A Note On the Use of Rank Ordered Logit Models for Ordered Response Categories

Timothy R. Johnson*

Department of Mathematics and Statistical Science
University of Idaho

January 11, 2022

Abstract

Models for rankings have been shown to produce more efficient estimators than comparable models for top choices. These models could also be applied to rankings of the categories of a response scale, but they are less suitable when response scale categories are ordered. This paper proposes modeling rankings of ordered response categories using a new version of the rank ordered logit model where the distribution of rankings is truncated to permit only those rankings that are admissible given the ordering of the categories. This model can also be applied to partial rankings, and can be adapted to a variety of cases where models of top choices have been used. Two simulation studies featuring models based on the stereotype regression model and the rating scale item response model show the greater efficiency of modeling (partial) rankings of ordered response categories in comparison to top choices. Some opportunities for generalizations of this model are also discussed.

Introduction

Chapman and Staelin (1982) recommended collecting and modeling rankings instead of first/top choices in multinomial logit models. They showed greater statistical efficiency of maximum likelihood estimators when modeling rankings in simulation studies. Beresteanu and Zincenko (2018) showed that rankings result in a lower asymptotic variance of maximum likelihood estimators. These studies also show that efficiency increases with the "depth" of the ranking when partial rankings are used. A variety of studies have used rankings as an alternative to first/top choices or ratings. Rank ordered logit models have been used to model rankings of, for example, the suitability of job applicants with résumé discrepancies (Kuhn, Johnson, & Miller, 2013), the perceived dependence of human traits on genetics (Shostak, Freese, Link, & Phelan, 2009), the importance of universal values to Canadian, Chinese, and Japanese respondents (Tafarodi et al., 2009), the importance of life domains to older adults (Hsieh, 2005), and blame of public officials for property damage and loss of life due to Hurricane Katrina (Malhotra & Kuo, 2008). See Allison and Christakis (1994) for a overview of the use of rank ordered logit models.

*Address: Department of Mathematics and Statistical Science, University of Idaho, 875 Perimeter Drive MS 1104, Moscow ID 83844-1104. Email: trjohns@uidaho.edu.

Models for first/top choices are closely related to regression models for ordered response categories that can be written as log-linear models for the odds for adjacent categories (e.g., Fullerton, 2009; Tutz, 2021). In these models only a single category is chosen. But suppose instead the respondents *rank* the ordered response categories from that which best represents their response to that which least represents their response. In rank ordered logit models, the alternatives are not assumed to be ordered in any particular way, and the model allows for rankings that are not consistent with an ordering of the alternatives. This paper suggests that it can be useful to elicit and model rankings of ordered response categories by restricting the admissible rankings to those that are consistent with the order of the categories. This can be done through a modification of the rank ordered logit model for unordered alternatives by restricting the set of possible rankings to those that are admissible given the ordering of the categories, and imposing a structure on the model like that used in ordinal regression models. As has been observed for models of unordered alternatives, this approach has the potential for greater statistical efficiency over models based only on the first/top category. Furthermore, the extra demands on the respondents in giving a ranking rather than a single choice is smaller than that for unordered alternatives. This is because when ordered categories are ranked sequentially, fewer individual choices may be necessary to determine a full ranking due to constraints on the set of admissible rankings.

The rest of this paper is organized as follows. After a brief review of the rank ordered logit model for unordered alternatives, a modified model is proposed for rankings of ordered response categories. This includes versions of the model for partial rankings. Two simulation studies are then summarized that feature a stereotype regression model for rankings of ordered response categories, and rating scale item response model for rankings of item categories. These studies investigate the statistical efficiency of these models for full and partial rankings, including models where only the top choice is elicited.

Rankings of Unordered Alternatives

Assume a set of K alternatives labeled by the integers $1, 2, \dots, K$, and let $\boldsymbol{y} = (y_1, y_2, \dots, y_K)'$ be an ordering vector of the set of alternatives such that y_k is the alternative ranked k-th. For example, the ordering vector $\boldsymbol{y} = (2, 3, 1)'$ means that the second alternative is ranked first, the third alternative is ranked second, and the first alternative is ranked third. A regression model for $\boldsymbol{y}$ would specify the probabilities of all $K!$ possible ordering vectors as a function of vector of covariates $\boldsymbol{x}$. To simplify notation and where there is no loss of clarity an observation index will be omitted until estimation is considered later, but it should be understood that $\boldsymbol{y}$, $\boldsymbol{x}$, and quantities that are functions thereof may vary over observations.

One regression model for rankings that has been extensively discussed in the literature and widely applied in practice is the rank ordered or "exploded" logit model (Beggs, Cardell, & Hausman, 1981; Chapman & Staelin, 1982; Punj & Staelin, 1978). This model can be viewed as an application of the Plackett-Luce model for rankings (Luce, 1959; Plackett, 1975), and as an extension of the logit model for discrete choice data (McFadden, 1974). Let Y_k and y_k denote the k-th element of the ordering vector and its realization, respectively. The rank ordered logit model can be written as

$$P(Y_1 = y_1, Y_2 = y_2, \dots, Y_K = y_K) = \prod_{k=1}^{K-1} \frac{e^{\eta_{y_k}}}{\sum_{l \in \mathcal{S}_k} e^{\eta_l}}, \tag{1}$$

where $\mathcal{S}_k = \{y_k, y_{k+1}, \dots, y_K\}$, and η_k is a function of a vector of covariates $\boldsymbol{x}$. If the covariates do not vary over the alternatives then η_k might be specified as $\eta_k = \alpha_k + \boldsymbol{\beta}_k'\boldsymbol{x}$. This model can be easily modified for partial rankings where only the top $K' < K - 1$ alternatives are ranked so that

$$P(Y_1 = y_1, Y_2 = y_2, \ldots, Y_{K'} = y_{K'}) = \prod_{k=1}^{K'} \frac{e^{\eta_{y_k}}}{\sum_{l \in \mathcal{S}_k} e^{\eta_l}}. \tag{2}$$

This includes as special cases both discrete choice models where only the alternative ranked first is observed so that $K' = 1$, and complete rankings where $K' = K - 1$.

Frequently in applications ranking data are elicited sequentially by asking respondents to first indicate their top choice, followed by their second choice, and so on, so that when a respondent is making the k-th choice they are choosing among the $K - k + 1$ alternatives in $\mathcal{S}_k$. For a full ranking of K alternatives this requires $K - 1$ sequential choices, while for a partial ranking the number of choices required is K', provided that $K' < K - 1$.

Rankings of Ordered Categories

Typically, logit models for (partial) rankings do not assume that the alternatives are ordered. The indices of the alternatives are arbitrary. But here we consider the case where the alternatives are a set of K ordered categories that make up a discrete rating scale, and the indices of these alternatives/categories represent their ordinal position on that scale. Typically when using such scales respondents are asked to select only one category, which is effectively a partial ranking where only the first/top choice is elicited. But suppose the respondent is asked to *rank* the K response categories from the category that best represents their response to that which least represents their response. Given the ordering of the categories it seems reasonable that not all rankings should be admissible. For example, if there are $K = 5$ response categories and fourth category is ranked first so that $y_1 = 4$. Then category ranked second should be either the third or fifth category so that $y_2 = 3$ or $y_2 = 5$. It would not be reasonable for any other category to be ranked second because these are the two categories that are "closest" to the category ranked first. Similarly, the categories that might be selected third should be restricted based in which categories were selected first and second, and which categories might be ranked fourth should depend on the categories that are ranked first, second, and third. For a more concrete example, suppose that the categories correspond to ordered levels of frequency of a behavior: *very rarely*, *rarely*, *sometimes*, *often*, and *very often*. If the respondent selected *often* as the category that best represents their response, and was then asked to select which of the remaining categories they would select if they had to select a different category from the first, then their response would have to be either *sometimes* or *very often*. If they then selected *sometimes* the category ranked third should be either *rarely* or *very often*. If instead they ranked the *very often* second, then they would necessarily have to rank the *sometimes*, *rarely*, and *very rarely* categories third, fourth, and fifth, respectively.

