

Marley, Anthony; Islam, Towhidul

Article

Conceptual relations between expanded rank data and models of the unexpanded rank data

Journal of Choice Modelling

Provided in Cooperation with:

Journal of Choice Modelling

Suggested Citation: Marley, Anthony; Islam, Towhidul (2012) : Conceptual relations between expanded rank data and models of the unexpanded rank data, Journal of Choice Modelling, ISSN 1755-5345, University of Leeds, Institute for Transport Studies, Leeds, Vol. 5, Iss. 2, pp. 38-80

This Version is available at:

<https://hdl.handle.net/10419/66825>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.



<http://creativecommons.org/licenses/by-nc/2.0/uk/>

Conceptual Relations Between Expanded Rank Data and Models of the Unexpanded Rank Data

A.A.J. Marley^{1,2*} T. Islam^{3,†}

¹ Department of Psychology, University of Victoria, PO Box 3050 STN CSC, Victoria BC V8W 3P5, Canada

² Department of Marketing and Consumer Studies, University of Guelph, Ontario, Canada

³ Centre for the Study of Choice (CenSoC), University of Technology Sydney, PO Box 123 Broadway, Sydney, NSW 2007, Australia

Received 20 September 2011, revised version received 25 March 2012, accepted 21 June 2012

Abstract

Louviere et al. (2008, *J. of Choice Modelling*, 1, 126-163) present two main empirical examples in which a respondent rank orders the options in various choice sets by repeated best, then worst, choice. They expand the ranking data to various “implied” choices in subsets and fit the expanded data in various ways; they do not present models of the original rank data, except in one case (that of the rank ordered logit). We build on that work by constructing models of the original rank data that are consistent with the “weights” implied by the data expansions. This results in two classes of models: the first includes the reversible ranking model and has useful “score” properties; the second includes the rank ordered logit model and has natural “process” interpretations. We summarize known and new results on relations between the two classes of models and present fits of the models to the data of a case study concerning micro-generation of electricity using solar panels - that is, where individual households generate electricity using a renewable energy technology.

Keywords: best-worst choice; choice models; micro-generation; probabilistic choice; rank expansions; ranking; renewable energy

1 Introduction

Louviere and Woodworth (1990) and Finn and Louviere (1992) developed a discrete choice task in which a person is asked to select both the best and the worst

*ajmarley@uvic.ca

†islam@uoguelph.ca



option in an available (sub)set of choice alternatives (options); they also presented and evaluated a probabilistic model of their data. Recent research related to this topic by Louviere and his collaborators has begun to make a clear distinction between best-worst choice as a method, called Best Worst Scaling (BWS), and models that are applied to data obtained using BWS. For instance, Scarpa et al. (2011) obtained rank orders of options by asking respondents to make a sequence of best, then worst, choices - each choice set had 5 alternatives, and a rank order was obtained by a respondent making the following sequence of choices: the most preferred alternative out of 5; the least preferred of the remaining 4; the most preferred of the remaining 3; the least preferred of the final 2. Scarpa et al. fit the data by a model based on repeated best choices (namely, an extension of the rank-ordered logit) that includes three variance terms, each of which depends on one aspect of the design: the placement in the survey of the current choice set; the size of the current choice set; and the order in which the subsets of the current choice set are considered in producing the rank order. The important thing to note is that the choices were obtained by BWS, i.e., a sequence of best, then worst, choices, yet the model was formulated and tested as a sequence of 4 best choices. The model did have variance terms to account for the difference between the order in which the options were selected and the order in which they appear in the model. It remains for future research to determine whether this approach is reasonable, or whether it is preferable to apply models that include representations of best and worst choices and the order in which they are applied - as in Lancsar and Louviere (2009) and Collins and Rose (2011). Section 7 presents a case study that indirectly addresses the above question.

A related issue, if one treats BWS as a data collection method, rather than a model, is: what model should be fit to the resulting (ranking) data? This question is explored in Louviere et al. (2008), with emphasis on fitting models to the best (first) choices, only, of individuals, even when partial or full rank order data were collected by, say, repeated best, then worst. Specifically, the two main empirical examples in that paper had respondents rank order sets (of size 3, 4, or 5) by repeated best, then worst, choice. These rank data were then expanded to “implied” best choices on various subsets, and the expanded data used, in various ways, to fit the original best choice data with a multinomial logit (MNL) model; except in one case (that of the rank ordered logit), rank order models were neither presented, or fit, by Louviere et al. We build on their work by constructing models of the unexpanded rank data that are consistent with the “weights” proposed by the Louviere et al. expansions; these rank models belong to a general class, which we call *weighted utility ranking (WUR) models*, the general properties of which can be stated independently of the Louviere et al. expansions.

Full details on the use of expansions in fitting best choice data are given in Louviere et al. (2008); Islam, Louviere and Pihlens (2009); and Ebling, Frischknecht and Louviere (2010). In this paper, we only provide sufficient detail of those methods to illustrate particular weighted utility ranking models.

The remainder of the paper is as follows: Section 2 illustrates the data expansion methods in Louviere et al. (2008); Section 3 introduces the notation; Section 4 introduces the class of weighted utility ranking models, some of which, as shown in Section 5, can be partially motivated by the Louviere et al. expansions. Section 6 introduces ranking models based on repeated best and/or worst choice; the rank ordered logit is the best known model in this class; this section also presents known and new results relating weighted utility and repeated best and/or worst ranking models. Section 7 present fits of the models to the data of a case study concerning micro-generation of electricity using solar panels. Finally, Section 8 presents a discussion and conclusion. Appendices A and B include some technical material and Appendix C is a list of acronyms that appear in the paper.

2 Previous Work: Data Expansion Methods

The following example, taken from Louviere et al.(2008), illustrates their methods and results; all such methods and results generalize to larger examples. After introducing this material, we develop related probabilistic ranking models.

Consider drink combinations described by type of drink (coffee, tea, coke, water), price (\$1.00, \$1.50, \$2.00, \$2.50) and container (bottle, can). Construct 16 choice options using an orthogonal main effects plan (OMEF), where the 1st (respectively, 2nd) attribute is type of drink (respectively, price), each with 4 levels 0, 1, 2, 3, and the 3rd is container with 2 levels 0, 1 (Table 1).

Table 1. Descriptions of the options.

	Attribute		
	1 st	2 nd	3 rd
Option:	Levels:		
1	0	0	0
2	1	1	0
3	0	2	1
4	0	3	1
5	1	0	0
6	1	1	0
7	1	2	1
8	1	3	1
9	2	0	1
10	2	1	1
11	2	2	1
12	2	3	0
13	3	0	1
14	3	1	1
15	3	2	0
16	3	3	0

We assume that a balanced incomplete block design (BIBD) is used to put the 16 drink combinations of Table 1 into the 20 choice sets of Table 2, with each choice set of size 4; each drink combination occurs in 5 choice sets; and each pair of drink combinations occurs in one choice set.

Table 2. BIBD of 20 choice sets of size 4.

Choice Set:	Options in choice set:			
1	2	5	8	14
2	1	5	6	7
3	5	9	12	16
4	4	5	11	15
5	3	5	10	13
6	1	2	3	4
7	2	6	9	11
8	2	7	13	16
9	2	10	12	15
10	1	8	9	10
11	6	8	13	15
12	4	7	8	12
13	3	8	11	16
14	1	14	15	16
15	3	6	12	14
16	7	10	11	14
17	4	9	13	14
18	1	11	12	13
19	4	6	10	16
20	3	7	9	15

Suppose a participant has, separately, rank ordered the four options in each of the 20 choice sets of Table 2 from best to worst, and “...consider what happens if we “pretend” that we have more data than only the ranks, but do this in a very systematic and structured way.” (Louviere et al., 2008, p. 138). We now rephrase the details of this “pretence” (given on p. 138-9 of Louviere et al., 2008) and then, in Section 5, use the results as a motivation for probabilistic models of the unexpanded rank data. There, we also describe the similarities and differences between our approach (ranking models applied to rank data) and Louviere et al.’s approach (MNL models for best choices in rank data, with that rank data expanded to “implied” best choices on subsets).

Let $X = \{w, x, y, z\}$ denote a typical one of the 20 choice sets in Table 2, and assume that the participant has given the rank order $w \succ x \succ y \succ z$, where \succ means “is preferred to” - i.e., the ranking is from best (first) to worst (last)¹. Now

¹For the remainder of the paper, best (respectively, worst) refers to the first (respectively, last) option in such a (partial or full) rank order.

consider all two- and three-element subsets of X - Table 3 shows them, where, for each column, 1 (respectively, 0) means that the option at the head of that column is (is not) in the choice set in the current row. Following Louviere et al., we now “pretend” that because we know the preference order $w \succ x \succ y \succ z$, we can apply that order to predict the best choice for each 2- and 3-element subset of $X = \{w, x, y, z\}$. For example, choice set 10 in Table 3 is the 3-element subset $\{x, y, z\}$ and the “implied” best choice is x as it precedes both y and z in the rank order $w \succ x \succ y \succ z$. The final row of Table 3, Total, gives the “implied” number of best choices of each element of X for the given (data) rank order.

Comment Louviere et al. include subsets with one element (singletons) in their expansion of Table 3, as they do in all their expansions; this gives Totals of 8,4,2,1 in Table 3. The ranking probabilities of each of the ranking models that we motivate by such expansions are the same, with or without the inclusion of such singleton sets; so we do not mention this point again.

Table 3. “Implied” best choices for each 2- and 3-element subset of $\{w, x, y, z\}$ given the rank order $w \succ x \succ y \succ z$.

	Rank Position				
	1 st	2 nd	3 rd	4 th	
Preference Order:	w	x	y	z	
Choice Set:					Implied Choice:
1	1	1	0	0	w
2	1	0	1	0	w
3	1	0	0	1	w
4	0	1	1	0	x
5	0	1	0	1	x
6	0	0	1	1	y
7	1	1	1	0	w
8	1	1	0	1	w
9	1	0	1	1	w
10	0	1	1	1	x
11	1	1	1	1	w
Total	7	3	1	0	

A parallel expansion is applied to the participant’s data rank order for each of the 20 choice sets in Table 2. Louviere et al. then use the (natural log of the) Total (score, weight) for each option in each of the 20 choice sets as the dependent variable in fitting an MNL model to the best (first) choices in the rank order data for those sets. They use various estimation methods, such as weighted least squares (WLS), weighted conditional logit (WCL), and maximum likelihood (ML), and obtain similar estimates (up to a scale factor) with each method. The most recent related work (Ebling et al., 2010) obtains excellent in- and out-of-sample fits of an MNL model to the best choices in partial rank

data using a WLS method based on an expansion of that partial rank data - see Section 5.2 for related expansions.

As noted above, our main purpose is to further explore the use of such “expanded” best choices, and the resultant “implied” choice totals, in the motivation of probabilistic models of the unexpanded rank data. The details of the arguments appear in Section 5, with a brief summary here. Specifically, using Table 3 as illustration, we agree with Louviere et al. that, when choices are deterministic, the rank order $w \succ x \succ y \succ z$ is equivalent to (contains the same information as) the (implied) best choice for each of the choice sets in Table 3. However, we reason further that, when choices are probabilistic, the assumption of this equivalence (or other related ones discussed below) can be interpreted as specifying a particular (deterministic) functional relation between the probability of a ranking of a set and the best choice probabilities on subsets of that set; the details of such functional relations are given in Section 5. Said another way, in the probabilistic case, the (sample) rank order of a set and the (sample) best choices in the non-empty subsets of that set are only “equivalent” for a model satisfying a particular functional relation between the probabilities of ranking and of best choice. Importantly, that functional relation is not, in general, compatible with the multinomial logit (MNL) model holding exactly for the best choices in the rank orders.

Thus, our perspective is rather different than that of Louviere et al. (2008) and Ebling et al. (2010). Generally, their work involves using the expanded data in fitting the best choices, only, in the rank data, with those best choices assumed to satisfy an MNL model; the exception is their fitting the rank data by the rank ordered logit - see Section 6 for that model. Hence, it seems that they are implicitly assuming that the expanded data (choice totals) provide estimates of best choices that are improved relative to those obtained using the unexpanded rank data. We, on the other hand, focus on modeling the original rank data, but with ranking models motivated, to some extent, by the Louviere et al. expansions (choice totals). In fact, the likelihoods given by the two approaches are such that they yield the same sufficient statistic (Section 5.1.1).

