Statistical Modelling 2022; **22**(**5**): 409–429

Multivariate ordinal random effects models including subject and group specific response style effects

Gunther Schauberger[1,2] and Gerhard Tutz[2]
[1]Chair of Epidemiology, Department of Sport and Health Sciences, Technical University of Munich, Germany
[2]Department of Statistics, LMU Munich, Munich, Germany

Abstract: Common random effects models for repeated measurements account for the heterogeneity in the population by including subject-specific intercepts or variable effects. They do not account for the heterogeneity in answering tendencies. For ordinal responses in particular, the tendency to choose extreme or middle responses can vary in the population. Extended models are proposed that account for this type of heterogeneity. Location effects as well as the tendency to extreme or middle responses are modelled as functions of explanatory variables. It is demonstrated that ignoring response styles may affect the accuracy of parameter estimates. An example demonstrates the applicability of the method.

Key words: adjacent categories model, cumulative model, multivariate ordinal response, random effects models, response styles

Received 12 June 2019; **revised** 5 October 2020; **accepted** 12 October 2020

1 Introduction

Random effects models are a common tool for the modelling of heterogeneity and dependence of responses in repeated measurements and clustered data. Binary and ordinal random effects models account for heterogeneity by including subject-or cluster-specific effects in the linear predictor. The most widely used model is the random intercept model which allows the overall tendency to positive responses to vary over subjects. In more general models also slopes can depend on the subject. However, there is one subject-specific trait that is usually ignored, the tendency to select extreme or middle categories. This tendency can be considered as a response style, which is a consistent pattern of responses that is independent of the content of a response (Johnson, 2003). The problem with response styles is that they can affect the validity of parameter estimates because estimates may be biased if the response style is ignored. Therefore, models should account for the response style to avoid misleading estimates. Response styles have been investigated in particular in the survey literature, but the phenomenon occurs whenever subjects give ratings on an ordinal scale.

Address for correspondence: Gunther Schauberger, Department of Sport and Health Sciences, Technical University of Munich, Georg-Brauchle-Ring 56, 80992 München, Germany.
E-mail: gunther.schauberger@tum.de

 10.1177/1471082X20978034

Further examples could be if a medical diagnosis refers to categories such as normal, borderline and abnormal or when patients rate their pain on a five-point scale. The phenomenon is found in all assessed ordinal categorical variables, which arise when an assessor processes an unknown amount of information, leading to the judgement of the grade of the ordered categorical scale (Anderson, 1984).

Response styles have been investigated in particular in the social sciences, see, for example, Messick (1991), Baumgartner and Steenkamp (2001), Marin et al. (1992) and Meisenberg and Williams (2008), and in latent trait modelling in psychometrics, see Bolt and Newton (2011), Johnson (2003), Eid and Rauber (2000), Böckenholt (2017) and Tutz et al. (2018). An overview was given by Van Vaerenbergh and Thomas (2013).

Binary and ordinal random effects models without the modelling of response style are well established tools. In particular, the case of binary response variables has been investigated thoroughly, see, for example, Hedeker and Gibbons (1994), Hinde (1982), Anderson and Aitkin (1985), Liu and Pierce (1994), Pinheiro and Bates (1995) and Booth and Hobert (1999). Also for ordinal responses, several modelling strategies and estimation methods have been proposed. Harville and Mee (1984), Jansen (1990) and Tutz and Hennevogl (1996) considered cumulative type random effects models; adjacent categories type models were considered by Hartzel et al. (2001).

Models that explicitly account for response styles are able to reduce the bias. They account for additional heterogeneity in the population, which in some applications is itself of interest, in particular if it is linked to explanatory variables. In this article it is argued that simultaneous modelling of response styles is recommended if response styles are present, more concisely, if individuals have a tendency to prefer middle or extreme categories. According to the classification of different types of response style by Van Vaerenbergh and Thomas (2013), these tendencies refer to a continuum between the so-called mid-point response style (MRS) and extreme response style (ERS). If response styles that vary over individuals are present but ignored, the accuracy of parameter estimates may suffer. Moreover, the models, that are proposed allow researchers to investigate which variables determine the preference for middle or extreme categories. It should be noted that methods that account for response tendencies have been used before in regression for univariate ordinal responses, see Tutz and Berger (2016, 2017). However, in univariate ordinal models, response styles are not modelled as a subject-specific trait but as a tendency that is determined by covariates only. Consequently, estimation methods that are used for univariate ordinal responses are quite different from the methods used here, where response style is meant to represent a consistent pattern of responses independent of the content. In a multivariate setting, response styles were incorporated by Tutz et al. (2018) within the partial credit model. However, their model lacks the possibility to include covariates.

