Journal of Choice Modelling 50 (2024) 100475

Contents lists available at ScienceDirect

Journal of Choice Modelling

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

A Bayesian generalized rank ordered logit model

Haotian Cheng [a,*], John N. Ng'ombe [b], Dayton M. Lambert [c]

[a] School of Natural Resources and the Environment, University of Arizona, Tucson, AZ, 85721, USA
[b] Department of Agribusiness, Applied Economics and Agriscience Education, North Carolina A&T State University, Greensboro, NC, 27411, USA
[c] Department of Agricultural Economics, Oklahoma State University, Stillwater, OK, 74078, USA

ARTICLE INFO

JEL classification:
C11
D9
L65
Q1

Keywords:
Best-worst survey
Generalized ranked order logit regression
Bayesian analysis
Preference heterogeneity

ABSTRACT

Using rank-ordered logit regression, researchers typically analyze consumer preference data collected with Best-Worst Scaling (BWS) surveys. We propose a generalized rank-ordered logit (GROL) model that allows flexibility in modeling preference heterogeneity. The GROL and mixed rank-ordered logit model (MROL) accommodate preference heterogeneity. However, the GROL also allows one to model heterogeneity as a function of demographic or environmental variables. A Monte Carlo experiment compares the estimates of accuracy and precision of the proposed GROL estimation with the MROL specification. Simulation results suggest that the GROL model performs comparatively well when the GROL or the MROL is the true data-generating process (dgp). Coefficient and willingness-to-pay estimates of the GROL are more precise and accurate compared to the MROL when the MROL is the true dgp. We surmise that the increased precision of the GROL estimator arises from the added flexibility for modeling different sources of heterogeneity. An empirical application analyzes a BWS survey on consumer preferences for single-use eating-ware (SUEW) products made from biobased materials. Findings suggest that consumers value most product degradability and using non-plastic materials to fabricate SUEW. Consumers also valued the rapidity of product degradability and using non-plastic materials to make SUEW plates. Respondent attentiveness also affected willingness-to-pay (WTP) estimates across attributes. Results suggest attentive respondents were about $3.00 more WTP for biodegradable SUEW than inattentive respondents.

1. Introduction

Conjoint experiments are commonly used to elicit consumer preferences (Lusk et al., 2006; Melstrom and Lupi, 2013; Wu et al., 2020; Thomas et al., 2021; Lambert et al., 2022; Cheng et al., 2021). Alternative approaches include scale-based, max-diff, or best-worst scaling (BWS) survey designs. Finn and Louviere (1992) extended Richardson (1938)'s max-diff preference ranking method to BWS. BWS generates information on respondent preferences by collecting data on a good's most and least preferred qualities (Scarpa et al., 2011; Ryan et al., 2003; Auger et al., 2006). A Scopus search for the keywords "best-worst-scaling" identified 646 studies that used the method since 2006. Ninety-six percent of the studies were published from 2010 to 2022. Medicine (34 percent) and the social sciences (20 percent) were the most prominent fields using BWS. The agricultural and biological sciences (21 percent), business, management, and accounting (15 percent), economics, econometrics and finance (14 percent), and environmental science (12 percent) followed (Scopus, 2022).

* Corresponding author.
E-mail address: haotiancheng@arizona.edu (H. Cheng).

https://doi.org/10.1016/j.jocm.2024.100475
Received 26 July 2022; Received in revised form 26 January 2024; Accepted 31 January 2024
Available online 5 February 2024

There are three versions of BWS. In version 1, respondents evaluate the degree of importance of a set of objects ranked by a rating scale (Finn and Louviere 1992). Version 2 is the BWS 'profile case.' Version 2 arranges profiles in combinations, where after respondents indicate their 'best' and 'worst' choices that correspond with an attribute (Flynn, 2010; Cheung et al., 2016; Mühlbacher et al., 2020; Aizaki and Fogarty, 2019; Cheng, Feuz, and Lambert, 2023a). The third version of BWS pertains to choice experiments where responders are offered a sequence of choice sets, each including at least three profiles (Louviere et al., 2000). Respondents select each set's 'best' and 'worst' item profiles. Previous studies that used BWS version 3 include Mühlbacher et al. (2020), Cheng et al. (2023b), and Adamsen et al. (2013). BWS data are usually analyzed using rank-ordered logit (ROL) regression (Scarpa et al., 2011). Like its cousin, the multinomial logistic (MNL) model, the ROL model maintains the Independence of Irrelevant Alternatives (IIA) property and does not accommodate preference or scale heterogeneity.

The contributions of this study are methodological and topical. We extend the research on choice-based modeling using BWS by developing a generalized ROL model (hereafter, GROL). The GROL extends the ROL by parametrizing scale heterogeneity as a function of demographic and environmental variables and residual taste (preference) heterogeneity. The GROL is an extension of Calfee et al. (2001)'s mixed-effect ROL (MROL) and Fiebig et al. (2010)'s generalized multinomial logit (GMNL). Previous applications using the MROL to analyze BWS data include Fok et al. (2012), Chang et al. (2016), and Naji et al. (2018). Fiebig et al. (2010)'s GMNL has garnered attention in the broader choice modeling literature (e.g., Hole and Kolstad, 2012; Keane and Wasi, 2013; Greene and Hensher, 2010; Balogh et al., 2016; Wright et al., 2018; Liu et al., 2019; Cheng et al., 2021), but the GMNL has been part of an ongoing debate whether scale and preference heterogeneity are separable (Hess and Rose, 2012; Hess and Train, 2017). We are less preoccupied with this relatively unproductive debate and more interested in comparing the performance of competing estimators in terms of accuracy and precision in controlled, experimental settings that are representative of the methods researchers use to collect choice experiment data. We also adopt the position of Hess and Rose that the GMNL allows for additional flexibility in modeling distributions to explain variation in tastes. Feibig et al. (2010) also highlighted this advantage of the GMNL. Moreover, like Hess and Train (2017), we view the GMNL as a tool for estimating the combined effects of all sources of covariance through the scale terms and not as a tool for partitioning sources of heterogeneity into categorically distinct effects.

A feature distinguishing the GROL (or GMNL) from other random parameter logit models is that heterogeneity in preferences due to shifts in scale can be parametrized as a function of exogenous, demographic, or environmental variables. We view the usefulness of modeling scale heterogeneity and why it is of specific interest in modeling choice experiments like this. Some researchers might include demographic or other exogenous variables directly into the right-hand side of the deterministic component of utility, in addition to attribute levels and prices. Arguably, the only variables that should be included in the right-hand side of the linear utility model are attribute levels, including price, the variables used to construct d-optimal experimental designs. It follows that introducing respondent characteristics such as age, education, or other variables omitted from the design matrix of an experiment is inconsistent with the experimental design itself and leads to model misspecification. Suppose one acknowledges that individual-specific factors other than the variables included in the experimental design affect choices and introduce heteroscedasticity. In that case, a natural question is how to introduce this information into a choice model without compromising an experimental design's d-optimality. Parameterizing the scaling factors as a function of individual-specific variables accomplishes this task without compromising the functioning role of a good experimental design. Parameterization of the scaling factor as a function of respondent/demographic characteristics allows researchers to partition the economic effects of the attributes included in the choice experiment from other exogenous factors that determine scale heterogeneity as it relates to choice. This argument also holds for the scaled MNL and the GROL estimator proposed here. We consider this level of generality and flexibility as a potential advantage over the MROL or other competing random parameter models regarding statistical precision and additional information on the relationship between variation in preferences and respondent demographic characteristics.

This paper also contributes to improving methods for comparing competing choice model estimators. Researchers often make conclusions about the comparative performance of choice models using a sample of $N = 1$. In other words, one or a couple of example data sets are used to estimate preferences with competing models. Therefore, conclusions about the performance of these models rest on a single realization of one or a few samples. In other instances, the researcher might conduct a Monte Carlo study. However, how simulated data sets are generated is frequently inconsistent with what practitioners do when they collect data. In other words, attributes and prices included in a Monte Carlo study of choice models should be structured the way one typically structures a choice experiment. In the limited number of studies that have compared the performance of estimators typically used in choice modeling through MC simulation, none used d-optimal experimental designs to anchor underlying data-generating processes (dgp). We address these issues in our Monte Carlo (MC) simulation that compares the accuracy and precision of estimates from GROL with the MROL specifications. The MC study conducted here uses a 100 percent d-optimal design as part of the underlying dgp. The MC results suggest that the GROL model performs well regarding the accuracy and precision of coefficients and WTP when either the GROL or the MROL specification is the true dgp. We also find that the mean squared error of attribute coefficients and WTP estimates are smaller for the GROL, even when the true dgp was the MROL.

An empirical application analyzes BWS data on consumer preferences for purchasing single-use eating-ware (SUEW) products made with biobased materials. SUEW may be durable, less expensive, flexible, and readily available, but they are also ubiquitous sources of pollution (Bläsing and Amelung, 2018; Hodson et al.,2017; Xanthos and Walker, 2017). Due to broader environmental concerns, producing SUEW with smaller environmental footprints has drawn worldwide commitment from various stakeholders and government agencies (Wang et al., 2018). For example, United Nations members' 2015 Sustainable Development Goals include targets to manage and mitigate waste pollution from SUEW made from plastic (United Nations, 2015). The empirical results reported here provide insight into consumer preferences for the attributes that support reducing SUEW's environmental footprint. Findings suggest that consumers value product degradability and using non-plastic, biobased materials to fabricate SUEW products.

2. Methods and procedures

The notation used to derive the generalized ROL follows Beggs et al. (1981)'s and Allison and Christakas (1994)'s exposition of the ROL, and Calfee et al. (2001)'s mixed-effects (or random parameter) ROL (MROL). These studies begin with a linear, indirect utility function with the stochastic terms of utility (ϵ) generated from the type I extreme value distribution. Utility is comprised of deterministic and random components as $v_{ijt} = \mathbf{x}_{ijt}\boldsymbol{\beta} - \beta_m \bullet p + \epsilon_{ijt}$, where for individual i, v_{ijt} is the utility derived from choice occasion t, $\mathbf{x}_{ijt}\beta$ is the deterministic component, $\mathbf{x}_{ijt}$ includes $j = 1, \ldots J$ attributes of the good, β is a J by 1 vector of attribute coefficients, β_m is the marginal utility of income, p is the price, and ϵ_{ijt} is an unobserved random disturbance term with an expected value of zero and variance σ_ϵ^2. Typically, researchers assume that the variance term equals one. For the type-I extreme value distribution, the constant is $\pi^2/6$.