3 General Notation

Let T with $|T| \geq 2$ denote the finite set of potentially available choice options, and let $D(T)$ denote the *design*, i.e., the set of (sub)sets of choice alternatives that occur in the study. For any subset $X \subseteq T$, with $|X| \geq 2$, $B_X(x)$ denotes the probability that the alternative x is chosen as best² in X , $W_X(y)$ the probability that the alternative y is chosen as worst in X , and $BW_X(x, y)$ the probability that the alternative x is chosen as best in X and the alternative $y \neq x$ is chosen

²Most previous work using similar mathematical notation has used $P_X(x)$ or $P(x|X)$ where we use $B_X(x)$. We use the latter for best, and $W_X(x)$ for worst, to distinguish clearly between such *Best* and *Worst* choice probabilities.

as worst in X . At this point, we do not consider whether x is chosen as best in X , then y is chosen as worst in $X - \{x\}$; or y is chosen as worst in X , then x is chosen as best in $X - \{y\}$; or x is chosen as best in X and $y \neq x$ is chosen as worst in X “simultaneously”; we do consider such issues when we present specific models of repeated best and/or worst choice. Thus

$$0 \leq B_X(x), W_X(y), BW_X(x, y) \leq 1$$

and

$$\sum_{x \in X} B_X(x) = \sum_{y \in X} W_X(y) = \sum_{\substack{x, y \in X \\ x \neq y}} BW_X(x, y) = 1.$$

Much of the following material uses the *multinomial logit (MNL) model for best choices*: there is a difference scale³ u such that for all $y \in Y \subseteq T$,

$$B_Y(y) = \frac{e^{u(y)}}{\sum_{z \in Y} e^{u(z)}}. \tag{1}$$

The representation restricted to $Y \subseteq T$, $|Y| = 2$, is the *binary MNL* model. The natural parallel for worst choices is the *multinomial logit (MNL) model for worst choices*: there is a difference scale v such that for all $y \in Y \subseteq T$,

$$W_Y(y) = \frac{e^{v(y)}}{\sum_{z \in Y} e^{v(z)}}.$$

Marley and Louviere (2005) present a theoretical argument for the case where $v = -u$, i.e., we have

$$W_Y(y) = \frac{e^{-u(y)}}{\sum_{z \in Y} e^{-u(z)}}. \tag{2}$$

Thus, the worst scale value is the negative of the best scale value; this is not a necessary restriction. One interpretation of this constraint is that, for the MNL model, it yields the result that the probability that $y \in Y$ is selected as *best* in a set Y with scale values $u(z)$, $z \in Y$, is equal to the probability that $y \in Y$ is selected as *worst* in a set Y with scale values $-u(z)$, $z \in Y$. In particular, when (1) and (2) both hold, we have that for all $x, y \in X$, $x \neq y$, $B_{\{x, y\}}(x) = W_{\{x, y\}}(y)$, and write the common value as $p(x, y)$; empirically, this relation may not always hold (Shafir, 1993). Collins and Rose (2011) present good fits of generalizations of (1) and (2) to ranking data obtained by repeated best, then worst, choices.

Next we need the idea of a ranking of a set X from its best (most preferred) to its worst (least preferred) element, and similarly of a ranking of X from its worst

³Psychologists might call this a *utility scale* or a *utility function*. We use the simpler term and, to avoid confusion, use the term *scale factor* when we are referring to changes in the variability of the random component of a random utility models.

(least preferred) to its best (most preferred) element. We refer to these as *best to worst*, and *worst to best*, rankings, respectively. Such rankings may be empirical, i.e., resulting from a person's judgments, but of at least equal importance in this paper is the theoretical role such rankings play in the development of probabilistic choice models, including those for best (worst) choices. For a set X , $X \subseteq T$, with $|X| = n$, $\rho = \rho_1\rho_2\dots\rho_{n-1}\rho_n$ denotes a typical best to worst rank order of X , and $R(X)$ denotes the set of rank orders of X ; as already noted, sometimes we write such a rank order in the form $\rho_1 \succ \rho_2 \succ \dots \succ \rho_n$. $B_{R(X)}(\rho)$ denotes the probability that ρ occurs as a best to worst rank order, and $W_{R(X)}(\rho)$ denotes the probability that ρ occurs as a worst to best rank order - thus, in the former case ρ_1 is the best element in the rank order, in the latter case it is the worst element. The assumption that $B_{R(X)}(\rho)$ and $W_{R(X)}(\rho)$ are probabilities and that a ranking occurs at each choice opportunity is summarized by: for each $\rho \in R(X)$,

$$0 \leq B_{R(X)}(\rho), W_{R(X)}(\rho) \leq 1$$

and

$$\sum_{\rho \in R(X)} B_{R(X)}(\rho) = 1 = \sum_{\rho \in R(X)} W_{R(X)}(\rho).$$

The *reverse* of a rank order $\rho = \rho_1\rho_2\dots\rho_{n-1}\rho_n$ is the rank order $\rho^* = \rho_n\rho_{n-1}\dots\rho_2\rho_1$. There is no necessary theoretical or empirical relation between $B_{R(X)}(\rho)$ and $W_{R(X)}(\rho^*)$, though it is often of interest to consider the constraint $B_{R(X)}(\rho) = W_{R(X)}(\rho^*)$.

The above notation does not explicitly indicate whether a ranking is obtained by, say, a sequence of best and/or worst choices; repeated best choices; or repeated worst choices. We introduce such notation in Section 6; otherwise we write $B_{R(X)}(\rho)$ for the probability of the rank order ρ , where ρ_1 is best, ρ_2 second best, etc., without reference to how it was obtained.

The above general notation has to be extended in various ways to handle the three main cases of BWS, defined in Louviere, Flynn, and Marley (2012): Case 1: (repeated) best and worst choice among 'things' - these things are usually generic objects (such as brands of detergent) or social issues (such as concern for the environment); Case 2: (repeated) best and worst choice of attributes in a profile of attributes - such as the best and worst attribute-level in a health state; Case 3: (repeated) best and worst choice among profiles - such as computers with different levels on various attributes. In Case 3, it is desired to measure the utility of individual attribute-levels, and of their combination in giving the overall utility of a profile (Marley & Pihlens, 2012). Louviere et al.'s data expansions have been applied mainly to Case 3, possibly because this case is the most closely related to revealed preference market data where a person purchases a multiattribute good or service. However, the basic expansion methods and ranking models can be described independent of whether Case 1, 2 or 3 is being studied, and that is what we do here.

4 New Approach: Weighted Utility Ranking Models

We now develop a class of ranking models that involve weights similar to those that arise in the Louviere et al. expansions. Certain of the models in this class can be indirectly motivated by those expansions - as is done Section 5. However, the results of the current section are in no way dependent on those expansions.

Definition 1 *Assume that, for every choice set $X \in D(T)$, there is a “local” importance weight $r_X(i)$ for position i , $i = 1, \dots, |X|$, of a rank order of the options in X , and that the weights satisfy the ordinal constraints*

$$r_X(1) \geq r_X(2) \geq \dots \geq r_X(|X| - 1) \geq r_X(|X|), \quad (3)$$

with $r_X(1) > r_X(|X|)$. Also assume that there is a (utility) scale u defined on the options of a set T . Then the ranking probabilities on the sets $X \in D(T)$, $|X| \geq 2$, satisfy a **weighted utility ranking (WUR) model** provided

$$B_{R(X)}(\rho) = \frac{e^{\sum_{i=1}^{|X|} r_X(i) u(\rho_i)}}{\sum_{\eta \in R(X)} e^{\sum_{i=1}^{|X|} r_X(i) u(\eta_i)}}. \quad (4)$$

Comment 1 When we have data on full rank orders, it is reasonable to assume that the inequalities in (3) are strict, in particular that $r_X(|X| - 1) > r_X(|X|)$. Now consider the case where all the choice sets X in the design $D(T)$, are of the same size k , and r_X is independent of X , i.e., $r_X(i) = r(i)$, and define, for $i = 1, \dots, k$,

$$r'(i) = \frac{r(i) - r(k)}{r(k-1) - r(k)}$$

and

$$u'(\rho_i) = u(\rho_i)[r(k-1) - r(k)].$$

Then it is easily checked that the ranking probabilities given by (4) are identical to those given by a model with the above weights and scale values. In particular, we have

$$r'_X(|X| - 1) = 1, r'_X(|X|) = 0. \quad (5)$$

We assume these two fixed weights in our fits in Section 7 of WUR models to full rank data.

Comment 2 It is easily checked that the ranking probabilities given by (4) for any set of weights are identical to those of a model with the weights: for each choice set X , and $i = 1, \dots, |X|$,

$$r'_X(i) = r_X(i) - \frac{1}{|X|} \left(\sum_{i=1}^{|X|} r_X(i) \right).$$

The latter weights satisfy the condition: for each choice set $X \in D(T)$,

$$\sum_{i=1}^{|X|} r'_X(i) = 0. \quad (6)$$

We generally do not impose the latter constraint as the representation of most of the (ranking) models are simpler without it. However, we do use it in proving a property of certain “scores” for the class of WUR models (see *Comment 1* later in this section).

Corollary 2 For $x \in X$ and $\rho \in R(X)$, define $\rho(x) = i$ if x is in rank position i in ρ ; and for convenience, put $r_X[\rho(x)] = 0$ whenever $x \notin X$. Then $r_X[\rho(x)]$ is well-defined and (4) becomes

$$B_{R(X)}(\rho) = \frac{e^{\sum_{x \in X} r_X[\rho(x)]u(x)}}{\sum_{\eta \in R(X)} e^{\sum_{x \in X} r_X[\eta(x)]u(x)}}. \quad (7)$$

Theorem 3 For each set $X \subseteq T$ in the design⁴ $D(T)$, and rank order $\rho \in R(X)$, let

$$s_X(\rho) = \begin{cases} 1 & \text{if } \rho \text{ is selected from } R(X) \\ 0 & \text{otherwise} \end{cases}. \quad (8)$$

Then, given the representation (7), equivalently (4), the set of scores, for $x \in T$,

$$s(x) = \left(\sum_{X \in D(T)} \sum_{\rho \in R(X)} s_X(\rho) r_X[\rho(x)] \right), \quad (9)$$

is a sufficient statistic.

⁴We assume each set $X \in D(T)$ is presented once. The results generalize easily to the case where each set X is presented N_X times.

Proof. The likelihood of the data is

$$\begin{aligned}
 & \prod_{X \in D(T)} \prod_{\rho \in R(X)} B_{R(X)}(\rho)^{s_X(\rho)} \tag{10} \\
 &= \prod_{X \in D(T)} \prod_{\rho \in R(X)} \left(\frac{e^{\sum_{x \in T^{r_X}[\rho(x)]u(x)}}}{\sum_{\eta \in R(X)} e^{\sum_{x \in T^{r_X}[\eta(x)]u(x)}}} \right)^{s_X(\rho)} \\
 &= \prod_{X \in D(T)} \left(\frac{1}{\sum_{\eta \in R(X)} e^{\sum_{x \in T^{r_X}[\eta(x)]u(x)}}} \right) \\
 &\quad \times \prod_{X \in D(T)} e^{\sum_{\rho \in R(X)} \sum_{x \in T^{s_X(\rho)r_X}[\rho(x)]u(x)}} \\
 &= \left(\frac{1}{\prod_{X \in D(T)} \sum_{\eta \in R(X)} e^{\sum_{x \in T^{r_X}[\eta(x)]u(x)}}} \right) \\
 &\quad \times e^{\sum_{x \in T} \left(\sum_{X \in D(T)} \sum_{\rho \in R(X)} s_X(\rho) r_X[\rho(x)] \right) u(x)}. \tag{11}
 \end{aligned}$$

So the set of *scores* defined by (9) is a sufficient statistic. ■

Comments

1. Using general language (with undefined terms in quotation marks), the following states a result due to Huber (1963) in such a way that it applies to WUR models; Appendix A states the terms and results exactly. Assume that one is interested in the rank order, only, of the (utility) scale values in the WUR model (7). An *acceptable loss function* is a “penalty” function with a value that remains constant under a common permutation of the scores and the scale values, and that increases if the ranking is made worse by misordering a pair of scale values. Let T be a master set with $n \geq 2$ elements and assume that, for some k with $n \geq k \geq 2$, every subset of T with exactly k elements appears in the design⁵ $D(T)$. Then, given the model in (7), ranking the scale values in descending order of the scores, breaking ties at random, has “minimal average loss” amongst all (“permutation invariant”) ranking procedures that depend on the data only through the set of scores.

2. Given the properties of the scores stated in *Comment 1.*, they are likely useful starting values in estimating the maximum likelihood values of the utilities $u(x)$, $x \in T$. In fact, various empirical work on the maximum difference (maxdiff) model for best-worst choice (a special case of the representation in (7) - see below) - gives a linear relation between the (best minus worst) scores and the (maximum likelihood) estimates of the utilities⁶ (Louviere et al., 2008, 2012).

⁵Further work is needed to extend the theoretical result to, say, balanced incomplete block (BIBD) designs. See Marley & Pihlens (2012) for related discussions of *connected designs*.

⁶Assume that the the maxdiff model holds, and a balanced incomplete design (BIBD) is used for the survey. If the utility (scale) values are in a small range - say, [-1,1] - then the above linear

Thus, WUR models have relatively simple representations for ranking probabilities, with very nice score properties. However, as with the special case of the reversible ranking model (Section 6.1, Def. 4), the best and worst choice probabilities are complex. In particular, let $B_X(x)$ for $x \in X$ be the sum of the rank order probabilities, given by (7), in which x is ranked first, and let $W_X(x)$ for $x \in X$ be the sum of the rank order probabilities, given by (7), in which x is ranked last. Then the representations of the best and worst choice probabilities are relatively complex, and have no “simple” process interpretation (other than as a sum of rank orders, with each rank order generated as above); in this sense, the rank orders can be considered basic, and the best and worst choice probabilities derived.

Next we show that the representation (7) includes various standard models of best, worst, and best-worst choice as special cases:

- i. The *MNL model for best choices* is the special case of (7) where: for $z \in T$

$$r_X[\rho(z)] = \begin{cases} 1 & \text{if } \rho(z) = 1 \\ 0 & \text{otherwise} \end{cases} .$$

Thus, we only have discriminating information (from the participant) on the first option in the ranking (“best”). So replacing the rank notation $B_{R(X)}(\rho)$ by the best choice notation $B_X(x)$ gives

$$B_X(x) = \frac{e^{u(x)}}{\sum_{r \in X} e^{u(r)}} .$$

- ii. The *MNL model for worst choices* is the special case of (7) where: for $z \in T$

$$r_X[\rho(z)] = \begin{cases} -1 & \text{if } \rho(z) = |X| \\ 0 & \text{otherwise} \end{cases} .$$

Thus, we only have discriminating information (from the participant) on the last option in the ranking (“worst”). So replacing the rank notation $B_{R(X)}(\rho)$ by the worst choice notation $W_X(x)$, gives

$$W_X(x) = \frac{e^{-u(x)}}{\sum_{r \in X} e^{-u(r)}} .$$

- iii. The *maxdiff model for best-worst choice* is the special case of (7) where: for $z \in T$

$$r_X[\rho(z)] = \begin{cases} 1 & \text{if } \rho(z) = 1 \\ -1 & \text{if } \rho(z) = |X| \\ 0 & \text{otherwise} \end{cases} .$$

relation holds under a first-order Taylor expansion of the maxdiff choice probabilities (Marley & Schlereth, unpublished).