The article is structured as follows. In Section 2, we introduce multivariate ordinal response models. In Section 3, we develop an approach to include response styles into multivariate ordinal response models, while in Section 4, we give further details of the estimation procedure. Section 5, illustrates the performance of the proposed method using a simulation study. Finally, in Section 6, the method is applied to data from the German Longitudinal Election Study (GLES).

2 Models for multivariate ordinal responses

Regression models for a single ordinal response form the constituents of multivariate models. Therefore, we first briefly consider models that are in common use for ordinal responses and then consider extensions to multivariate responses.

2.1 Ordinal response models

Interesting models for ordinal responses are in particular the cumulative model and the adjacent-categories model. With the response Y_i taking values from $\{1, \ldots, k\}$, the *cumulative model* has the form

$$P(Y_i \leq r|\boldsymbol{x}_i) = F(\gamma_{0r} + \boldsymbol{x}_i^T \boldsymbol{\gamma}),$$

where $F(.)$ is a distribution function, $\boldsymbol{x}_i^T = (x_{i1}, \ldots, x_{ip})$ is a covariate vector with $\boldsymbol{\gamma}$ as the corresponding coefficient vector, and $-\infty = \gamma_{00} \leq \gamma_{01} \leq \cdots \leq \gamma_{0k} = \infty$ are category-specific intercepts. The model is usually derived from an underlying regression model. Let $\tilde{Y}_i = -\boldsymbol{x}_i^T \boldsymbol{\gamma} + \varepsilon_i$ be a regression model which contains the latent response variable $\tilde{Y}_i$ and a noise variable ε_i, which has cumulative distribution function $F(.)$. It is assumed that one cannot observe the variable $\tilde{Y}_i$ but a coarser categorized version; more precisely, one observes $Y_i = r$ if $\gamma_{0,r-1} < \tilde{Y}_i \leq \gamma_{0r}$. It is easily seen that one obtains the cumulative model. The intercepts $\gamma_{01}, \ldots, \gamma_{0k}$ are the thresholds on the latent continuum that determine the categorization. Thus, they have to be ordered.

Among the class of cumulative models, the most widely used model is the *proportional odds model* which uses the logistic distribution $F(\eta) = \exp(\eta)/(1 + \exp(\eta))$ yielding

$$\text{logit}\, P(Y_i \leq r|\boldsymbol{x}_i) = \gamma_{0r} + \boldsymbol{x}_i^T \boldsymbol{\gamma}.$$

With $\psi_{rj}(x) = P(Y_i \leq r|(x_1, \ldots, x_j = x, \ldots, x_p)/P(Y_i > r|(x_1, \ldots, x_j = x, \ldots, x_p))$ denoting the cumulative odds when the jth variable has value x_j (all others kept fixed), one obtains for the parameters the simple form

$$\mathrm{e}^{\gamma_j} = \psi_{rj}(x+1)/\psi_{rj}(x),$$

that is, e^{γ_j} is the change in cumulative odds if x_j increases by one unit, and the change does not depend on the category, which makes it a proportional odds model.

An alternative model is the *adjacent-categories model*, which has the form

$$P(Y_i = r+1|Y_i \in \{r, r+1\}, \boldsymbol{x}_i) = F(\gamma_{0r} + \boldsymbol{x}_i^T \boldsymbol{\gamma}),$$

where $F(.)$ is again a distribution function. For the logistic distribution function, one obtains

$$\log\left(\frac{P(Y_i = r+1)}{P(Y_i = r)}\right) = \gamma_{0r} + \boldsymbol{x}_i^T \boldsymbol{\gamma}.$$

With $\phi_{rj}(x) = P(Y_i = r+1|(x_1, \ldots, x_j = x, \ldots, x_p))/P(Y_i = r|(x_1, \ldots, x_j = x, \ldots, x_p))$ denoting the local odds that compare categories $r, r+1$, one obtains

$$e^{\gamma_j} = \phi_{rj}(x+1)/\phi_{rj}(x),$$

which is the change in local odds if x_j increases by one unit. As in the proportional odds model, the increase does not depend on r.

Also the so-called sequential model is used in ordinal regression but is less used in multivariate settings. For an overview on ordinal regression see Agresti (2009), Tutz (2012) and Peyhardi et al. (2015).

