homepage: www.elsevier.com/locate/jmp

A representation theorem for finite best–worst random utility models

Hans Colonius

University of Oldenburg, Germany

ARTICLE INFO

Article history:
Received 7 September 2020
Received in revised form 20 August 2021
Accepted 30 August 2021
Available online xxxx

This paper is dedicated to the memory of Tony Marley. Without his encouragement, it would not exist

Keywords:
Best–worst choices
Random utility theory
Block–Marschak polynomials

ABSTRACT

This paper investigates the representation of best–worst choice probabilities (picking the best and the worst alternative from an offered set). It is shown that non-negativity of best–worst Block–Marschak polynomials is necessary and sufficient for the existence of a random utility representation. The representation theorem is obtained by extending proof techniques for a corresponding result on best choices (picking the best alternative from an offered set) developed by Falmagne (1978).

Choosing an element from an offered set of alternatives is arguably the most basic paradigm of preference behavior. Typically, if the same set is offered several times, the chosen element will not always be the same. This is often attributed to the participant's preference fluctuating over time due to the effect of various alternatives to be compared, or to the difficulty of distinguishing between similar alternatives. Theories of choice behavior try to account for the probability $P_B(a)$ of choosing an alternative a, say, from an offered set B, which is a subset of the base set A (*best-choices paradigm*). In a related development, Finn and Louviere (1992) proposed a discrete choice task in which a participant is asked to select both the best and the worst option in an available (sub)set of options. For an offered set $B \subseteq A$, let $BW_B(a, b)$ be the probability that a participant chooses a as best and b as worst alternative in set B (*best–worst choices paradigm*). Best–worst choices contain a great deal of information about the person's ranking of options. As observed in Marley and Louviere (2005), if there are, e.g., 4 items in A, one obtains information about the best option in 9 out of 11 possible non-empty, non-singleton subsets of A. Applications of this best–worst choice paradigm have strongly increased over the last years.[1] In particular, best–worst scaling is being used as a method of collecting ranking data which is then modeled in various ways related to the multinomial logit for best choices or to weighted versions of the rank-ordered logit for repeated best choices (for details, see the monograph by Louviere, Flynn, & Marley, 2015, p.11 pp.)

Marley and colleagues have developed various models for the best–worst choices paradigm (Marley & Louviere, 2005; see also Marley & Regenwetter, 2017, p. 400 pp.). However, one problem has apparently remained unsolved up to now: What are necessary and sufficient conditions on probabilities $BW_B(a, b)$ for the existence of a random utility representation? After giving a precise statement of this problem, we present a complete answer to the question. The solution turns out to be a modification of an approach to the analogous problem for the best-choice paradigm developed by Falmagne (1978). He showed that non-negativity of certain linear combinations of the best-choice probabilities (so-called *Block–Marschak polynomials*) is both necessary and sufficient for a random utility representation of best-choice probabilities. In this note, we extend Falmagne's method to obtain a solution to the above problem. Because large parts of the proof are completely parallel to those in Falmagne (1978), we relegate most details and proofs to an online supplement and only present relevant definitions and a sketch of the proof in this note.

1. Some definitions

For a finite set A, we write $|A|$ for the number of elements in A, $\mathscr{P}(A)$ for the power set of A, $\mathscr{P}(A, i)$ for the set of all subsets of A containing exactly i elements. For any nonempty set A, finite or not, let $\Phi(A)$ be the set of all finite, nonempty subsets of A.

Definition 1. Let A be a nonempty set of n elements ($n \geq 2$). For any $B \subseteq A$, with $|B| \geq 2$, and any $a, b \in B$, $a \neq b$, let $BW_B(a, b) \mapsto [0, 1]$ denote the probability that a and b are respectively chosen as best and worst elements in the subset B of

E-mail address: hans.colonius@uni-oldenburg.de.
URL: https://uol.de/en/hans-colonius/.

[1] A nearly exponential increase in papers and citations since 2000, with 118 publications and 2365 citations in 2020 (Web-of-Science, August 2021).

https://doi.org/10.1016/j.jmp.2021.102596

A. Let $\mathbb{BWP} = \{BW_B \mid B \subseteq A, BW_B(a,b) \mapsto [0,1], a,b \in B, a \neq b\}$ be the collection of all those probability measures. Suppose that

$$\sum_{a,b \in B,\, a \neq b} BW_B(a,b) = 1 \qquad (BW_B \in \mathbb{BWP}).$$

Then $(A, \mathbb{BWP})$ is called a *finite system of best–worst choice probabilities*, or more briefly, a *system*.

Definition 2. Let $(A, \mathbb{BWP})$ be a finite system of best–worst choice probabilities. Suppose that there is collection of jointly distributed random variables $\{U_c \mid c \in A\}$, with U_c measuring the variable value of alternative $c \in A$. For any subset $B \subseteq A$ and $a, b \in B$, with $a \neq b$, define the two events

$$\mathbf{B}_{a,B} = \{U_a = \max_{c \in B} U_c\}$$
$$\mathbf{W}_{b,B} = \{U_b = \min_{c \in B} U_c\}.$$

Thus, for the subset B, $\mathbf{B}_{a,B}$ is the event that random variable U_a has the maximum value compared to all other elements in the set B, and $\mathbf{W}_{b,B}$ is the corresponding event that random variable U_b has the minimum value compared to all other elements in the set B.

Definition 3. A finite system $(A, \mathbb{BWP})$ is called *a best–worst random utility system* if there exists a collection of jointly distributed random variables $\{U_c \mid c \in A\}$ such that for all $B \subseteq A$ and $a, b \in B$, $a \neq b$,

$$BW_B(a,b) = P(\mathbf{B}_{a,B} \cap \mathbf{W}_{b,B})$$

The collection $\{U_c \mid c \in A\}$ will be *called a random representation of* $(A, \mathbb{BWP})$.

From now on, we will always assume the system to be finite without mentioning it specifically. Consider the 2-element case $\{a, b\}$. A simple, but important implication of Definition 3 is the following: for any distinct $a, b \in A$

$$P(U_a = U_b) = 0.$$

Thus, a random utility representation of best–worst choice probabilities does not allow for ties.

2. Best–worst Block–Marschak polynomials

In analogy to Falmagne's Block–Marschak polynomials, we define a best–worst version of them.

Definition 4. For any $B \subset A$, $B \neq A$, $a, b \in A \setminus B$, $a \neq b$, in a system $(A, \mathbb{BWP})$,

$$K_{ab,B} = \sum_{i=0}^{|B|} (-1)^i \sum_{C \in \mathscr{P}(B,|B|-i)} BW_{A\setminus C}(a,b). \tag{1}$$

The $K_{ab,B}$ are called *best–worst Block–Marschak polynomials of* $(A, \mathbb{BWP})$, or *best–worst BM polynomials*, for short.

Observe that

$$\begin{aligned}
K_{ab,\emptyset} &= \sum_{i=0}^{0} (-1)^i \sum_{C \in \mathscr{P}(\emptyset,0-i)} BW_{A\setminus C}(a,b) = BW_A(a,b);\\
K_{ab,\{c\}} &= BW_{A\setminus\{c\}}(a,b) - BW_A(a,b)\\
&= BW_{A\setminus\{c\}}(a,b) - K_{ab,\emptyset};\\
K_{ab,\{c,d\}} &= BW_{A\setminus\{c,d\}}(a,b) - [BW_{A\setminus\{c\}}(a,b) + BW_{A\setminus\{d\}}(a,b)]\\
&\quad + BW_A(a,b)\\
&= BW_{A\setminus\{c,d\}}(a,b) - K_{ab,\{c\}} - K_{ab,\{d\}} - K_{ab,\emptyset}.
\end{aligned}$$

This suggests a heuristic interpretation of the $K_{ab,B}$ as the probability of all rankings of A with a as best and b as worst while ignoring all alternatives that are in B. The following lemma is implied by a Möbius inversion argument.