Thus, we only have discriminating information (from the participant) on the first option in the ranking (“best”) and the last option in the ranking (“worst”). So replacing the rank notation $B_{R(X)}(\rho)$ by the best-worst notation $BW_X(x, y)$, $x \neq y$, gives

$$B_X(x, y) = \frac{e^{[u(x)-u(y)]}}{\sum_{\substack{r,s \in X \\ r \neq s}} e^{[u(r)-u(s)]}} ,$$

i.e., the *maxdiff model* of best-worst choice (Marley & Louviere, 2005). For each set $X \subseteq T$ in the design $D(T)$, and $z \in X$, let⁴

$$bw_X(z) = \begin{cases} 1 & \text{if } z \text{ is selected as best in } X \\ -1 & \text{if } z \text{ is selected as worst in } X \\ 0 & \text{otherwise} \end{cases} .$$

For convenience, also set $bw_X(z) = 0$ if $z \notin X$, and define: for $z \in T$,

$$bw(z) = \sum_{X \in D(T)} bw_X(z),$$

i.e., $bw(z)$ is the number of times across the design that z is chosen as best minus the number of times z is chosen as worst. Then, as a special case of the development of (9) and the results that follow from it, the set of scores $bw(z)$, $z \in T$, is a sufficient statistic, and ranking in descending of these scores is optimal in the sense of minimizing the expected loss.

- iv. *The reversible ranking model* (Section 6.1, Def 4) is the special case where: for $z \in T$,

$$r_X[\rho(z)] = |X| - i \quad \text{if } \rho(z) = i .$$

In this case, (9) is the *Borda score*, which has important desirable properties in voting (e. g., see Saari, 2008).

5 Relations Between Previous Data Expansion Methods and Weighted Utility Ranking Models

We now describe how the weights are derived in various of Louviere et al.’s expansions in the deterministic case for full and partial rank data; extend those motivations to the nondeterministic (probabilistic) case; and present specific WUR models motivated by that framework. Except where noted, all results on WUR models in this section and Section 6 are new, though we do use the Louviere et al. expansions to motivate specific models.

5.1 Full rank data

5.1.1 Expansion to all binary subsets

We continue using the example of Section 2. Louviere et al. (2008) do not present the following binary expansion, but it is presented in Louviere et al. (2012). As do Louviere et al., we begin by supposing that all choices are deterministic. Let $\rho_1 \succ \rho_2 \succ \rho_3 \succ \rho_4$ be a typical (data) rank order for each of the 4-element sets in Table 2. Louviere’s expansion of each such rank order to binary choices is as follows:

- i) replace the rank order by the “implied” binary choices $\rho_1 \succ \rho_2, \rho_1 \succ \rho_3, \rho_1 \succ \rho_4, \rho_2 \succ \rho_3, \rho_2 \succ \rho_4, \rho_3 \succ \rho_4$;
- ii) “pretend” (Louviere et al, 2008, p. 138) that those binary choices are independent;
- iii) “predict” the total number of implied best (first) choices for each option in the set $\{\rho_1, \rho_2, \rho_3, \rho_4\}$ from all possible binary comparisons of options in that set, given the rank order $\rho_1 \succ \rho_2 \succ \rho_3 \succ \rho_4$; this gives the Total in the final row of Table 4;
- iv) apply the same procedure to each data rank order obtained for each of the 20 sets in Table 2; this gives the “expanded ” data that is then used in fitting an MNL model to the best choices in the original rank data. For instance, as already mentioned, Louviere et al. use the following estimation methods: weighted least squares (WLS); weighted conditional logit (WCL); and maximum likelihood (ML).

Table 4. “Implied” best choices for each 2-element subset of $\{\rho_1, \rho_2, \rho_3, \rho_4\}$ given the rank order $\rho_1 \succ \rho_2 \succ \rho_3 \succ \rho_4$.

	Rank Position				
	1 st	2 nd	3 rd	4 th	
Preference					
Order →	ρ_1	ρ_2	ρ_3	ρ_4	
Choice Set:					Implied Choice:
1	1	1	0	0	ρ_1
2	1	0	1	0	ρ_1
3	1	0	0	1	ρ_1
4	0	1	1	0	ρ_2
5	0	1	0	1	ρ_2
6	0	0	1	1	ρ_3
Total	3	2	1	0	

Now we extend the above approach to the relation between probabilistic binary choices and probabilistic rankings (Table 5). We assume that the “expanded” binary choices are controlled by the binary choice probabilities in the

final column of Table 5 and that those probabilities satisfy the MNL; for instance, in row 4, ρ_2 is “chosen” over ρ_3 with probability $p(\rho_2, \rho_3)$. The MNL assumption is crucial in extending Louviere’s weight results (though not the method) from the deterministic to the probabilistic case⁷. We also assume that there is a set of best choice probabilities $P_X(x)$, $x \in X$, $X \in D(T)$, $|X| \geq 3$, that are generated by the same scale values as the binary choice probabilities. For now we leave it unspecified whether or not these are the theoretical best choice probabilities for sets with 3 or more options. Ultimately, Louviere et al. (2008) and Ebling et al. (2010) *do* assume that those best choice probabilities on larger sets (3 or more options) satisfy an MNL model with those scale values. Specifically, they use the data expansion methods in fitting an MNL model to the best choices in the rank data, not as a model of that (full) rank data. In contrast, we use the expansions as a way to motivate models of the rank orders. In particular, our models force the (marginal) best choice probabilities on larger sets (3 or more options) to *not* satisfy an MNL model.

Comment The general statements in the previous paragraph apply to all our uses in this paper of the Louviere et al. expansions and weights. We do not repeat those statements in later sections.

Table 5. Choice probabilities of the “implied” best choices for each 2-element subset of $\{\rho_1, \rho_2, \rho_3, \rho_4\}$ given the rank order $\rho_1 \succ \rho_2 \succ \rho_3 \succ \rho_4$.

	Rank Position					
	1 st	2 nd	3 rd	4 th		
Preference						
Order \rightarrow	ρ_1	ρ_2	ρ_3	ρ_4		
Choice Set:					Implied Choice:	Choice Probability:
1	1	1	0	0	ρ_1	$p(\rho_1, \rho_2)$
2	1	0	1	0	ρ_1	$p(\rho_1, \rho_3)$
3	1	0	0	1	ρ_1	$p(\rho_1, \rho_4)$
4	0	1	1	0	ρ_2	$p(\rho_2, \rho_3)$
5	0	1	0	1	ρ_2	$p(\rho_2, \rho_4)$
6	0	0	1	1	ρ_3	$p(\rho_3, \rho_4)$
Total	3	2	1	0		

The probability of the pattern of “implied” choices in Table 5 (which includes the deterministic case as a special case) is

$$p(\rho_1, \rho_2)p(\rho_1, \rho_3)p(\rho_1, \rho_4)p(\rho_2, \rho_3)p(\rho_2, \rho_4)p(\rho_3, \rho_4),$$

which is

$$\prod_{1 \leq i < j \leq 4} p(\rho_i, \rho_j). \tag{12}$$

⁷The method “works” for models other than the MNL, but the results are complex, and do not, in general, produce “weights” (total scores).

Note that this probability equals 1 (i.e., the deterministic case) only if each of the stated binary choice probabilities in the final column of Table 5 equals 1 - i.e., $p(\rho_i, \rho_j) = 1$ for $1 \leq i < j \leq 4$; these probabilities can approach 1 for the MNL model, (1), if each of $u(\rho_i) - u(\rho_{i+1})$, $i = 1, 2, 3$, is large and positive. However, except in the deterministic case (which is not exactly achievable by an MNL model), the sum of the probabilities (12) over all the choice patterns obtained by expanding all possible rank orders of a choice set X can be less than 1 (see next), and hence (12) does not produce a ranking model.

To show that the sum of (12) over all possible rank orders of a current choice set $X = \{w, x, y, z\}$ can be less than 1, assume every binary choice probability is in the open interval $(0, 1)$ and consider a participant making independent choices of the best element in each binary subset of X , of which there are 6. Then the probability over all possible patterns of such independent 6 choices is 1; includes the terms in (12) for all possible rank orders of X ; but also includes, for example, the term $p(w, x)p(y, w)p(z, w)p(x, y)p(x, z)p(y, z)$, which has its value in the open interval $(0, 1)$. However, that pattern includes the “cycle” of choices $w \succ x, x \succ y, y \succ z, z \succ w$, which is not compatible with the Table 5 expansion of any rank order of X . Hence this pattern of binary choices has positive probability, but is never produced by Table 5. Therefore, the sum of the terms in (12) over all rank orders is less than 1, and hence those terms do not correspond to a model of the data rank orders. So we next turn to such a ranking model, suggested by the representation in (12).

Comment If we use the notation $I_{R(X)}(\rho)$ for the expression in ((12)), then, with the assumption of a common MNL model across the binary choice probabilities, (12) becomes

$$I_{R(X)}(\rho) = \frac{e^{[3u(\rho_1)+2u(\rho_2)+u(\rho_3)]}}{C[(u(\rho_i), i \in \{1, 2, 3, 4\})]} \tag{13}$$

where

$$C[(u(\rho_i), i \in \{1, 2, 3, 4\})] = [e^{u(\rho_1)} + e^{u(\rho_2)}][e^{u(\rho_1)} + e^{u(\rho_3)}][e^{u(\rho_1)} + e^{u(\rho_4)}] \\ \times [e^{u(\rho_2)} + e^{u(\rho_3)}][e^{u(\rho_2)} + e^{u(\rho_4)}][e^{u(\rho_3)} + e^{u(\rho_4)}]$$

has the same value for every rank order on the given set.

Starting with (12), one “natural” way to generate a probability distribution over the rank orders is given by the following process (hinted at above): Given a presented set X , the respondent makes all binary comparisons of (distinct) pairs in that set; in our example, there are 6 such comparisons. If the resulting comparisons are compatible with a rank order, that rank order is selected; otherwise, the process starts over. This gives the probability of rank order $\rho \in R(X)$ as

$$B_{R(X)}(\rho) = \frac{\prod_{1 \leq i < j \leq 4} p(\rho_i, \rho_j)}{\sum_{\eta \in R(X)} \prod_{1 \leq i < j \leq 4} p(\eta_i, \eta_j)}. \tag{14}$$

In particular, when the binary choice probabilities satisfy a common MNL model, (14) can be rewritten in the form

$$B_{R(X)}(\rho) = \frac{e^{[3u(\rho_1)+2u(\rho_2)+u(\rho_3)]}}{\sum_{\eta \in R(X)} e^{[3u(\eta_1)+2u(\eta_2)+u(\eta_3)]}}. \quad (15)$$

Equation (14) is the *reversible ranking model* (Marley, 1968, and Section 6.1, Def. 4) for choice sets of size 4.

Results and questions parallel to those in the remainder of this subsection apply to all the expansions and related WUR models in this paper. We include them only for the present case.

Comments

1. Let X be a typical set with $|X| = 4$, then the ranking model (15) is a WUR model, (7), with weights $r_X(i) = |X| - i$ and sufficient statistic given by the scores in (9).

2. Note that (13) only differs from (15) in the form of the denominator which, in both forms, is independent of the data rank order. Now change the likelihood function (10) in the following way: replace each $B_{R(X)}(\rho)$ by $I_{R(X)}(\rho)$ as given by (13). Then a routine check shows that the argument in the proof of Theorem 3 showing that the scores in (9) is a sufficient statistic still holds. Remember, though, that the representations in (13) are a result of a Louviere et al. expansion, and are not a model of the original rank order data.

3. Define a set of best choice probabilities P_X for choice set X in the design by: for $x \in X$,

$$P_X(x) = \frac{e^{u(x)}}{\sum_{z \in X} e^{u(z)}}, \quad (16)$$

Then the representation (15) is equivalent to the form

$$B_{R(X)}(\rho) = \frac{P_X(\rho_1)^3 P_X(\rho_2)^2 P_X(\rho_3)}{\sum_{\eta \in R(X)} P_X(\eta_1)^3 P_X(\eta_2)^2 P_X(\eta_3)},$$

which can be thought of as a normalized weighted conditional logit (WCL) model. Related (non-normalized) WCL models are used by Louviere et al. (2008) in fitting the best choices in the rank order data. However, Louviere et al. assume that those best choices satisfy an MNL model, which is not the case for the above normalized WCL model (see the next comment).

4. The following results are from Marley (1968). The representation in (15) across subsets $X \subseteq T$ is not compatible with a random utility model. The probability $\tilde{B}_X(w)$ of choosing a particular option w from $X = \{w, x, y, z\}$ as best in the rank orders of X is given by the sum of the relevant ranking probabilities given by (15) - that is,

$$\tilde{B}_X(w) = \sum_{\eta \in R(X - \{w\})} B_{W_X}(w\eta).$$

Importantly, the marginal best choice probabilities \tilde{B}_X do *not* satisfy an MNL model for best choices. This leads to the question of what is the correct interpretation of the probabilities \tilde{B}_X - for instance: are they the probabilities of best choices only when a rank order is asked of a participant; or are they also the probabilities of best choices when a single best choice is asked of a participant. Finally, note that the model in (15) can be stated, and motivated, without the involvement of the (“dummy”) binary MNL model actually used in its development above - though its development from binary choice probabilities underscores its interpretation as a process model.