All of these models can be given in matrix form as multivariate generalized linear models (GLMs) for categorical responses. For observation i, one obtains

$$g(\boldsymbol{\pi}_i) = \boldsymbol{X}_i\boldsymbol{\beta} \quad \text{or} \quad \boldsymbol{\pi}_i = h(\boldsymbol{X}_i\boldsymbol{\beta}),$$

where $\boldsymbol{\pi}_i^T = (\pi_{i1}, \ldots, \pi_{iq})$, $\pi_{ir} = P(Y_i = r|\boldsymbol{x}_i)$, is the vector of response probabilities of length $q = k-1$, g is the (multivariate) link function, $h = g^{-1}$ is the inverse link function and $\boldsymbol{X}_i$ is the corresponding design matrix with components of $\boldsymbol{\eta}_i = \boldsymbol{X}_i\boldsymbol{\beta}$ having the form

$$\eta_{ir} = \gamma_{0r} + \boldsymbol{x}_i^T\boldsymbol{\gamma} = (0, \ldots, 0, 1, 0, \ldots, 0, \boldsymbol{x}_i^T)\boldsymbol{\beta}, \quad r = 1, \ldots, q,$$

where $\boldsymbol{\beta}^T = (\gamma_{01}, \ldots, \gamma_{0q}, \boldsymbol{\gamma}^T)$, see Fahrmeir and Tutz (2001) and Tutz (2012).

2.2 Random effects models

Random effects models aim at explicitly modelling the heterogeneity of clustered responses. A cluster can be any statistical unit for which repeated measurements are available. In our applications, a cluster typically refers to a person and repeated measurements refer to responses on a set of items. For such clustered data, let the ordinal response $Y_{it} \in \{1, \ldots, k\}$ denote measurement t in cluster i, $i = 1, \ldots, n$, $t = 1, \ldots, T_i$. In random effects models, one assumes that the corresponding model for observation Y_{it} has the form

$$g(\boldsymbol{\pi}_{it}) = \boldsymbol{X}_{it}\boldsymbol{\beta} + \boldsymbol{Z}_{it}\boldsymbol{b}_i, \tag{2.1}$$

where $\boldsymbol{\pi}_{it}^T = (\pi_{it1}, \ldots, \pi_{itq})$ denotes the vector of response probabilities with $\pi_{itr} = P(Y_{it} = r|\boldsymbol{X}_{it}, \boldsymbol{Z}_{it}, \boldsymbol{b}_i)$, $\boldsymbol{X}_{it}$ and $\boldsymbol{Z}_{it}$ are design matrices linked to the ith cluster, $\boldsymbol{\beta}$ is a fixed coefficient vector, and $\boldsymbol{b}_i$ is an s-dimensional cluster or subject-specific random effect, for which a distribution, for example $\boldsymbol{b}_i \sim N(\boldsymbol{0}, \boldsymbol{Q})$, is assumed. The corresponding linear predictor $\boldsymbol{\eta}_{it} = \boldsymbol{X}_{it}\boldsymbol{\beta} + \boldsymbol{Z}_{it}\boldsymbol{b}_i$ has components

$$\eta_{itr} = \gamma_{0tr} + \boldsymbol{x}_{it}^T\boldsymbol{\gamma} + \boldsymbol{z}_{it}^T\boldsymbol{b}_i,$$

where $\boldsymbol{x}_{it}^T = (x_{it1}, \ldots, x_{itp})$ is the covariate vector associated with fixed effects and $\boldsymbol{z}_{it}^T = (z_{it1}, \ldots, z_{its})$ the covariate vector associated with random effects.

The simplest random effects model is a model that includes random intercepts only. It has the linear predictor

$$\eta_{itr} = \gamma_{0tr} + \boldsymbol{x}_{it}^T\boldsymbol{\gamma} + b_i,$$

where $b_i \sim N(0, \sigma^2)$. Thus, each respondent is assumed to have its own response level. The random effect model accounts for the association between responses and allows to investigate the heterogeneity of individuals captured in the variance of the random effect. It models the response to items given the covariates and the individual random effects and therefore the response behaviour on the level of the individual, in contrast to marginal models, which are population averaged models.

3 Accounting for response styles

3.1 Including response styles

In both multivariate ordinal models, the cumulative and the adjacent categories model, the intercepts $\gamma_{0t1} \ldots \gamma_{0tq},\ q = k - 1$, can be seen as thresholds. As has been shown, the cumulative model can even be derived from a latent continuous variable which is observed in categories that are determined by thresholds on the latent continuum. Usually the thresholds are considered as less interesting nuisance parameters. They do not depend on covariates and just indicate the basic preference for specific categories. For example, in the cumulative model if $\gamma_{0t,r-1} = \gamma_{0tr}$, one obtains $P(Y_{it} = r) = 0$. If the difference between adjacent categories, $\gamma_{0tr} - \gamma_{0t,r-1}$ is large the probability of $Y_{it} = r$ will be comparatively large. This property of the thresholds, namely to determine the basic preference for specific categories, will be used in the following to model the subject-specific tendencies to choose specific categories.

Cumulative models

In the cumulative model, the ordering of thresholds, $\gamma_{0t1} \leq \cdots \leq \gamma_{0tq}$, is indispensable. That means the introduction of response styles has to retain the order. For this purpose, a reparameterization is necessary.