Lemma 5. *Let $(A, \mathbb{BWP})$ be a system of best–worst choice probabilities. Then, for all $B \subset A$, $B \neq A$, and $a, b \in A \setminus B$,*

$$BW_{A\setminus B}(a,b) = \sum_{C \in \mathscr{P}(B)} K_{ab,C}. \tag{2}$$

3. The main theorem

We can now state the main theorem of this note.

Theorem 6. *A finite system of best–worst choice probabilities is a best–worst random utility system if and only if the best–worst Block–Marschak polynomials are nonnegative.*

Necessity follows from the definition of best–worst BM polynomials and its proof is found in the online supplement.

3.1. Sufficiency proof (sketch)

We have to show that if

$$K_{ab,B} \geq 0$$

for all $B \subset A$, $B \neq A$ and $a, b \in A\setminus B$, then $(A, \mathbb{BWP})$ is a best–worst random utility system. A critical step in the proof is based on the following variant of a well-known result by Block and Marschak (1960).

Lemma 7. *Let Π be the set of all permutations of base set A and $S(aBb)$ the subset of all permutations of A such that a is chosen as best and b as worst element of the offered subset B. $(A, \mathbb{BWP})$ is a best–worst random utility system if and only if there exists a probability measure $\mathrm{P}[\,.\,]$ on $\mathscr{P}(\Pi)$ satisfying*

$$BW_B(a,b) = \mathrm{P}[S(aBb)] \tag{3}$$

for all $B \in \Phi(A)$ and $a, b \in B$.

By this lemma, the remainder of the proof consists in constructing $\mathrm{P}[\,.\,]$ on $\mathscr{P}(\Pi)$. We introduce a generalization of the subsets $S(aBb)$.

Definition 8. Let $|A| = n$ with $n > 2$ and $B \subseteq A$ with $|B| = n - m$ $(n \geq m)$; for $B \in \Phi(A)$ and distinct $b_1, b_2, \ldots, b_{k'}, b_{k'+1}, b_{k'+2}, \ldots, b_k \in B$ $(1 \leq k' < k \leq n - m)$, define

$$\begin{aligned}
&\mathrm{S}(b_1 b_2 \ldots b_{k'} B b_{k'+1} b_{k'+2} \ldots b_k) =\\
&\quad \{\pi \in \Pi \mid \pi(b_1) > \pi(b_2) > \cdots > \pi(b_{k'}) > \pi(b)\\
&\quad > \pi(b_{k'+1}) > \pi(b_{k'+2}) > \cdots > \pi(b_k),\\
&\quad \text{for all } b \in B \setminus \{b_1, \ldots, b_k\}\}.
\end{aligned}$$

In a first step, we define subsets of Π by

$$\mathrm{S}(a_1 a_2 \ldots a_{k'} A a_{k'+1} a_{k'+2} \ldots a_k),$$

containing all rankings π of A such that $\pi(a_1) > \pi(a_2) > \cdots > \pi(a_{k'}) > \pi(a) > \pi(a_{k'+1}) > \pi(a_{k'+2}) > \cdots > \pi(a_k)$ for all $a \in A \setminus \{a_1, \ldots, a_k\}$. Using non-negativity of the best–worst Block–Marschak polynomials, a function is defined recursively on these subsets in analogy to the construction of Falmagne (1978), where recursion is with respect to k. Summation over k' and an induction argument result in a measure $\mathrm{P}[\,.\,]$ being defined on singletons, and this is extended to $\mathscr{P}(\Pi)$ in a standard way. Lemma 7 then implies the existence of a random utility representation $(A, \mathbb{BWP})$. For details, see the online supplement.

4. Discussion

This paper adds to the theoretical underpinnings of the best–worst choice paradigm: non-negativity of certain linear combinations of best–worst choice probabilities (i.e. the best–worst BM polynomials) is shown to be necessary and sufficient for a random utility representation of these choice probabilities. Most results on this paradigm up to now, are contained in Marley and Louviere (2005) relating models of best choices, worst choices, and best–worst choices, based on random ranking and random utility, to each other and pointing to open problems. Recently, de Palma, Kilani, and Laffond (2017) presented additional relations between these paradigms under slightly stricter random utility representations and derived various expressions for independent and generalized extreme value distributed utilities.

Non-negativity of the best–worst BM polynomials seems like a rather abstract condition. However, as the theorem shows, it is completely equivalent to postulating a random utility scale to underlie the best–worst choices without making any distributional assumptions, except for the exclusion of ties. In realistic choice experiments, violations of non-negativity allow for a strict test of the random utility hypothesis. For example, the second polynomial

$$K_{ab,\{c\}} = BW_{A\setminus\{c\}}(a, b) - BW_A(a, b),$$

expresses a *regularity condition*: the probability of choosing a as best and b as worst does not increase when other elements are added to the choice set, in analogy to the best-choice situation (see, e.g., Corbin & Marley, 1974; Fishburn, 1998).

For best-choice data, testing non-negativity of the BM polynomials has been in the focus of work by McCausland and colleagues (McCausland, Davis-Stober, Marley, Park, & Brown, 2019). They tested the random utility hypothesis by directly testing the full set of Falmagne‘s BM conditions using observed choices from all choice subsets (with base set of size 5) whose choice distributions are constrained by these conditions. Computing Bayes factors in favor of random utility, they found support for the hypothesis for a large majority of their participants. It would be interesting to perform an analogous study testing best–worst BM polynomials. Related interesting work on best choices has appeared in signal detection theory for recognition memory (see, e.g. Kellen, Winiger, Dunn, & Singmann, 2018). A starting point in all cases could be testing the regularity condition for best–worst choices.

As our results can be considered a straightforward extension of Falmagne‘s work on representing best choices, it should be noted that Fiorini (2004) gave an alternative proof of Falmagne's result using polyhedral combinatorics. His proof is very short and elegant reducing the representation theorem to a complete linear description of the *multiple choice polytope*. In view of this, it seemed obvious to look for an analogous description of the best–worst choice polytope (see Marley & Regenwetter, 2017, for definitions), as has been undertaken in unpublished work of Doignon, Fiorini, Guo, et al. (2015) showing the necessity part of the representation theorem, but we are not aware of a solution to the sufficiency part of the representation problem using these techniques.

A possible extension of the best–worst choice paradigm, suggested by a reviewer, is to consider (i, j)-choices where i and j $(i < j)$ are two positions in the ranking, so that best–worst would correspond to the $(1, n)$-choice. We conjecture that a random representation could be derived by an appropriate extension of Definition 8 of subsets of permutations and a corresponding generalization of the BM polynomials, but we have not yet followed up on this.

Finally, Falmagne (1978) presented some results on the uniqueness of the random utility representation for best choices (see also Colonius, 1984, for additional results). We leave it as another open problem to derive corresponding properties for the case of the best–worst random utility representation developed here.

Acknowledgments

I am especially grateful to Jean-Claude Falmagne and Tony Marley for feedback on early versions of this paper. Adele Diederich and the reviewers also gave very helpful comments improving the presentation. Any errors and deficiencies, of course, remain the author's responsibility. This work was supported in part by DFG, Germany grant CO 94/6-1 to H. Colonius and A. A. J. Marley (U. of Victoria, Canada).

Appendix A. Supplementary data

Supplementary material related to this article can be found online at https://doi.org/10.1016/j.jmp.2021.102596.