When a respondent is restricted to give a ranking of a set of rating scale categories that is consistent with the order of the categories as described above, this restriction significantly limits the size of the set of possible rankings that could be observed. If the categories were not ordered then there are $K!$ possible rank orders. But if the rankings are restricted to only those that are admissible based on the order of the categories, then the number of possible rankings is significantly less. Consider that given that category y_1 is ranked first that there are $y_1 - 1$ categories "lower" and $K - y_1$ categories "higher" than category y_1. The number of orderings of the remaining $K-1$ categories is then

$$\frac{(K-1)!}{(y_1-1)!(K-y_1)!} = \binom{K-1}{y_1-1},$$

so the total number of admissible rank orders of K ordered categories is

$$\sum_{y_1=1}^{K} \binom{K-1}{y_1-1} = 2^{K-1}.$$

This is considerably less than the number of rank orders of unordered categories if $K > 2$. For $K = 3$, 5, or 7 categories, for example, there are 6, 120, and 5040 possible orderings, respectively, whereas for ordered categories there are only 4, 16, and 64 possible orderings, respectively. Furthermore, if the respondent is providing their ranking by ordering the categories through a series of sequential choices, starting with the category ranked first, then the number of required choices is less than for unordered categories. For both ordered and unordered alternatives the number of sequential choices necessary to provide a complete ranking is no larger than $K-1$, but for ordered alternatives it may be smaller depending on the ordering. This is because for some orderings the full admissible ranking is determined after one or more but fewer than $K-1$ choices. For example, if $K = 5$ and the first two elements of the ordering vector $\boldsymbol{y}$ are $y_1 = 4$ and $y_2 = 5$, then the ordering must be $\boldsymbol{y} = (4, 5, 3, 2, 1)'$ as there are no other admissible rankings with the same first two elements of the ordering vector. In general, given that the category ranked first is y_1, the minimum number of additional choices required to produce a complete rank ordering is $\min(y_1 - 1, K - y_1)$, and the maximum number of additional choices required is $K - 2$ if $1 < y_1 < K$, and zero otherwise. For a partial ranking of the top K' categories where $1 < K' < K - 1$, the minimum number of additional choices is $\min(y_1 - 1, K - y_1, K' - 1)$, and the maximum number of additional choices is $K' - 2$ if $1 < y_1 < K$, and zero otherwise. So in comparison to rankings of unordered alternatives, rankings of ordered categories will tend to require fewer individual choices by the respondent.

The rank ordered logit model described earlier can be modified for the (partial) ranking of a set of ordered categories where the set of admissible rankings is restricted as described above. To do this $\mathcal{S}_k$ is replaced with $\mathcal{S}_k^*$ in Equation 2 where $\mathcal{S}_1^* = \{1, 2, \ldots, K\}$ and

$$\mathcal{S}_k^* = \left\{ l \in \mathcal{S}_k : l = \min \mathcal{S}_k' - 1 \text{ or } l = \max \mathcal{S}_k' + 1 \right\} \tag{3}$$

for $k > 1$, where $\mathcal{S}_k'$ denotes the complement of $\mathcal{S}_k$ so that $\mathcal{S}_k' = \{y_1, y_2, \ldots, y_{k-1}\}$. This creates a kind of "truncated" version of the rank ordered logit model. The set of categories $\mathcal{S}_k^*$ is then the one or two categories that are "adjacent" to $\mathcal{S}_k'$. Consider, for example, the ordering vector $\boldsymbol{y} = (4, 3, 5, 2, 1)'$, so that the fourth category is ranked first and the first category is ranked last. Then

$$\begin{aligned}
\mathcal{S}_1 &= \{4, 3, 5, 2, 1\}, \\
\mathcal{S}_2 &= \{3, 5, 2, 1\}, \\
\mathcal{S}_3 &= \{5, 2, 1\}, \\
\mathcal{S}_4 &= \{2, 1\},
\end{aligned}$$

so that $\mathcal{S}'_2 = \{4\}$, $\mathcal{S}'_3 = \{4,3\}$, and $\mathcal{S}'_4 = \{4,3,5\}$, and thus

$$\begin{aligned}
\mathcal{S}^*_1 &= \{4,3,5,2,1\},\\
\mathcal{S}^*_2 &= \{3,5\},\\
\mathcal{S}^*_3 &= \{5,2\},\\
\mathcal{S}^*_4 &= \{2\}.
\end{aligned}$$

Note that $\mathcal{S}_k$ denotes the categories that could be ranked k-th for a model for unordered alternatives, while $\mathcal{S}^*_k$ denotes the set of categories that could be ranked k-th to produce an admissible ranking for ordered alternatives. For both ordered or unordered alternatives $\mathcal{S}'_k$ is the set of alternatives that have not yet been ranked. The sets $\mathcal{S}_k$, $\mathcal{S}^*_k$, and $\mathcal{S}'_k$ and how they evolve as k increases are depicted in Figure 1 for the example show above. For another example, consider an ordering vector $\boldsymbol{y} = (4,5,3,2,1)'$. Here $\mathcal{S}_1$, $\mathcal{S}^*_1$, $\mathcal{S}_2$, and $\mathcal{S}^*_2$ are the same as before, but now $\mathcal{S}_3 = \{5,2,1\}$, $\mathcal{S}_4 = \{2,1\}$, $\mathcal{S}'_3 = \{4,3\}$, $\mathcal{S}'_4 = \{4,5,3\}$, $\mathcal{S}^*_3 = \{3\}$, and $\mathcal{S}^*_4 = \{1\}$. These sets are depicted in Figure 2. Another example is given in Figure 3 where $\boldsymbol{y} = (5,4,3,2,1)'$ so that now $\mathcal{S}_2 = \{4,3,2,1\}$, $\mathcal{S}_3 = \{3,2,1\}$, $\mathcal{S}_4 = \{2,1\}$, $\mathcal{S}'_2 = \{4\}$, $\mathcal{S}'_3 = \{4,5\}$, $\mathcal{S}'_4 = \{4,5,3\}$, $\mathcal{S}^*_2 = \{4\}$, $\mathcal{S}^*_3 = \{3\}$, and $\mathcal{S}^*_4 = \{2\}$. Note also how the number of sequential choices necessary for each ordering vector differs between these examples. In general, once the set of ranked categories (i.e., $\mathcal{S}'_k$) includes either the first or last category, the number of categories in the choice set (i.e., $\mathcal{S}^*_k$) will then always be one, and so all remaining ranks are determined. In the first example the ranking can stop after the third choice, in the second example ranking can stop after the second choice, and in the third example the full ranking is determined after the first choice. Eliciting a rank order through sequential choices can be viewed as a kind of "branching question" design, but the branching is fundamentally different for ordered versus unordered alternatives. For unordered alternatives the k-th question is to choose from among the $K-k+1$ alternatives that have not yet been selected. But for ordered alternatives/categories, the k-th question is to either select from two categories that are adjacent to the set of categories already selected, or the questioning is terminated because the rank order of the remaining categories is determined. If $K' > 1$ another useful way to write the rank ordered logit model for ordered categories is

$$P(Y_1 = y_1, Y_2 = y_2, \dots, Y_{K'} = y_{K'}) = \frac{e^{\eta_{y_1}}}{\sum_{l=1}^{K'} e^{\eta_l}} \prod_{\substack{k=2\\ 1\le a_k\\ b_k \le K}}^{K} \frac{e^{\eta_{y_k}}}{e^{\eta_{a_k}} + e^{\eta_{b_k}}},$$

where $a_2 = y_1 - 1$, $b_2 = y_2 + 1$, and for $k > 2$

$$a_k = \begin{cases} a_{k-1} - 1, & y_k < y_{k-1}, \\ a_{k-1}, & \text{otherwise}, \end{cases} \quad \text{and} \quad b_k = \begin{cases} b_{k-1} + 1, & y_k > y_{k-1}, \\ b_{k-1}, & \text{otherwise}. \end{cases}$$

Every category choice after the first is either a paired comparison between the categories a_k and b_k, which are the two categories that are adjacent to the set of categories already ranked, or the remaining rankings are determined.