Let first the number of categories k be *even*. The model with random effects has the predictor $\eta_{itr} = \gamma_{0tr} + \boldsymbol{x}_{it}^T\boldsymbol{\gamma} + \boldsymbol{z}_{it}^T\boldsymbol{b}_i$. We use for the thresholds γ_{0tr} a reparameterization that is centred at the middle. With $m = k/2$, the threshold γ_{0tm} distinguishes between the categories $\{1, \ldots, m\}$ and $\{m + 1, \ldots, k\}$. This is the only threshold that is kept.

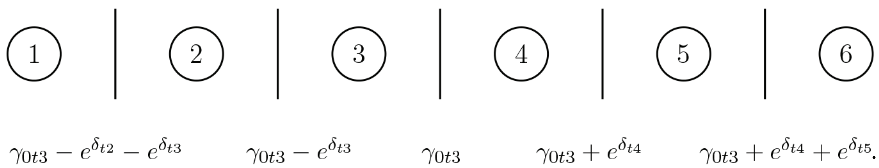
All others are reparameterized by using

$$\gamma_{0tr} = \gamma_{0tm} + \sum_{j=m+1}^{r} \exp(\delta_{tj}),\ r > m, \quad \gamma_{0tr} = \gamma_{0tm} - \sum_{j=r+1}^{m} \exp(\delta_{tj}),\ r < m.$$

The parameters $e^{\delta_{tr}}$ represent the differences between thresholds, which is seen from $\gamma_{0tr} - \gamma_{0t,r-1} = e^{\delta_{tr}}$. That means, in particular, that all differences between adjacent categories are positive. For illustration, let us consider the case $k = 6$, for which one obtains

$$\gamma_{0t1} = \gamma_{0t3} - e^{\delta_{t2}} - e^{\delta_{t3}},\ \gamma_{0t2} = \gamma_{0t3} - e^{\delta_{t3}},\ \gamma_{0t4} = \gamma_{0t3} + e^{\delta_{t4}},\ \gamma_{0t5} = \gamma_{0t3} + e^{\delta_{t4}} + e^{\delta_{t5}},$$

which yields the structure



The new threshold parameters are given by $\gamma_{0t3}, \delta_{t2}, \delta_{t3}, \delta_{t4}, \delta_{t5}$. In general, the new parameters, which can take any value without any restriction, are given by $\gamma_{0tm}, \delta_{t2}, \ldots, \delta_{tq}$. The 'middle' threshold γ_{0tm} is kept but all others are reparameterized by using the parameters $\delta_{t2}, \ldots, \delta_{tq}$.

Let a_i be a subject-specific response style parameter. It is included in the predictor by substituting $\delta_{tr} + a_i$ for δ_{tr}. Then the difference between adjacent linear predictors is no longer $e^{\delta_{tr}}$ but

$$\eta_{itr} - \eta_{it,r-1} = e^{\delta_{tr}}\, e^{a_i}\,. \tag{3.1}$$

If $a_i \to -\infty$, the predictors coincide and the probability mass is concentrated in the extreme response categories 1 and k. If a_i is large, all the differences become large and the probability mass is concentrated in the middle categories m and $m + 1$. It is straightforward to derive that for $a_i \to \infty$, the probabilities of categories m and $m + 1$ sum up to one. In summary, the subject-specific parameter a_i modifies the thresholds such that it accounts for a respondents tendency to middle or extreme categories. Figure 1 shows the effect of the response style parameter on the probabilities. Positive values of a_i increase the probabilities of the middle categories, while negative values increase the probabilities of the extreme categories.

For *k odd*, the reparameterization has a slightly different form. Then $m = (k + 1)/2$ denotes the middle category, which is between thresholds $\gamma_{0t,m-1}$ and $\gamma_{0t,m}$. Centring

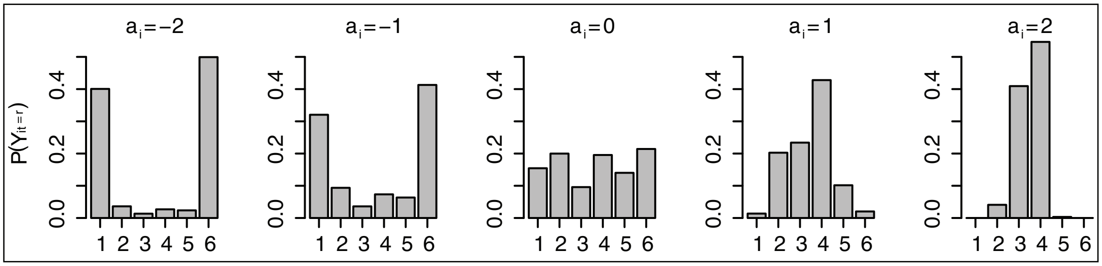


Figure 1 Probabilities of single categories (for an example with $k = 6$) depending on different values of response style parameter a_i for the cumulative model.