References

Block, H., & Marschak, J. (1960). Random orderings and stochastic theories of responses. In I. Olkin, S. Ghurye, W. Hoeffding, W. Madow, & H. Mann (Eds.), *Contributions to probability and statistics*. Stanford University Press.
Colonius, H. (1984). *Stochastische Theorien individuellen Wahlverhaltens [in German]*. Springer-Verlag.
Corbin, R., & Marley, A. (1974). Random utility models with equality: an apparent, but not actual, generalization of random utility models. *Journal of Mathematical Psychology*, *11*, 274–293.
Doignon, J. P., Fiorini, S., Guo, Y., et al. (2015). On a probabilistic model of best-worst choice. Unpublished Manuscript.
Falmagne, J. C. (1978). A representation theorem for finite random scale systems. *Journal of Mathematical Psychology*, *18*, 52–72.
Finn, A., & Louviere, J. (1992). Determining the appropriate response to evidence of public concern: the case of food safety. *Journal of Public Policy and Marketing*, *11*, 12–25.
Fiorini, S. (2004). A short proof of a theorem of Falmagne. *Journal of Mathematical Psychology*, *48*, 80–82.
Fishburn, P. C. (1998). Stochastic utility. In S. Barbera, P. Hammond, & C. Seidl (Eds.), *Handbook of utility theory, Vol. 1* (pp. 273–319). Dordrecht Boston London: Kluwer Academic Publ., (chapter 7).
Kellen, D., Winiger, S., Dunn, J. C., & Singmann, H. (2018). Testing the foundations of signal detection theory in recognition memory. http://dx.doi.org/10.31234/osf.io/p5rj9.
Louviere, J., Flynn, T., & Marley, A. (2015). *Best-worst scaling*. Cambridge, UK: Cambridge University Press.
Marley, A., & Louviere, J. (2005). Some probabilistic models of best, worst, and best-worst choices. *Journal of Mathematical Psychology*, *49*, 464–480.
Marley, A., & Regenwetter, M. (2017). Choice, preference, and utility: probabilistic and deterministic representations. In W. Batchelder, H. Colonius, E. N. Dzhafarov, & J. Myung (Eds.), *New handbook of mathematical psychology Volume 1 foundations and methodology* (pp. 374–453). Cambridge, UK: Cambridge University Press.
McCausland, W., Davis-Stober, C., Marley, A., Park, S., & Brown, . (2019). Testing the random utility hypothesis directly. *The Economic Journal*, *130*, 183–207.
de Palma, A., Kilani, K., & Laffond, G. (2017). Relations between best, worst, and best-worst choices for random utility models. *Journal of Mathematical Psychology*, *76*, 51–58.

Online Supplement to "A Representation Theorem for Finite Best-Worst Random Utility Models"

Hans Colonius

University of Oldenburg

---

1. Online Supplement

Introduction

Choosing an element from an offered set of alternatives is arguably the most basic paradigm of preference behavior. Typically, if the same set is offered several times, the participant's choice will not always be the same. This is often attributed to the participant's preference fluctuating over time due to the effect of various alternatives to be compared, or to the difficulty of distinguishing between similar alternatives. Theories of choice behavior try to account for the probability $P_A(a)$ of choosing an alternative a, say, from an offered set B, which is a subset of the base set A. This intrinsic randomness leads naturally to postulating the existence of a random variable U_b, say, for each alternative $b \in B$ representing the momentary strength of preference for alternative a. The participant is supposed to choose a from B if the momentary (sampled) value of U_a exceeds that of any other alternative U_b, $b \in B \setminus \{a\}$. Such a *random utility representation* can be traced back to the

Email address: `hans.colonius@uni-oldenburg.de` (Hans Colonius)
URL: `https://uol.de/en/hans-colonius/` (Hans Colonius)

beginnings of psychophysics (Fechner, 1860; Thurstone, 1927; see Falmagne, 1985, and Dzhafarov and Colonius, 2011) and may perhaps be considered a cornerstone of both early and contemporary theories of choice and decision making ("discrete choice") in psychology, economics, statistics, and beyond (Luce, 1959; Block and Marschak, 1960; Tversky, 1972; Corbin and Marley, 1974; Manski and McFadden, 1981; Fishburn, 1998; Louviere et al., 2000; Hess and Daly, 2014).

An alternative to the best-choice paradigm is the *ranking paradigm*: the participant is asked to rank-order all elements of an offered set from best to worst resulting in a probability distribution over all possible rankings. Many statistical models have been proposed for this paradigm (e.g. Critchlow et al., 1991) and some are directly connected to models of best choice. For example, a classic result of Block and Marschak (1960) shows that, under specified conditions, the existence of a probability distribution over all possible rankings is necessary and sufficient for a random utility representation of best choices. The present paper concerns a relatively recent choice paradigm that, in terms of complexity, lies somewhat in-between the paradigms of choice and ranking.

Marley (1968) developed the *reversible ranking* model where a ranking is obtained by a sequence of best and/or worst choices (which Marley called superior and/or inferior). Motivated, in part, by their familiarity with that work, Finn and Louviere (1992) proposed a discrete choice task in which a participant is asked to select both the best and the worst option in an available (sub)set of options. For an offered set B, subset of a base set A, let $BW_B(a, b)$ be the probability that a participant chooses a as best and b as worst alternative in the set B. As observed in Marley and Louviere (2005), if there are, e.g., 4 items in A, one obtains information about the best option in 9 out of 11 possible non-empty, non-singleton subsets of A. Thus, best-worst choices contain a great deal of information about the person's ranking of options. Applications of this best-worst choice paradigm have strongly increased over the years. In particular, best-worst scaling is being used as a method of collecting ranking data which is then modeled in various ways related to the multinomial logit for best choices or to weighted versions of the rank ordered logit for repeated best choices (for details, see the monograph by Louviere et al., 2015, p.11 pp.)

Marley and colleagues have developed various random ranking and random utility models for the best-worst paradigm (Marley and Louviere, 2005; see also Marley and Regenwetter, 2017, p. 400 pp.). However, one problem has apparently remained unsolved up to now: What are necessary and sufficient conditions on probabilities $BW_B(a, b)$ for the existence of random variables U_a, U_b, U_c such that the probability to choose a as best and b as worst from a subset B can be represented as

$$P(U_a \geq U_c \geq U_b, c \in B \setminus \{a, b\})$$

for any $B \subseteq A$? For a formal definition, see below.

The solution to this problem turns out to be a minor modification of an approach to the analogous problem for the choice paradigm developed by Falmagne (1978). He was able to show that the non-negativity of certain linear combinations of choice probabilities (so-called *Block-Marschak polynomials*) is both necessary and sufficient for a random utility representation of best choices. An important part of Falmagne's ingenious proof is the construction of a probability measure on the set of rankings which, via the above-mentioned result by Block and Marschak (1960), implies the existence of a random utility representation. Falmagne's construction can be extended to find sufficient conditions for the existence of a random utility representation for best-worst choices as well. Readers are also advised to consult the latter paper for further details.

This online supplement is organized as follows. After some basic notation, we give a formal definition of a system of best-worst choice probabilities and its corresponding random utility representation. First, a necessary condition for this representation in the form of linear inequalities of best-worst choices is provided, closely following arguments from Falmagne (1978) for best choices. Section 1 introduces best-worst Block-Marschak polynomials, shows how to recover best-worst choice probabilities from them using the Moebius inversion (Theorem 6), and states the main representation theorem. Section 1 investigates the structure of rankings (permutations) compatible with best-worst choices including some counting results. Section 1 contains the best-worst probability version of the Block-Marschak result, that is, a probability measure on rankings is necessary and sufficient for best-worst random utility representations. A probability measure on rankings is developed in the subsequent section such that best-worst choice probabilities are defined in terms of the probability measure on appropriate subsets of rankings, completing the proof of the main theorem. For the discussion we point to the theoretical note accompanying this supplement.