The ROL assumes individuals rank choice alternatives in an order of most to least preferred. Let $r_{ijt} = (r_{i1t}, \ldots, r_{iJt})$ indicate an individual's choice ranking in descending order of preference. The probability that an individual orders $\mathbf{r}$ in any particular sequence is $\Pr[v_{it}(r_{i1t}) > v_{it}(r_{i2t}) > \cdots v_{it}(r_{iJt})]$. Calfee et al. (2001) show how the J-dimensional vector of ordered preferences is decomposable into J - 1 binary (0, 1) variables that indicate preferred alternatives, subject to the censoring of choice set more preferred elements. A closed-form term for this probability is (Calfee et al., 2001):

$$\Pr[v_{it}(r_{i1t}) > v_{it}(r_{i2t}) > \cdots v_{it}(r_{iJt})] = \prod\nolimits_{t=1}\prod\nolimits_{i=1}\prod_{h=1}^{J-1}\frac{\exp\left(\mathbf{x}_{ijt}(r_h)\boldsymbol{\beta}\right)}{\sum_{m=h}^{J}\exp\left(\mathbf{x}_{ijt}(r_m)\boldsymbol{\beta}\right)} \tag{1}$$

where $\mathbf{x}_{ijt}(r_h)$ includes an alternative's attributes that receive rank h in the ordered set. The log-likelihood follows by taking the natural log of Equation (1).

The MROL version assumes that individual utility partworths are a function of population effects and a random term; for example, $\beta_i = \overline{\beta} + \eta_i$, with $\overline{\beta}$ a population average response and η_i a random deviation from the population means denoting preference (or 'residual taste') heterogeneity. The deviations are distributed as random normal variables with a constant variance. Additional details for the MROL model and its estimation are in Srinivasan et al. (2006), Fok et al. (2012), and Chang et al. (2016).

2.1. Generalized rank ordered logit

Parameterization of the GROL follows Fiebig et al.'s (2010) scale and preference heterogeneity treatment in their GMNL model. The GROL parameterizes individual-specific random parameters for choice attributes as:

$$\beta_i = \sigma_i \bullet \overline{\beta} + \gamma \bullet \eta_i + (1-\gamma) \bullet \sigma_i \bullet \eta_i \tag{2}$$

where σ_i is an observation-level scale parameter. The parameter $\gamma \in [0, 1]$ is estimable and determines the trade-off between scale effects and differences in taste (Fiebig et al., 2010). According to Fiebig et al. (2010), this parametrization of individual-specific effects is a flexible way of modeling heterogeneity that arises from two sources because it nests both the scaled-MNL and normal-MIXL. Hess and Train (2017) and Hess and Rose (2012) argue that the GMNL cannot partition the sources of heterogeneity into scale or preferences. Their observation is evident in equation (2), which suggests that scale and preference heterogeneity are multiplicative and confounded. (What may be of substantive interest is the partial effect of σ_i and η_i concerning the estimation of the β_i.) When γ equals one, then the GMNL reduces to what Feibig et al. call the GMNL-I, which parameterizes individual-specific preferences as $\beta_i = \sigma_i \overline{\beta} + \eta_i$. When γ equals zero, the individual preferences are $\beta_i = \sigma_i (\overline{\beta} + \eta_i)$, which Feibig et al. call the GMNL-II.

The scale term is parametrized as a function of demographic and environmental characteristics included in $\mathbf{z}$:

$$\sigma_i = \exp(\overline{\sigma} + \mathbf{z}_i\theta + \tau \bullet \varepsilon_{0i}) \tag{3}$$

where the constant $\overline{\sigma}$ is parameterized as $-0.5 \bullet \tau^2$ such that E $(\sigma_i) = 1$ when $\theta = \mathbf{0}$ (Gu, Hole, and Knox, 2013). The vector θ weights the importance of individual-level demographic and environmental characteristics and their influence on preference heterogeneity. The parameter τ is also estimable and controls the magnitude of the scale parameter, which is increasing in τ. The term ε_{0i} is a standardized normal random variable.

The GROL reduces to an ROL when the IIA assumption holds (Cameron and Trivedi, 2005), which happens when $\sigma_i = 1$ and $\eta_i = \mathbf{0}$. A scaled ROL results when $\eta_i = \mathbf{0}$, with results in shared preferences across the population. Calfee et al.'s MROL is nested in Eq. (2) when scale heterogeneity is absent and $\sigma_i = 1$. The GROL nests each of these models. Of particular interest is the estimation of the conditional individual-specific WTP values. For all specifications, marginal willingness to pay (WTP) for attribute k is computed as follows: Take draws from the posterior distributions of the attribute coefficient (β_k) and price coefficient (β_m) across many MCMC post-burn-in samples (s). For each s, calculate the WTP as $WTP_s = \frac{\overline{\beta}_k^s}{\overline{\beta}_m^s}$. The WTP is then the distribution of WTP_s across all the samples. Taking the mean of this distribution gives the overall mean WTP. Computing WTP in this way accounts for the uncertainty in the model parameters reflected in the posterior distributions, providing a more accurate estimation of WTP in a Bayesian framework.

2.2. Bayesian estimation of the GROL

Simulated maximum likelihood is typically used to maximize the GMNL log-likelihood function because ε_{0i} is a random, unobserved variable (Hole, 2007; Gu, Hole, and Knox, 2013). Due to the complexity of the GROL log-likelihood and the nonlinearity of its parameters, we use a Bayesian approach to generate posterior distributions of the population average effects. There are several advantages to this approach. First, Bayesian estimation of the GROL parameters is flexible because it increases the likelihood that theoretical restrictions on parameters are maintained through judicious selection of prior distributions (details follow). Second, Bayesian techniques are exact and valid for any sample size and can account for uncertainty in the data and parameters (Gelman et al., 2013; McElreath, 2020). This feature is particularly important because economic and behavioral data can be scarce or limited, making it difficult sometimes to obtain accurate parameter estimates. Thus, in such cases, priors can be used to incorporate information from previous studies or expert opinion into likelihood functions, thereby increasing the accuracy and precision of estimates and the likelihood of model convergence for difficult problems. Third, using maximum likelihood to estimate complex models like the GROL can be computationally demanding and may require more data than is available since some parameters are only identified by the system's nonlinearities. In these cases, priors can help facilitate convergence and reduce computational burden (Gelman et al., 2013; McElreath, 2020). Moreover, Bayesian priors 'shrink' parameter estimates towards a plausible range of values, thus potentially reducing the risk of overfitting (van de Schoot et al., 2021). This aspect is especially important when working with economic-behavioral data, which are subject to measurement errors and other types of noise.

Given a random sample of $i = 1, \ldots N$ individuals facing identically distributed and independent ϵ_{ijt}'s, the GROL likelihood function is:

$$P(\mathbf{y}_n; \beta_i) = \prod_{i=1}^{N} \prod\nolimits_{t=1} \prod\nolimits_{i=1} \prod_{h=1}^{J-1} \frac{\exp\left(\mathbf{x}_{ijt}(r_h)\beta_i\right)}{\sum_{m=h}^{J} \exp\left(\mathbf{x}_{ijt}(r_m)\beta_i\right)} \tag{4}$$

where $\mathbf{y}_n$ is the decision-maker's choice, $\mathbf{x}_{ijt}(r_h)$ denotes an alternative's attributes that receive rank h in the ordered set and β_i was defined in Equation (2). A frequentist approach would use simulated maximum likelihood to optimize the log-likelihood function of Equation (4) over the sample to obtain parameter estimates. A Bayesian approach is used here to estimate the GROL model and to compare its performance with the MROL model, also estimated using Bayesian procedures.

Bayesian estimation of the GROL proceeds as follows. Let $\pi = [\boldsymbol{\beta}, \tau, \boldsymbol{\eta}, \boldsymbol{\sigma}, \boldsymbol{\varepsilon}, \theta]$, where $\boldsymbol{\beta}, \tau, \boldsymbol{\eta}, \boldsymbol{\sigma}, \boldsymbol{\varepsilon}$ and θ are as defined in Equations (2) and (3). Equation (4) and prior distributions, $P(\pi)$, for all parameters to be estimated are required. By Bayes' rule, the posterior distribution for the parameters of interest is:

$$P(\pi|\mathbf{y}_n) = \frac{p(\mathbf{y}_n|\pi)p(\pi)}{\int p(\mathbf{y}_n|\boldsymbol{\Sigma\pi})p(\pi)d\pi} \propto p(\mathbf{y}_n|\pi)p(\pi) \tag{5}$$

Exact inference of the posterior distribution for the parameters in Equation (5) is challenging because $\int p(\mathbf{y}_n|\pi)p(\pi)d\pi$ does not have a closed-form solution. However, developments in Bayesian statistics have resulted in more efficient Markov Chain Monte Carlo (MCMC) methods that can be used to estimate the posterior distribution of the parameters in Equation (5). More details of MCMC are found in Gelman et al. (2013) and McElreath (2020). In this study, we recover the posterior distributions of the model's parameters using R-Stan's Hamiltonian Monte Carlo and No U-turn Sampler (HMC-NUTS). The HMC-NUTS sampler is superior to Gibbs Metropolis-Hastings algorithms regarding the number of iterations typically required to achieve convergence (Hoffman and Gelman 2014).

We model residual taste heterogeneity (η) using the Cholesky factorization form of the covariance between the η_i. The lower triangular matrix of the covariance matrix is $\mathbf{L}$, which is parameterized as $\mathbf{R\omega}$, where $\mathbf{R}$ is the factorization of the covariance term's correlation matrix, and ω are standard deviations. Individual-specific deviations around the population-level effects are then formulated as $\eta_i = \mathbf{L}\mathbf{z}_i$, with the $\mathbf{z}_i$ random draws from the standardized normal distribution.

The priors for the population effects are:

$$\overline{\beta}_{attributes} \sim N(0, 2)$$

$$\theta \sim N(0, 0.25)$$

Priors for the parameters on the attributes are Gaussian and centered on zero with a standard deviation of 2. The prior for the scale parameters are also centered on zero with a standard deviation of 0.25.

The priors used to model the *combined effects* of preference heterogeneity are:

$$\gamma \sim \text{Beta}(2, 2)$$

$$(\tau, \omega) \sim \text{Exponential}(1)$$

$$\mathbf{R} \sim \text{Lkj}(4)$$

$$(z_i, \varepsilon_{0i}) \sim \text{Normal}(0, 1)$$

where Lkj is the Lewandowski-Kurowicka-Joe distribution with a rate parameter of four (Lewandowski et al., 2009). When the rate parameter is set to two, the density function tends to a uniform distribution. The density function tends toward the normal distribution when the rate is six. Thus, the average number of four was selected, which maintains the shape of the normal but with slightly wider tails. The prior for the γ parameter is Beta (2, 2) because it is restricted to be inside the (0,1) interval. The exponential distribution with a rate parameter of one is the prior for τ. Set this way, the exponential prior carries no more information than an average deviation around zero (McElreath, 2020).