5. Assume that the design and Louviere expansion of this section are used to fit an MNL model to the best choices in the rank order data, using maximum likelihood applied to the likelihood expressions (13) - that is, for a typical set $X = \{w, x, y, z\}$, if the participant responds with the rank order $\rho = wxyz$, then the term in the likelihood function for that choice is given by (13) with $\rho_1 = w$, $\rho_2 = x$, $\rho_3 = y$, $\rho_4 = z$. Based on the results in various papers on the use of expansion methods (e.g., Louviere et al., 2008; Islam et al., 2009; Ebling et al., 2010), we expect a good fit of the estimated MNL model to the best choices in the data rank orders. However, we have pointed out in 4., above, that the best choices given by the related reversible ranking model in (15) do *not* satisfy an MNL model. Thus, it is of interest for future research to study: i. the extent to which the probability of best choices given by the reversible ranking model can deviate from, or be approximated by, an MNL model, and ii. whether the parameter estimates of the best-fitting MNL model for the best choices using the expansion methods yield a good fit of the related reversible ranking model to the rank order data.

5.1.2 Expansion to all non-empty, non-singleton subsets

The second expansion of full rank order data that we, and Louviere et al. (2008), consider is based on the assumption that a single sample rank order $\rho = \rho_1\rho_2\rho_3\rho_4$ for a set X can be expanded, in a consistent way, to all non-empty and non-singleton subsets of X (Table 6). We assume that the “expanded” choices are controlled by the best choice probabilities in the final column of Table 6 and that those probabilities satisfy the MNL model; for instance, in row 7, ρ_1 is “chosen” over ρ_2 and ρ_3 with probability $P_{\{\rho_1, \rho_2, \rho_3\}}(\rho_1)$. Remember, we are using the expansion as a motivation for a ranking model; as pointed out in previous sections, this is not its use in Louviere et al.

As before, the probability of the pattern of choices in Table 6 can approach 1 (i.e., be near deterministic) if each of the stated choice probabilities in the final column of Table 6 approaches 1, which they can for the MNL model, (1), if each of $u(\rho_i) - u(\rho_{i+1})$, $i = 1, 2, 3, 4$, is large and positive. However, as for the expansion of Table 5, except in the deterministic case (which is not exactly achievable with the assumed MNL representations), the sum of the probabilities of the product of the expressions in Table 6 over all (partial or full) rank orders can be less than

1.

Thus, assuming the choice probabilities in Table 6 satisfy an MNL model, and proceeding in a manner paralleling that in the previous section, we obtain the probability of rank order $\rho \in R(X)$ as

$$B_{R(X)}(\rho) = \frac{e^{[7u(\rho_1)+3u(\rho_2)+u(\rho_3)]}}{\sum_{\eta \in R(X)} e^{[7u(\eta_1)+3u(\eta_2)+3u(\eta_3)]}}. \tag{17}$$

Table 6. Choice probabilities of the “implied” best choices for each 2- and 3-element subset of $\{\rho_1, \rho_2, \rho_3, \rho_4\}$ given the rank order $\rho_1 \succ \rho_2 \succ \rho_3 \succ \rho_4$.

	Rank Position					
	1 st	2 nd	3 rd	4 th		
Preference						
Order \rightarrow	ρ_1	ρ_2	ρ_3	ρ_4		
Choice					Implied	Choice
Set:					Choice:	Probability:
1	1	1	0	0	ρ_1	$p(\rho_1, \rho_2)$
2	1	0	1	0	ρ_1	$p(\rho_1, \rho_3)$
3	1	0	0	1	ρ_1	$p(\rho_1, \rho_4)$
4	0	1	1	0	ρ_2	$p(\rho_2, \rho_3)$
5	0	1	0	1	ρ_2	$p(\rho_2, \rho_4)$
6	0	0	1	1	ρ_3	$p(\rho_3, \rho_4)$
7	1	1	1	0	ρ_1	$P_{\{\rho_1, \rho_2, \rho_3\}}(\rho_1)$
8	1	1	0	1	ρ_1	$P_{\{\rho_1, \rho_2, \rho_4\}}(\rho_1)$
9	1	0	1	1	ρ_1	$P_{\{\rho_1, \rho_3, \rho_4\}}(\rho_1)$
10	0	1	1	1	ρ_2	$P_{\{\rho_2, \rho_3, \rho_4\}}(\rho_2)$
11	1	1	1	1	ρ_1	$P_{\{\rho_1, \rho_2, \rho_3, \rho_4\}}(\rho_1)$
Total	7	3	1	0		

Results parallel to the *Comments* of Section 5.1.1 also hold for the above WUR model. In particular, the “dummy” MNL best probabilities in the final column of Table 6, used in motivating the ranking model, are *not* the (marginal) best choice probabilities of that model.

5.2 Partial rank data

Let $\{w, x, y, z\}$ be the set of available options and suppose that w is selected as best and z as worst. If we now check each (sub)set of size 2, 3, 4 of $\{w, x, y, z\}$ in turn, then we see that the subsets $\{x, y\}$ and $\{x, y, z\}$ are the only ones where the best element is not determined either by the information that w is best or by the information that z is worst. We next present the (Louviere) expansions for this case in a way that parallels the approach in Ebling et al (2010), though the weights we use differ slightly from theirs; at the end of the section, we present, and discuss, a WUR model with their weights.

5.2.1 Expansion to implied binary choice sets

As before, we assume that the binary choice probabilities satisfy the MNL model; the related data expansion to all binary choice sets (vis, Table 7) is mentioned in Louviere et al. (2008), but is not tested there. Denote a typical (data) partial rank order of $X = \{w, x, y, z\}$ by $\rho_1 \succ \{\rho_2, \rho_3\} \succ \rho_4$, with the expansion to all binary choice sets in Table 7; the dashes (–) in row 4 indicate that we have no relevant data for those rows. Remember, we are using that expansion as a motivation for a ranking model; as pointed out in previous sections, this is not its use in Louviere et al.

Table 7. Choice probabilities of the “implied” best choices for each 2-element subset of $\{\rho_1, \rho_2, \rho_3, \rho_4\}$ given the partial rank order $\rho_1 \succ \{\rho_2, \rho_3\} \succ \rho_4$.

	Rank Position					
	1 st	–	–	4 th		
Choice Set →	ρ_1	ρ_2	ρ_3	ρ_4		
Choice					Implied	Choice
Subset:					Choice:	Probability:
1	1	1	0	0	ρ_1	$p(\rho_1, \rho_2)$
2	1	0	1	0	ρ_2	$p(\rho_1, \rho_3)$
3	1	0	0	1	ρ_2	$p(\rho_1, \rho_4)$
4	0	1	1	0	–	–
5	0	1	0	1	y	$p(\rho_2, \rho_4)$
6	0	0	1	1	z	$p(\rho_3, \rho_4)$
Total	3	1	1	0		

Proceeding as in previous sections, the probability of the pattern of “implied” choices in Table 7 is

$$\begin{aligned}
 & p(\rho_1, \rho_2)p(\rho_1, \rho_3)p(\rho_1, \rho_4)p(\rho_2, \rho_4)p(\rho_3, \rho_4) \\
 & = p(\rho_1, \rho_2)p(\rho_1, \rho_3)p(\rho_1, \rho_4)[p(\rho_2, \rho_3) + p(\rho_3, \rho_2)]p(\rho_2, \rho_4)p(\rho_3, \rho_4).
 \end{aligned}$$

Adding the assumption of a common MNL model across the probabilities, this becomes

$$\frac{e^{[3u(\rho_1)+2u(\rho_2)+u(\rho_3)]} + e^{[3u(\rho_1)+2u(\rho_3)+u(\rho_2)]}}{C[(u(\rho_i), i \in \{1, 2, 3, 4\})]}$$

where

$$\begin{aligned}
 C[(u(\rho_i), i \in \{1, 2, 3, 4\})] & = [e^{u(\rho_1)} + e^{u(\rho_2)}][e^{u(\rho_1)} + e^{u(\rho_3)}][e^{u(\rho_1)} + e^{u(\rho_4)}] \\
 & \quad \times [e^{u(\rho_2)} + e^{u(\rho_3)}][e^{u(\rho_2)} + e^{u(\rho_3)}][e^{u(\rho_3)} + e^{u(\rho_4)}]
 \end{aligned}$$

has the same value for each of the rank orders on the given set.

As before, the probability of this pattern of choices can approach 1 (i.e., be near deterministic) if each of the stated binary choice probabilities in the final

column of Table 7 approaches 1, which they can for the MNL model, (1), if each of $u(\rho_1) - u(\rho_i)$, $i = 2, 3, 4$, and $u(\rho_i) - u(\rho_4)$, $i = 2, 3$, is large and positive. However, as for the expansion in Table 5, except in the deterministic case (which is not exactly achievable), the sum of the probabilities over all (partial or full) rank orders given by the above expression is less than 1.

The following process does generate a probability distribution over the observed best-worst choices (partial rank orders). Assume that the participant makes all 6 binary comparisons of distinct options in the 4-element set; if those choices give a rank order, the participant reports the first (respectively, last) option in that rank order as best (respectively, worst); otherwise, the participant resamples until a rank order is obtained. As we know only the best and worst choice, we replace the notation $B_{R(X)}(\rho)$ with $BW_X(\rho_1, \rho_4)$, and the above process gives the best-worst choice probability

$$\begin{aligned} BW_X(\rho_1, \rho_4) &= \frac{e^{[3u(\rho_1)+2u(\rho_2)+u(\rho_3)]} + e^{[3u(\rho_1)+2u(\rho_3)+u(\rho_2)]}}{\sum_{\eta \in R(X)} e^{[3u(\eta_1)+2u(\eta_2)+u(\eta_3)]}} \\ &= B_{R(X)}(\rho_1\rho_2\rho_3\rho_4) + B_{R(X)}(\rho_1\rho_3\rho_2\rho_4), \end{aligned}$$

where the terms in the second expression are those of the reversible ranking model, (15).

5.2.2 Expansion to implied non-empty, non-singleton subsets

As usual, we assume that all the choice probabilities in Table 8 satisfy the MNL model. Now we consider the expansion of the (data) partial rank order $\rho_1 \succ \{\rho_2, \rho_3\} \succ \rho_4$ to best choices on all non-empty, non-singleton subsets; this parallels, but is not identical to, the case presented in Ebling et al. (2010); we present their representation later in this section. Remember, we are using the expansion as a motivation for a ranking model; as pointed out in previous sections, this is not its use in Louviere et al.

The observed best-worst pair (ρ_1 best, ρ_4 worst) gives no information of the relative rank of ρ_2 and ρ_3 in the set $\{\rho_2, \rho_3\}$ (row 4) or in the set $\{\rho_2, \rho_3, \rho_4\}$ (row 10) - though, in the latter set, we do know that one of them has to be best as ρ_4 is worst in the choice set $\{\rho_1, \rho_2, \rho_3, \rho_4\}$. Also, if we assume that there is an underlying rank order⁸, then either $\rho_2 \succ \rho_3$, which means we place $p(\rho_2, \rho_3)$ for choice set 4 and $P_{\{\rho_2, \rho_3, \rho_4\}}(\rho_2)$ for choice set 10, or $\rho_3 \succ \rho_2$, which means we place $p(\rho_3, \rho_2)$ for choice set 4 and $P_{\{\rho_2, \rho_3, \rho_4\}}(\rho_3)$ for choice set 10.

As before, the probability of the pattern of choices in Table 8 can approach 1 (i.e., be near deterministic) if each of the stated choice probabilities in the final column of Table 8 approaches 1, which they can for the MNL model, (1), if each

⁸We could develop a model that uses only the partial order information of Table 3. However, the resulting representation does not simplify in the way that the model based on full rank orders does. In particular, it does not produce a model with the weights (8,3,3,1) that appear in Ebling et al. (2010) WLS expansion.

of $u(\rho_1) - u(\rho_i)$, $i = 2, 3, 4$, and $u(\rho_i) - u(\rho_4)$, $i = 2, 3$, is large and positive. However, as for previous expansions, except in the deterministic case (which is not exactly achievable with the assumed MNL representations), the sum of the probabilities over all (partial or full) rank orders given of the product of the expressions in Table 8 is less than 1.

Table 8. Choice probabilities of the “implied” best choices for each 2- and 3-element subset of $\{\rho_1, \rho_2, \rho_3, \rho_4\}$ given the partial rank order $\rho_1 \succ \{\rho_2, \rho_3\} \succ \rho_4$.