Some definitions and basic results

For a finite set A, we write $|A|$ for the number of elements in A, $\mathscr{P}(A)$ for the power set of A, $\mathscr{P}(A, i)$ for the set of all subsets of A containing exactly i elements. For any nonempty set A, finite or not, let $\Phi(A)$ be the set of all finite, nonempty subsets of A.

Definition 1. Let A be a nonempty set of n elements ($n \geq 2$). For any $B \subseteq A$, with $|B| \geq 2$, and any $a, b \in B$, $a \neq b$, let $BW_B(a, b) \mapsto [0, 1]$ denote the probability that a and b are respectively chosen as best and worst elements in the subset B of A. Let $\mathbb{BWP} = \{BW_B \,|\, B \subseteq A, BW_B(a, b) \mapsto [0, 1], a, b \in B, a \neq b\}$ be the collection of all those probability measures. Suppose that

$$\sum_{a,b \in B,\, a \neq b} BW_B(a, b) = 1 \qquad (BW \in \mathbb{BWP}).$$

Then $(A, \mathbb{BWP})$ is called a *finite system of best-worst choice probabilities*, or more briefly, a *system*.

The quantities $BW_B(a, b)$ will be referred to as *best-worst choice probabilities* indicating that an alternative a is judged as best and b as worst, when subset B is available.

Definition 2. Let $(A, \mathbb{BWP})$ be a finite system of best-worst choice probabilities. Suppose that there is collection of jointly distributed random variables $\{U_c \,|\, c \in A\}$, with U_c measuring the variable value of alternative $c \in A$. For any subset $B \subseteq A$ and $a, b \in B$, with $a \neq b$, define the two events

$$\mathbf{B}_{a,B} = \{U_a = \max_{c \in B} U_c\}$$

$$\mathbf{W}_{b,B} = \{U_b = \min_{c \in B} U_c\}.$$

So, for the subset B, $\mathbf{B}_{a,B}$ is the event that the random variable U_a has the maximum value compared to all other elements in the set B, and $\mathbf{W}_{b,B}$ is the corresponding event that the random variable U_b has the minimum value compared to all other elements in the set B.

Definition 3. A finite system $(A, \mathbb{BWP})$ is called *a best-worst random utility system* if there exists a collection of jointly distributed random variables $\{U_c \,|\, c \in A\}$ such that for all $B \subseteq A$ and $a, b \in B$, $a \neq b$,

$$BW_B(a,b) = P(\mathbf{B}_{a,B} \cap \mathbf{W}_{b,B})$$

The collection $\{U_c \,|\, c \in A\}$ will be *called a random representation of* $(A, \mathbb{BWP})$. From now on, we will always assume the systems to be finite without mentioning it specifically.

Consider the 2-element case $\{a, b\}$. A simple, but important implication of Definition 3 is the following: for any distinct $a, b \in A$

$$\begin{aligned} 1 &= BW_{\{a,b\}}(a,b) + BW_{\{a,b\}}(b,a) \\ &= P(U_a > U_b) + P(U_a = U_b) + P(U_b > U_a) + P(U_b = U_a). \end{aligned}$$

Since we also have

$$1 = P(U_a > U_b) + P(U_a = U_b) + P(U_b > U_a),$$

we obtain

$$P(U_a = U_b) = 0.$$

Thus, a random utility representation of best-worst choice probabilities does not allow for ties. Let us introduce two abbreviations:

$$\mathrm{M}_B^+ = \max\{U_c \,|\, c \in B\} \quad \text{and} \quad \mathrm{M}_B^- = \min\{U_c \,|\, c \in B\}.$$

We conclude that with $a, b \in B$

$$BW_B(a,b) = P(U_a \geq \mathrm{M}_B^+ \geq \mathrm{M}_B^- \geq U_b) = P(U_a > \mathrm{M}_{B\setminus\{a,b\}}^+ \geq \mathrm{M}_{B\setminus\{a,b\}}^- > U_b).$$

Now let $B_0, B_1, \ldots, B_n \in \Phi(A)$; for $a, b \in B_0$,

$$BW_{B_0}(a,b) = P(U_a \geq \mathrm{M}_{B_0}^+ \geq \mathrm{M}_{B_0}^- \geq U_b).$$

Denote

$$E_i = \{U_a \geq M_{B_i}^+ \geq M_{B_i}^- \geq U_b\}$$

for $0 \leq i \leq n$, and observe that, for $0 \leq i, j \leq n$,.

$$\begin{aligned} E_i \cap E_j &= \{U_a \geq M_{B_i}^+ \geq M_{B_i}^- \geq U_b\} \cap \{U_a \geq M_{B_j}^+ \geq M_{B_j}^- \geq U_b\} \\ &= \{U_a \geq M_{B_i \cup B_j}^+ \geq M_{B_i \cup B_j}^- \geq U_b\}. \end{aligned}$$

For $a, b \in B_0$, this implies

$$BW_{B_0 \cup B_1 \cup \cdots B_n}(a,b) = P\left(\bigcap_{i=0}^{n} E_i\right).$$

From this follows, for example,

$$BW_{B_0}(a,b) - [BW_{B_0 \cup B_1}(a,b) + BW_{B_0 \cup B_2}(a,b)] + BW_{B_0 \cup B_1 \cup B_2}(a,b) \geq 0, \quad (1)$$

since

$$P(E_0) - [P(E_0 \cap E_1) + P(E_0 \cap E_2)] + P(E_0 \cap E_1 \cap E_2) \geq 0,$$

holds for arbitrary events E_0, E_1, E_2 in a probability space. The following theorem is completely parallel to the one in Falmagne (1978, Theorem 1).

Theorem 4. *Let* $(A, \mathbb{BWP})$ *be a best-worst random utility system. For any* $a, b \in B_0 \in \Phi(A)$*, and any finite collection* $\mathscr{B} = \{B_j \,|\, j \in J, B_j \subset A \text{ or } B_j = \emptyset\}$*, we have*

$$\sum_{i=0}^{|J|}(-1)^i \sum_{C \in \mathscr{P}(J,i)} BW_{B_0 \cup \mathscr{B}(C)}(a, b) \geq 0, \tag{2}$$

where $\mathscr{B}(C) = \bigcup_{j \in C} B_j$*.*

Note that the case $\mathscr{B} = \{B_1, B_2\}$ corresponds to Equation 1, while $\mathscr{B} = \{\emptyset\}$, $\mathscr{B} = \{B_1\}$ yield, respectively,

$$\begin{aligned} &BW_{B_0}(a, b) \geq 0, \\ &BW_{B_0}(a, b) \geq BW_{B_0 \cup B_1}(a, b). \end{aligned} \tag{3}$$

The last inequality expresses a *regularity condition* for best-worst random utility systems: the probability of choosing a as best and b as worst does not increase when other elements are added to the choice set, in analogy to the best-choice situation (see, e.g., Corbin and Marley, 1974; Fishburn, 1998).