The warmup series (or 'burn-in') of the HMC sampler included 10,000 iterations, followed by an additional 10,000 iterations. We set the maximum tree depth to 15 and the average probability of accepting a posterior to 0.95 to improve sampling efficiency. We ran four chains in parallel and thinned the chains by 10, resulting in 1000 posterior samples for each chain. Gelman and Rubin (1992)'s $\widehat{R}$ statistic is used to verify chain convergence for each parameter. A $\widehat{R}$ diagnostic approaching one indicates convergence. An Effective Sample Size (ESS) is another diagnostic tool used to assess convergence of the MCMC. A higher ESS (e.g., upwards of 1000 in this case) suggests that the chain has converged to the target distribution, indicating that the sampled values are representative of the posterior distribution. All estimations were performed *rstan*, an R software interface (Stan Development Team, 2023).

2.3. Monte Carlo experiment

We compare the GROL's accuracy and precision with the MROL in a Monte Carlo experiment. We formulate two dgp's; one is the GROL, and the other is the MROL. We compare the performance of each estimator under both dgp's. We simulate 1000 replicate data sets for two cases. The first case considers the GROL specification as the true dgp. The second case considers the MROL to be the true dgp.

There are 100 and 500 'individuals' considered in the simulation. Each individual answers 12 questions, resulting in 12 choice scenarios. Each scenario includes $J = 3$ alternatives ranked from most preferred to least preferred. The alternatives are included in $\mathbf{x}_{ijt}$, a vector including two attributes and a price. Attributes 1 and 2 have two levels, '0' and '1', while the price attribute includes discrete values of 2, 4, and 5. As with most empirical surveys on consumer preferences, we hold attribute and price levels fixed across cases and individuals but structure them such that the design matrix is 100 percent d-optimal.

Concerning the simulation of preference rankings, options are randomly ranked from 'most preferred' to 'least preferred' and based on the MROL or GROL dgp. To parameterize scale heterogeneity in the GROL model, four binary 'demographic' variables drawn from a uniform (0, 1) distribution are included in $\mathbf{z}_i$. If the random number drawn was less than 0.5, then the demographic variable received a '0'; otherwise, it received '1'. The true values of the corresponding scale parameters are $\theta = [-0.2, 0.2, -0.3, \text{and } 0.4]$, the scale parameter τ is set to 0.4, and $\gamma = 0.5$.

For the GROL dgp, the Monte Carlo experiment proceeds as follows.

a) Draw $n = 100 \times 12 \times 3 = 3600$ or $n = 500 \times 12 \times 3 = 18{,}000$ independent and identically distributed random errors from the Gumbel distribution $\epsilon_{ijt} \sim Gumbel\left(0, \frac{\pi^2}{6}\right)$;
b) Draw 100 or 500 random errors from the Gaussian distribution $\varepsilon_{0i} \sim N(0,1)$;
c) Calculate $\sigma_i = \exp(\overline{\sigma} + \mathbf{z}_i\theta + \tau \bullet \varepsilon_{0i})$, where $\overline{\sigma} = -0.5 \bullet \tau^2$;
d) Simulate η_i from the multivariate Gaussian distribution $\eta_i \sim MVN(\mathbf{0}, \Sigma)$, where Σ is the covariance matrix of the attribute parameters. The covariance matrix was set to:

$$\Sigma = \begin{pmatrix} 0.15 & 0.15 & 0.10 \\ 0.15 & 0.25 & 0.10 \\ 0.10 & 0.10 & 0.20 \end{pmatrix}.$$

e) Calculate β_i as $\beta_i = \sigma_i \bullet \overline{\beta} + \gamma \bullet \eta_i + (1-\gamma) \bullet \sigma_i \bullet \eta_i$ with true population means for attribute 1, attribute 2, and price parameters $\overline{\beta} = [1, 0.5, -0.5]$;
f) Compute utilities as $v_{ijt} = \mathbf{x}_{ijt}\beta_i + \epsilon_{ijt}$ and rank "most preferred" to "least preferred";
g) Estimate the GROL and MROL models using Bayesian HMC to recover the parameter estimates;
h) Repeat steps (a) through (g) 1000 times.

The same procedure was used for the MROL dgp except, but with $\theta = \mathbf{0}$, the scale parameter $\tau = 0$, and $\sigma_i = 1$..

2.4. Evaluation of the parameter estimates

Interest is in the accuracy and precision of the GROL and MROL under the different dgps. The MC experiment generates four sets of parameters: 1) GROL estimates under the GROL dgp, 2) MROL estimates under the GROL dgp, 3) GROL estimates under the MROL dgp, and 4) MROL estimates under the MROL dgp. Mean squared error (MSE) and bias are calculated for comparisons. Mean squared error (MSE) is computed as:

$$MSE(\widehat{\mathbf{A}}) = \frac{1}{M}\sum_{m=1}^{M}[\widehat{\mathbf{A}}_m - \mathbf{A}]^2 \tag{6}$$

where $\mathbf{A}$ includes the model parameters, $m = 1{,}000$, and $\widehat{\mathbf{A}}$ is the posterior mean estimate of $\mathbf{A}$, the true attribute values. Bias is calculated as:

$$Bias(\widehat{\mathbf{A}}) = \frac{1}{M}\sum_{m=1}^{M}[\widehat{\mathbf{A}}_m - \mathbf{A}] \tag{7}$$

3. Data for the empirical application

Survey data on preferences for SUEW were collected using an online BWS survey hosted by Qualtrics, December 2019.[1] Respondents were 18 or older and randomly drawn from a nationally representative sampling frame of US households. Qualtrics stratifies their sample according to United States Census regions, income levels, gender, and age.[2] Individuals were invited to participate in the survey via computer or cell phone. Respondents who completed the survey were compensated with coupons. Sampling corresponded with a three percent margin of error with a 95 percent confidence interval.

After responding to a survey consent question, respondents were asked to answer a series of screening questions to identify the subgroup of consumers who would most likely define the SUEW market. The series of screening questions included 1) if the respondent was responsible for preparing and serving food in the household; 2) if they shopped for groceries; 3) if they planned and organized home entertainment events; and (4) if their household used SUEW, were they the person making the purchase decision. Non-consenting respondents and respondents answering 'no' to screening questions 1 to 4 did not continue the survey. There were 345 completed surveys.

Demographic information collected included respondent gender, age, educational attainment, residential location, household income, and household size. Fifty-one percent of the respondents were male (49 percent in 2010 US Census) (Table 1). The average age of respondents was 46 years (the 2010 median age from the 2010 US Census is 37 years). Forty-three percent of the respondents had a college degree. On average, there were 2.8 persons living in a household (2.6 persons in 2010 US Census). Thirty-three percent of respondents lived in rural areas, according to the US Census Bureau's definition of rurality (McGeeney et al., 2019). Eighteen percent of the respondents lived in the northeastern region (18 percent in the 2010 US Census), 21 percent in the Midwest (22 percent in the 2010 US Census), and 37 in the south (37 percent in the 2010 US Census), with the remainder living in western states. Respondents reported their 2018 household income before taxes in eight brackets. The $25,000 to $49,999 range had the most respondents (23 percent), followed by $50,000 to $74,999 range (18 percent). Respondents were asked about their political orientation (strong conservative, moderately conservative, lean towards conservative, independent, lean toward liberal, strong liberal). Respondents were also asked to indicate their residential status (single homeowner, rent, apartment, mobile home) and if they lived in an urban or rural location.

3.1. Information screens

The survey focused on the use of biobased products in the manufacture of SUEW. Respondents received background information on what defined an SUEW product before they completed BWS tasks. There were five information screens on 1) definitions of single-use eating ware products and the term 'biobased'; 2) product degradability (Figure A1 in the Appendix); 3) the contribution of biobased products to the US economy (Figure A2 in the Appendix); 4) product content certification (Figure A3 in the Appendix); and 5) the use of wheat straw (a biobased material) for fabricating bioplastic molds (Figure A4 in the Appendix).

The first information screen included definitions of 'single-use eating-ware products' and 'biobased.' Respondents were provided with examples of products made with biobased inputs, including shopping bags (which can be made from corn starch); drinking straws (which can be made from bamboo or wheat straw); bowls, cartons, containers, and plates (which can be made from sugar cane, paper, or molded wheat straw). The first information screen included the following text:

We consume single-use products every day when we shop for food, eat at restaurants, and entertain. For example.

- We use disposable bags to carry groceries.
- Leftover food we take home after eating-out is placed in a bag or box.
- If food is delivered to our home or eaten at a restaurant, it might be packaged in a container or wrapping.
- We might use disposable utensils, bowls, plates, or cups when we entertain.
- We might use disposable utensils, bowls, or plates for everyday use.

These single-use products can be made from materials such as petroleum-based plastics, recycled products, paper made from trees,

[1] XXXX University IRB Application AG-19-9.

[2] The four census divisions of the lower 48 US states are the Northeast (ME, NH, VT, NY, PA, MA, RI, CT, NJ, DE, MD, and DC), South (DE, MA, VA, WV, KY, NC, SC, TN, GA, FL, AL, MI, AR, LO, TX, and OK), Midwest (including ND, SD, NE, KS, MH, IA, MO, WI, IL, IN, MI, and OH), and West (all other states).

Table 1
Variable names and summary statistics.

Variable Name	Description	Mean	Standard Deviation	Min	Max
Demographics:					
Age	Respondents age (years)	46.40	16.31	19	90
Male	1 if male, otherwise 0	0.51		0	1
Mw	1 if in Midwest, otherwise 0	0.21		0	1
Ne	1 if in Northeast, otherwise 0	0.18		0	1
So	1 if in south, otherwise 0	0.37		0	1
Recycle	1 if recycles on a regular basis, otherwise 0	0.82		0	1
Envir	1 if member of any environmental organization, otherwise 0	0.12		0	1
College	1 = had college or higher, otherwise 0	0.43		0	1
Famil	1 = unfamiliar, 8 = very familiar	5.00		1	8
Rural	1 if rural, otherwise 0	0.32	0.47	0	1
Seconds	Time to finish the survey (seconds)	1 411	1878	390	44864
	1 if less than $25,000	0.19		0	1
	2 if $25,000 to $49,999	0.23		0	1
	3 if $50,000 to $74,999	0.18		0	1
Hhi	4 if $75,000 to $99,999	0.12		0	1
	5 if $100,000 to $149,999	0.16		0	1
	6 if $150,000 to $200,000	0.05		0	1
	7 if $200,000 or more	0.03		0	1
N = 345					

or plant fibers from agricultural crops.