	Rank Position					
	1 st	2 nd	3 rd	4 th		
Preference						
Order \rightarrow	ρ_1	ρ_2	ρ_3	ρ_4		
Choice					Implied	Choice
Set:					Choice:	Probability:
1	1	1	0	0	ρ_1	$p(\rho_1, \rho_2)$
2	1	0	1	0	ρ_1	$p(\rho_1, \rho_3)$
3	1	0	0	1	ρ_1	$p(\rho_1, \rho_4)$
4	0	1	1	0	–	–
5	0	1	0	1	ρ_2	$p(\rho_2, \rho_4)$
6	0	0	1	1	ρ_3	$p(\rho_3, \rho_4)$
7	1	1	1	0	ρ_1	$P_{\{\rho_1, \rho_2, \rho_3\}}(\rho_1)$
8	1	1	0	1	ρ_1	$P_{\{\rho_1, \rho_2, \rho_4\}}(\rho_1)$
9	1	0	1	1	ρ_1	$P_{\{\rho_1, \rho_3, \rho_4\}}(\rho_1)$
10	0	1	1	1	–	–
11	1	1	1	1	ρ_1	$P_{\{\rho_1, \rho_2, \rho_3, \rho_4\}}(\rho_1)$
Total	7	1	1	0		

Turning to the related WUR representation, we know only that ρ_1 is best and ρ_4 is worst, so we replace $B_{R(X)}(\rho)$ with $BW_X(\rho_1, \rho_4)$, and reasoning as in the previous section, we obtain the representation

$$\begin{aligned}
 BW_X(\rho_1, \rho_4) &= \frac{e^{[7u(\rho_1)+3u(\rho_2)+u(\rho_3)]} + e^{[7u(\rho_1)+3u(\rho_3)+u(\rho_2)]}}{\sum_{\eta \in R(X)} e^{[7u(\eta_1)+3u(\eta_2)+3u(\eta_3)]}} \\
 &= B_{R(X)}(\rho_1\rho_2\rho_3\rho_4) + B_{R(X)}(\rho_1\rho_3\rho_2\rho_4),
 \end{aligned}$$

where the terms in the second expression are those of the reversible ranking model, (17).

The above representation can be rewritten as

$$BW_X(\rho_1, \rho_4) = \frac{e^{[8u(\rho_1)+4u(\rho_2)+2u(\rho_3)+u(\rho_4)]} + e^{[8u(\rho_1)+4u(\rho_3)+2u(\rho_2)+u(\rho_4)]}}{\sum_{\eta \in R(X)} e^{[8u(\eta_1)+4u(\eta_2)+2u(\eta_3)+u(\rho_4)]}}. \tag{18}$$

A weighted utility representation similar to those above, but based on Ebling et al.'s (2010) weights (8, 3, 3, 1), is

$$BW_X(\rho_1, \rho_4) = \frac{e^{[8u(\rho_1)+3u(\rho_2)+3u(\rho_3)+u(\rho_4)]}}{\sum_{\eta \in R(X)} e^{[8u(\eta_1)+3u(\eta_2)+3u(\eta_3)+u(\eta_4)]}}.$$

Appendix B shows that this representation can have $BW_X(\rho_1, \rho_4)$ approach 1, i.e., deterministic choice, by making each of the differences $u(\rho_1) - u(\rho_i)$, $i = 2, 3, 4$, and $u(\rho_i) - u(\rho_4)$, $i = 2, 3$, large and positive.

Note that the pattern of weights (8, 3, 3, 1) is the average of the (8, 4, 2, 1) weights for the rankings $\rho_1\rho_2\rho_3\rho_4$ and $\rho_1\rho_3\rho_2\rho_4$ in the representation (18). However, I do not think arguments paralleling those we have given yield this ‘‘average’’ model, even if one introduces singleton choice sets to Table 8.

6 Models of Ranking by Repeated Best and/or Worst Choices, and Their Relations to WUR Models

The classic expression for a ranking probability in terms of best choice probabilities, usually developed with the best choice probabilities satisfying the MNL model, (1), has the form: for $\rho = \rho_1\rho_2 \dots \rho_{|X|-1}\rho_{|X|}$,

$$B_{R(X)}(\rho) = B_X(\rho_1)B_{X-\{\rho_1\}}(\rho_2) \dots B_{\{\rho_{|X|-1}, \rho_{|X|}\}}(\rho_{|X|-1}). \quad (19)$$

However, the above expansion does *not* require the choice probabilities to satisfy the MNL model. We now summarize the classic argument for using the MNL in this context, then develop various ranking models in terms of repeated best and/or worst choices.

Assume that the (best) choice probabilities and the ranking probabilities on all subsets of a set T satisfy a *random utility model* in the sense that there are random variables $\mathbf{U}_z, z \in T$, such that for all $x \in X \subseteq T$ with $|X| \geq 2$, $\rho = \rho_1\rho_2 \dots \rho_{|X|}$,

$$B_{R(X)}(\rho) = \Pr(\mathbf{U}_{\rho_1} > \mathbf{U}_{\rho_2} > \dots > \mathbf{U}_{\rho_{|X|-1}}),$$

and

$$B_X(x) = \Pr(\mathbf{U}_x > \mathbf{U}_y, \forall y \in X - \{x\}).$$

Then a classic result is that a random utility model satisfies (19) if and only if the preference probabilities satisfy the MNL model, (1), (see Luce & Suppes, 1965, Th. 50).

As we have already noted, the relation in (19) between ranking and choice probabilities does *not* imply that the choice probabilities satisfy the MNL model.

This is obvious in the sense that one can have a set of choice probabilities on the subsets $X \subseteq T$ that do not satisfy the MNL model and then *define* the ranking probabilities on those subsets by (19). We now take such an approach to developing ranking models based on repeated best and/or worst choice.

For 4 options, there are 8 distinct patterns of best and/or worst questions, any one of which an experimenter can ask a participant to use for the generation of a rank order on those options; and, also, at least 8 possible (MNL-based) models of best and/or worst choice for the final rank order. Thus, if one truly believes that BWS is a method, not a model, then there are 64 possible combinations of method and model for the ranking of 4 options. Given such combinatorial possibilities, it is likely highly desirable to constrain the order in which a participant can produce a rank order - for instance, by instructing the participant in the order of best and/or worst choices to be used; removing each selected option from view once it has been selected; and modeling the data with the same sequence of best and/or worst choices - this is the approach in Scarpa and Marley (2011). Here, we introduce notation for 4 such models, illustrated with 4-element sets. First, we emphasize that we are using the notation $\rho = \rho_1\rho_2\rho_3\rho_4$ for the rank order in which ρ_i , $i = 1, 2, 3, 4$, is the option in rank position i , and the ranking is: best, $i = 1$; 2nd best, $i = 2$; 3rd best, $i = 3$; 4th best, $i = 4$ - thus, this notation does *not* carry information on the choices leading to this rank order; the latter information is introduced next.

Let r and s be two indicators that can, independently, take on the values b (for **best**) and w (for **worst**). Let $P_{R(X)}^{r,s}(\rho_1\rho_2\rho_3\rho_4)$ denote the probability of the rank order $\rho = \rho_1\rho_2\rho_3\rho_4$ (from best to worst), when the rank order is obtained by: the first choices is of type r ; the second of type s ; the third, again, of type r . Then we can have:

i. repeated best:

$$P_{R(X)}^{b,b}(\rho) = B_X(\rho_1)B_{X-\{\rho_1\}}(\rho_1)B_{\{\rho_3,\rho_4\}}(\rho_3). \quad (20)$$

ii. repeated worst:

$$P_{R(X)}^{w,w}(\rho) = W_X(\rho_4)W_{X-\{\rho_4\}}(\rho_3)W_{\{\rho_1,\rho_2\}}(\rho_2). \quad (21)$$

iii. repeated best-worst:

$$P_{R(X)}^{b,w}(\rho) = B_X(\rho_1)W_{X-\{\rho_1\}}(\rho_4)B_{\{\rho_2,\rho_3\}}(\rho_2). \quad (22)$$

iv. repeated worst-best:

$$P_{R(X)}^{w,b}(\rho) = W_X(\rho_4)B_{X-\{\rho_4\}}(\rho_1)W_{\{\rho_2,\rho_3\}}(\rho_3). \quad (23)$$

For each of (20) and (22), it is easily checked that the sum of the rank orders in which ρ_1 is first equals $B_X(\rho_1)$; and for each of (21) and (23), the sum of the rank orders in which ρ_4 is last equals $W_X(\rho_4)$. With slightly more work - assuming, say MNLs of the form given in (1) and (2) - one can show that, in general, for each of (20) and (22), the sum of the rank orders in which ρ_4 is last does *not* equal $W_X(\rho_4)$; and for each of (21) and (23), the sum of the rank orders in which ρ_1 is first does *not* equal $B_X(\rho_1)$.

The natural first assumption in testing these models is to assume that the best (respectively, worst) choice probabilities satisfy the MNL model (1) (respectively, (2)), and, as needed by data, generalizations of those models that include a scale factor that depends on the location of the current choice set in the rank order; this scale factor relates to the possibly changing variability (“consistency”) of the choices across sets. An example of such a generalization of the MNL model for best choices, (1) is: there is a nonnegative scale factor s defined for each integer $2, 3, \dots, |T|$, and a difference scale u such that for all $y \in Y \subseteq T$,

$$B_Y(y) = \frac{e^{s(|X|)u(y)}}{\sum_{z \in Y} e^{s(|X|)u(z)}}. \quad (24)$$

We then define a *generalized rank ordered logit model (GROL)* as a set of rank orders that satisfy (19) with the best choice probabilities satisfying (24). The form (24) is a special case of what Vermut and Magidson (2005, Section 2.4) call an MNL with *replication-specific scale factor* and of Fiebig et al.’s (2010) *generalized multinomial logit model (GMNL)*; the latter model also includes scale (variance) heterogeneity across individuals.

Scarpa et al. (2011) collected ranking data by repeated best, then worst, choices and fit that data quite successfully with a model based on repeated best choices, with those choices satisfying a generalization of the model in (24) that included dependencies of the scale factor s on aspects of the design that took account of the difference between the data collection method (repeated best, then worst) and the model (repeated best); it would be interesting to see if their data could be better fit by a model that matched their data collection method - i.e., the model of case iii., above. Collins and Rose (2011) fit related models to stated preference data on dating choices. Also, as required by data, the MNL in the above can be replaced by other models, such as the GMNL for best choices, also adapted to worst choices (Fiebig et al., 2010). And one can consider latent class extensions of these models.

In Section 7, in addition to fitting various WUR models, we fit the above rank ordered logit (ROL) and the generalized rank ordered logit (GROL).

6.1 Known relations between models of repeated best and/or worst choice and WUR models

As throughout the paper: T denotes a finite set of options; we use the convention that, for each $x \in T$, $B_{\{x\}}(x) = W_{\{x\}}(x) = 1$; and we write $p(x, y)$ for $B_{\{x,y\}}(x)$

and, if needed, $p^*(x, y)$ for $W_{\{x, y\}}(x)$. We also assume that the binary choice probabilities are *transitive* in the sense that, for any $r, s, t \in T$, $p(r, s) = p(s, t) = 1$ implies that $p(r, t) = 1$, and $p^*(r, s) = p^*(s, t) = 1$ implies that $p^*(r, t) = 1$; transitivity ensures that the various choice probabilities defined below are, in fact, probabilities.

1. The rank order logit

The rank order logit (ROL) is the ranking model (20) with the best choice probabilities satisfying the MNL model. Lam, Konig & Franses (2011) have developed a fascinating WUR model approximation to the ranking probabilities given by the ROL model, using a Taylor expansion around the scale value $u(x_o)$ of a referent option x_o . The authors present applications of the resulting WUR model to several data sets, and show that their interpretations of the data from the WUR model approximation to the ROL are similar to those obtained by fitting the ROL. An interesting theoretical question is whether, given a ROL model, there is a WUR model that gives *exactly* the same choice probabilities as that ROL model on one set or multiple sets. Answering this question will likely involve the fact that the ROL is a random utility model, and thus satisfies the constraints of such a model, whereas, in general, a WUR model is not a random utility model.

2. The reversible ranking model

The following results from Marley (1968) are needed for a result that we develop for WUR models that are also models based on repeated best and/or worst choices.

Definition 4 (Marley, 1968, Def. 6) *A reversible ranking model is a set of choice and ranking probabilities on all the non-empty subsets of a finite set T that satisfies (20), (21), and, for each $\rho \in R(X)$, $X \subseteq T$, $P_{R(X)}^{b,b}(\rho) = P_{R(X)}^{w,w}(\rho)$.*

Definition 5 (Marley, 1968, Def. 2) *A concordant choice model is a set of choice probabilities on all the non-empty subsets of a finite set T such that for each $x, y \in X \subseteq T$, $x \neq y$,*

$$B_X(x)W_{X-\{x\}}(y) = W_X(y)B_{X-\{y\}}(x). \tag{25}$$

Theorem 6 (Marley, 1968, Th. 8). *Assume a set of choice and ranking probabilities on all the non-empty subsets of a finite set T for which the binary choice probabilities are transitive. Then the following two conditions are equivalent:*

- (i) *the probabilities form a reversible ranking model;*
- (ii) *the probabilities form a concordant choice model and for each $\rho = \rho_1\rho_2\dots\rho_{|X|} \in R(X)$, $X \subseteq T$,*

$$P_{R(X)}^{b,b}(\rho) = P_{R(X)}^{w,w}(\rho) = \frac{\prod_{1 \leq i < j \leq |X|} p(\rho_i, \rho_j)}{\sum_{\eta \in R(X)} \prod_{1 \leq i < j \leq |X|} p(\eta_i, \eta_j)}. \tag{26}$$

Marley (1968) did not note the fact that the probability of a rank order (from best to worst) in the reversible ranking model is the same *whatever* pattern of repeated best and/or worst choices is used in its generation; we now give a short proof of that fact.