Proof of Theorem 4

Writing E_0 for the event $\{U_a \geq U_c \geq U_b,\, c \in B_0 \setminus \{a, b\})\}$ and $E_j, j \in J$, for the event $\{U_a \geq M^+_{B_j} \geq M^-_{B_j} \geq U_b\}$, we get

$$BW_{B_0 \cup \mathscr{B}(C)}(a, b) = P\left[\bigcap_{j \in C}(E_0 \cap E_j)\right].$$

The theorem follows from the fact that, for any finite collection$\{E_j \,|\, j \in J\}$ of events and any event E_0 in a probability space, we have

$$\sum_{i=0}^{|J|}(-1)^i \sum_{C \in \mathscr{P}(J,i)} P\left[\bigcap_{j \in C}(E_0 \cap E_j)\right] \geq 0. \tag{4}$$

Indeed, (4) certainly holds if $P(E_0) = 0$; while if $P(E_0) \neq 0$, dividing on both sides by $P(E_0)$, (4) is equivalent, by *Poincaré's identity*, to

$$1 \geq P\left(\bigcup_{j \in J} E_j \mid E_0\right).$$

Best-Worst Block-Marschak polynomials: the main theorem

Suppose $(A, \mathbb{BWP})$ is a system. Consider the following expressions:

$$BW_A(a,b),$$

$$BW_{A\setminus\{c\}}(a,b) - BW_A(a,b),$$

$$BW_{A\setminus\{c,d\}}(a,b) - [BW_{A\setminus\{c\}}(a,b) + BW_{A\setminus\{d\}}(a,b)] + BW_A(a,b),$$

$$BW_{A\setminus\{c,d,e\}}(a,b) - [BW_{A\setminus\{c,d\}}(a,b) + BW_{A\setminus\{c,e\}}(a,b) + BW_{A\setminus\{d,e\}}(a,b)] + [BW_{A\setminus\{c\}}(a,b) + BW_{A\setminus\{d\}}(a,b) + BW_{A\setminus\{e\}}(a,b)] - BW_A(a,b),$$

etc.

Each of these expressions is a case of the one in the left member of (2). In analogy to Falmagne's (1978) terminology, we introduce a compact notation.

Definition 5. *For any $B \subset A$, $B \neq A$, $a, b \in A \setminus B, a \neq b$, in a system $(A, \mathbb{BWP})$, we define*

$$K_{ab,B} = \sum_{i=0}^{|B|} (-1)^i \sum_{C \in \mathscr{P}(B,|B|-i)} BW_{A\setminus C}(a,b). \tag{5}$$

The $K_{ab,B}$ are called best-worst Block-Marschak polynomials of $(A, \mathbb{BWP})$, *or* best-worst BM polynomials, *for short.*

Observe that

$$
\begin{aligned}
K_{ab,\emptyset} &= \sum_{i=0}^{0}(-1)^i \sum_{C\in\mathscr{P}(\emptyset,0-i)} BW_{A\setminus C}(a,b)\\
&= BW_A(a,b);\\
K_{ab,\{c\}} &= BW_{A\setminus\{c\}}(a,b) - BW_A(a,b)\\
&= BW_{A\setminus\{c\}}(a,b) - K_{ab,\emptyset};\\
K_{ab,\{c,d\}} &= BW_{A\setminus\{c,d\}}(a,b) - [BW_{A\setminus\{c\}}(a,b) + BW_{A\setminus\{d\}}(a,b)] + BW_A(a,b)\\
&= BW_{A\setminus\{c,d\}}(a,b) - K_{ab,\{c\}} - K_{ab,\{d\}} - K_{ab,\emptyset}.
\end{aligned}
$$

Similar computations show that

$$
\begin{aligned}
K_{ab,\{c,d,e\}} = BW_{A\setminus\{c,d,e\}}(a,b) &- K_{ab,\{c,d\}} - K_{ab,\{c,e\}} - K_{ab,\{d,e\}}\\
&- K_{ab,\{c\}} - K_{ab,\{d\}} - K_{ab,\{e\}} - K_{ab,\emptyset}
\end{aligned}
$$

These examples suggest the following result.

Theorem 6. *Let $(A,\mathbb{BWP})$ be a system of best-worst choice probabilities. Then, for all $B\subset A, B\neq A$, and $a,b\in A\setminus B$,*

$$BW_{A\setminus B}(a,b) = \sum_{C\in\mathscr{P}(B)} K_{ab,C}. \tag{6}$$

Proof of this theorem is omitted here since it is completely analogous to the one in Falmagne (1978, Theorem 2, pp. 57–8) by replacing the "ordinary" BM polynomials by best-worst BM polynomials. Alternatively, with the same polynomial replacement, it is also analogous to the one given in Colonius (1984, pp. 58–60), using *Möbius inversion* (for the latter definition see, e.g. van Lint and Wilson, 2001). We can now state the main theorem of this paper.

Theorem 7. *A finite system of best-worst choice probabilities is a best-worst random utility system if and only if the best-worst Block-Marschak polynomials are nonnegative.*

The necessity follows from Theorem 4 and the definition of best-worst Block-Marschak polynomials (Section 1). The rest of the paper concerns the sufficiency, that is, to show that if

$$K_{ab,B} \geq 0$$

for all $B \subset A, B \neq A$ and $a, b \in A \setminus B$, then $(A, \mathbb{BWP})$ is a best-worst random utility system. The proof requires an analysis of the Boolean algebra of the sets of permutations on A in a system $(A, \mathbb{BWP})$ of best-worst choice probabilities.

Sets of rankings and counting results

The next definition is our basic tool for constructing the probability measure on the rankings (permutations) in the subsequent section. It is illustrated by a number of examples. Moreover, a counting lemma and a partitioning lemma needed for the construction are presented here.

For any $B \subseteq A$, we write Π_B for the set of $|B|!$ permutations on B. For simplicity, we abbreviate $\Pi_A = \Pi$. Let $\geq$ be an arbitrarily chosen simple order on A. As usual, we write, for any $a, b \in A, a < b$ iff not $a \geq b$, and $a > b$ iff not $b \geq a$.

Definition 8. Let $|A| = n$ with $n > 2$ and $B \subseteq A$ with $|B| = n - m$ $(n \geq m)$; for $B \in \Phi(A)$ and distinct $b_1, b_2, \ldots, b_{k'}, b_{k'+1}, b_{k'+2}, \ldots, b_k \in B$

$(1 \leq k' < k \leq n - m)$, define

$\mathrm{S}(b_1 b_2 \ldots b_{k'}; B; b_{k'+1} b_{k'+2} \ldots b_k) =$

$\{\pi \in \Pi \,|\, \pi(b_1) > \pi(b_2) > \cdots > \pi(b_{k'}) > \pi(b) > \pi(b_{k'+1}) > \pi(b_{k'+2}) > \cdots > \pi(b_k),$

for all $b \in B \setminus \{b_1, \ldots, b_k\}\}$.

For simplicity, we write $b_1 b_2 \ldots b_k$ for the ranking (defined by $>$) corresponding to permutation π with $\pi(b_1) > \pi(b_2) > \cdots > \pi(b_k)$. Moreover, if no confusion arises we also omit the semicolons around set B and write $\mathrm{S}(b_1 b_2 \ldots b_{k'} B b_{k'+1} b_{k'+2} \ldots b_k)$. The following examples illustrate the properties of the sets defined above.

Example 9. *Let* $|A| = n$ *and* $B = \{b\}$ *for* $a, c \in A \setminus B$

$$\mathrm{S}(aA \setminus Bc) = \mathrm{S}(baAc) + \mathrm{S}(aAcb) + \mathrm{S}(aAc)$$

Example 10. *Let* $|A| = n$ *and* $B = \{b, d\}$*; for* $a, c \in A \setminus B$

$$\begin{aligned}\mathrm{S}(aA \setminus Bc) &= \mathrm{S}(baAc) + \mathrm{S}(aAcb) + \mathrm{S}(daAc) + \mathrm{S}(aAcd) + \mathrm{S}(bdaAc) \\ &\quad + \mathrm{S}(dbaAc) + \mathrm{S}(aAcbd) + \mathrm{S}(aAcdb) + \mathrm{S}(baAcd) + \mathrm{S}(daAcb) + \mathrm{S}(aAc).\end{aligned}$$

Example 11. *Let* $A = \{p, q, r, u, v\}$ *and* $B = \{p, q\}$ *then*

$$\mathrm{S}(pBq) = \{\pi \in \Pi_A \,|\, \pi(p) > \pi(q)\}.$$

With $|\Pi_A| = 5! = 120$*, it follows that* $|\mathrm{S}(pBq)| = 60$ *since exactly half of the permutations have* p *ranked before* q*.*

The next lemma determines the number of elements in the sets of Definition 8.