The case of a partial ranking of $K' < K-1$ ordered categories has already been discussed. But another kind of partial ranking that will also be considered here is based on the "best and worst" choices — i.e., the categories ranked one and K, respectively. Here the respondent provides a partial ranking by only indicating the categories that most and least represent their response. This approach has the advantage of being relatively simple and quick for the respondent, while

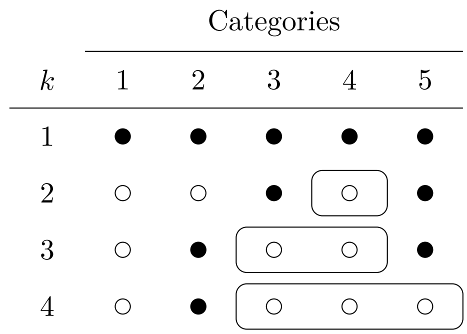

Figure 1: Category sets for the ordering vector $\boldsymbol{y} = (4, 3, 5, 2, 1)'$. The filled dots are in $\mathcal{S}_k^*$, the categories that are enclosed are in $\mathcal{S}_k'$, and the categories that are not enclosed are in $\mathcal{S}_k$.

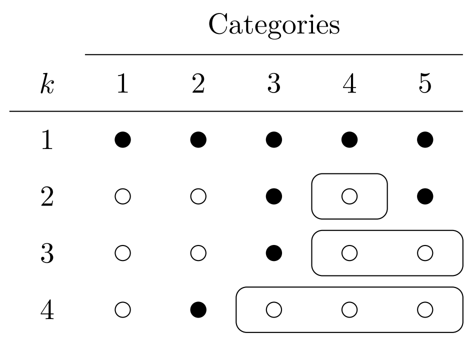

Figure 2: Category sets for the ordering vector $\boldsymbol{y} = (4, 5, 3, 2, 1)'$. The filled dots are in $\mathcal{S}_k^*$, the categories that are enclosed are in $\mathcal{S}_k'$, and the categories that are not enclosed are in $\mathcal{S}_k$.

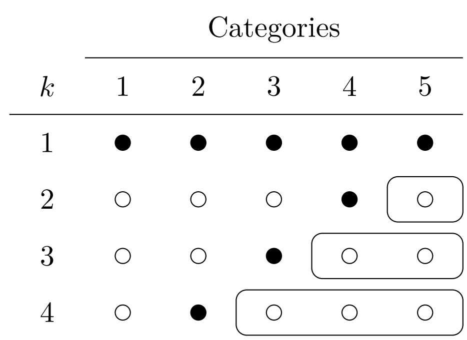

Figure 3: Category sets for the ordering vector $\boldsymbol{y} = (5, 4, 3, 2, 1)'$. The filled dots are in $\mathcal{S}_k^*$, the categories that are enclosed are in $\mathcal{S}_k'$, and the categories that are not enclosed are in $\mathcal{S}_k$.

also providing more information than a single choice. The collection and modeling of best and worst choices has been discussed extensively for unordered alternatives (see Flynn & Marley, 2014, and Marley & Flynn, 2015, for reviews). Here we can adapt the model described above to data where the respondent selects the categories that most and least represent their response when the categories are ordered. Assuming the same underlying model for the full ordering, the probability of a partial ranking based on the best and worst categories is

$$P(Y_1 = y_1, Y_K = y_K) = \sum_{\boldsymbol{y} \in \mathcal{U}(y_1, y_K)} P(Y_1 = y_1, Y_2 = y_2, \ldots, Y_K = y_K),$$

where $\mathcal{U}(y_1, y_K)$ is the set of all admissible ordering vectors where the first and last elements are y_1 and y_K, respectively. The value of y_K will always be either the first or last category so that $y_K = 1$ or $y_K = K$. The size of $\mathcal{U}(y_1, y_K)$ is

$$\binom{K-2}{|y_1 - y_K| - 1}.$$

For example, if $K = 7$ and the last category is ranked last so that $y_7 = K$, then there is one admissible ordering if $y_1 = 1$ or $y_1 = 6$, five admissible orderings if $y_1 = 2$ or $y_1 = 5$, and ten admissible orderings if $y_1 = 3$ or $y_1 = 4$.

In specifying the model described above for ordered categories, it may be desirable that the dependence on any covariates also reflect the ordering of the categories. This is not the case for the typical discrete choice model without choice-specific covariates where $\eta_k = \alpha_k + \boldsymbol{\beta}_k' \boldsymbol{x}$. In discrete choice models for unordered alternatives the indexing of the categories is arbitrary. Changes in the indices only results in a reparameterization of the model. But a special case of this model known sometimes as the "stereotype model" due to Anderson (1984) does reflect a specified ordering of the categories. The one-dimensional stereotype model restricts $\boldsymbol{\beta}_k$ to be proportional across categories so that $\boldsymbol{\beta}_k = \phi_k \boldsymbol{\beta}$. The categories are ordered with respect to their indices provided that $\phi_k \leq \phi_{k+1}$ or $\phi_k \geq \phi_{k+1}$ for all $k < K - 1$. Then the logarithm of an odds ratio for categories k and k' is proportional to $\phi_k - \phi_{k'}$. Other adjacent-category logit models can be viewed as special cases where $\phi_k = k$ so that the odds for adjacent categories are proportional (e.g., Fullerton & Xu, 2018). See Anderson (1984) for a discussion of other ordering properties of the stereotype model. Other models might be considered for how the η_k depend on the covariates, but here the focus will be on the stereotype regression model and a related rating scale item response model. These models are featured in the next two sections.