around the middle category is obtained by using the reparameterization

$$\gamma_{0t,m-1} = \gamma_{0t} - \exp(\delta_{tm})/2, \quad \gamma_{0t,m} = \gamma_{0t} + \exp(\delta_{tm})/2,$$

$$\gamma_{0tr} = \gamma_{0tm} + \sum_{j=m+1}^{r} \exp(\delta_{tj}), r > m, \quad \gamma_{0tr} = \gamma_{0t,m-1} - \sum_{j=r+1}^{m-1} \exp(\delta_{tj}), r < m-1.$$

One obtains again $\eta_{itr} - \eta_{it,r-1} = e^{\delta_{tr}}$. After substituting $\delta_{tr} + a_i$ for δ_{tr}, one obtains $\eta_{itr} - \eta_{it,r-1} = e^{\delta_{tr}} e^{a_i}$ as in the case where k is even. Now the parameters are given by $\gamma_{0t}, \delta_{t2}, \ldots, \delta_{tq}$. For $a_i \to \infty$, the probability of the middle category m becomes one; if $a_i \to -\infty$, the predictors coincide and the probability mass is concentrated in the extreme response categories 1 and k.

The parameter a_i is a subject-specific effect that models the individual tendency to extreme or middle categories. It is important that the response style effects a_i vary over individuals but not over the repeated measurements. When response styles are understood as a consistent pattern of responses that is independent of the content of a response, the parameters should be the same in all the items or measurements that are answered or taken on one individual.

Adjacent categories model

In the adjacent categories model, the intercept γ_{0tr} determines which of the categories $r, r+1$ is preferred. It is seen from the model

$$P(Y_{it} = r+1 | Y_{it} \in \{r, r+1\}, \boldsymbol{x}_{it}) = F(\gamma_{0tr} + \boldsymbol{x}_{it}^T \boldsymbol{\gamma} + b_i), \quad r = 1, \ldots, k-1$$

that γ_{0tr} determines the basic level of a response in category $r+1$ given $Y_{it} \in \{r, r+1\}$. The parameters can again be seen as thresholds, however, they are not necessarily ordered. It is not as obvious as in the cumulative model but also for the adjacent categories model, the difference between adjacent thresholds determines the probability of observing a specific categorical value.

Therefore, for adjacent categories models also, response style effects can be included by modifying the thresholds. The basic idea is to increase or decrease the

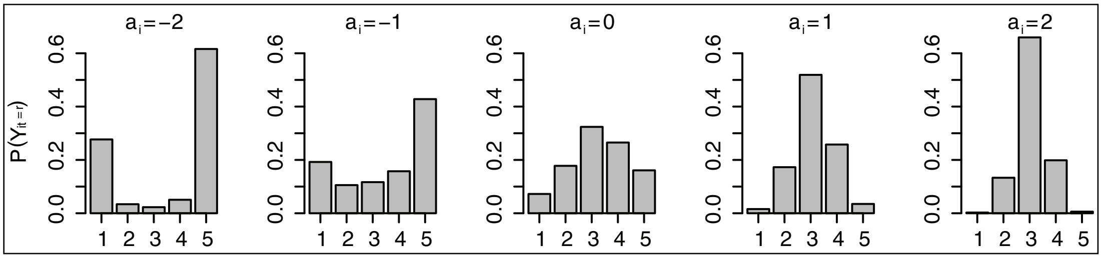


Figure 2 Probabilities of single categories (for an example with $k = 5$) depending on different values of response style parameter a_i in the adjacent categories model.

difference between thresholds with a centring at the middle category (see also Tutz and Berger (2016)). In the predictor $\eta_{itr} = \gamma_{0tr} + \boldsymbol{x}_{it}^T\boldsymbol{\gamma} + b_i$ for the tth variable, we propose to replace the threshold γ_{0tr} by

$$\gamma_{0tr} + (k/2 - r)a_i,$$

where a_i is a subject-specific parameter. It is seen that the difference between adjacent predictors has the form

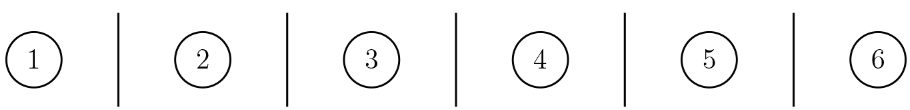
$$\eta_{itr} - \eta_{it,r-1} = \gamma_{0tr} - \gamma_{0t,r-1} - a_i. \tag{3.2}$$

The weights $(k/2 - r)$ are chosen such that the difference between predictors changes by a_i. If a_i is positive the difference decreases, if it is negative the difference increases with the introduction of the subject-specific parameter. Figure 2 illustrates how different values of the parameter a_i affect the probabilities of the single response categories. Positive values of a_i increase the probabilities of the middle categories, while negative values increase the probabilities of the extreme categories.