Lemma 12. *Let $|A| = n$ with $n > 2$ and $B \subseteq A$ with $|B| = n - m$; for $1 \le k' < k \le n - m$ the number of elements contained in*

$$\mathrm{S}(b_1 b_2 \dots b_{k'} B\, b_{k'+1} b_{k'+2} \dots b_k)$$

equals[1]*,*

$$\begin{aligned}|\mathrm{S}(b_1 b_2 \dots b_{k'} B b_{k'+1} b_{k'+2} \dots b_k)| =& (n-m-k)! \prod_{i=1}^{m} [n - m + i)] \\ =& \frac{(n-m-k)!\; n!}{(n-m)!}.\end{aligned}$$

Note that the result does not depend on k'. While an elaborate proof is available, the result is easily obtained via the following argument[2] : The elements of $\mathrm{S}(b_1 b_2 \dots b_{k'} B b_{k'+1} b_{k'+2} \dots b_k)$ are a subset of the $n!$ permutations of A. Moreover, for the m elements in $A \setminus B$ there are

$$n \times (n-1) \times \cdots \times (n - m + 1) = \frac{n!}{(n-m)!}$$

ways to arrange them. There are $n - m - k$ elements in B that are not the k elements that have a fixed order, and they can be ordered in $(n - m - k)!$ ways, leading to the result of Lemma 12. We consider a few examples for illustration.

Example 11 (continuing from p. 12)**.** *With $A = \{p, q, r, u, v\}$ and $B = \{p, q\}$, we have $n = 5$, $m = 3$, and $k = 2$ for $\mathrm{S}(pBq)$. Thus, by Lemma 12*

$$|\mathrm{S}(pBq)| = (5 - 3 - 2)!(5 - 2)(5 - 1)5 = 1!60 = 60,$$

as inferred before by a different argument.

---

[1] for $m = 0$, the Π term is set to 1, so that $|\mathrm{S}(b_1 b_2 \dots b_{k'} A b_{k'+1} b_{k'+2} \dots b_k)| = (n - k)!$

[2] I am most grateful to a reviewer for pointing this out.

Example 9 (continuing from p. 12)**.** *With $|A| = n$ and $B = \{b\}$, we need to show that, for $a, c \in A \setminus B$,*

$$|\mathrm{S}(aA \setminus Bc)| = |\mathrm{S}(baAc)| + |\mathrm{S}(aAcb)| + |\mathrm{S}(aAc)|,$$

because all sets on the right are pairwise disjoint. For the left hand side, $m = 1$ and $k = 2$, thus

$$|\mathrm{S}(aA \setminus Bc)| = (n - 1 - 2)!(n - (1 - 1)) = (n - 3)!\, n.$$

For the first two sets on the right hand side, $m = 0$ and $k = 3$,thus

$$|\mathrm{S}(baAc)| = |\mathrm{S}(aAcb)| = (n - 0 - 3)!\, 1 = (n - 3)!;$$

and, for the third set $m = 0$ and $k = 2$, thus

$$|\mathrm{S}(aAc)| = (n - 0 - 2)!\, 1 = (n - 2)!;$$

summing up the numbers from the right, $2(n - 3)! + (n - 2)! = (n - 3)!\, [2 + n - 2] = (n - 3)!\, n$, which equals the number on the left.

Lemma 13. *For any $B \subset \Phi(A)$ and $a, b \in A \setminus B$*

$$S(aA \setminus Bb) = \sum_{C \in \mathscr{P}(B)} \sum_{\pi \in \Pi_C} \sum_{(\pi_1 \pi_2) = \pi} S(\pi_1 aAb\pi_2); \qquad (7)$$

here, the last summation over $(\pi_1\pi_2) = \pi$ means that ranking π is split into all possible pairs, for example, cde is split into

$$(cde)(), (cd)(e), (c)(de), ()(cde)$$

so that $(cd)(e)$ corresponds to $S(cdaAbe)$, $()(cde)$ to $S(aA; bcde)$, etc.

Proof of Lemma 13

First, we show that the union in the right member of Eq. 13 is disjoint. For $(\pi_1, \pi_2) = \pi \neq \pi' = (\pi'_1, \pi'_2)$, the sets $S(\pi_1 aAb\pi_2)$ and $S(\pi'_1 aAb\pi'_2)$ are clearly disjoint, so that disjointness remains to be shown for the two first summation signs in (13). For now, let us abbreviate $S(\pi_1 aAb\pi_2)$ as $S_\pi(aAb)$. For $|B| = 1$, the lemma's claim is implicit in Example 9, and for $|B| = 2$, Example 11 (Table in Appendix C) demonstrates the partition.

Let $|B| \geq 3$ take $C, C' \in \mathscr{P}(B)$, $\pi \in \Pi_C, \pi' \in \Pi_{C'}$, $\pi \neq \pi'$. Suppose $C = C'$. Then $|C| \geq 2$ (otherwise, $\Pi_C = \{\pi\}$, contradicting $\pi \neq \pi'$), and there will be at least two elements $d, e \in C$ such that such that $\xi(d) < \xi(e)$ for all $\xi \in S_\pi(aAb)$, while $\xi'(e) < \xi'(d)$ for all $\xi' \in S_{\pi'}(aAb)$. Thus

$$S_\pi(aAb) \cap S_{\pi'}(aAb) = \emptyset. \tag{8}$$

Then case $C \neq C'$ is similar. For example, suppose $d \in C \setminus C'$, then $\xi(d) > \xi(a)$ or $\xi(b) > \xi(d)$ for all $\xi \in S_\pi(aAb)$, while $\xi'(a) > \xi'(d) > \xi'(b)$ for all $\xi' \in S_{\pi'}(aAb)$, entailing again Eq. 8 .

We turn to the proof of equality, and write $G(a, b, A, B)$ for the right member of 13 . Assume $\xi \in S(aA \setminus Bb)$. Then, either $\xi(a) > \xi(c) > \xi(b)$ for all $c \in B$, implying $\xi \in S(aAb) \subset G(a, b, A, B)$; or, there are $c_1, c_2, \ldots, c_j \in B$ such that

$$\xi(c_1) > \xi(c_2) > \ldots > \xi(c_{j_1}) > \xi(a) > \xi(b) > \xi(c_{j_1+1}) > \ldots > \xi(c_j).$$

This yields $\xi \in S(c_1 c_2 \ldots c_{j_1}\, aAb\, c_{j_1+1} \ldots c_j) \subset G(a, b, A, B)$. We conclude that $S(aA \backslash Bb) \subset G(a, b, A, B)$.The converse implication follows from the fact that for any choice of $C \in \mathscr{P}(B)$ and $\pi \in \Pi_C$, we have $S_\pi(aAb) \subset S(aA \backslash Bb)$.

The following example illustrates the lemma.