Stereotype Model for Rankings of Ordered Categories

Here we consider a small simulation study to compare the performance of estimators for the parameters of a stereotype model for rankings of ordered response categories. The comparisons are made between full rankings and partial rankings, including the case when only the top category is observed. Based on studies of models for unordered alternatives, the expectation here is that inferences will be more precise when based on (partial) rankings rather than on only top choices. Consider a rating scale with K ordered categories, and let $\boldsymbol{y}_i = (y_{i1}, y_{i2}, \ldots, y_K)'$ denote the i-th respondent's ordering of the categories. Similarly, for a partial rank ordering of the top K' categories then $\boldsymbol{y}_i = (y_{i1}, y_{i2}, \ldots, y_{K'})'$, which includes when only the top category is given so that $\boldsymbol{y}_i = y_{i1}$. We also consider partial orderings based on best-worst choices so that $\boldsymbol{y}_i = (y_{i1}, y_{iK})'$. The estimators used here are maximum likelihood estimators. Assuming the mutual independence of the rankings across n respondents, the likelihood function for full rank orderings as well as partial rank orderings based on the top K' categories can be written as

$$L(\boldsymbol{\theta}) = \prod_{i=1}^{n} \prod_{k=1}^{K'} \frac{e^{\eta_{iy_{ik}}}}{\sum_{l \in \mathcal{S}_{ik}^*} e^{\eta_{il}}},$$

and the log-likelihood function can then be written as

$$\log L(\boldsymbol{\theta}) = \sum_{i=1}^{n} \sum_{k=1}^{K'} \eta_{iy_{ik}} - \sum_{i=1}^{n} \sum_{k=1}^{K'} \log \left(\sum_{l \in \mathcal{S}_{ik}^*} e^{\eta_{il}} \right).$$

For partial rank orderings based on the best and worst categories, the likelihood function can be written as

$$L(\boldsymbol{\theta}) = \prod_{i=1}^{n} P(Y_{i1} = y_{i1}, Y_{iK} = y_{iK}),$$

where

$$P(Y_{i1} = y_{i1}, Y_{iK} = y_{iK}) = \sum_{\boldsymbol{y} \in \mathcal{U}(y_{i1}, y_{iK})} P(Y_{i1} = y_1, Y_{i2} = y_2, \ldots, Y_{iK} = y_K)$$

and

$$P(Y_{i1} = y_1, Y_{i2} = y_2, \ldots, Y_{iK} = y_K) = \prod_{k=1}^{K} \frac{e^{\eta_{iy_{ik}}}}{\sum_{l \in \mathcal{S}_k^*} e^{\eta_{il}}}.$$

Note that $\eta_{iy_{ik}}$ and $\mathcal{S}_{ik}^*$ can vary over respondents since $\eta_{iy_{ik}} = \alpha_k + \phi_k \boldsymbol{\beta}' \boldsymbol{x}_i$, where $\boldsymbol{x}_i$ is a vector of covariates, and $\mathcal{S}_{ik}^*$ is determined by $\boldsymbol{y}_i$ as given in Equation 3. Here $\boldsymbol{\theta}$ includes $\alpha_2, \alpha_3, \ldots, \alpha_K$, $\phi_2, \phi_3, \ldots, \phi_{K-1}$, and $\boldsymbol{\beta}$, where for identification $\alpha_1 = 0$, $\phi_1 = 0$, and $\phi_K = 1$ are fixed.

A total of 10,000 samples of $n = 200$ rankings each were generated for each of four data-generating models with $K = 2$, 3, 5, and 7 categories. For each sample the covariate vectors were defined as $\boldsymbol{x}_i = x_i$, where x_i is an equally-spaced sequence of values from $x_1 = -3$ to $x_{200} = 3$. The parameter values used in the data-generating model for each value of K are shown in Table 1. The parameters were selected so that the points where the probabilities for adjacent categories are equal are equidistant, the distance between adjacent points equals $6/(K-1)$, and these points are centered at $x = 0$. Figure 4 shows the probability of each category being ranked first as a function of x for each data-generating model. For each sample estimates were obtained by maximizing the likelihood function based on partial rankings of the first K' choices where $K' < K - 1$, partial rankings based on best and worst categories, and complete rankings (i.e., $K' = K - 1$). Three measures estimated from the samples of parameter estimates were used to assess the accuracy of the estimators: the root mean squared error (RMSE), bias, and standard error (SE).

Table 2 shows the results for the model with $K = 3$ response categories. Note that when $K = 3$ the partial rankings based on first choices (i.e., $K' = 1$) and best/worst choices are equivalent and so are not shown separately as they are for larger values of K'. Table 3, Table 4, and Table 5 show the results for models with $K = 4$, 5, and 7 categories, respectively. Overall, the root mean squared error, bias, and standard error all tend to improve (i.e., get smaller) as K' increases, with the results for the best/worst choices falling usually somewhere between those for $K' = 2$ and $K' = 3$. As expected, ranking ordered categories results in more accurate estimators than just eliciting the top category (i.e., $K' = 1$). The bias in the estimators is typically small

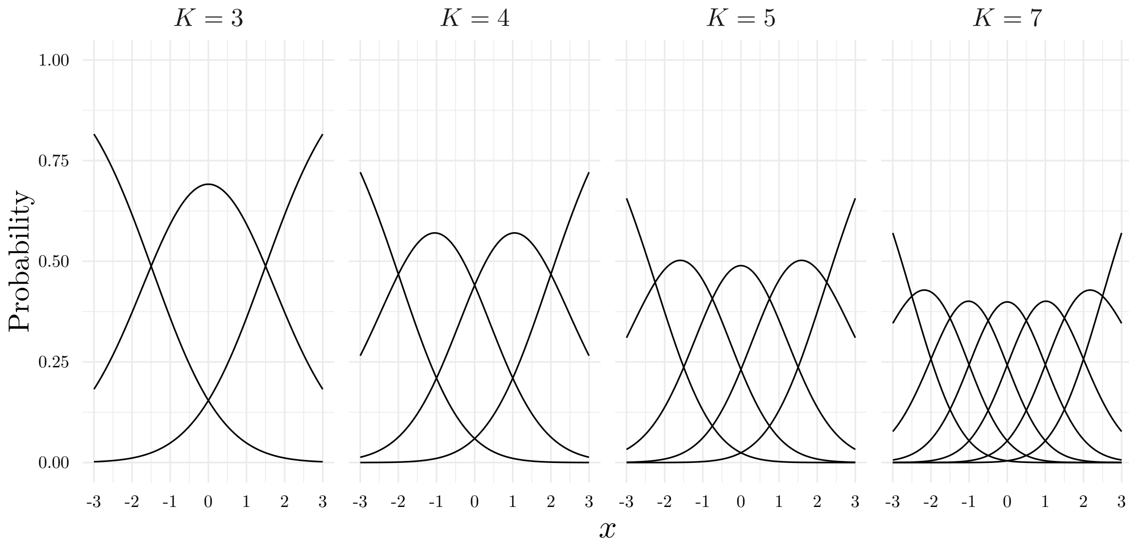

Figure 4: Probabilities for the category ranked first as a function of x for $K = 3$, 4, 5, and 7 categories based on the stereotype regression model.

and often negligible relative to the standard error, so most of the estimation error reflected in the mean squared error is being driven by the variability in the sampling distribution.[1] When comparing the models for a different number of categories, the accuracy of the full ranking relative to using only the top category tends to increase with the number of categories. For the estimator of the β parameter, for example, the standard error when $K' = 1$ is about 20%, 35%, 54%, and 85% larger than when using the full ranking for the models with $K = 3$, 4, 5, and 7 categories, respectively. Assuming the variances of the maximum likelihood estimators are approximately inversely proportional to the sample size, and recalling that the sample size used in the simulation was $n = 200$, the effective sample size of the full rankings is about 289, 366, 473, and 686 observations for $K = 3$, 4, 5, and 7 categories, respectively.