For illustration, let us consider two specific cases, both with an odd ($k = 5$) and an even ($k = 6$) number of categories. For $k = 5$, one obtains the following thresholds

$$\gamma_{0t1} + 1.5a_i \quad \gamma_{0t2} + 0.5a_i \quad \gamma_{0t3} - 0.5a_i \quad \gamma_{0t4} - 1.5a_i.$$

For $k = 6$, one obtains the thresholds



$$\gamma_{0t1} + 2a_i \quad \gamma_{0t2} + a_i \quad \gamma_{0t3} \quad \gamma_{0t4} - a_i \quad \gamma_{0t5} - 2a_i.$$

It should be noted that the modification of thresholds differs for the two types of models. In the cumulative model, the thresholds have to be ordered, therefore the subject-specific parameter modifies the difference between adjacent thresholds in a way that retains the order. The modification of differences between thresholds is multiplicative as is seen from equation (3.1). In the adjacent categories model, thresholds do not need to be ordered. Therefore, one can use a parameterization that changes the differences between thresholds in an additive way, see equation (3.2). The modification by adding subject-specific parameters has the advantage that one remains in the GLM framework, however, it can not be used for cumulative models. For cumulative models, an additive term a_i would destroy the ordering of thresholds if it were assumed to be a random effect following a normal distribution.

3.2 Including covariates that determine the response style

In the general model proposed here, the effect of explanatory variables is included by letting the parameters a_i depend on them. Covariates that determine the response style can be included by using the response style term $a_i + \boldsymbol{z}_{it}^T\boldsymbol{\alpha}$ instead of the simple subject-specific effect a_i. Then, in the adjacent categories representation, one obtains

$$\eta_{itr} = \gamma_{0tr} + (k/2 - r)(a_i + \boldsymbol{z}_{it}^T\boldsymbol{\alpha}) + \boldsymbol{x}_{it}^T\boldsymbol{\gamma} + b_i. \tag{3.3}$$

In the cumulative model, one obtains (for k even)

$$\eta_{itr} = \gamma_{0tm} + \sum_{j=m+1}^{r} \mathrm{e}^{\delta_{tj}}\,\mathrm{e}^{a_i+\boldsymbol{z}_{it}^T\boldsymbol{\alpha}} + \boldsymbol{x}_{it}^T\boldsymbol{\gamma} + b_i, \quad r > m,$$
$$\eta_{itr} = \gamma_{0tm} - \sum_{j=r+1}^{m} \mathrm{e}^{\delta_{tj}}\,\mathrm{e}^{a_i+\boldsymbol{z}_{it}^T\boldsymbol{\alpha}} + \boldsymbol{x}_{it}^T\boldsymbol{\gamma} + b_i, \quad r < m.$$

The linear predictor is constructed by modifying the thresholds, which corresponds to the increase or decrease of differences between adjacent thresholds. The concept works for the cumulative and the adjacent-categories model. The corresponding models contain two effects, the location effect captured in $\boldsymbol{x}_{it}^T\boldsymbol{\gamma}$ and the response style effect contained in $\boldsymbol{z}_{it}^T\boldsymbol{\alpha}$. The variables $\boldsymbol{x}_{it}$ and $\boldsymbol{z}_{it}$ can be distinct, overlapping or identical.

For $\boldsymbol{z}_{it} = \boldsymbol{x}_{it}$, it is instructive to investigate the structure of the linear predictor to make the link to more general ordinal models. With $\boldsymbol{z}_{it} = \boldsymbol{x}_{it}$, the predictor (3.3) can be rewritten as

$$\begin{aligned}\eta_{itr} &= (k/2 - r)a_i + b_i + \gamma_{0tr} + \boldsymbol{x}_{it}^T((k/2 - r)\boldsymbol{\alpha} + \boldsymbol{\gamma})\\ &= (k/2 - r)a_i + b_i + \gamma_{0tr} + \boldsymbol{x}_{it}^T\boldsymbol{\beta}_r,\end{aligned}$$

where $\boldsymbol{\beta}_r = (k/2 - r)\boldsymbol{\alpha} + \boldsymbol{\gamma}$, $r = 1, \ldots, k-1$. Thus, one has an ordinal model which allows covariate effects to vary across categories, constraining the parameters such that a parameterization $\boldsymbol{\beta}_r = (k/2 - r)\boldsymbol{\alpha} + \boldsymbol{\gamma}$ exists. In the cumulative model, the model considered here cannot be seen as a sub-model of the general model with category-specific effects because the predictor has a more complicated form to ensure that thresholds are ordered.