A definition of 'biobased' followed, providing information on the potential use of biobased inputs in the manufacture of products: All of the single-use items previously mentioned can also be made partly or entirely from **biobased materials**. Products made from bio-based materials are called '**biobased products**'.

The remaining information screens provided additional context on SUEW biobased product attributes, such as biodegradability, where the product is made, and input composition (Figure A2, A3, A4).

3.2. Cheap talk screen and trap question I

A cheap talk screen followed the information screens. Previous research suggests that cheap talk screens help reduce hypothetical bias (List et al., 2006; Loomis, 2013). Respondents were asked to reflect on their usual budget allocated for this expense as they completed the best-worst ranking tasks. In internet-based surveys, respondents may not dedicate similar levels of attention to answering questions. Inattentiveness may affect the quality of responses, especially from distracted respondents. We control for this scenario by identifying inattentive respondents with a trap question (Malone and Lusk, 2018). The trap question was included in the cheap talk script as:

> *In surveys like this, people often do not pay much attention to the actual prices shown because they don't really have to pay the cost of the plate they prefer. Instead, they simply notice that one price is higher than another. When answering the survey questions on the next screen, please closely examine the prices and consider these in comparison to your household's budget before choosing a particular plate attribute.* ***To show that you have read the instructions, please answer the question below about "What color is the sky according to the above paragraph?" by checking "none of the above" as your answer*** [emphasis added].

Respondents who correctly answered the question continued to the next survey section. Respondents who incorrectly answered the trap question were asked to re-read the paragraph. Respondents who incorrectly answered the question on the second try were coded as inattentive (='1', '0' otherwise). Of the 345 completed surveys, 90 percent of the respondents correctly answered the first trap question (Table 2). Of the respondents who incorrectly answered the first trap question, 34 percent revised their answer to the correct response.

3.3. Best-worst choice section

The following attributes were used to develop SUEW profiles: a) product degradability (3 levels; not degradable, degradable in 6 months (*Degrade6*), degradable in 24 months (*Degrade24*)); b) origin (2 levels; made in the US, or made elsewhere (*Origin*)); c) product content certification (2 levels; no or yes, (*Label*)); d) material source (3 levels; plastic, paper (*Paper*), or wheat straw (*Wheat*)); and e) a price for a 25-count of 10-inch size SUEW plates (6 levels; $2.27, $3.82, $5.36, $6.91, $8.45, or $10.00) (Table 3).[3] Price points were determined from a review of 20 SUEW products. The highest price was $10.00 for a 25-count package of 10-inch plates. The lowest

[3] Prices were collected from Amazon, June 2019. The link to the $10.00 package of 25 single use food plates is: https://www.amazon.com/10/25counts. The link to the $2.27 package of single use food plate is: https://www.amazon.com/2.27/25counts.

Table 2
Trap question summary.

Trap Question	Answer	Number of respondents	Percent
First trap question, first attempt	correct	311	90.0 %
	incorrect	34	10.0 %
First trap question, second attempt	correct	11	3.3 %
	incorrect	23	6.3 %
Second trap question, first attempt	correct	286	83.0 %
	incorrect	59	17.0 %
Second trap question, second attempt	correct	18	5.1 %
	incorrect	41	11.8 %
Inattentive respondents		56	16.2 %
N		345	

Table 3
Choice experiment levels and attributes.

Attribute	Level 1	Level 2	Level 3	Level 4	Level 5	Level 6
Degradability	Not degradable	6 months	24 months			
Content certification	No	Yes				
Material	Plastic	Paper	Wheat straw			
Origin	Made in the US	Made elsewhere				
Price ($/25 count)	$2.27	$3.82	$5.63	$6.91	$8.45	$10.00

price for the same quantity and plate size was $2.27. These lower- and upper-bound prices were used to determine the other three price levels. The prices were uniformly distributed between the lower and upper price bounds.

The choice experiment's design space included $6 \times 3 \times 3 \times 2 \times 2 = 216$ profiles, which makes 23,220 choice combinations ($C^2_{216} =$ 23,220). The SAS macro %*mktex* was used to generate a balanced fractional factorial orthogonal design that was used to structure choice tasks (SAS, 9.4; Lentner and Bishop, 1986). The d-optimal efficiency score was 91 percent, with 12 choice tasks per respondent. The total number of observations available for analysis was 20,700 (12 tasks $\times$ 5 products $\times$ 345 respondents).

Respondents were asked to assume that the SUEW products were identical in all ways (including functionality) except for the attributes under evaluation. Respondents viewed a screen with the five attributes, from which they selected the most and least attractive attribute from among the choices (Fig. 1). The most and least attractive attributes were removed from the set, leaving three attributes to rank most or least preferred in a second round. Completion of the second round resulted in an attribute ranking for a task. Thus, each choice set ended up with a ranking of '1' (most preferred) to '5' (least preferred).

3.4. Debriefing questions and trap question II

Debriefing questions followed the BWS choice experiment. The debriefing questions asked where respondents would most likely purchase SUEW plates (big box stores, warehouse clubs, convenience stores, online), how much they typically spent on disposable plates in the last six months, and the importance of each attribute as related to their purchasing decision (Table 4). Respondent views on environmental issues were also collected from a series of Likert questions. Respondents were asked if they 'strongly disagreed' or 'strongly agreed' on a five-interval scale regarding their outlook on causes of environmental problems or issues and potential solutions to them (Table 5).

The second trap question was embedded in Likert questions, covering respondents' views on the environment. Respondents were asked, "*Do you live in the United States?*" with a correct answer of 'strongly agree.' Respondents answering correctly advanced to the survey's next section. Respondents who incorrectly answered the question received a second chance to revise their answers. Respondents incorrectly answering the question on the second try were coded as inattentive (='1', '0' otherwise). For the second trap question, 83 percent of the respondents responded with the correct answer on their first try. On the second try, 38 percent revised their answer to the correct response.

In total, enumerating the responses from traps one and two, 16 percent of the respondents were coded as 'inattentive' with a '1' ('0' otherwise) (Table 2). The inattentive dummy variable was interacted with each of the product attributes, and price, to control for respondent inattention (discussed in next section).

4. Results

Table 6 presents Monte Carlo results for the GROL and MROL when the sample size of simulated respondents is 100. Table 7 presents the MC results when the sample size of simulated respondents is 500.

In both tables, the first part presents the GROL and MROL results when the true dgp is the GROL, while the second part presents the results when the dgp is the MROL. The true parameter vector is listed in the last column.

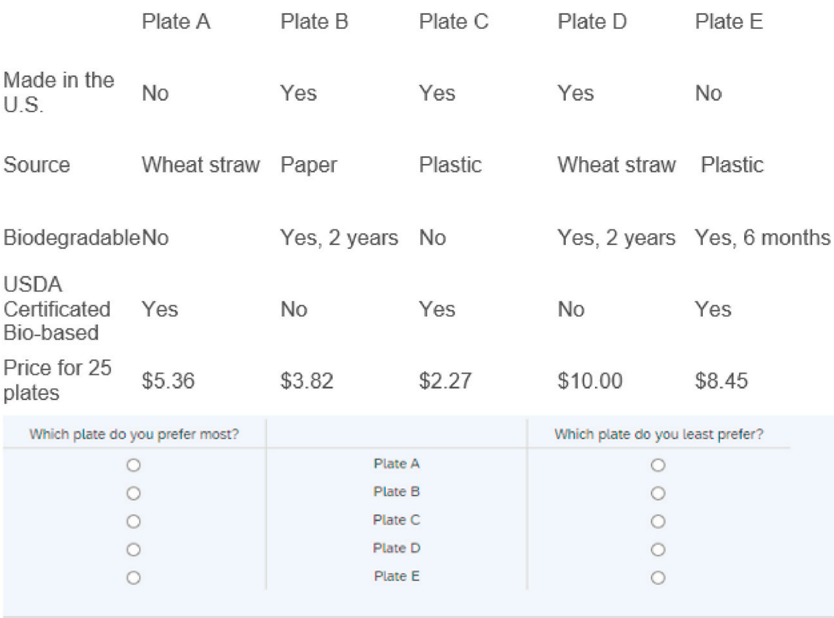

	Plate A	Plate B	Plate C	Plate D	Plate E
Made in the U.S.	No	Yes	Yes	Yes	No
Source	Wheat straw	Paper	Plastic	Wheat straw	Plastic
Biodegradable	No	Yes, 2 years	No	Yes, 2 years	Yes, 6 months
USDA Certificated Bio-based	Yes	No	Yes	No	Yes
Price for 25 plates	$5.36	$3.82	$2.27	$10.00	$8.45

Which plate do you prefer most?		Which plate do you least prefer?
○	Plate A	○
○	Plate B	○
○	Plate C	○
○	Plate D	○
○	Plate E	○

Fig. 1. Best-worst question example.

Table 4
Respondent perceptions and viewpoints on product attributes and shopping habits.

Statements (1)	Mean	Std. Dev.	Min	Max
How important were each of the following attributes to you in making your choices?				
The plate was made from wheat straw	2.71	1.34	1	5
The plate was USDA-certified biobased	3.14	1.29	1	5
The plate was made in the United States	3.11	1.39	1	5
The plate's biodegradability	3.76	1.24	1	5
The plate's price	3.91	1.15	1	5
Compared to a low price, please rate the importance of the following attributes for disposable plates or utensils:				
Recyclable	3.69	1.31	1	5
Made from renewable source	3.47	1.29	1	5
Sturdy	4.08	1.01	1	5
Biodegradable	4.24	0.93	1	5
Appealing appearance	3.78	1.21	1	5
Safe to use	3.18	1.35	1	5
In the past 6 months, about how much did you spend on disposable plates?				
$0.00	0.08		0	1
$1.00-$1.99	0.03		0	1
$2.00-$2.99	0.05		0	1
$3.00-$3.99	0.08		0	1
$4.00-$4.99	0.07		0	1
$5.00-$5.99	0.10		0	1
$6.00-$6.99	0.06		0	1
$7.00-$9.99	0.10		0	1
$10.00-$19.99	0.18		0	1
$20.00-$29.99	0.11		0	1
$30.00 or more	0.13		0	1
Where do you most often purchase disposable plates?				
Big Box Stores	0.44		0	1
Retail Grocery Stores	0.21		0	1
Warehouse Clubs	0.15		0	1
Discount Store	0.15		0	1
Online	0.03		0	1
Convenience Stores	0.01		0	1
Other	0.01		0	1
N = 345				

Notes.
(1) Likert scale: 1 = 'strongly disagree', 2 = 'somewhat disagree', 3 = 'neither agree or disagree', 4 = 'somewhat agree', 5 = 'strongly agree'.