Rating Scale Model for Rankings of Ordered Categories

Here we consider a simulation study of a rating scale model for rankings. The original rating scale model (Andrich, 1978) is an item response model for ordered response categories, but where the response is just a single category. Let Y_{ij} denote the category selected by the i-th respondent when responding to the j-th item with K categories such that $Y_{ij} = 1, 2, \dots, K$. The rating scale model can be written as

$$P(Y_{ij} = y) \propto \exp\left[(y-1)(\zeta_i - \delta_j) - \textstyle\sum_{k=1}^{y} \tau_k\right]$$

where ζ_i is a respondent-specific parameter, while δ_j and τ_k are relative "location" parameters for the j-th item and the k-th category, respectively. The rating scale model can be written as

[1]Recall that the mean squared error can be decomposed into the sum of the squared bias and the squared standard error (i.e., the variance).

a stereotype model with a random effect such that

$$P(Y_{ij} = y) \propto \exp\left[\alpha_y + \gamma_y(\zeta_i + \boldsymbol{\beta}'\boldsymbol{x}_j)\right],$$

where $\alpha_y = -\sum_{k=1}^{y} \tau_k$, $\gamma_y = y - 1$, $\boldsymbol{\beta} = (\delta_1, \delta_2, \ldots, \delta_m)'$, and $\boldsymbol{x}_j$ is a m-dimensional vector where the j-th element is -1 and all other elements are zero so that $\boldsymbol{\beta}'\boldsymbol{x}_j = \delta_j$. To identify the model, let $\tau_1 = \tau_2 = 0$ which implies that $\alpha_1 = \alpha_2 = 0$. This model could be viewed as a kind of hierarchical mixed effects generalization of the stereotype regression model featured in the previous section when applied to situations where the respondent *ranks* the response categories for each item so that Y_{ijk} is the category that the i-th respondent ranks k-th for the j-th item. The random effects are due to the ζ_i parameters. Treating $\zeta_1, \zeta_2, \ldots, \zeta_n$ as missing data, the complete-data log-likelihood function for a full ranking or partial ranking of the top K' categories of each item could be written as $L(\boldsymbol{\theta}, \zeta_1, \zeta_2, \ldots, \zeta_n) = \sum_{i=1}^{n} \log L_i(\boldsymbol{\theta}, \zeta_i)$ where

$$\log L_i(\boldsymbol{\theta}, \zeta_i) = \sum_{j=1}^{m}\sum_{k=1}^{K'} \eta_{ijy_{ijk}} - \sum_{j=1}^{m}\sum_{k=1}^{K'} \log\left(\sum_{l \in \mathcal{S}^*_{ijk}} e^{\eta_{ijl}}\right),$$

and where $\eta_{ijk} = \alpha_y + \gamma_y(\zeta_i + \boldsymbol{\beta}'\boldsymbol{x}_j)$ and the choice set $\mathcal{S}^*_{ijk}$ is based on the ordering vector of the i-th respondent for the j-th item, $\boldsymbol{y}_{ij} = (y_{ij1}, y_{ij2}, \ldots, y_{ijK'})'$. Assuming that all ζ_i are independently and identically distributed as $N(0, \sigma)$, the marginal likelihood function is

$$L(\boldsymbol{\theta}, \sigma) = \prod_{i=1}^{n} \int L_i(\boldsymbol{\theta}, \zeta) f(\zeta, \sigma) d\zeta,$$

where $f(\zeta, \sigma)$ is the probability density function of ζ_i. Similarly, a marginal likelihood function can also be derived for a partial ranking based on the best/worst categories of each item.

The design of the simulation study used here is like that used in the previous section, but only a $K = 5$ category model was considered. A total of 10,000 samples of $n = 200$ observations each were generated for a $m = 3$ item model with category parameters $\tau_1 = \tau_2 = 0$, $\tau_3 = 1$, $\tau_4 = 2$, $\tau_5 = 3$, item parameters $\delta_1 = -2.5$, $\delta_2 = -1.5$, and $\delta_3 = -0.5$, and $\sigma = 1$. Figure 5 shows the probabilities for the category ranked first for each item. Parameter estimates were obtained by maximizing the logarithm of the marginal likelihood function with respect to $\boldsymbol{\theta}$ and σ, where the integral was approximated using Gauss-Hermite quadrature. Estimates were obtained based on the complete ranking as well as partial rankings of the top $K' = 1$, 2, and 3 categories, as well as the best/worst categories. Table 6 shows the root mean squared error, bias, and standard error of each parameter based on the complete and partial rankings. The trends in the accuracy measures are similar to those seen in the regression model in the previous section with $K = 5$ categories. The estimators had negligible bias but the root mean squared errors and the standard errors decreased with K', with the best/worst partial ranking performing between $K' = 1$ and $K' = 2$. However, the relative effect of the depth of the partial ranking was less than that observed for the stereotype model with the same number of categories. Here the standard errors for the item and category parameters ranged from about 20% to 38% higher when $K' = 1$ in comparison to a complete ranking.

For an item response model it is also useful to consider the compare the information functions for $\zeta_.$. The item information function for the j-th item can be written as

$$I_j(\zeta) = \sum_{r \in \mathcal{R}} I_{jr}(\zeta) P_{jr}(\zeta),$$

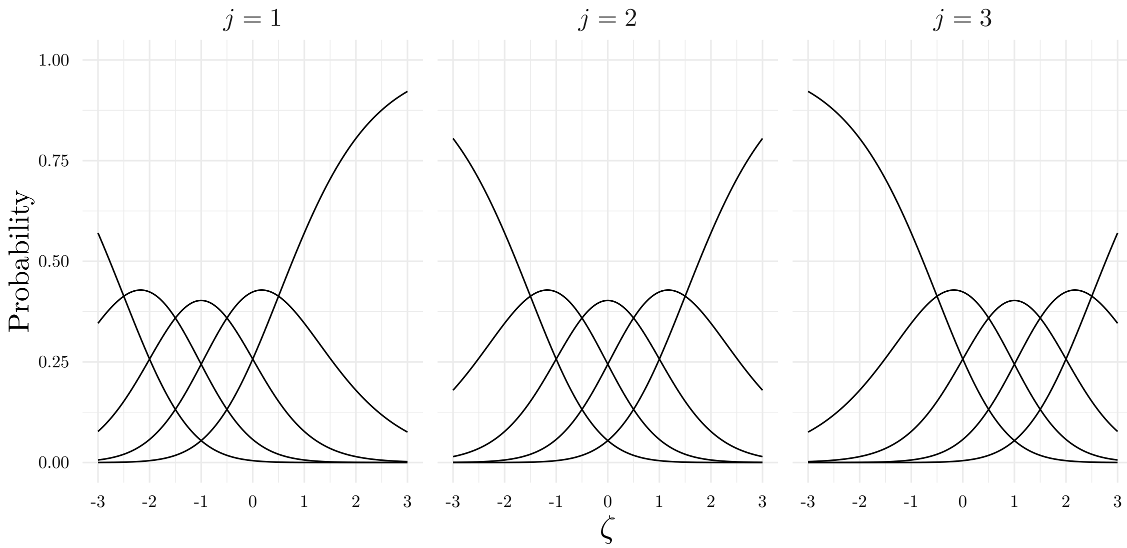

Figure 5: Probabilities for the category ranked first for each item as a function of ζ based on the rating scale model.

where $\mathcal{R}$ is the set of all admissible (partial) rankings of the categories of the j-th item,

$$I_{jr}(\zeta) = -\frac{\partial^2 \log P_{jr}(\zeta)}{\partial \zeta^2}$$

is the item information function of the r-th ranking of the j-th item, and

$$P_{jr}(\zeta) = \prod_{k=1}^{K'} \frac{e^{\eta_{jy_{jk}}}}{\sum_{l \in \mathcal{S}^*_{rk}} e^{\eta_{jl}}}$$

is the probability of the r-th ranking of the categories of the j-th item. In contrast to item information functions for polytomous item response models, where the summation would be over all response categories (e.g., Baker & Kim, 2004), here the summation is over all (partial) rankings of the categories. Figure 6 shows the information functions of the second item for each of the five models considered based on the item parameters for the data-generating model used in the simulation study. The information functions for first and third items are similar in shape, but shifted so that the functions are symmetric around $\zeta = -1$ and $\zeta = 1$, respectively. For the complete rankings and the partial rankings of the first K' categories, item information tends to increase with K', although the information functions tend to converge as $|\zeta|$ gets larger. This seems consistent with the relationship between ζ and the rankings. For extremely low or high values of ζ, the probability becomes concentrated on an ordering vector of $(1, 2, 3, 4, 5)'$ or $(5, 4, 3, 2, 1)'$, respectively. In this case greater values of K' provide less information as the category ranked first is increasingly likely to be the first or last, and the ranking of the remaining categories is determined. But for moderate values of ζ a "deeper" ranking with larger K' is more informative. The relationship between the information function for the best/worst partial ranking and the other partial rankings is a bit more complex. It is higher than that of the