Example 11 (continuing from p. 12)**.** *With* $A = \{p, q, r, u, v\}$ *and* $B = \{p, q\}$*, let us consider* $S(uA \setminus Bv)$*; thus,* $n = 5$*,* $|A \setminus B| = 3 = n - m$*, so* $m = 2$*, and* $k = 2$*. From Lemma 12 ,*

$$|S(uA \setminus Bv)| = \frac{(n-m-k)!\, n!}{(n-m)!} = \frac{1!\, 5!}{3!} = 20.$$

These 20 rankings are listed in the first column of the table in 1, partitioned into the additive components $S(\pi_1 uAv\pi_2)$ *(right-hand side in Lemma 13). The second column shows the corresponding subsets* C *of* B*. For instance, with* $C = \emptyset$*,* $S(\pi_1 uAv\pi_2) = S(uAv)$ *and* $|S(uAv)| = 6$*; and with* $C = \{p, q\}$*,* $\Pi_C = \{pq, qp\}$ *and for each permutation there are 3 ways to split them into* (π_1, π_2)*:* $(pq)(), ()(pq), (p)(q)$ *and* $(qp)(), ()(qp), (q)(p)$*.*

Lemma 14. *For all* $1 \leq k' < k \leq n$*,*

$$\begin{aligned} &\sum_{a \in A \setminus \{a_1, \ldots, a_k\}} S(a_1, \ldots a_{k'}\, aA\, a_{k'+1} \ldots a_k) \\ &\qquad = \sum_{a \in A \setminus \{a_1, \ldots, a_k\}} S(a_1, \ldots a_{k'}\, A\, a\, a_{k'+1} \ldots a_k) \\ &\qquad = S(a_1, \ldots a_{k'}\, A\, a_{k'+1} \ldots a_k) \end{aligned}$$

The straightforward proof is omitted. Next, we define the union of certain sets of rankings which the probability measure will ultimately be constructed on. For $k \geq 2$ and distinct $a_1, a_2, \ldots, a_k \in A$,

$$S_A(a_1, \ldots, a_k) = \sum_{k'=1}^{k-1} S(a_1 \ldots a_{k'}\, A\, a_{k'+1} \ldots a_k). \tag{9}$$

A Block-Marschak type lemma

This lemma provides a critical step in our proof. It is a variant of the well-known result by Block & Marschak (1960) (see also Marley & Louviere 2005, section on random ranking models).

Lemma 15. $(A, \mathbb{BWP})$ *is a best-worst random utility system if and only if there exists a probability measure* $\mathrm{P}[\,.\,]$ *on* $\mathscr{P}(\Pi)$ *satisfying*

$$BW_B(a, b) = \mathrm{P}[S(aBb)] \tag{10}$$

for all $B \in \Phi(A)$ *and* $a, b \in B$.

Proof. (Lemma 15)
For simplicity, we set $A = \{1, 2, \ldots, n\}$ and take $\geq$ as the natural order of the reals.

(Necessity) Let $\{U_i \,|\, 1 \leq i \leq n\}$ be a random representation of $(A, \mathbb{BWP})$, with joint probability measure P satisfying (10). For any $\pi \in \Pi$, define

$$p(\{\pi\}) = \mathrm{P}[U_{\pi^{-1}(1)} < U_{\pi^{-1}(2)} < \cdots < U_{\pi^{-1}(n)}].$$

It is easy to verify that p is a probability distribution on Π that can be extended to a probability measure P on $\mathscr{P}(\Pi)$, satisfying (10).

(Sufficiency) Conversely, suppose that (10) holds for some probability measure P on $\mathscr{P}(\Pi)$. Define the joint distribution of a collection $\{U_i \,|\, 1 \leq i \leq n\}$ of random variables by

$$\mathrm{P}[U_1 = \xi_1, U_2 = \xi_2, \ldots, U_n = \xi_n] = \begin{cases} \mathrm{P}(\pi) & \text{if } \pi(i) = \xi_i,\ 1 \leq i \leq n, \\ 0 & \text{otherwise} \end{cases} \tag{11}$$

for all n-tuples $\xi_1, \xi_2, \ldots, \xi_n$ of real numbers. It can be checked that then

$$BW_B(i, j) = \mathrm{P}[U_i \geq M_B^+ \geq M_B^- \geq U_j] \tag{12}$$

for any $B \in \Phi(A)$, $i, j \in B$. □

Defining a probability measure on the rankings of A

This section completes the sufficiency part of the main theorem (Theorem 7). Thus, we assume the best-worst BM polynomials to be non-negative. The first step is to find a function on the sets

$$S(a_1, \dots, a_{k'} A a_{k'+1}, \dots, a_k) \subset \Pi_A,$$

with $|A| = n$ and $1 \leq k' < k \leq n$. To this end, a function F' is defined inductively. For $k = 2$ (thus, $k' = 1$), define

$$\mathrm{F}'[S(a_1 A a_2)] := K_{a_1 a_2, \emptyset}. \tag{13}$$

For $k \geq 3$, $k' < k$, we define

$$\mathrm{F}'[S(a_1 \dots a_{k'} A\, a_{k'+1} \dots a_k)] :=$$

$$\frac{\mathrm{F}'[S(a_1 \dots a_{k'-1} A\, a_{k'+2} \dots a_{k-1})] \;\; K_{a_{k'} a_{k'+1}, \{a_1, \dots, a_{k'-1}, a_{k'+2}, \dots, a_{k-1}\}} / (n-k)!}{\sum\limits_{\pi \in \Pi_{\{a_1, \dots, a_{k'-1}, a_{k'+2}, \dots, a_{k-1}\}}} \mathrm{F}'[S(\pi(a_1) \dots \pi(a_{k'-1}) A\, \pi(a_{k'+2}) \dots \pi(a_{k-1})]} \tag{14}$$

assuming the denominator > 0, and set $\mathrm{F}' = 0$ otherwise.

Lemma 16. *(a)* $\mathrm{F}' \geq 0$*;*

(b)

$$\sum_{\pi \in \Pi_B} \mathrm{F}'[S(\pi_1 a A b\, \pi_2)] = K_{ab,B}/(n-k)!$$

for any $B \subset A$ *and* $a, b \in A \setminus B$*, and* π *of the form* $(\pi_1 \pi_2)$*.*

For $k = 2$ *(thus,* $k' = 1$*), define*

$$\mathrm{F}'[S(a_1 A a_2)] := K_{a_1 a_2, \emptyset}. \tag{15}$$

Non-negativity of F′ *(a)* follows from assuming non-negative best-worst BM polynomials and *(b)* is immediate from the above recursive definition. Note that $(n-k)!$ is number of elements in $S(\pi_1 a A b\, \pi_2)$.

This suggests the following heuristic interpretation of the $K_{ab,B}$ as the probability measure of all rankings of A with a as best and b as worst ignoring all alternatives that are in B.

Next, we extend F′ to a function F on the sets $S_A(a_1, \ldots, a_k)$ $(k \leq n)$ by defining:

$$\begin{aligned} \mathrm{F}[S_A(a_1, \ldots, a_k)] &\equiv \mathrm{F}\left[\sum_{k'=1}^{k-1} S(a_1 \ldots a_{k'}\, A\, a_{k'+1} \ldots a_k)\right] \\ &:= \sum_{k'=1}^{k-1} \mathrm{F}'[S(a_1 \ldots a_{k'}\, A\, a_{k'+1} \ldots a_k)] \end{aligned} \tag{16}$$

For F to be a probability distribution on $\mathscr{P}(\Pi)$, we need to show

(i) $\mathrm{F} \geq 0$;

(ii) $\sum_{\pi \in \Pi} \mathrm{F}[S(\pi(a_1), \ldots, \pi(a_n)] = 1$.