4 Estimation of the random effects model

Let the data be given by $(Y_{it}, \boldsymbol{x}_{it}, \boldsymbol{z}_{it})$, $i = 1, \ldots, n$, $t = 1, \ldots, T_i$, where $\boldsymbol{x}_{it}$ denotes the vector of explanatory variables that determine the location and $\boldsymbol{z}_{it}$ denotes the explanatory variables that determine the response style.

The model contains two random effects, the subject-specific intercept b_i and the subject-specific response style effect a_i. For them, a normal distribution $(b_i, a_i) \sim N(\mathbf{0}, \boldsymbol{\Sigma})$ is assumed. Maximization of the marginal log-likelihood can be obtained by integration techniques. The marginal likelihood is

$$L(\boldsymbol{\beta}, \boldsymbol{\Sigma}) = \prod_{i=1}^{n} \int \int P(\{Y_{i1}, \ldots, Y_{iT_i}\}) f(b_i, a_i) db_i da_i,$$

where $f(b_i, a_i)$ now denotes the two-dimensional density of the person parameters, and $\boldsymbol{\beta}$ collects all fixed parameters. The diagonals of the matrix $\boldsymbol{\Sigma}$ contain the variance of random intercepts σ_b^2 and the response style parameters, σ_a^2, the off diagonals are the covariances between intercepts and the response style, cov_{ba}. The corresponding log-likelihood is

$$l(\boldsymbol{\beta}, \boldsymbol{\Sigma}) = \sum_{i=1}^{n} \log \left(\int \int \prod_{t=1}^{T_i} \prod_{r=1}^{k} \{P(Y_{it} = r|\boldsymbol{\beta}, b_i, a_i)\}^{y_{itr}} f(b_i, a_i) db_i \, da_i \right),$$

where $y_{itr} = 1$ if $Y_{it} = r$ and $y_{itr} = 0$ otherwise.

For the adjacent categories model, one has

$$P(Y_{it} = r|\boldsymbol{\beta}, b_i, a_i) = \frac{\exp(\sum_{l=1}^{r-1} \{\gamma_{0tl} + (k/2 - l)(a_i + \boldsymbol{x}_{it}^T \boldsymbol{\alpha}) + \boldsymbol{x}_{it}^T \boldsymbol{\gamma} + b_i\})}{\sum_{s=1}^{k} \exp(\sum_{l=1}^{s-1} \{\gamma_{0tl} + (k/2 - l)(a_i + \boldsymbol{x}_{it}^T \boldsymbol{\alpha}) + \boldsymbol{x}_{it}^T \boldsymbol{\gamma} + b_i\})},$$

for the cumulative model, one has to build the differences $P(Y_{it} = r|\boldsymbol{\beta}, b_i, a_i) = P(Y_{it} \leq r|\boldsymbol{\beta}, b_i, a_i) - P(Y_{it} \leq r-1|\boldsymbol{\beta}, b_i, a_i)$. For the adjacent categories model, the embedding into the framework of generalized mixed models allows to use methods that have been developed for this class of models. One strategy is to use joint maximization of a penalized log-likelihood with respect to parameters and random effects appended by estimation of the variance of random effects, see Breslow and Clayton (1993) and McCulloch and Searle (2001). However, joint maximization algorithms tend to underestimate the variances and, therefore, the true values of the random effects. An alternative strategy, which is used here, is numerical integration by Gauss–Hermite integration methods. Early versions for univariate random effects date back to Hinde (1982) and Anderson and Aitkin (1985). For an overview on estimation methods for generalized mixed model see McCulloch and Searle (2001) and Tutz (2012). For the cumulative model one cannot rely on generalized mixed models, we used a modified Gauss–Hermite procedure.

5 Simulation

A small simulation study is conducted to evaluate the performance of the method and the possible consequences of ignoring the response style. Before presenting the results, we first describe the general settings and the parameters that are used in the simulation scenarios.