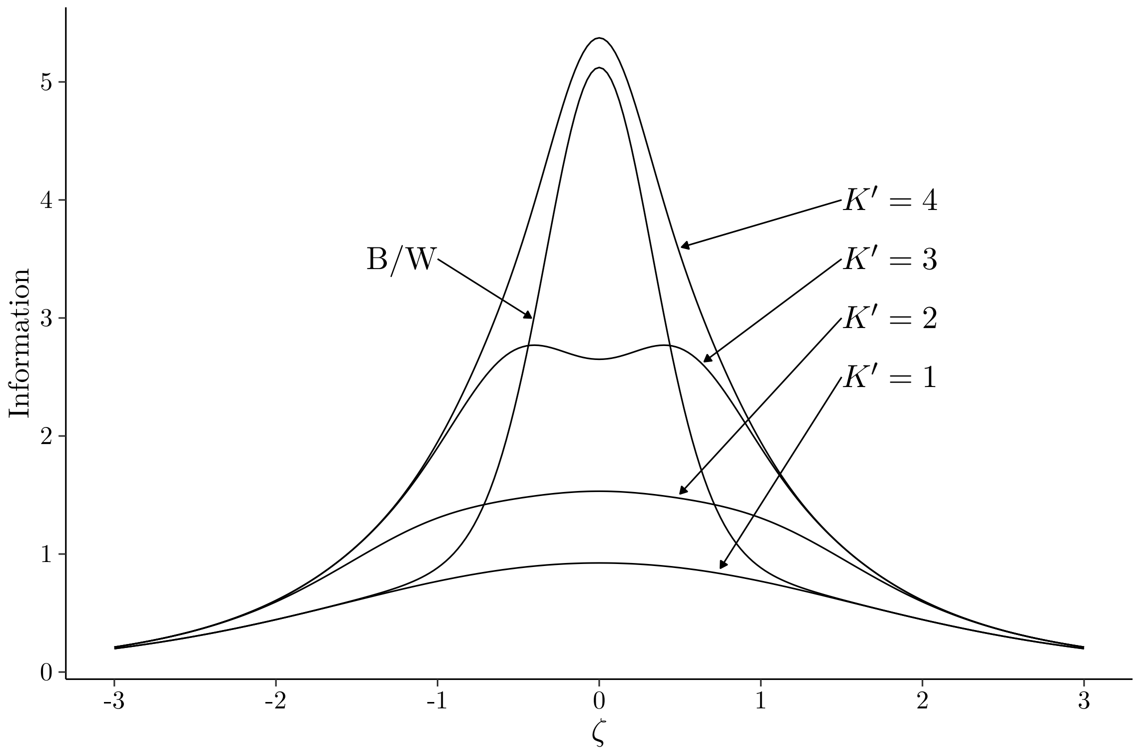

Figure 6: Item information functions for rating scale models for full and partial rankings.

partial rankings for $K' > 1$ for moderate values of ζ, but lower for extreme values of ζ. But as would be expected the information function of the best/worst partial ranking is always less than that of the complete ranking, but greater than the partial ranking based that gives only the top category (i.e., $K' = 1$). Taken together these results show that (partial) rankings can provide more efficient inferences for ζ over having respondents give only their top response category.

Discussion

As has been shown for models of unordered alternatives, models for rankings of ordered alternatives, such as the ordered response categories of a scale, can result in more efficient estimators than comparable models for only the single top category. And while more effort is required on the part of a respondent to produce a ranking rather than a single choice, this is less so in the case of rankings of ordered response categories. There are far fewer admissible rankings for ordered alternatives, and when the ranking is done sequentially the first few initial choice(s) will often determine the complete ranking. These benefits are also true for models for partial rankings of ordered response categories such as the top two or more categories, or the best and worst categories.

This paper featured models based on a kind of truncated rank ordered logit model which could be extended in several ways. The stereotype regression model and rating scale item response model were the basis of the models featured in this paper, but other generalizations could be considered by using other models for first choices or rankings. Fullerton and Xu (2018) discuss a broad family of adjacent-category models for first choices that could be extended to models for rankings. Another useful generalization is through the specification of random effects based on related models with random effects for first choices or unordered alternatives. The rating scale model can be viewed as a kind of random effects model for multilevel data. But there are other possibilities. Skrondal and Rabe-Hesketh (2003) discuss a general framework for how random effects can be specified for the rank ordered logit model. They show how to specify respondent-specific random effects, as was done for the rating scale model. This can be useful for multilevel data when multiple rankings are observed for each respondent. But they also discuss specifying alternative-specific random effects. Böckenholt (2001) also discussed this kind of random effects structures and also latent classes for ranked ordered logit models. These approaches could be useful adapted for models for rankings of ordered response categories. The use of category-specific random effects can be useful for accounting for some violations of the independence of irrelevant alternatives assumption that is implicit in the model. Johnson (2007) proposed a generalized stereotype model for ordered response categories that specifies category-specific random effects for discrete choice models to account for individual differences in response scale use. A similar approach could be used to extend the stereotype regression model for rankings of ordered categories.

Another generalization would be to consider modifications of the random utility formulation of the rank ordered logit model. Ranking models can be motivated by assuming unobserved random utilities for the alternatives, $U_{i1}, U_{i2}, \dots, U_{iK}$, such that the ordering of these latent variables implies the observed ordering of the alternatives. The rank ordered logit model for unordered alternatives can be derived by assuming that $U_{ik} = \eta_{ik} + \epsilon_{ik}$, where the errors $\epsilon_{i1}, \epsilon_{i2}, \dots, \epsilon_{iK}$ are independently and identically distributed standard Gumbel random variables (Luce & Suppes, 1965). An alternative approach that relaxes the assumption of independently and identically distributed errors is to assume that they have a multivariate normal distribution so that the utilities may have different variances and non-zero covariances. For unordered alternatives this implies the Thurstonian model for rankings (Böckenholt, 1992, 1993). The Thurstonian model is not consistent with a model for ordered response categories since it can give rise to non-admissible rankings. But this approach could maybe be modified to limit the possible rankings to the admissible rankings through a truncated joint distribution of the utilities, or through some other joint distribution of the utilities where non-admissible rankings have zero probability.

The generalizations described above could also be applied to partial rankings of the first K' categories, but other models might be considered for best and worst category partial rankings. Marley and Louviere (2005) discuss broad classes of models for best-worst choices based on separate but interrelated models for the best and worst choices. Some of these models might be adapted to situations where the partial rankings of the model are restricted to those that are admissible given the ordering of the response categories.

Software

The C++ and R (R Core Team, 2021) source code to estimate the models discussed in this paper is available from the author.[2] The C++ source code uses the Armadillo C++ library (Sanderson & Curtin, 2016) distributed with the **RcppArmadillo** R package (Eddelbuettel & Sanderson, 2014), and C++ classes and functions from the **roptim** R package (Pan, 2020) for using C source code underlying the `optim` function in R. The **Rcpp** R package (Eddelbuttel,

[2]The code will be made available in a R package hosted on GitHub.