Corollary 7 *The probability of each rank order (from best to worst) given by the reversible ranking model is independent of the pattern of repeated best and/or worst choices that is used in its generation.*

Proof. This result is easily shown by a repeated use of (25) to convert any such pattern to a rank order generated by, say, all best choices. As above, let $\rho = \rho_1\rho_2\dots\rho_{|X|}$ be a rank order of X with the elements listed from first best to second best...to last best. If the final choice in obtaining the rank order is a worst choice, then it must be of the form $W_{\{\rho_i, \rho_{i+1}\}}(\rho_{i+1})$ for some $1 \leq i < N$. But $W_{\{\rho_i, \rho_{i+1}\}}(\rho_{i+1}) = B_{\{\rho_i, \rho_{i+1}\}}(\rho_i)$ and so that worse choice can be replaced by a best choice. By induction, if some other stage in obtaining the rank order involves a worst choice, then it is of the form $W_{\{\rho_i, \rho_{i+1}, \dots, \rho_{i+j}\}}(\rho_{i+j})$ for some $j > 1$, and all choices in the rank order on subsets of $\{\rho_i, \rho_{i+1}, \dots, \rho_{i+j}\}$ are best choices. Using the property (25) (repeatedly), we can move the worst choice to the binary choice $W_{\{\rho_{i+j-1}, \rho_{i+j}\}}(\rho_{i+j})$ and replace it by $B_{\{\rho_{i+j-1}, \rho_{i+j}\}}(\rho_{i+j-1})$. ■

Finally, suppose that the binary choices for a set $X \subseteq T$, $|X| \geq 2$, satisfy the MNL model - that is, for $x, y \in T$, $x \neq y$,

$$p(x, y) = \frac{e^{u(x)}}{e^{u(x)} + e^{u(y)}}. \tag{27}$$

Then it is easily checked that a set of ranking probabilities on X satisfy the reversible ranking model, (26), with those binary choice probabilities, iff those ranking probabilities satisfy a WUR model, (4), with $r_X(i) = |X| - i$ - that is, with the Borda score (Saari, 2008).

7 An Application

We present, and fit, the results of a case study concerning micro-generation of electricity using solar panels - that is, where individual households generate electricity using a renewable energy technology. This is a potentially very significant energy source as individual households account for one third of all energy consumption in North America. Growing energy demand, finite fossil fuel supplies, worries about energy security, and environmental concerns, are all factors encouraging the increasing use of renewable resources for electricity generation. A Discrete Choice Experiment (DCE) was conducted to examine trade-offs of features of solar (photo-voltaic, PV) panels for household level electricity generation among Canadian households. Data for the final survey were collected from June 1 to June 15, 2011. A total of 372 respondents were approached from a random sample of households in Ontario, Canada, 74 of whom either declined

or did not complete the survey; so the final sample for estimation included 298 completed questionnaires, a response rate of 80%. This sample size was considered adequate for the purpose of the study. The average completion time for those who completed the survey tasks was 25 minutes. Data were collected by Pureprofile (a large online panel provider in Australia, North-America, and other countries), and study participants were screened based on owning a house in Ontario, Canada. We carried out a pilot study on a convenience sample of 145 households in Guelph, Ontario, to determine the attributes, survey design, and feedback, to use in the final survey.

7.1 Experimental design

We used a relatively new approach to constructing DCEs proposed and applied by Louviere et al. (2008) - see the example in Section 2. Specifically, they proposed first constructing a set of profiles using an experimental design suitable for identifying particular forms of indirect utility functions, and then using a balanced incomplete block design (BIBD) to assign the profiles to choice sets of a fixed size. We constructed 16 profiles (choice options) using an orthogonal main effects plan (OMEP) sampled from the 64 x 64 factorial (3 attributes with 4 levels, i.e., $4^3 = 64$; 6 attributes with 2 levels, i.e., $2^6 = 64$). We then used a BIBD to assign the 16 profiles to 20 choice sets, each set having 4 options. Table 9 lists the attributes and attribute-levels in the profiles; these attributes and attribute-levels were chosen based on extensive research, reviews of product/retailer ads/claims, websites, pilot study, etc. Each respondent was shown the list of attributes and their ranges before the choice task, and informed that the stated savings in energy cost and carbon emission were on an annual basis.

Table 9. Attributes and their levels for solar technology for household electricity production

Attributes	Attributes levels varied in subsequent survey questions			
Total initial investment including installation and connection to national grid [*] .	\$20,000	\$ 25,000	\$30,000	\$35,000
Energy cost saving	10%	20%	30%	40%
Saving in carbon emission	0	1.0 tonnes of CO ₂	2 tonnes of CO ₂	3 tonnes of CO ₂
Payback period	5 year	10 years		
Tax Incentives & subsidy/rebates	Grant \$2500	Refund of HST ^{**}		
Export reward as per micro-FIT program (pass all or excess capacity to national grid)	64 cents/KWh ^{***}	80 cents/KWh		
Yearly inflation on fossil fuel cost	3%	6%		
Possibility of government policy changes about green energy technologies	No	Yes		
% of local households already adopted one of these technologies	5%	10%		

^{*}The cost varied in this research is based on 3KWh capacity. This 3KWh generation can meet average household demand.

^{**}Harmonized Sales Tax.

^{***}64 cents/KWh for Ground Mounted Panels and 80 cents/KWh for Roof Mounted Panels.

The response task was framed as a sequential choice process, with respondents instructed to choose the most preferred alternative out of four (Q1), then the least preferred out of the remaining three (Q2), and, finally, the most preferred out of the remaining two (Q3); Figure 1 shows the actual questions. Respondents were also asked to indicate (Q4, Fig. 1) if they would choose none of the 4 options; we do not analyze the answers to Q4 in this paper. This method of preference elicitation provides more information than, say, a best choice on each choice set; in particular, assuming a rank ordered logit for the ranking choices, it requires less than 30% of the sample size required to obtain the same precision using best choices, only (Scarpa et al., 2011).

Figure 1 shows a sample screen shot of the survey (choice set 1 out of 20 choice sets). Each time a respondent selected a profile, that profile disappeared from the screen so as to restrict the respondent’s next choice to the remaining profiles.

Figure 1. A sample screen shot from the survey (choice set 1 of 20)

Features	Option A	Option B	Option C	Option D
Total initial investment including installation and connection to national grid (3KWh Capacity)	\$35,000	\$25,000	\$20,000	\$20,000
Energy cost saving	40%	10%	10%	20%
Carbon emission saving	0	1 tonne of CO2	0	1 tonne of CO2
Payback period	5 years	10 years	10 years	5 years
Tax Incentives & subsidy/rebates	Grant \$2,500	Refund of HST	Grant \$2,500	Grant \$2,500
Export reward as per micro-FIT program (pass all or excess capacity to national grid)	80 cents/KWh, Roof Mounted	80 cents/KWh, Roof Mounted	64 cents/KWh, Ground Mounted	80 cents/KWh, Roof Mounted
Yearly inflation on fossil fuel cost	3%	3%	3%	6%
Possibility of government policy changes about green energy technologies	No	No	Yes	No
% of local households already adopted one of these technologies	5%	5%	10%	10%

Q1. Which of the four options would you MOST likely choose?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Q2. Which of the remaining three options would you LEAST likely choose?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Q3. Which of the two remaining options would you MOST likely choose?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Q4. If you would choose none of the four options, check the box to the right: <input type="checkbox"/>				

Scenario 1 of 20

7.2 Models

The ranking data was collected by repeated best, then worst, choices. However, as this is a case study and we are using the data as a preliminary test of WUR models, we focus on two WUR models motivated by the Louviere expansions - with preassigned weights (for choice sets of size 4) of (3, 2, 1, 0) and (7, 3, 1, 0). We also fit WUR models where we estimate the weights, rather than preassigning them; in this case, we use the property of (5) to set the 3rd and 4th weight to 1 and 0, respectively, and estimate the 1st and 2nd weight ("scale factor") following standard procedures (e.g., Bateman et. al., 2011, and Scarpa et. al., 2011).

Interestingly, for this data set, the estimated weights are then (4.804, 2.647, 1, 0), which are reasonably close to (3, 2, 1, 0). Therefore, in the following, we present the detailed fits of WUR models with preassigned weights (3, 2, 1, 0) and with estimated weights (4.804, 2.647, 1, 0). For comparison, we also fit the ROL model, i.e., (20) with the best choices satisfying the MNL model, (1), and the GROL model, i.e., (20) with the best choices satisfying (24); as for the WUR models, we fit with preassigned weights (“scale factors”)⁹ (3, 2, 1, 0) and with estimated weights. Note that the best choices in each of the ROL and the GROL satisfy an MNL model, which is what Louviere et al. (2008) and Ebling et al. (2010) assume in their data fits based on the Louviere expansions.

7.3 Model estimation results

All fits are by maximum likelihood (ML) using Gauss code. In all the model fits, the (utility) scale value of each profile (with a typical value $u(x)$) is the sum of the effects coded estimates¹⁰ of its attribute-level (β) values. The estimates of the WUR model with preassigned weights (3, 2, 1, 0) and for estimated weights are summarized in Table 10. We also estimated the WUR model with various “strange” weights (i.e., not monotonic decreasing from first to last) and in these cases model fit was very poor; for example, the log likelihood of the WUR model with weights (1, 8, 2, 4) is -18791.6, an increase in log likelihood of 1333.1 points compared to that for preassigned weights (3, 2, 1, 0).

Table 10 shows that, for each level of each attribute, the estimates have intuitively reasonable properties. For instance, the estimated utility decreases as the cost of a solar installation increases; increases as energy cost saving increases; increases as emission savings increase. The estimated benefits of solar panels as an investment also behave reasonably: a shorter payback time is preferred; a grant is preferred to a HST refund; a higher export reward is preferred. Three attributes do not have significant effects on the estimated utility: inflation of fossil fuel; government policy change; and percent of household adopted. Even though the t-ratios of these attributes are low, the orthogonal design ensures that their presence (or absence) in the model will not affect the significance of the other parameters. Figure 2 shows the fit of the 480 (20 choice sets x 24 rank orders) predicted rank order probabilities to the observed rank order probabilities and Figure 3 plots the mean centered maximum likelihood (ML) estimates for options (y-axis) against the mean centered weighted scores for options (x-axis) for the WUR model with preassigned weights (3, 2, 1, 0).

The highly linear (proportional) relation in Figure 3 is in agreement with related results in the literature (see Section 4), showing the extremely useful

⁹The 4th weight, 0, is not relevant as there is no set in the GROL to which it is applied. The 3rd weight applies to all binary choice sets. Assuming that the scale factor $s(3)$ in the formula (24) is non zero allows us to replace it by $s'(3) = 1$ and the scale u by $u' = s(3).u$

¹⁰That is, for each attribute, the omitted (β) attribute-level estimate in the tables of fits is the negative of the sum of the presented attribute level (β) estimates.

properties of the (mean centred) weighted scores for preliminary interpretation of the data. The non-parametric correlation, i.e. Spearman's ρ , is 1.0 between ML estimates and scores, in (exact) agreement with the theoretical result in Section 4, again demonstrating the value of these simple data statistics.

Table 10. Parameter estimates, and log likelihood, for the WURM with pre-assigned weights (3,2,1,0) and with estimated weights

Effect	Weights (3,2,1,0) (Log L = -17458.5)			Weights Estimated (Log L = -17420.8)			
	β	SE(β)	t-ratios	β	SE(β)	t-ratios	
Cost	\$25,000	0.100	0.011	9.45	0.062	0.009	7.08
	\$30,000	-0.095	0.010	-9.13	-0.057	0.009	-6.65
	\$35,000	-0.301	0.011	-27.83	-0.190	0.020	-9.52
Energy Cost Saving 20%	20%	-0.091	0.011	-8.68	-0.059	0.008	-7.07
	30%	0.061	0.011	5.75	0.038	0.007	5.09
	40%	0.288	0.011	26.28	0.181	0.019	9.74
Saving CO2 Emission	1 Tonne	-0.035	0.011	-3.35	-0.025	0.007	-3.57
	2 Tonnes	0.061	0.011	5.77	0.037	0.007	5.01
	3 Tonnes	0.138	0.011	12.60	0.078	0.010	7.56
Payback Period	10 Years	-0.055	0.006	-9.03	-0.035	0.005	-6.98
Tax Incentives & Subsidy	HST Refund	-0.026	0.006	-4.18	-0.019	0.004	-4.68
Export Reward	0.80\$/ KWH	0.098	0.006	15.90	0.059	0.007	8.24
Inflation of Fossil Fuel	6%	0.001	0.006	0.19	-0.005	0.005	-1.04
GOVT Policy Change	Yes	-0.006	0.006	-0.98	-0.008	0.004	-1.97
% of Household adopted	10%	0.007	0.006	1.09	0.007	0.004	1.74
				Weight 1	4.804	0.444	10.88
				Weight 2	2.647	0.223	11.80
				Weight 3	1.00 ¹	.	.
				Weight 4	0.000 ¹	.	.