Table 5
Respondent perceptions and viewpoints on environmental issues.

Statements (1)	Mean	Std. Dev.	Min	Max
This survey could encourage producers of single-use food containers to use bio-based materials.	4.01	0.99	1	5
Consumers affect the environment with their product choices.	4.30	0.89	1	5
My personal actions have no impact on environmental problems.	2.50	1.43	1	5
Science and technology will find ways to solve environmental problems.	3.71	1.03	1	5
Most people are unwilling to make sacrifices to address environmental problems.	3.70	1.03	1	5
Government policy is needed to solve environmental problems.	3.79	1.12	1	5
Private industry will develop ways to minimize environmental problems.	3.60	1.09	1	5
Protecting the world's forests is critical to maintaining healthy environment.	4.38	0.91	1	5
Protecting the world's oceans is critical to maintaining healthy environment.	4.40	0.88	1	5
There is no urgent need to slow climate change.	2.42	1.44	1	5
Reducing the amount of single-use plastic pollution is important.	4.01	0.99	1	5
There is no urgent need to reduce greenhouse gas emissions.	4.30	0.89	1	5
We have a responsibility to protect the environment for future generations.	2.50	1.43	1	5
I do not have enough knowledge to make well-informed decisions on environmental issues.	3.71	1.03	1	5
N = 345				

Notes: (1) Likert scale: 1 = 'strongly disagree', 2 = 'somewhat disagree', 3 = 'either agree or disagree', 4 = 'somewhat agree', 5 = 'strongly agree'.

Table 6
Posterior means, bias and mean square errors, and WTP estimates of the GROL and mixed ROL models under Alternative's dgp when the sample size = 100.

GROL DGP

	GROL				Mixed ROL				
Variable Name	Posterior Mean	Bias	MSE	$\widehat{R}$	Posterior Mean	Bias	MSE	$\widehat{R}$	True
Price	−0.4820	0.0180	0.0003	1.001	−0.4783	0.0217	0.0005	1.001	−0.5000
Attr 1	0.9882	−0.0118	0.0001	1.001	0.9818	−0.0182	0.0003	1.001	1.0000
Attr 2	0.5186	0.0186	0.0004	1.001	0.5121	0.0121	0.0002	1.001	0.5000
WTP 1	2.0502	0.0502	0.0026	1.001	2.0527	0.0527	0.0028	1.001	2.0000
WTP 2	1.0759	0.0759	0.0060	1.001	1.0707	0.0707	0.0053	1.001	1.0000
γ	0.4714	−0.0286	0.0142	1.001					0.5000
τ	0.4515	0.0515	0.0028	1.001					0.4000
η_{Price}	0.1651	0.0151	0.0004	1.001	0.1801	0.0301	0.0011	1.001	0.1500
η_1	0.2686	0.0186	0.0004	1.001	0.2740	0.0240	0.0009	1.001	0.2500
η_2	0.2015	0.0015	$5.27e^{-5}$	1.001	0.2056	0.0056	0.0002	1.001	0.2000
θ_1	−0.2413	−0.0413	0.0018	1.001					−0.2000
θ_2	0.1619	−0.0381	0.0015	1.001					0.2000
θ_3	−0.3275	−0.0375	0.0008	1.001					−0.3000
θ_4[a]	0.3432	−0.0568	0.0033	1.001					0.4000

Mixed ROL DGP

	GROL				Mixed ROL				
Variable Name	Posterior Mean	Bias	MSE	Rhat	Posterior Mean	Bias	MSE	Rhat	True
Price	−0.4802	0.0198	0.0004	1.001	−0.4797	0.0203	0.0004	1.001	−0.5000
Attr 1	0.9863	−0.0137	0.0002	1.001	0.9858	−0.0142	0.0002	1.001	1.0000
Attr 2	0.5014	0.0014	$3.61e^{-6}$	1.001	0.5013	0.0013	$3.54e^{-6}$	1.001	0.5000
WTP 1	2.0539	0.0539	0.0030	1.001	2.0550	0.0550	0.0032	1.001	2.0000
WTP 2	1.0441	0.0441	0.0020	1.001	1.0450	0.0450	0.0021	1.001	1.0000
γ	0.5385			1.001					0.0000
τ	0.0081			1.001					0.0000
η_{Price}	0.1547	0.0047	$2.54e^{-5}$	1.001	0.1548	0.0048	$2.66e^{-5}$	1.001	0.1500
η_1	0.2607	0.0107	0.0001	1.001	0.2613	0.0113	0.0001	1.001	0.2500
η_2	0.2035	0.0035	$1.40e^{-5}$	1.001	0.2039	0.0039	$1.70e^{-5}$	1.001	0.2000
θ_1	−0.0101			1.001					0.0000
θ_2	−0.0102			1.001					0.0000
θ_3	−0.0102			1.001					0.0000
θ_4[b]	−0.0086			1.001					0.0000

[a] To conserve space, we do not report the results for $\widehat{\sigma}_i$ as it would mean reporting means and credible intervals for every observation in the analysis. These are available from authors upon request.

[b] To conserve space, we do not report the results for $\widehat{\sigma}_i$ as it would mean reporting means and credible intervals for every observation in the analysis. These are available from authors upon request.

Table 7
Bias and mean square errors, and WTP estimates of the GROL and mixed ROL models under the Alternative's dgp when the sample size = 500.

GROL DGP									
	GROL				Mixed ROL				
	Posterior Mean	Bias	MSE	$\widehat{R}$	Posterior Mean	Bias	MSE	$\widehat{R}$	True
Price	−0.4855	0.0145	0.0002	1.001	−0.4853	0.0147	0.0002	1.001	−0.5000
Attr 1	0.9917	−0.0083	$6.90e^{-5}$	1.001	0.9912	−0.0088	$7.76e^{-5}$	1.001	1.0000
Attr 2	0.5129	0.0129	0.0002	1.001	0.5098	0.0098	0.0001	1.001	0.5000
WTP 1	2.0426	0.0426	0.0033	1.001	2.0424	0.0424	0.0108	1.001	2.0000
WTP 2	1.0564	0.0564	0.0033	1.001	1.0505	0.0505	0.0019	1.001	1.0000
γ	0.4824	−0.0176	0.0014	1.001					0.5000
τ	0.4716	0.0716	0.0046	1.001					0.4000
η_{Price}	0.1497	−0.0003	$7.01e^{-6}$	1.001	0.1527	0.0027	$5.73e^{-6}$	1.001	0.1500
η_1	0.2498	−0.0002	$1.70e^{-6}$	1.001	0.2522	0.0022	$1.08e^{-5}$	1.001	0.2500
η_2	0.1999	−0.0001	$3.34e^{-6}$	1.001	0.2007	0.0007	$6.57e^{-6}$	1.001	0.2000
θ_1	−0.2529	−0.0529	0.0028	1.001					−0.2000
θ_2	0.1246	−0.0754	0.0057	1.001					0.2000
θ_3	−0.3550	−0.0550	0.0030	1.001					−0.3000
θ_4[a]	0.3162	−0.0838	0.0070	1.001					0.4000

Mixed ROL DGP									
	GROL				Mixed ROL				
Variable Name	Posterior Mean	Bias	MSE	Rhat	Posterior Mean	Bias	MSE	Rhat	True
Price	−0.4855	0.0145	0.0002	1.001	−0.4854	0.0146	0.0002	1.001	−0.5000
Attr 1	0.9916	−0.0084	$3.96\ e^{-7}$	1.001	0.9912	−0.0088	$4.76e^{-7}$	1.001	1.0000
Attr 2	0.5144	0.0144	0.0018	1.001	0.5125	0.0125	0.0002	1.001	0.5000
WTP 1	2.0424	0.0424	0.0036	1.001	2.0420	0.0420	0.0018	1.001	2.0000
WTP 2	1.0595	0.0595	0.0002	1.001	1.0558	0.0558	0.033	1.001	1.0000
γ	0.4851								0.0000
τ	0.0118								0.0000
η_{Price}	0.1498	0.0002	$7.85e^{-6}$	1.001	0.1506	0.0006	$3.15e^{-5}$	1.001	0.1500
η_1	0.2498	0.0002	$4.36e^{-6}$	1.001	0.2503	0.0003	$7.98e^{-6}$	1.001	0.2500
η_2	0.1999	−0.0001	$3.18e^{-6}$	1.001	0.2007	0.0007	$1.63e^{-5}$	1.001	0.2000
θ_1	−0.0748								0.0000
θ_2	−0.0734								0.0000
θ_3	−0.0666								0.0000
θ_4[b]	−0.0727								0.0000

[a] To conserve space, we do not report the results for $\widehat{\sigma}_i$ as it would mean reporting means and credible intervals for every observation in the analysis. These are available from authors upon request.
[b] To conserve space, we do not report the results for $\widehat{\sigma}_i$ as it would mean reporting means and credible intervals for every observation in the analysis. These are available from authors upon request.

When the true dgp was the GROL and the sample size was 100, the MROL coefficients and WTP bias were similar to the bias for the same variables estimated with the GROL (Table 6). In general, the MSE was similar to the MROL coefficients and WTP. The bias and MSE of preference heterogeneity parameters estimated with the MROL were larger than the GROL estimates. When the true dgp was the MROL, the bias and MSE of the coefficient and WTP estimates were similar to the GROL. The bias and MSE of the preference heterogeneity parameters were larger when estimated with the MROL.

Next, we report the results when the sample size was 500 (Table 7). When the true dgp was GROL, the bias and MSE of the MROL coefficient and WTP estimates were similar to those of the GROL were. The bias and MSE of the preference heterogeneity parameters were smaller for the GROL. When the true dgp was the MROL, the bias and MSE of the MROL coefficient and WTP estimates were similar to those of the GROL. The bias and MSE of the preference heterogeneity parameters were smaller for the GROL. Researchers are typically interested in the accuracy and precision of WTP estimates and the coefficients used to estimate WTP, not the precision or accuracy of parameters that measure heterogeneity. The bias and MSE of the preference heterogeneity parameters were generally larger for the MROL. This modest MC experiment suggests that the GROL performs as well as its MROL cousin in terms of the bias and MSE of the population-level parameters used to estimate WTP. This finding may be due to the added flexibility in modeling preferences as a mixture of distributions the GROL affords.