2013; Eddelbuettel & Balamuta, 2018; Eddelbuettel & Francois, 2011) was used to construct R functions to interface with the C++ functions.

References

Allison, P. D. & Christakis, N. A. (1994). Logit models for sets of ranked items. *Sociological Methodology*, *24*, 199–228.

Anderson, J. A. (1984). Regression and ordered categorical variables (with discussion). *Journal of the Royal Statistical Society. Series B*, *46*, 1–30.

Andrich, D. (1978). A rating formulation for ordered response categories. *Psychometrika*, *43*, 561–573.

Baker, F. B., & Kim, S. H. (2004). *Item response theory: Parameter estimation techniques*. CRC Press.

Beggs, S., Cardell, S., & Hausman, J. (1981). Assessing the potential demand for electric cars. *Journal of Econometric*, *16*, 1–19.

Beresteanu, A. & Zincenko, F. (2018). Efficiency gains in rank-ordered multinomial logit models. *Oxford Bulletin of Economics and Statistics*, *80*, 122–134.

Böckenholt, U. (1992). Thurstonian representation for partial ranking data. *British Journal of Mathematical and Statistical Psychology*, *45*, 31–49.

Böckenholt, U. (1993). Applications of Thurstonian models to ranking data. In M. A. Flinger & J. S. Verducci (Eds.), *Probability models and statistical analyses for ranking data* (pp. 157–172). Springer-Verlag.

Böckenholt, U. (2001). Mixed-effects analyses of rank-ordered data. *Psychometrika*, *66*, 45–62.

Chapman, R. G. & Staelin, R. (1982). Exploiting rank ordered choice set data with the stochastic utility model. *Journal of Marketing Research*, *19*, 288–301.

Eddelbuettel, D. (2013). *Seamless R and C++ integration with Rcpp*. Springer.

Eddelbuettel, D., & Balamuta, J. J. (2018). Extending R with C++: A brief introduction to Rcpp. *The American Statistician*, *72*, 28–36.

Eddelbuettel, D., & Francois, R. (2011). Rcpp: Seamless R and C++ integration. *Journal of Statistical Software*, *40(8)*, 1–18.

Eddelbuettel, D., & Sanderson, C. (2014). RcppArmadillo: Accelerating R with high-performance C++ linear algebra. *Computational Statistics and Data Analysis*, *71*, 1054–1063.

Flynn, T. N., & Marley, A. A. (2014). Best-worst scaling: theory and methods. In S. Hess & A. Daly (Eds.), *Handbook of choice modelling*. Edward Elgar Publishing.

Fullerton, A. S. (2009). A conceptual framework for ordered logistic regression models. *Sociological Methods & Research*, *38*, 306–347.

Fullerton, A. S., & Xu, J. (2018). Constrained and unconstrained partial adjacent category logit models for ordinal response variables. *Sociological Methods & Research*, *47*, 169–206.

Hsieh, C. (2005). Age and relative importance of major life domains. *Journal of Aging Studies*, *19*, 503–512.

Johnson, T. R. (2007). Discrete choice models for ordinal response variables: A generalization of the stereotype model. *Psychometrika*, *72*, 489–504.

Kuhn, K. M., Johnson, T. R., & Miller, D. (2013). Applicant desirability influences reactions to discovered résumé embellishments. *International Journal of Selection and Assessment*, *21*, 111–120.

Luce, R. (1959). *Individual choice behavior*. Wiley.

Luce, R. D., & Suppes, P. (1965). Preference, utility, and subjective probability. In R. D. Luce, R. R. Bush, & E. Galanter (Eds.), *Handbook of mathematical psychology* (Vol. III, pp. 235–406). Wiley.

Malhotra, N., & Kuo, A. G. (2008). Attributing blame: The public's response to Hurricane Katrina. *The Journal of Politics*, *70*, 120–135.

Marley, A. A. J., & Flynn, T. N. (2015). Best worst scaling: theory and practice. In J. D. Wright (Ed.), *International encyclopedia of the social & behavioral sciences* (2nd ed.), Elsevier.

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

McFadden, D. (1974). Conditional logit analysis of qualitative choice behaviour. In P. Zarembka (Ed.), *Frontiers in econometrics*. Academic Press.

Pan, Y. (2020). roptim: General purpose optimization in R using C++. R package version 0.1.5. https://CRAN.R-project.org/package=roptim

Plackett, R. L. (1975). The analysis of permutations. *Applied Statistics*, *24*, 193–201.

Punj, G. N., & Staelin, R. (1978). The choice process for graduate business schools. *Journal of Marketing Research*, *15*, 588–598.

R Core Team (2021). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/.

Sanderson, C., & Curtin, R. (2016). Armadillo: A template-based C++ library for linear algebra. *Journal of Open Source Software*, *1(2)*, 26.

Shostak, S., Freese, J., Link, B. G., & Phelan, J. C. (2009). The politics of the gene: Social status and beliefs about genetics for individual outcomes. *Social Psychology Quarterly*, *72*, 77–93.

Skrondal, A., & Rabe-Hesketh, S. (2003). Multilevel logistic regression for polytomous data and rankings. *Psychometrika*, *68*, 267–287.

Tafarodi, R. W., Shaughnessy, S. C., Lee, W. W. S., Leung, D. Y. P., Ozaki, Y., Morio, H., & Yamaguchi, S. (2009). Disregard for outsiders: A cultural comparison. *Journal of Cross-Cultural Psychology*, *40*, 567–583.

Tutz, G. (2021). Ordinal regression: A review and a taxonomy of models. *WIREs Computational Statistics*, e1545. https://doi.org/10.1002/wics.1545

Table 1: Parameters of the data-generating models for the simulation study of regression models for rankings of ordered categories.

	K			
Parameter	3	4	5	7
α_1	0	0	0	0
α_2	1.5	2	2.25	2.5
α_3	0	2	3	4
α_4		0	2.25	4.5
α_5			0	4
α_6				2.5
α_7				0
ϕ_1	0	0	0	0
ϕ_2	0.5	1/3	0.25	1/6
ϕ_3	1	2/3	0.5	1/3
ϕ_4		1	0.75	0.5
ϕ_5			1	2/3
ϕ_6				5/6
ϕ_7				1
β	2	3	4	6

Table 2: Root mean squared error (RMSE), bias, and standard error (SE) of the estimators of the parameters of the regression models for the rankings of $K = 3$ response categories including partial rankings of the first $K' = 1$ categories.

		K'	
Measure	Parameter	1	2
RMSE	α_2	0.319	0.248
	α_3	0.417	0.248
	γ_2	0.130	0.107
	β	0.141	0.116
Bias	α_2	0.046	0.032
	α_3	−0.008	0.001
	γ_2	0.000	0.002
	β	0.033	0.023
SE	α_2	0.315	0.245
	α_3	0.417	0.248
	γ_2	0.130	0.107
	β	0.137	0.114

Table 3: Root mean squared error (RMSE), bias, and standard error (SE) of the estimators of the parameters of the regression models for the rankings of $K = 3$ response categories including partial rankings of the first $K' = 1, 2$ categories and best and worst (B/W) choices.