Obviously, (i) follows from the non-negativity of F′. We obtain (ii) as a special case of the general result that for $2 \leq j \leq n$ $(k' < j)$

$$\sum_{C \in \mathscr{P}(A,j)} \sum_{\substack{\pi \in \Pi_C \\ C = \{a_1, \ldots, a_j\}}} \mathrm{F}[S_A(\pi(a_1), \ldots, \pi(a_j)] = 1; \tag{17}$$

(ii) is then obtained from (17) for $j = n$. Equation 17 is proved by induction on j. For $j = 2$, we have

$$\sum_{C \in \mathscr{P}(A,2)} \sum_{\substack{\pi \in \Pi_C \\ C=\{a_i,a_\ell\}}} \mathrm{F}[S_A(\pi(a_i), \pi(a_\ell))] = \sum_{\substack{a_i,a_\ell \in A \\ a_i \neq a_\ell}} \mathrm{F}[S_A(a_i, a_\ell)]$$

$$= \sum_{\substack{a_i,a_\ell \in A \\ a_i \neq a_\ell}} \mathrm{F}'[S(a_i A a_\ell)] = \sum_{\substack{a_i,a_\ell \in A \\ a_i \neq a_\ell}} K_{a_i a_\ell, \emptyset} = \sum_{\substack{a_i,a_\ell \in A \\ a_i \neq a_\ell}} BW_A(a_i, a_\ell) = 1.$$

Now assume that (17) holds for all j with $2 \leq j \leq k-1 < n$ $(k' < k-1)$; then

$$\sum_{C \in \mathscr{P}(A,k)} \sum_{\substack{\pi \in \Pi_C \\ C=\{a_1,\ldots,a_k\}}} \mathrm{F}[S_A(\pi(a_1), \ldots, \pi(a_k)]$$

$$= \sum_{C \in \mathscr{P}(A,k)} \sum_{\substack{\pi \in \Pi_C \\ C=\{a_1,\ldots,a_k\}}} \sum_{k'=1}^{k-1} \mathrm{F}'[S(a_1 \ldots a_{k'}\, A\, a_{k'+1} \ldots a_k)]$$

$$= \sum_{C' \in \mathscr{P}(A,k-1)} \sum_{\substack{\pi' \in \Pi'_C \\ C'=\{a_1,\ldots,a_{k-1}\}}} \sum_{k'=1}^{k-2} \sum_{a \in A \setminus C'} \mathrm{F}'[S(\pi'(a_1) \ldots \pi'(a_{k'})\, aA\, \pi'(a_{k'+1}) \ldots \pi'(a_{k-1})]$$

$$= \sum_{C' \in \mathscr{P}(A,k-1)} \sum_{\substack{\pi' \in \Pi_{C'} \\ C'=\{a_1,\ldots,a_{k-1}\}}} \sum_{k'=1}^{k-2} \mathrm{F}'[S(\pi'(a_1) \ldots \pi'(a_{k'})\, A\, \pi'(a_{k'+1}) \ldots \pi'(a_{k-1})]$$

$$= \sum_{C' \in \mathscr{P}(A,k-1)} \sum_{\substack{\pi' \in \Pi'_C \\ C'=\{a_1,\ldots,a_{k-1}\}}} \mathrm{F}[S_A(\pi'(a_1), \ldots, \pi'(a_{k-1})]$$

$$= 1$$

by the induction hypothesis. Thus, (17) holds for $j = n$. We extend the probability distribution on Π in a standard way to obtain a probability measure P on $\mathscr{P}(\Pi)$.

In view of Lemma 15, we need to show, finally, that

$$BW_B(a, b) = \mathrm{P}[S(aBb)] \tag{18}$$

for all $B \in \Phi(A)$ and $a, b \in B$. Now,

$$\begin{aligned}
\mathrm{P}[S(aBb)] &= \mathrm{P}\left[\sum_{C\in\mathscr{P}(A\setminus B)}\sum_{\pi\in\Pi_C}\sum_{\pi=(\pi_1\pi_2)} S(\pi_1 aA\, b\, \pi_2)\right] && \text{by Lemma 13}\\
&= \sum_{C\in\mathscr{P}(A\setminus B)} K_{ab,C} && \text{by Lemma 16}\\
&= BW_B(a,b) && \text{by Theorem 6}
\end{aligned}$$

completing the proof of Theorem 7.

Table for Example 11

$S(uA \setminus Bv)$	$S(\pi_1 uAv\pi_2)$	$\lvert S(\pi_1 uAv\pi_2)\rvert$
$upqrv$		$n=5; m=0; k=2$
$uprqv$		$\frac{(n-m-k)!\,n!}{(n-m)!}=6$
$uqprv$	$C=\emptyset$	
$uqrpv$		
$urpqv$		
$urqpv$		
$puqrv$		$n=5; m=0; k=3$
$purqv$	$C=\{p\}$	$\frac{(n-m-k)!\,n!}{(n-m)!}=2$
$uqrvp$		
$urqvp$		
$quprv$		$n=5; m=0; k=3$
$qurpv$	$C=\{q\}$	$\frac{(n-m-k)!\,n!}{(n-m)!}=2$
$uprvq$		
$urpvq$		
$pqurv$		$n=5; m=0; k=4$
$qpurv$		$\frac{(n-m-k)!\,n!}{(n-m)!}=1$
$urvpq$	$C=\{p,q\}$	
$urvqp$		
$purvq$		
$qurvp$		

Table .1: Example 11, for explanation, see text.

References

Block, H., Marschak, J., 1960. Random orderings and stochastic theories of responses, in: Olkin, I., Ghurye, S., Hoeffding, W., Madow, W., Mann, H. (Eds.), Contributions to probability and statistics. Stanford University Press, Stanford, CA.

Colonius, H., 1984. Stochastische Theorien individuellen Wahlverhaltens [in German]. Springer-Verlag.

Corbin, R., Marley, A., 1974. Random utility models with equality: an apparent, but not actual, generalization of random utility models. Journal of Mathematical Psychology 11, 274–293.

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

Dzhafarov, E.N., Colonius, H., 2011. The Fechnerian idea. American Journal of Psychology 124, 127–140.

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

Falmagne, J.C., 1985. Elements of psychophysical theory. Oxford University Press, New York, NY.

Fechner, G.T., 1860. Elemente der Psychophysik [Elements of psychophysics]. Breitkopf & Härtel., Leipzig, Germny.

Finn, A., Louviere, J., 1992. Determining the appropriate response to evidence of public concern: the case of food safety. Journal of Public Policy and Marketing 11, 12–25.

Fishburn, P., 1998. Stochastic utility, in: S., B., Hammond, P., Seidl, C. (Eds.), Handbook of utility theory. Kluwer Academic Publ., Dordrecht Boston London. volume 1. chapter 7, pp. 273–319.

Hess, S., Daly, A., 2014. Handbook of choice modelling. Edward Elgar Publishing.

van Lint, J., Wilson, R., 2001. A course in combinatorics. 2nd ed., Cambridge University Press.

Louviere, J., Flynn, T., Marley, A., 2015. Best-worst scaling. Cambridge University Press, Cambridge, UK.

Louviere, J., Hensher, D., Swait, J., 2000. Stated choice methods: analysis and applications. Cambridge University Press.

Luce, R., 1959. Individual choice behavior: a theoretical analysis. John Wiley & Sons.

Manski, C.F., McFadden, D., 1981. Structural analysis of discrete data with econometric applications. MIT Press, Cambridge, MA.

Marley, A., 1968. Some probabilistic models of simple choice and ranking. Journal of Mathematical Psychology 5, 311–332.

Marley, A., Louviere, J., 2005. Some probabilistic models of best, worst, and best-worst choices. Journal of Mathematical Psychology 49, 464–480.

Marley, A., Regenwetter, M., 2017. Choice, preference, and utility: probabilistic and deterministic representations, in: Batchelder, W., Colonius, H., Dzhafarov, E.N., Myung, J. (Eds.), New Handbook of Mathematical Psychology Vol. I Foundations and Methodology. Cambridge University Press, Cambridge, UK, pp. 374–453.

Thurstone, L.L., 1927. A law of comparative judgment. Psychological Review 34, 273–286.

Tversky, A., 1972. Elimination by aspects: a theory of choice. Psychological Review 79, 281–299.