Table 8
Bayesian estimation results of the GROL and MROL model of single-use plate.

Variable	The GROL Model			The MROL Model		
	Posterior Mean	95 % Credible Interval		Posterior Mean	95 % Credible Interval	
Preference Heterogeneity						
Price	−0.1353	−0.2038	−0.0871	−0.0684	−0.0769	−0.0601
Paper	0.7377	0.4650	1.1176	0.3644	0.3150	0.4133
Wheat	0.5353	0.3371	0.8223	0.2727	0.2266	0.3181
Degrade6	0.6426	0.4070	0.9736	0.3174	0.2674	0.3674
Degrade24	0.3822	0.2275	0.5992	0.1881	0.1404	0.2347
Label	0.1641	0.0782	0.2850	0.0872	0.0474	0.1268
Origin (Made in US)	0.0533	0.0078	0.1097	0.1514	0.1130	0.1887
d × *Price*	0.1431	−0.1321	0.4350	0.0320	0.0114	0.0529
d × *Paper*	0.0229	−0.2307	0.2720	0.0571	−0.0641	0.1761
d × *Wheat*	−0.3798	−0.7090	−0.1176	−0.0028	−0.1181	0.1115
d × *Degrade6*	−0.4316	−0.7735	−0.1683	−0.2143	−0.3370	−0.0938
d × *Degrade24*	−0.1346	−0.3592	0.0664	−0.2275	−0.3421	−0.1112
d × *Label*	−0.0068	−0.2133	0.2005	−0.0670	−0.1583	0.0248
d × *Origin*	0.0533	0.0078	0.1097	−0.0231	−0.1161	0.0682
Scale Heterogeneity						
Age	−0.0014	−0.0070	0.0041			
Hhi	−0.0078	−0.0541	0.0345			
Recycle	−0.1327	−0.3258	0.0681			
Envir	−0.3350	−0.6200	−0.0784			
Male	−0.0571	−0.2224	0.1043			
MW	−0.3243	−0.5320	−0.1135			
NE	−0.1877	−0.4133	0.0392			
S	−0.2577	−0.4491	−0.0839			
College	−0.1263	−0.2925	0.0416			
Rural	−0.0905	−0.2630	0.0771			
Familiar	−0.0367	−0.0772	0.0036			
τ	0.3405	0.1080	0.4906			
γ	0.2060	0.0211	0.6160			
Log Likelihood	−21,951			−21,953		
N	345			345		

Notes: 95 % Credible Intervals excluding zero imply significance.

4.1. *Empirical results*

Table 8 presents the estimation results for the GROL and MROL models. The Gelman-Rubin $\widehat{R}$ statistics for all parameter estimates were close to one. All ESS statistics exceeded 1000. (Chain plots are included in the Appendix Figure A5 [4]). There were no divergences detected in the GROL and MROL estimates. The 95 % credible intervals are reported next to the posterior means of each model in Table 8, and Fig. 2 shows the population-level coefficient distributions of the attributes and price. Parameter estimates are significant when the 95 % credible intervals exclude zero. Results from Table 8 indicate that the preference heterogeneity parameters are insignificant for either model. This finding implies that if one were to estimate an MROL using this data, one would conclude that preferences for SUEW are commonly shared across this sample of respondents.

We discuss parameter estimates whose credible intervals do not include zero. Inspection of Table 8 indicates that the scale variables, *Envir*, *MW*, *NE*, and *S*, are significant. This result suggests considerable variation exists around an individual's average choice responses. The effects of the demographic variables on scale heterogeneity can also be summarized as an elasticity. The percent change in scale, given a 1-unit change in a demographic variable, is $\frac{\partial\widehat{\sigma}}{\partial\overline{z}_k}\bullet\frac{\overline{z}_k}{\sigma}=\widehat{\gamma}_k\bullet\overline{z}_k$ where $\overline{z}_k$ is the mean of the *k* th covariate included in the scale function of Equation (3). The variable *Envir* has a negative effect on variability across choices. For the individuals who are members of an environmental group or organization, the variation around their SUEW choices decreased by 0.50 percent. Response heterogeneity was lower for individuals who lived in the *MW*, *NE*, and *S* regions. For these individuals, the variation around their choices decreased by 0.46 percent, 0.30 percent, and 0.37 percent, respectively.

The γ parameter governs the trade-off between sources of preference heterogeneity; γ was 0.37 (Table 8). According to Feibig et al.'s classification scheme, this means the estimated model is a mixture of what they call "GMNL-I" ($\beta_i=\sigma_i\overline{\beta}+\eta_i$) and "GMNL-II" ($\beta_i=\sigma_i(\overline{\beta}+\eta_i)$), but it leans more towards the GMNL-I form.

4.2. *Attribute WTP premium*

Table 9 presents the WTP estimates for both attentive and inattentive respondents. These were estimated using both the GROL and

[4] We thinned 1000 out of 10,000 draws (10,000 iterations/10 thins = 1000 samples).

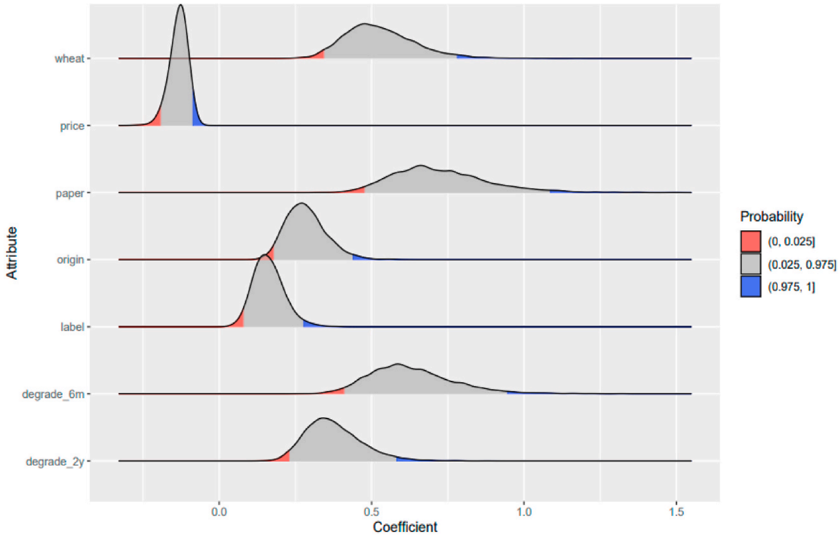

Fig. 2. Distribution of attribute and price coefficients.

Table 9
Willingness to pay estimates from GROL and MROL models considering attention and inattention of respondents.

Attributes	GROL			MROL		
	Premium	95 % Credible Interval		Premium	95 % Credible Interval	
Attention						
Paper	5.45	4.56	6.54	5.32	4.44	6.32
Wheat	3.96	3.15	4.85	3.98	3.19	4.90
Degrade6	4.75	3.93	5.72	4.63	3.80	5.54
Degrade24	2.83	2.07	3.65	2.75	1.97	3.53
Label(Certified biobased)	1.21	0.61	1.86	1.27	0.71	1.90
Origin (Made in US)	2.11	1.50	2.77	2.21	1.62	2.89
Inattention						
$d \times$ *Price*	0.39	0.03	0.69	0.47	−0.17	0.74
$d \times$ *Paper*	1.06	−0.90	3.23	0.83	−0.83	2.57
$d \times$ *Wheat*	0.17	−1.62	2.09	−0.04	−1.68	1.59
$d \times$ *Degrade6*	−2.81	−4.72	−0.88	−3.13	−4.91	−1.27
$d \times$ *Degrade24*	−3.19	−5.15	−1.39	−3.32	−5.02	−1.54
$d \times$ *Label*	−0.99	−2.64	0.49	−0.98	−2.40	0.38
$d \times$ *Origin*	−0.05	−1.55	1.54	−0.34	−1.67	1.03
Log *Likelihood*	−21,951			−21,953		
N	345			345		

MROL models. For attentive respondents, under the GROL model, the extrinsic attribute with the highest premium is '*source*' (*Paper*, \$5.45 and *Wheat*, \$3.96) followed by '*degradability*' (*Degrade6,* \$4.75 and *Degrade24,* \$2.83). Similar results are observed with the MROL model, though the WTP estimates are slightly different for the GROL model. Noticeably, the WTP estimates among attentive respondents from both models are significant and have the same signs. Concerning inattentive respondents, and under the GROL model, the premiums of *Degrade6* and *Degrade24* are -\$2.81 and -\$3.19 (both significant), respectively. From the MROL model, the premium of *Degrade6* and *Degrade24* are \$4.63 and \$2.75 (both significant), respectively. However, inattentive respondents were willing to pay for *Degrade6* (*Degrade24*) by \$3.13 (\$3.32) less than attentive respondents. The findings are similar to what Arjunan et al. (2010) concluded, indicating that product biodegradability is an important concern for consumers.

Moreover, based on Table 9, the attribute *Origin*'s premium is \$2.12 (under GROL) and \$2.21 (under the MROL) for attentive respondents. The SUEW products made in the US are more highly valued, similar to Barnes et al.'s (2011) conclusion that consumers prefer locally sourced products. The attribute *Label* (\$1.21 for the GROL, and \$1.27 for the MROL) exhibits the lowest premium for attentive respondents, suggesting that respondents ranked the biobased certification label lowest regarding product features. The labeling premium for inattentive respondents is -\$0.99 and -\$0.98 under the GROL and MROL, respectively, but insignificant. Respondents are willing to pay for attributes related to environmental issues, labeling, and materials required to make SUEW.

5. Conclusions

This study extended Calfee et al. (2001)'s mixed ranked ordered logit (ROL) model and Fiebig et al. (2010)'s generalized multinomial logit (GMNL) to a Bayesian generalized ranked ordered logit (GROL) model for estimating consumer WTP for extrinsic attributes while controlling for inattention bias. We modified the mixed ranked ordered logit (MROL), frequently used to model best-worst data, to include variation in scale and preference heterogeneity. A Monte Carlo experiment compared the proposed estimator to the MROL. The GROL is similar in accuracy to MROL estimators of WTP regarding precision and bias when the true dgp was the GROL or the MROL. However, the GROL outperformed the MROL estimator regarding heterogeneity when the true dgp was the GROL or the MROL. We surmise that the extra gain in precision from the GROL in heterogeneity is due to its ability to model preference heterogeneity as a mixture of sources.