		K'			
Measure	Parameter	1	2	3	B/W
RMSE	α_2	0.523	0.414	0.320	0.378
	α_3	0.532	0.392	0.310	0.373
	α_4	0.692	0.533	0.305	0.311
	γ_2	0.197	0.171	0.134	0.150
	γ_3	0.199	0.171	0.134	0.150
	β	0.145	0.115	0.105	0.121
Bias	α_2	0.095	0.065	0.045	0.054
	α_3	0.090	0.062	0.043	0.049
	α_4	0.001	0.009	0.001	0.000
	γ_2	0.000	0.001	−0.001	−0.003
	γ_3	0.002	0.003	0.002	0.006
	β	0.043	0.029	0.023	0.029
SE	α_2	0.515	0.409	0.317	0.374
	α_3	0.524	0.387	0.307	0.370
	α_4	0.692	0.533	0.305	0.311
	γ_2	0.197	0.171	0.134	0.150
	γ_3	0.199	0.171	0.134	0.150
	β	0.138	0.112	0.102	0.118

Table 4: Root mean squared error (RMSE), bias, and standard error (SE) of the estimators of the parameters of the regression models for the rankings of $K = 5$ response categories including partial rankings of the first $K' = 1, 2, 3$ categories and best and worst (B/W) choices.

		K'				
Measure	Parameter	1	2	3	4	B/W
RMSE	α_2	0.738	0.585	0.459	0.363	0.540
	α_3	0.749	0.573	0.455	0.370	0.511
	α_4	0.772	0.602	0.419	0.345	0.526
	α_5	1.025	0.796	0.595	0.344	0.369
	γ_2	0.276	0.238	0.206	0.162	0.208
	γ_3	0.265	0.221	0.185	0.140	0.168
	γ_4	0.276	0.238	0.203	0.161	0.207
	β	0.150	0.117	0.100	0.094	0.127
Bias	α_2	0.129	0.085	0.058	0.050	0.069
	α_3	0.159	0.106	0.073	0.062	0.094
	α_4	0.121	0.083	0.056	0.047	0.061
	α_5	−0.011	−0.001	−0.003	−0.001	−0.001
	γ_2	−0.002	−0.001	−0.001	−0.001	−0.009
	γ_3	−0.004	−0.001	−0.001	−0.001	−0.002
	γ_4	0.000	0.001	0.000	0.001	0.012
	β	0.052	0.035	0.027	0.023	0.035
SE	α_2	0.727	0.579	0.455	0.360	0.535
	α_3	0.732	0.563	0.449	0.364	0.502
	α_4	0.763	0.596	0.415	0.342	0.523
	α_5	1.025	0.796	0.595	0.344	0.369
	γ_2	0.276	0.238	0.206	0.162	0.207
	γ_3	0.265	0.221	0.185	0.140	0.168
	γ_4	0.276	0.238	0.203	0.161	0.207
	β	0.140	0.111	0.096	0.091	0.122

Table 5: Root mean squared error (RMSE), bias, and standard error (SE) of the estimators of the parameters of the regression models for the rankings of $K = 7$ response categories including partial rankings of the first $K' = 1, 2, 3, 4, 5$ categories and best and worst (B/W) choices.

		K'						
Measure	Parameter	1	2	3	4	5	6	B/W
RMSE	α_2	1.233	0.990	0.804	0.638	0.510	0.421	0.934
	α_3	1.256	0.971	0.794	0.641	0.521	0.448	0.868
	α_4	1.276	0.991	0.790	0.645	0.531	0.463	0.869
	α_5	1.278	0.993	0.803	0.608	0.506	0.441	0.865
	α_6	1.351	1.059	0.849	0.665	0.454	0.398	0.891
	α_7	1.740	1.348	1.099	0.844	0.628	0.387	0.475
	γ_2	0.428	0.372	0.325	0.273	0.233	0.194	0.339
	γ_3	0.405	0.333	0.290	0.250	0.206	0.165	0.255
	γ_4	0.410	0.336	0.297	0.247	0.202	0.162	0.231
	γ_5	0.411	0.337	0.295	0.252	0.207	0.167	0.260
	γ_6	0.437	0.378	0.332	0.278	0.235	0.195	0.344
	β	0.167	0.128	0.108	0.096	0.087	0.084	0.146
Bias	α_2	0.215	0.139	0.103	0.076	0.056	0.058	0.110
	α_3	0.315	0.205	0.149	0.111	0.091	0.090	0.196
	α_4	0.347	0.224	0.163	0.126	0.104	0.102	0.225
	α_5	0.314	0.204	0.149	0.113	0.093	0.092	0.195
	α_6	0.205	0.134	0.102	0.076	0.056	0.056	0.103
	α_7	−0.026	−0.022	−0.015	−0.005	−0.008	−0.001	0.001
	γ_2	−0.003	−0.003	−0.001	−0.002	−0.004	−0.002	−0.023
	γ_3	−0.005	−0.003	−0.002	−0.002	−0.003	−0.001	−0.014
	γ_4	−0.005	−0.004	−0.003	0.000	−0.001	0.000	0.000
	γ_5	−0.007	−0.006	−0.005	−0.002	−0.003	−0.001	0.011
	γ_6	−0.001	−0.002	−0.004	0.000	0.001	0.002	0.026
	β	0.073	0.049	0.036	0.029	0.025	0.024	0.053
SE	α_2	1.214	0.980	0.797	0.634	0.507	0.417	0.927
	α_3	1.216	0.949	0.780	0.631	0.513	0.438	0.845
	α_4	1.228	0.965	0.773	0.633	0.520	0.452	0.839
	α_5	1.239	0.972	0.789	0.597	0.498	0.432	0.843
	α_6	1.335	1.050	0.843	0.661	0.451	0.394	0.885
	α_7	1.740	1.348	1.099	0.844	0.628	0.387	0.475
	γ_2	0.428	0.372	0.325	0.273	0.233	0.194	0.339
	γ_3	0.405	0.333	0.290	0.250	0.206	0.165	0.255
	γ_4	0.410	0.336	0.297	0.247	0.202	0.162	0.231
	γ_5	0.410	0.337	0.295	0.252	0.207	0.167	0.259
	γ_6	0.437	0.378	0.332	0.278	0.235	0.195	0.343
	β	0.150	0.119	0.101	0.091	0.083	0.081	0.136

Table 6: Root mean squared error (RMSE), bias, and standard error (SE) of the estimators of the parameters of the rating scale models for based on complete rankings and partial rankings of the first $K' = 1, 2, 3$ categories and the best and worst (B/W) categories.

		K'				
Measure	Parameter	1	2	3	4	B/W
RMSE	τ_3	0.240	0.224	0.208	0.202	0.233
	τ_4	0.258	0.214	0.201	0.192	0.239
	τ_5	0.335	0.287	0.252	0.242	0.305
	δ_1	0.273	0.234	0.213	0.202	0.244
	δ_2	0.221	0.193	0.174	0.165	0.194
	δ_3	0.177	0.160	0.147	0.142	0.163
	σ	0.119	0.100	0.090	0.088	0.104
Bias	τ_3	0.016	0.010	0.010	0.009	0.013
	τ_4	0.025	0.017	0.013	0.012	0.021
	τ_5	0.037	0.027	0.022	0.020	0.032
	δ_1	−0.031	−0.023	−0.019	−0.017	−0.027
	δ_2	−0.021	−0.015	−0.012	−0.011	−0.018
	δ_3	−0.009	−0.006	−0.006	−0.006	−0.010
	σ	0.008	0.005	0.003	0.002	0.009
SE	τ_3	0.240	0.224	0.208	0.201	0.233
	τ_4	0.257	0.213	0.201	0.191	0.238
	τ_5	0.333	0.286	0.251	0.241	0.303
	δ_1	0.272	0.233	0.212	0.202	0.242
	δ_2	0.220	0.193	0.174	0.164	0.193
	δ_3	0.177	0.160	0.147	0.142	0.162
	σ	0.119	0.100	0.090	0.088	0.104