¹ fixed

Figure 2. Predicted versus observed rank order probabilities for the WURM model with weights (3,2,1,0)

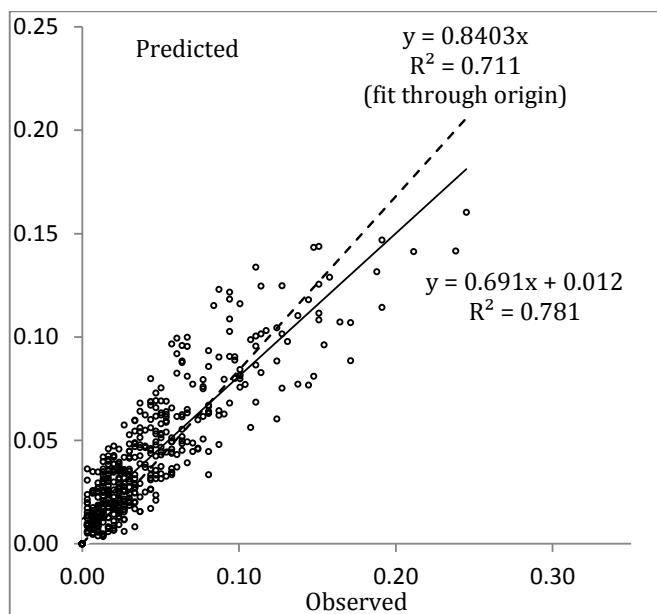


Figure 3. Mean centered weighted scores for options versus mean centered ML estimates for options for the WURM model with weights (3,2,1,0)

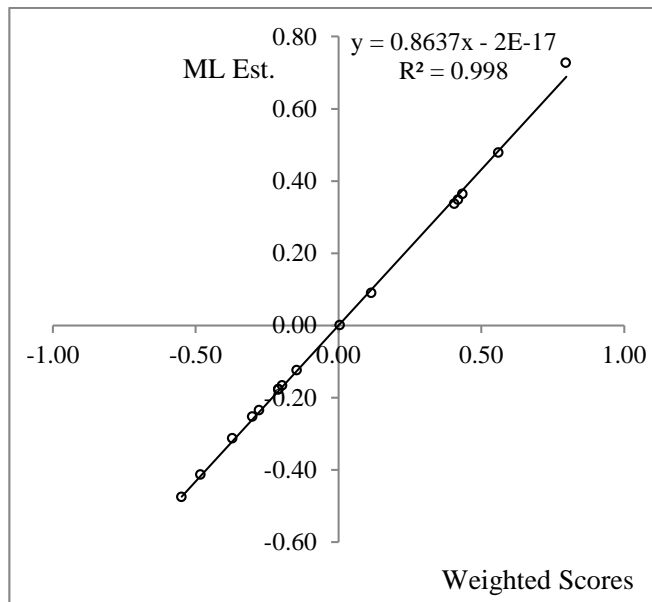
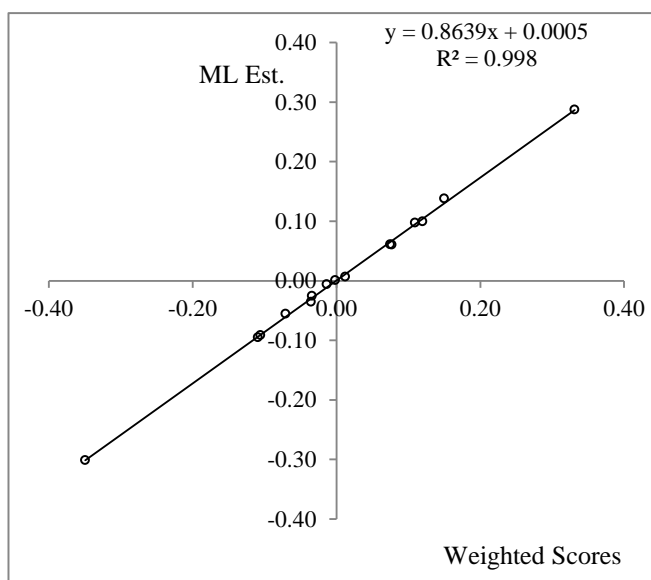


Figure 4 plots the corresponding highly linear relation between mean centered ML estimates for attribute-levels (y-axis) and the mean centered weighted scores for attribute-levels (x-axis) for the WUR model with preassigned weights (3, 2, 1, 0); the non-parametric correlation, i.e. Spearman’s ρ , is 0.999, between ML estimates and scores. We do not at this time have theoretical results predicting these relations.

We also fit the rank ordered logit (ROL) model and the generalized rank order logit (GROL) model, and obtained utility estimates similar to those for the corresponding WUR model. We omit the full details of the estimates and fits for the ROL and GROL models in the interests of space, plus the fact that the major purpose of the paper is to motivate and illustrate the properties of the WUR model; however, the following summarizes the main points. The (R^2) fits of the GROL model with preassigned weights (3, 2, 1, 0) are: (a) fit of observed and predicted rank order probabilities is 0.701 (0.772 when not constrained to fit through the origin); (b) fit of mean centered ML estimates for options versus mean centered weighted scores for options is 0.998, and (c) fit of mean centered ML estimates for attribute-levels versus mean centered scores for attribute-levels versus is 0.999. Thus, with these preassigned weights, the WUR model is slightly better than the GROL model in predicting observed rank order probabilities for this data set. As with the WUR model, the fits of the GROL model with “strange” weights are very poor. When we estimate weights for the GROL model, we obtain (3.42, 2.14, 1, 0) (1 and 0 fixed), which are close to (3, 2, 1, 0). In terms of LL (log likelihood), the fits of the GROL model and of the WUR model are

substantially better than those of the rank order logit (ROL). For example, the LL of the ROL model is -17587.8 compared to a LL of -17458.5 for the WUR model with preassigned weights (3, 2, 1, 0) and -17420.7 for the WUR model with estimated weights - thus, the improvements are 129 and 169 likelihood points for WUR model with preassigned and estimated weights. However, each of the WUR and GROL model with estimated weights has three additional parameters compared to the ROL model, and thus detailed comparisons of the three models will require use of fit measures that have penalties for (additional) parameters. Section 7.5 briefly discusses possible future comparisons of various extensions of the WUR model and its comparison with other models.

Figure 4. Mean centered weighted scores for attribute-levels versus mean centered ML estimates for attribute-levels for the WURM model with weights (3,2,1,0)



7.4 Within- and cross-sample model performance

To test the within and cross-sample model performance of the WUR model with preassigned weights (3, 2, 1, 0), we randomly split the 298 households into two groups (each sub-sample with 149 households). We evaluated the (R^2) fit of the model in- and out-of-sample on (a) rank order prediction, (b) option (i.e. BIBD profile) estimated values and (c) attribute-level estimated values. The fits, as summarized in Table 11, show only a slight drop from in- to out-of-sample fit, supporting the external validity for this data set of the WUR model with preassigned weights (3, 2, 1, 0).

Table 11. In and out of sample fit to the rank orders, option values, and attribute values, of the WURM with weights (3,2,1,0)

		Within Sample (R^2)		Cross Sample (R^2)
Rank Orders	Sample 1	0.774	Sample 2 \rightarrow 1 ¹	0.722
	Sample 2	0.781	Sample 1 \rightarrow 2	0.714
Options	Sample 1	0.998	Sample 2 \rightarrow 1	0.966
	Sample 2	0.999	Sample 1 \rightarrow 2	0.960
Attributes	Sample 1	0.948	Sample 2 \rightarrow 1	0.961
	Sample 2	0.999	Sample 1 \rightarrow 2	0.963

¹ For $i,j=1,2$, i not equal j , Sample $i \rightarrow j$ means sample i estimates used to predict sample j data.

7.5 Discussion of model estimation

We have successfully fit examples of the class of weighted utility ranking models, with these fits illustrating the usefulness of the fascinating score properties of these models, such as the observed linear (proportional) relation between the maximum likelihood estimates of the parameters and the scores. To the extent that these models differentially weight scale values at different stages of the ranking process, they include some scale effects of the kind considered in recent fits of the generalized rank ordered logit and related models (e.g., Scarpa et al, 2011). Further work, mainly empirical, is needed to extend WUR models to take into account both scale effects and preference heterogeneity (as in the models in Magidson & Vermunt, 2007; Fiebig et al., 2010).

8 Discussion and Conclusion

We have used the Louviere et al. (2008) expansion methods to motivate models of full and partial rank order data, several of which have been studied previously, others being new. This class (of weighted utility ranking models) includes standard choice models as special cases, including the multinomial logit for best (and worst) choice, the maxdiff model for best-worst choice, and the reversible ranking model (for rankings). Every model in this class has fascinating score properties, various of which can be shown theoretically (such as the scores providing a sufficient statistic and the “optimal” estimate of the rank order of the unknown scale values), others of which hold empirically (such as a linear relation between maximum likelihood estimates and the scores).

The weights suggested by the Louviere et al. (2008) expansion methods played an important role in motivating specific weighted utility ranking (WUR) models - in particular, each expansion (when the MNL model holds), and the corresponding ranking model, have a common sufficient statistic that depends on the weights in each approach. If one is interested in the rank order, only, of the scale values in the model, then, as noted above, assuming those scale values are ranked in the same order as the scores in the sufficient statistic is “optimal” (with optimal as

defined in Appendix A). To the extent that WUR models with different weights, but the same scale values, give the same rank order of the scores, they will give the same estimated rank order for the scale values. This fact, and the close relation between Louviere et al expansions and specific WUR models, might partially explain why various expansions give scale estimates related up to a scale factor (as in Louviere et al., 2008).

In general, there is a clear difference between the properties of WUR models and of ranking models based on repeated best and/or worst choices. The main advantage of the WUR models is that they have useful score properties for estimating the rank order of the scale values, and, empirically, of estimating the scale values up to a proportionality factor; the main advantage of the repeated best and/or worst models of ranking is that they provide natural process interpretations of the rankings and can be based on relatively standard (e.g., MNL) representations of the best and/or worst choice probabilities.

The reversible ranking model, with binary MNLs, belongs to both classes of representation. The advantage of this model lies in the fact that it has a simple sufficient statistic, based on scores, and that ranking in descending order of these scores is optimal in the sense of minimizing the expected loss; the disadvantage is its relatively complex forms for best, and worst, choice probabilities.

An open empirical and theoretical issue is the extent to which a WUR model can ‘mimic’ a repeated best and/or worst choice model, and vice-versa.

Given the diversity of ranking models presented in this paper, which are only a small subset of current approaches (e.g., Agarwal et al., 2009; Doignon et al., 2004), it is up to further empirical research to decide which ranking models are the most useful, in terms of, say, ease of use and in- and out-of-sample predictions. Models based on repeated best and/or worst choices can often be given (psychological) process interpretations, and those based on weights may be appropriate for policy making contexts, given their close links to scoring rules in voting theory (Garcia-Lapresta, Marley, & Martinez-Panero, 2010; Saari, 2008).

Appendices

Appendix A Score Properties (Based on Huber, 1963)

Preliminary comments

In order to apply Huber’s (1963) results under exactly the conditions he assumed, we need the scores in (9) to satisfy the constraints¹¹: for each choice set X ,

$$\sum_{i=1}^{|X|} r_X(i) = 0. \tag{28}$$

In this case, it is easily checked that the scores in (9) satisfy the property:

$$\sum_{x \in T} s(x) = 0. \tag{29}$$

¹¹Note that for every X and every rank order $\rho \in R(X)$, $\sum_{x \in X} r_X[\rho(x)] = \sum_{i=1}^{|X|} r_X(i)$.

When (28), and hence (29), hold, the conditions of our application agree with the assumptions in Huber (1963), especially the example in his Section 7. Therefore, we now show that these constraints can be derived from our general assumptions, with no impact on our other assumptions and conclusions.

So assume we have the scores as in (9) with the weights satisfying the ordinal constraint (3), i.e., for each $x \in T$ and choice set X ,

$$s(x) = \left(\sum_{X \in D(T)} \sum_{\rho \in R(X)} s_X(\rho) r_X[\rho(x)] \right),$$

and

$$r_X(1) \geq r_X(2) \geq \dots \geq r_X(|X| - 1) \geq r_X(|X|),$$

with $r_X(1) > r_X(|X|)$.

Now, for $i = 1, \dots, |X|$, define

$$r'_X(i) = r_X(i) - \frac{1}{|X|} \left(\sum_{i=1}^{|X|} r_X(i) \right),$$

and for $x \in T$, define

$$s'(x) = s(x) - \sum_{X \in D(T)} \frac{1}{|X|} \left(\sum_{i=1}^{|X|} r_X(i) \right). \tag{30}$$

Then it is routine to show that for each choice set X ,

$$\sum_{i=1}^{|X|} r'_X(i) = 0,$$

and for each $x \in T$,

$$s'(x) = \left(\sum_{X \in D(T)} \sum_{\rho \in R(X)} s_X(\rho) r'_X[\rho(x)] \right),$$

which in turn yield

$$\sum_{x \in T} s'(x) = 0.$$

Therefore, as required, the weights r'_X satisfy (28) and, consequently, the scores s' satisfy (29). Also, as already noted in the body of the paper, the ranking probabilities for the WUR with weights $r'_X(i)$, $i = 1, \dots, |X|$, are identical to those of the WUR with the weights $r_X(i)$; and, by (30), the rank order of the scores $s'(x)$, $x \in T$, for this common model is the same as that of the scores $s(x)$, $x \in T$. These are the conditions on the weights and scores that are required for our application of Huber's results.

In conclusion, we apply Huber’s results assuming a set of weights r_X satisfying (28), and hence scores s satisfying (29), even though in the main text we have not assumed that (28) holds.

The application of Huber’s (1963) results

Let T be a set with $|T| = n \geq 2$ elements and assume that, for some k with $n \geq k \geq 2$, every subset of T with exactly k elements appears in the design. Consider the group of permutations of the numbers 1 to n . For a vector of (utility) values $\mathbf{u} = (u(t_1), u(t_2), \dots, u(t_n))$ and a permutation σ , let $\mathbf{u}^\sigma = (u(t_{\sigma(1)}), u(t_{\sigma(2)}), \dots, u(t_{\sigma(n)}))$. The decision space D is the set of all possible rank orders; thus, a typical $d \in D$ has the form

$$d(1, 2, \dots, n) = (d(1), d(2), \dots, d(n)),$$

where $d(i)$, $i = 1, \dots, n$, is a permutation of the numbers 1 to n , and, thus, $d(i)$ is the rank of i in the ranking d . For $d \in D$ and permutation σ , let

$$d^\sigma(1, 2, \dots, n) = (d[\sigma(1)], d[\sigma(2)], \dots, d[\sigma(n)]).$$

We assume that the “true” ranking of the (utility) scale values (in the relevant models in the main text) is in descending order of their values $u(z)$, $z \in T$, and departures from the true ranking are punished by some real-valued loss; denote by $L(\mathbf{u}, d)$ the loss incurred when \mathbf{u} is the true value of the parameters and decision (rank order) $d \in D$ is taken.