The GROL is an extension of Fiebig et al. (2010)'s generalized multinomial logit (GMNL). The GMNL model has attracted attention in the choice modeling literature but has also received substantial criticism. Stated briefly, some persuasively argue that researchers hoping to decompose the sources of variation in tastes into 'preference' or 'scale' heterogeneity are misguided. This assertion is not controversial and does not diminish the value of the GMNLs contribution to the choice modeling toolbox.

Rather than contributing to this debate, we instead focused on developing a Monte Carlo procedure that fairly and accurately compared the precision and accuracy of the GROL to its generative cousin, the MROL. We hope that researchers interested in comparing competing estimators and those who develop novel choice estimators consider carefully the design and implementation of Monte Carlo procedures used to compare and evaluate competing models. Comparisons of estimators that use a single data set or Monte Carlo studies that use inefficient design matrices are of limited value when comparing discrete choice estimators in side-by-side appraisals. Highlighting the issues or problems of competing models is constructive, but verbal or mathematical-symbolic arguments should be supplemented with credible and relevant Monte Carlo experiments to demonstrate the superiority or limitations of one approach over another.

We consider the added level of flexibility and generality given by the GROL and its antecedents as one of the advantages of the proposed model, for which we recommend its use among researchers. In sum, the GROL model offers another option for researchers using ranked ordered regression to analyze best-worst survey data.

This study has the following limitations, which imply suggestions for further research. First, an obvious extension of this study is to consider comparing the performance of the GROL with more than one discrete choice model. Second, we assumed that σ_i followed a lognormal distribution. The lognormal distribution has a heavy tail, which could affect the estimation of important parameters. We were unable to determine how this choice affected model estimates. Competing heteroskedastic functions include the quadratic, linear, and absolute value (or Laplace) forms that might affect estimates differently. We leave this comparative exercise as a future research direction. Third, what is compelling, and to our knowledge not examined before, are the partial effects of the individual-specific preferences (the β_i) concerning σ_i and η_i. These partial effects would show how preferences change, given a change in preference or scale heterogeneity. In this way, sources of heterogeneity could be partitioned into discrete parts if that interests the researcher. Exploring this form of heterogeneity is another area of research. Lastly, the Monte Carlo conducted here only looked at accuracy (bias) and precision (mean squared error). An expanded simulation study would include power, size, and the small-sample properties of competing estimators.

CRediT authorship contribution statement

Haotian Cheng: Writing – original draft, Software, Methodology, Data curation, Conceptualization. **John N. Ng'ombe:** Writing – review & editing, Writing – original draft. **Dayton M. Lambert:** Writing – review & editing, Supervision, Funding acquisition, Conceptualization.

Declaration of competing interest

The authors declare no competing interests.

Data availability

Data will be made available on request.

Acknowledgments

This research was supported by the Sparks Chair in Agricultural Sciences & Natural Resources, Oklahoma State University, and the Oklahoma Agricultural Experiment Station.

Appendices.

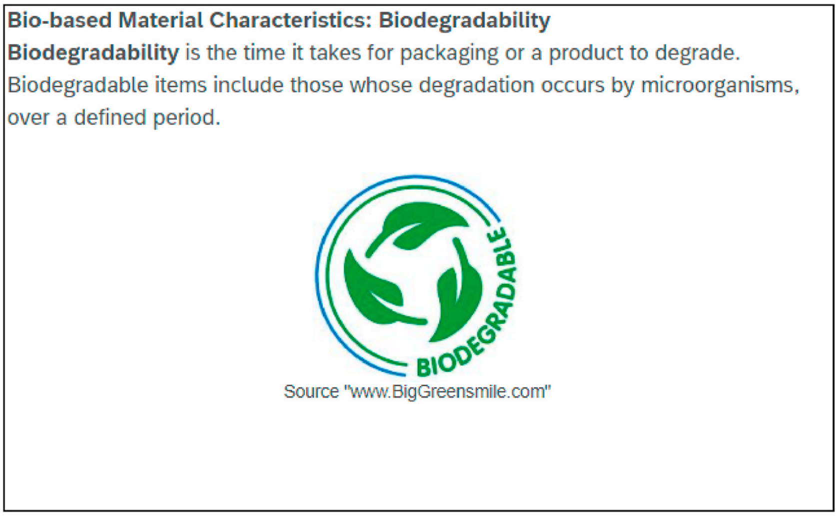

Fig. A1. Biobased Products and Degradability Information Screen.

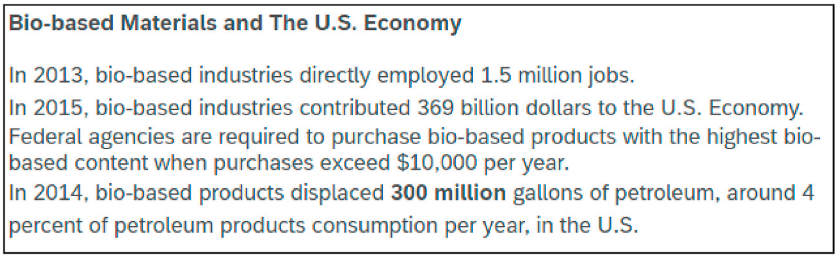

Fig. A2. Economic Contribution of Biobased Products to the US Economy Information Screen.

Fig. A3. Product Content Certification Information Screen.

What is Wheat Straw?

Wheat straw is a byproduct of producing wheat. Wheat straw is what remains after the wheat kernel is removed to make flour and cereal products. Wheat straw can also be used to make biobased products.

Fig. A4. Wheat Straw as A Biobased Input Information Screen

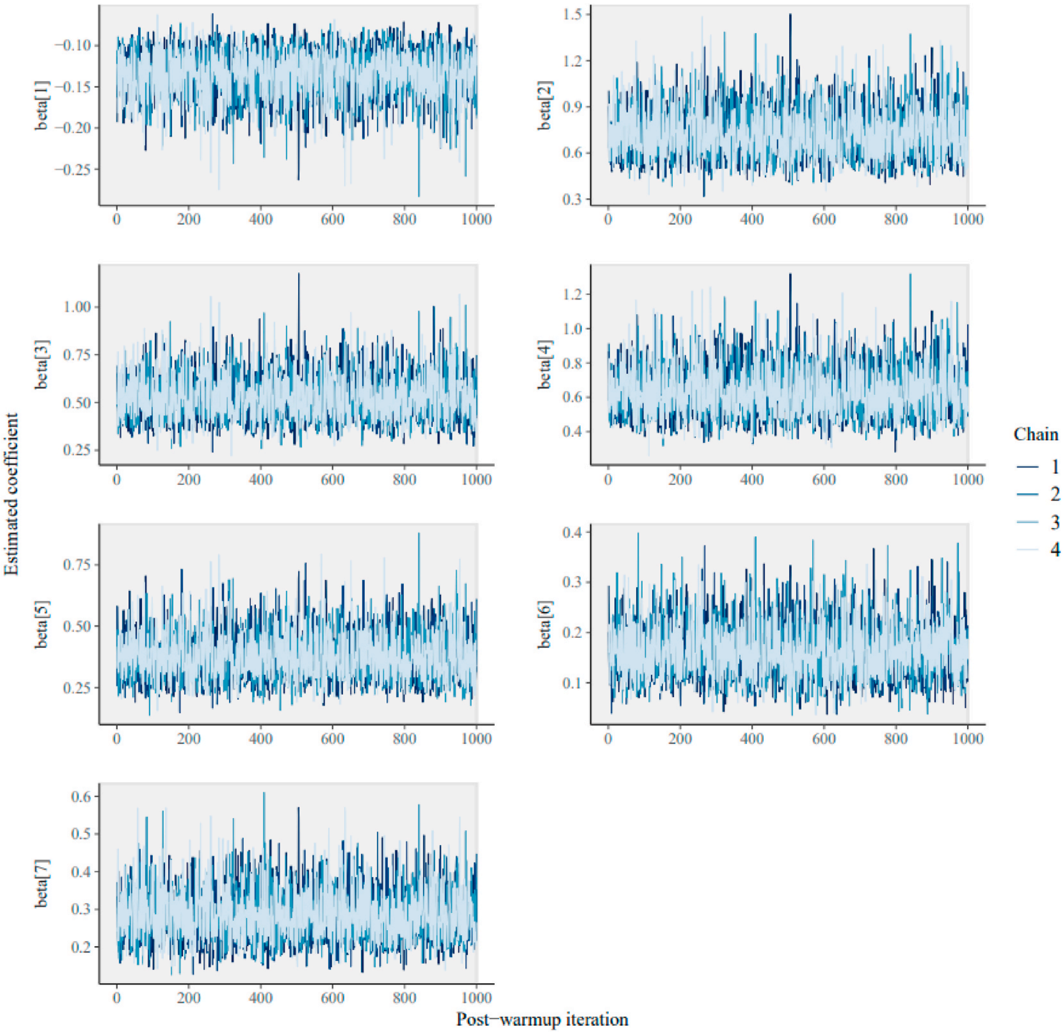

Fig. A5. Posterior Chains for Each Attribute

References

Adamsen, J.M., Rundle, S., Whitty, J.A., 2013. Best-worst scaling: reflections on presentation, analysis, and lessons learned from case 3 BWS experiments. Market& Social Research 21, 9–27. https://core.ac.uk/reader/143877740.

Aizaki, H., Fogarty, J., 2019. An R package and tutorial for case 2 best–worst scaling. Journal of Choice Modelling 32, 100171. https://doi.org/10.1016/j.jocm.2019.100171.

Allison, P.D., Christakis, N.A., 1994. Logit models for sets of ranked items. Socio. Methodol. 24, 199. https://doi.org/10.2307/270983.

Arjunan, S., Moir, C., Kirwan, K., Pink, D., 2010. *The greening of 'green' technology: adoption of bio-plastic PLA* [Paper presentation]. In: The Fourteenth International Consortium on Applied Bioeconomy Research (ICABR). Ravello, Italy.

Auger, P., Devinney, T.M., Louviere, J.J., 2006. Using best–worst scaling methodology to investigate consumer ethical beliefs across countries. J. Bus. Ethics 70 (3), 299–326. https://doi.org/10.1007/s10551-006-9112-7.

Balogh, P., Békési, D., Gorton, M., Popp, J., Lengyel, P., 2016. Consumer willingness to pay for traditional food products. Food Pol. 61, 176–184.

Barnes, M., Chan-Halbrendt, C., Zhang, Q., Abejon, N., 2011. Consumer preference and willingness to pay for non-plastic food containers in Honolulu, USA. J. Environ. Protect. 2 (9), 1264–1273. https://doi.org/10.4236/jep.2011.29146.

Beggs, S., Cardell, S., Hausman, J., 1981. Assessing the potential demand for electric cars. J. Econom. 17 (1), 1–19. https://doi.org/10.1016/0304-4076(81)90056-7.

Bläsing, M., Amelung, W., 2018. Plastics in soil: analytical methods and possible sources. Sci. Total Environ. 612, 422–435.

Calfee, J., Winston, C., Stempski, R., 2001. Econometric issues in estimating consumer preferences from stated preference data: a case study of the value of automobile travel time. Rev. Econ. Stat. 83 (4), 699–707. https://doi.org/10.1162/003465301753237777.

Cameron, C.A., Trivedi, P.K., 2005. Microeconometrics: Methods and Applications. Cambridge University Press.

Chang, F., Li, M., Xu, P., Zhou, H., Haque, M., Huang, H., 2016. Injury severity of motorcycle riders involved in traffic crashes in hunan, China: a mixed ordered logit approach. Int. J. Environ. Res. Publ. Health 13 (7), 714.

Cheng, H., Lambert, D.M., DeLong, K.L., Jensen, K.L., 2021. Inattention, availability bias, and attribute premium estimation for A biobased product. Agric. Econ. 53 (2), 274–288.

Cheng, H., Feuz, R., Lambert, D.M., 2023a. comparison of best-worst scaling marginal and rank methods. Appl. Econ. Lett. 1–4. https://doi.org/10.1080/13504851.2023.2187019.

Cheng, H., Zhang, T., Lambert, D.M., Feuz, R., 2023b. An empirical comparison of conjoint and best-worst scaling case III methods. Journal of Behavioral and Experimental Economics 106. https://doi.org/10.1016/j.socec.2023.102049.

Cheung, K.L., Wijnen, B.F.M., Hollin, I.L., Janssen, E.M., Bridges, J.F., Evers, S.M.A.A., Hiligsmann, M., 2016. Using best–worst scaling to investigate preferences in health care. Pharmacoeconomics 34 (12), 1195–1209. https://doi.org/10.1007/s40273-016-0429-5.

Fiebig, D.G., Keane, M.P., Louviere, J., Wasi, N., 2010. The generalized multinomial logit model: accounting for scale and coefficient heterogeneity. Market. Sci. 29 (3), 393–421. https://doi.org/10.1287/mksc.1090.0508.

Finn, A., Louviere, J.J., 1992. Determining the appropriate response to evidence of public concern: the case of food safety. J. Publ. Pol. Market. 11 (2), 12–25. https://doi.org/10.1177/074391569201100202.

Fok, D., Paap, R., Van Dijk, B., 2012. A rank-ordered logit model with unobserved heterogeneity in ranking capabilities. J. Appl. Econom. 27 (5), 831–846.

Flynn, T.N., 2010. Valuing citizen and patient preferences in health: recent developments in three types of best–worst scaling. Expert Rev. Pharmacoecon. Outcomes Res. 10 (3), 259–267. https://doi.org/10.1586/erp.10.29.

Gelman, A., Carlin, J.B., Stern, H.S., Rubin, D.B., 2013. Bayesian Data Analysis, third ed. Chapman and Hall/CRC.

Gelman, A., Rubin, D.B., 1992. Inference from iterative simulation using multiple sequences. Stat. Sci. 7 (4), 457–511. https://doi.org/10.1214/ss/1177011136.

Greene, W.H., Hensher, D.A., 2010. Does scale heterogeneity across individuals matter? An empirical assessment of alternative logit models. Transportation 37 (3), 413–428.

Gu, Y., Hole, A.R., Knox, S., 2013. Fitting the generalized multinomial logit model in stata. STATA J.: Promoting Communications on Statistics and Stata 13 (2), 382–397. https://doi.org/10.1177/1536867x1301300213.

Hess, S., Rose, J.M., 2012. Can scale and coefficient heterogeneity be separated in random coefficients models? Transportation 39 (6), 1225–1239.

Hess, S., Train, K., 2017. Correlation and scale in mixed logit models. Journal of choice modelling 23, 1–8.

Hoffman, M.D., Gelman, A., 2014. The No-U-turn sampler: adaptively setting path lengths in Hamiltonian Monte Carlo. J. Mach. Learn. Res. 15, 1593–1623. https://www.jmlr.org/papers/volume15/hoffman14a/hoffman14a.pdf.

Hodson, M.E., Duffus-Hodson, C.A., Clark, A., Prendergast-Miller, M.T., Thorpe, K.L., 2017. Plastic bag derived-microplastics as a vector for metal exposure in terrestrial invertebrates. Environ. Sci. Technol. 51 (8), 4714–4721.

Hole, A.R., 2007. Fitting mixed logit models by using maximum simulated likelihood. STATA J. 7 (3), 388–401.

Hole, A.R., Kolstad, J.R., 2012. Mixed logit estimation of willingness to pay distributions: a comparison of models in preference and WTP space using data from a health-related choice experiment. Empir. Econ. 42 (2), 445–469.

Keane, M., Wasi, N., 2013. Comparing alternative models of heterogeneity in consumer choice behavior. J. Appl. Econom. 28 (6), 1018–1045.

Lambert, L.H., Lambert, D.M., Ripberger, J.T., 2022. Public willingness to pay for farmer adoption of best management practices. J. Agric. Appl. Econ. 1–18.

Lewandowski, D., Kurowicka, D., Joe, H., 2009. Generating random correlation matrices based on vines and extended onion method. J. Multivariate Anal. 100 (9), 1989–2001.

List, J.A., Sinha, P., Taylor, M.H., 2006. Using choice experiments to value non-market goods and services: evidence from field experiments. B E J. Econ. Anal. Pol. 6 (2), 1–37. https://doi.org/10.2202/1538-0637.1132.

Liu, R., Gao, Z., Nayga Jr., R.M., Snell, H.A., Ma, H., 2019. Consumers' valuation for food traceability in China: does trust matter? Food Pol. 88, 101768.

Loomis, J., 2013. Strategies for overcoming hypothetical bias in stated preference surveys. J. Agric. Resour. Econ. 39 (1), 34–46. https://doi.org/10.22004/ag.econ.168258.

Louviere, J.J., Hensher, D.A., Swait, J.D., Adamowicz, W., 2000. Stated Choice Methods: Analysis and Applications, first ed. Cambridge University Press.

Lusk, J.L., Norwood, F.B., Pruitt, J.R., 2006. Consumer demand for a ban on antibiotic drug use in pork production. Am. J. Agric. Econ. 88 (4), 1015–1033.

Malone, T., Lusk, J.L., 2018. A simple diagnostic measure of inattention bias in discrete choice models. Eur. Rev. Agric. Econ. 45 (3), 455–462. https://doi.org/10.1093/erae/jby005.

McElreath, R., 2020. Statistical Rethinking: A Bayesian Course with Examples in R and STAN (Chapman & Hall/CRC Texts in Statistical Science), second ed. Chapman and Hall/CRC.

McGeeney, K., Kriz, B., Mullenax, S., Kail, L., Walejko, G., Vines, M., et al., 2019. 2020 Census Barriers Attitudes, and Motivators Study Survey Report: A New Design for the 21st Century. United States Census Bureau. Retrieved from. https://www2.census.gov/programs-surveys/decennial/2020/program-management/final-analysis-reports/2020-report-cbams-study-survey.pdf.

Melstrom, R.T., Lupi, F., 2013. Valuing recreational fishing in the great lakes. N. Am. J. Fish. Manag. 33 (6), 1184–1193.

Mühlbacher, A.C., Sadler, A., Lamprecht, B., Juhnke, C., 2020. Patient preferences in the treatment of hemophilia A: a best–worst scaling case 3 analysis. Value Health 23 (7), 862–869. https://doi.org/10.1016/j.jval.2020.02.013.

Naji, H., Xue, Q., Lyu, N., Wu, C., Zheng, K., 2018. Evaluating the driving risk of near-crash events using a mixed-ordered logit model. Sustainability 10 (8), 2868.

Richardson, M.W., 1938. Multimensional psychophysics. Psychol. Bull. 35, 659–660. https://ci.nii.ac.jp/naid/10009703610/.

Ryan, M., Watson, V., Amaya-Amaya, M., 2003. Methodological issues in the monetary valuation of benefits in healthcare. Expert Rev. Pharmacoecon. Outcomes Res. 3 (6), 717–727. https://doi.org/10.1586/14737167.3.6.717.

Scarpa, R., Notaro, S., Louviere, J., Raffaelli, R., 2011. Exploring scale effects of best/worst rank ordered choice data to estimate benefits of tourism in alpine grazing commons. Am. J. Agric. Econ. 93 (3), 813–828. https://doi.org/10.1093/ajae/aaq174.

Scopus, 2022. https://www.scopus.com/home.uri. (Accessed 29 May 2022).

Srinivasan, S., Bhat, C.R., Holguin-Veras, J., 2006. Empirical analysis of the impact of security perception on intercity mode choice: a panel rank-ordered mixed logit model. Transport. Res. Rec. 1942 (1), 9–15.
Stan Development Team, 2023. "RStan: the R Interface to Stan.". R package version 2.21.8. https://mc-stan.org/.
Thomas, M., Jensen, K.L., Lambert, D.M., English, B.C., Clark, C.D., Walker, F.R., 2021. Consumer preferences and willingness to pay for potting mix with biochar. Energies 14 (12), 3432.
van de Schoot, R., Depaoli, S., King, R., Kramer, B., Märtens, K., Tadesse, M.G., et al., 2021. Bayesian statistics and modelling. Nature Reviews Methods Primers 1 (1), 1.
Wang, J., Zheng, L., Li, J., 2018. A critical review on the sources and instruments of marine microplastics and prospects on the relevant management in China. Waste Manag. Res. 36 (10), 898–911.
Wright, S.J., Vass, C.M., Sim, G., Burton, M., Fiebig, D.G., Payne, K., 2018. Accounting for scale heterogeneity in healthcare-related discrete choice experiments when comparing stated preferences: a systematic review. The Patient-Patient-Centered Outcomes Research 11 (5), 475–488.
Wu, T., Henneberry, S.R., Ng'ombe, J.N., Melstrom, R.T., 2020. Chinese demand for agritourism in rural America. Sustainability 12 (7), 3052.
Xanthos, D., Walker, T.R., 2017. International policies to reduce plastic marine pollution from single-use plastics (plastic bags and microbeads): a review. Mar. Pollut. Bull. 118 (1–2), 17–26.