Definition 8 (*Adapted from Huber, 1963, Th. 1, p. 513*) *A loss function L is called **acceptable** if it satisfies the following two condition:*

(i) *L is permutation invariant: for all u, d and σ ,*

$$L(\mathbf{u}^\sigma, d^\sigma) = L(\mathbf{u}, d).$$

(ii) *L does not decrease if the ranking, d , is made worse by interchanging two items. More precisely, assume that $u(t_i) \geq u(t_j)$ and let $d^{(i,j)}$ be the ranking in which i and j are interchanged relative to their position in d . If d ranks item i before item j , $d(i) < d(j)$, then*

$$L(\mathbf{u}, d) \leq L(\mathbf{u}, d^{(i,j)}).$$

Property (ii) implies that ranking in descending order of the $u(t_i)$ minimizes the loss, as desired.

We now define the *risk* of a decision procedure. Let \mathcal{S} be the (discrete) space of a statistic \mathbf{s} ; define $s^\sigma(t_i) = s(t_{\sigma(i)})$, and let $p(\mathbf{s}|\mathbf{u})$ be the probability (for the assumed model) of the statistic \mathbf{s} given the parameter vector \mathbf{u} . Let φ_s be a randomized decision procedure that depends on the data only through \mathbf{s} , that is, a probability density $\varphi_s(d)$ on D such that $\sum_{d \in D} \varphi_s(d) = 1$. The risk is then:

$$R(\mathbf{u}, \varphi) = \sum_{d \in D} \sum_{\mathbf{s} \in \mathcal{S}} p(\mathbf{s}|\mathbf{u}) L(\mathbf{u}, d) \varphi_{\mathbf{s}}(d).$$

We have written the statement of the following theorem in terms of a discrete probability mass. The analogous theorem in Huber (1963) is proved for general (discrete or continuous) probability distributions.

In the theorem statement, $T = \{t_1, t_2, \dots, t_n\}$ is a set, $\mathbf{s} = (s(t_1), s(t_2), \dots, s(t_n))$ with $\sum_{i=1}^n s(t_i) = 0$ and $\mathbf{u} = (u(t_1), u(t_2), \dots, u(t_n))$ are vectors of length n , and for any vector \mathbf{r} of length n , $\mathbf{r}^{(i,j)}$, $i \neq j$, is the vector in which items r_i and r_j are interchanged.

Theorem 9 (Adapted from Huber, 1963, p. 513) Assume that the joint distribution $p(\mathbf{s}|\mathbf{u})$ of \mathbf{s} conditional on \mathbf{u} can be written as

$$p(\mathbf{s}|\mathbf{u}) = c(\mathbf{u}) f(\mathbf{u}, \mathbf{s}),$$

where f satisfies

- i. $f(\mathbf{u}, \mathbf{s}) = f(\mathbf{u}^{(i,j)}, \mathbf{s}^{(i,j)})$,
- ii. $f(\mathbf{u}, \mathbf{s}) \geq f(\mathbf{u}^{(i,j)}, \mathbf{s})$ whenever $u(t_i) \geq u(t_j)$ and $s(t_i) \geq s(t_j)$.

If L is an acceptable loss function, Def. 8, then ranking in descending order of the $s(t_i)$, breaking ties at random, has minimal risk among all permutation invariant ranking procedures that depend on the data through \mathbf{s} .

Property i. says that f is invariant under permutations and property ii. says that the size of f decreases if the order of the elements t_i and t_j is different in \mathbf{u} than \mathbf{s} .

We now rewrite the likelihood, (11), for a WUR, in a form that allows us to use the result of the above theorem.

Let $|T| = n$, $T = \{t_1, t_2, \dots, t_n\}$ (in no particular order), $\mathbf{u} = (u(t_1), u(t_2), \dots, u(t_n))$, and $\mathbf{s} = (s(t_1), s(t_2), \dots, s(t_n))$. Then the likelihood, (11), of the score vector \mathbf{s} of (9) - that is, the probability of \mathbf{s} conditional on \mathbf{u} - can be written as

$$p(\mathbf{s}|\mathbf{u}) = c(\mathbf{u}) f(\mathbf{u}, \mathbf{s}),$$

where

$$c(\mathbf{u}) = \left(\frac{1}{\prod_{X \in D(T)} \sum_{\eta \in R(X)} e^{\sum_{x \in T^r X} [\rho(x)] u(x)}} \right)$$

and

$$f(\mathbf{u}, \mathbf{s}) = \exp \sum_{i=1}^n s(t_i) u(t_i). \tag{31}$$

It is routine to check that f in (31) satisfies i. and ii. of Theorem 9, hence the result of that theorem holds and ranking in descending order of the $s(t_i)$, breaking ties at random, has minimal risk among all permutation invariant ranking procedures that depend on the data through s .

Appendix B Calculation of (Near) Deterministic Limit for the (8,3,3,1) WUR Model

Assume the utility values for the four options w, x, y, z have the values a, b, c, d , respectively, with $a > b \geq c > d$ - that is, $u(w) = a, u(x) = b, u(y) = c, u(z) = d$. Then, (4) becomes

$$\begin{aligned} BW_X(w, z) &= \frac{e^{[8a+3(b+c)+d]}}{\sum_{\eta \in R(X)} e^{[8u(\eta_1)+3u(\eta_2)+3u(\eta_3)+u(\rho_4)]}} \\ &= \frac{1}{1 + \sum_{\eta \in R(X)} e^{[8u(\eta_1)+3u(\eta_2)+3u(\eta_3)+u(\rho_4)] - [8a+3(b+c)+d]}}. \end{aligned}$$

Then, for $BW_X(w, z)$ to approach 1, we need to show that, for any permutation (rank order) (g, h, i, k) of the options in the set $\{w, x, y, z\}$, other than (w, x, y, z) or (w, y, x, z) , the quantity

$$[8u(g) + 3(u(h) + u(i)) + u(k)] - [8a + 3(b + c) + d]$$

can be made as negative as desired by making the differences $a - b, b - c, c - d$, and, therefore, $a - c, b - d$, and $a - d$, large and positive. That is, we need to show that, for any permutation (rank order) (g, h, i, k) of the options, the quantity

$$\begin{aligned} &[8a + 3(b + c) + d] - [8u(g) + 3(u(h) + u(i)) + u(k)] \\ &= 8(a - u(g)) + 3[(b + c) - (u(h) + u(i))] + (d - u(k)) \end{aligned}$$

can be made as positive as desired by making the differences $a - b, b - c, c - d$ large and positive.

Table B1 demonstrates this by full enumeration, where it can be checked that each term in the right hand column can be made as positive as desired under the constraints on a, b, c , and d given above.

Table B1. Conditions for positivity.

$u(g), u(h), u(i), u(k)$	$8(a - u(g)) + 3[(b + c) - (u(h) + u(i))] + (d - u(k))$
a, b, c, d	0
a, c, b, d	0
a, b, d, c	$2(c - d)$
a, d, b, c	$2(c - d)$
a, c, d, b	$2(b - d)$
a, d, c, b	$2(b - d)$
b, a, c, d	$5(a - b)$
b, c, a, d	$5(a - b)$
b, a, d, c	$5(a - b) + 2(c - d)$
b, d, a, c	$5(a - b) + 2(c - d)$
b, d, c, a	$7(a - b) + 2(b - d)$
b, c, d, a	$7(a - b) + 2(b - d)$
c, a, b, d	$5(a - c)$
c, b, a, d	$5(a - c)$
c, a, d, b	$5(a - c) + 2(b - d)$
c, d, a, b	$5(a - c) + 2(b - d)$
c, b, d, a	$7(a - c) + 2(c - d)$
c, d, b, a	$7(a - c) + 2(c - d)$
d, a, b, c	$5(a - d) + 2(c - d)$
d, b, a, c	$5(a - d) + 2(c - d)$
d, a, c, b	$5(a - d) + 2(b - d)$
d, c, a, b	$5(a - d) + 2(b - d)$
d, c, b, a	$7(a - d)$
d, b, c, a	$7(a - d)$

Appendix C Acronyms

- BIBD Balanced incomplete block design
- BWS Best-worst scaling
- GMNL Generalized multinomial logit
- GROL Generalized rank ordered logit
- ML Maximum likelihood
- MNL Multinomial logit
- OMEF Orthogonal main effects design
- ROL Rank ordered logit
- WCL Weighted conditional logit
- WLS Weighted least squares
- WUR Weighted utility ranking

Acknowledgments

This research has been supported by Natural Science and Engineering Research Council Discovery Grant 8124-98 to the University of Victoria for Marley. The work was carried out, in part, whilst Marley was a Distinguished Research Professor in the Centre for the Study of Choice, University of Technology, Sydney. We thank Terry Flynn and Riccardo Scarpa for general comments and Bart Frischknecht and Christine Eckert for their patient discussions of what we call the “Louviere expansions.” Any misinterpretations of those expansions are our own.

References

- Agarwal, S., C. Burgess, and K. Crammer (2009) *Advances in Ranking*. Workshop, Twenty-Third Annual Conference on Neural Information Processing Systems, Whistler, BC, December 11, 2009.
- Bateman, H., T. Islam, J. Louviere, S. Satchell and S. Thorp (2011) Retirement investor risk tolerance in tranquil and crisis periods: experimental survey evidence. *Journal of Behavioral Finance*, **12** 201-218.
- Collins, A. T. and J. M. Rose (2011) Estimation of stochastic scale with best-worst data. Manuscript, University of Sydney.
- Doignon, J.-P., A. Pekeć, and R. Regenwetter (2004) The repeated insertion model for rankings: missing link between two subset choice models. *Psychometrika*, **69** 33-54.
- Ebling, C., B. Frischknecht and J. Louviere (2010) Keep it simple: Easy ways to estimate choice models for single consumers. In P. Ballantine and J. Finsterwalder (Eds.) *Proceedings of the Australian and New Zealand Marketing Academy Conference*, pp. 1-7. University of Canterbury, Christchurch, New Zealand.
- Fiebig, D. G., M. P. Keane, J. Louviere, J. and N. Wasi (2010) The generalized multinomial logit model: Accounting for scale and coefficient heterogeneity. *Marketing Science*, **29** 393-421.
- Finn, A. and Louviere, J. 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.
- Huber, P. J. (1963) Pairwise comparison and ranking: optimum properties of the row sum procedure. *Annals of Mathematical Statistics*, **34** 511-520.
- Garcia-Lapresta, J. L., A. A. J. Marley and M. Martinez-Panero (2010) Characterizing best-worst voting systems in the scoring context. *Social Choice and Welfare*, **34** 487-496.
- Islam, T., J. Louviere, J. and Pihlens, D. (2009) *Aggregate choice and individual models: A comparison of top-down and bottom-up approaches*. In Proceedings of the Sawtooth Conference, March 2009. Sawtooth Software Inc.

Lam, K. Y., A. J. Koning and P. H. Franses, (2011) Estimating independent locally shifted ranking utility models of ranking data. *Multivariate Behavioral Research*, **46** 756-778.

Lancsar, E. and J. J. Louviere (2009) Estimating individual level discrete choice models and welfare measures using best worst choice experiments, CenSoC Working Paper No. 08-003, University of Technology, Sydney.

J. Louviere, T.N. Flynn and A. A. J. Marley (2012) Best-worst scaling: Theory, methods and applications. Manuscript, University of Technology, Sydney.

Louviere, J. J. and Islam, T. (2008) A Comparison of importance weights/measures derived from choice-based conjoint, constant sum scales and best worst scaling. *Journal of Business Research*, **61** 903-911.

Louviere, J. J. and G. Woodworth (1990) Best-worst scaling: A model for largest difference judgments. Working Paper, Faculty of Business, University of Alberta.

Louviere, J. J., D. Street, L. Burgess, N. Wasi, T. Islam and A. A. J. Marley, A. A. J. (2008) Modeling the choices of individual decision-makers by combining efficient choice experiment designs with extra preference information. *Journal of Choice Modelling*, **1** 126-163.

Luce, R. D. and P. Suppes (1965) Preference, utility, and subjective probability. In R. D. Luce, R. R. Bush and E. Galanter. (Eds.). *Handbook of Mathematical Psychology, Vol. III*. New York, NY: John Wiley and Sons, pp. 235-406.

Magidson, J. and Vermunt, J. K. (2007) Removing the scale factor confound in multinomial logit choice models to obtain better estimates of preference. *Sawtooth Software Conference Proceedings*.

Marley, A. A. J. (1968) Some probabilistic models of simple choice and ranking. *Journal of Mathematical Psychology*, **5** 311-332.

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

Marley, A. A. J., and D. Pihlens (2012) Models of best-worst choice and ranking among, and in, multiattribute options (profiles). *Journal of Mathematical Psychology*, **56** 24-34.

Saari, D. G. (2008) *Disposing dictators, demystifying voting paradoxes*. Cambridge University Press, Cambridge, UK.

Scarpa, R. and A. A. J. Marley (2011) Exploring the consistency of alternative best and/or worst ranking procedures. Presentation, *Second International Choice Conference*, Leeds, July 4-6, 2011.

Scarpa, R., S. Notaro, R. Raffelli, D. Pihlens and J. Louviere (2011) Exploring scale effects of best/worst rank ordered choice data to estimate visitors's benefits from alpine transhumance. *American Journal of Agricultural Economics*, **93** 809-824.

Shafir, E. (1993) Choosing versus rejecting: Why some options are both better and worse. *Memory and Cognition*, **21** 546-556.

Vermunt, J. K. and J. Magidson (2005) *Technical Guide for Latent GOLD*

Choice 4.0: Basic and Advanced. Belmont Massachusetts: Statistical Innovations.