5.1 Simulation settings

Per simulation, $n = 300$ observations on $T = 10$ items are used with each response variable having $k = 5$ categories. The data are simulated under the assumption that the models proposed in Section 3 hold. Therefore, separate simulations were performed for the cumulative model and for the adjacent categories model. For the adjacent categories model, the linear predictor in the data generating process (DGP) is defined as

$$\eta_{itr} = \gamma_{0tr} + (k/2 - r)(a_i + \boldsymbol{z}_{it}^T\boldsymbol{\alpha}) + \boldsymbol{x}_{it}^T\boldsymbol{\gamma} + b_i, \tag{5.1}$$

while for the cumulative model the linear predictor in the DGP is defined as

$$\begin{aligned} \eta_{itr} &= \gamma_{0t} + (\mathrm{e}^{\delta_{tm}}/2 + \sum_{j=m+1}^{r} \mathrm{e}^{\delta_{tj}})\, \mathrm{e}^{a_i+\boldsymbol{z}_{it}^T\boldsymbol{\alpha}} + \boldsymbol{x}_{it}^T\boldsymbol{\gamma} + b_i, \quad r \geq m, \\ \eta_{itr} &= \gamma_{0t} - (\mathrm{e}^{\delta_{tm}}/2 - \sum_{j=r+1}^{m-1} \mathrm{e}^{\delta_{tj}})\, \mathrm{e}^{a_i+\boldsymbol{z}_{it}^T\boldsymbol{\alpha}} + \boldsymbol{x}_{it}^T\boldsymbol{\gamma} + b_i, \quad r < m. \end{aligned} \tag{5.2}$$

The explanatory variables contained in $\boldsymbol{z}_{it}$ and $\boldsymbol{x}_{it}$ are equal. We use two continuous variables x_1 and x_2 drawn from a standard normal distribution and one binary variable x_3 drawn from a Bernoulli distribution with probability parameter $p = 0.5$. The location covariate effects are fixed to $\boldsymbol{\gamma} = (0.5, -0.2, -0.3)^T$.

The threshold parameters are set to fixed values. For the adjacent categories model, the matrix of the true threshold values is

$$\begin{pmatrix} \boldsymbol{\gamma}_{0\cdot 1} \\ \boldsymbol{\gamma}_{0\cdot 2} \\ \boldsymbol{\gamma}_{0\cdot 3} \\ \boldsymbol{\gamma}_{0\cdot 4} \end{pmatrix} = \begin{pmatrix} -0.84 & -0.80 & -1.26 & -0.16 & -1.07 & -0.60 & -1.01 & -2.18 & -1.76 & -0.85 \\ 0.32 & 0.94 & -1.08 & 0.07 & 0.82 & -0.29 & -0.60 & -2.00 & -0.66 & -0.26 \\ 0.90 & 1.11 & 1.47 & 0.71 & 1.48 & -0.14 & -0.47 & -0.64 & -0.29 & -0.14 \\ 1.23 & 1.38 & 2.22 & 1.22 & 1.71 & 0.95 & 1.42 & 1.50 & 0.24 & 0.14 \end{pmatrix}$$

while for the cumulative model the threshold parameters are

$$\begin{pmatrix} \boldsymbol{\gamma}_{0\cdot} \\ \boldsymbol{\delta}_{0\cdot 2} \\ \boldsymbol{\delta}_{0\cdot 3} \\ \boldsymbol{\delta}_{0\cdot 4} \end{pmatrix} = \begin{pmatrix} 0.61 & 1.02 & 0.19 & 0.39 & 1.15 & -0.21 & -0.53 & -1.32 & -0.47 & -0.20 \\ 0.15 & 0.55 & -1.71 & -1.47 & 0.64 & -1.17 & -0.89 & -1.71 & 0.10 & -0.53 \\ -0.54 & -1.77 & 0.94 & -0.45 & -0.42 & -1.90 & -2.04 & 0.31 & -0.99 & -2.12 \\ -1.11 & -1.31 & -0.29 & -0.67 & -1.47 & 0.09 & 0.64 & 0.76 & -0.63 & -1.27 \end{pmatrix}.$$

Two parameters are varied, namely the standard deviation of the random response style effect and the fixed response style effects corresponding to the explanatory variables. For the covariance matrix of the random effects $(b_i, a_i) \sim N(\mathbf{0}, \boldsymbol{\Sigma})$, we use

$$\boldsymbol{\Sigma} = \begin{pmatrix} 1 & 0 \\ 0 & \sigma_a^2 \end{pmatrix}$$

with three different effect strengths

$$\sigma_a = 0 \textit{ (Null)}, \quad \sigma_a = 0.5 \textit{ (Medium)}, \quad \sigma_a = 1 \textit{ (Strong)}.$$

For the response style effects $\boldsymbol{\alpha}$, we use $\boldsymbol{\alpha} = (0, 0, 0)^T$ *(Null)*, $\boldsymbol{\alpha} = (0.25, 0.125, -0.25)^T$ *(Medium)* and $\boldsymbol{\alpha} = (0.5, 0.25, -0.5)^T$ *(Strong)*, respectively.

In total, these different values constitute nine different settings for the cumulative model and the adjacent categories model, respectively. These nine settings indicate different strengths of response styles, both for the response style caused by explanatory variables and caused by the random effects. Per setting, 100 replications are conducted. Each dataset is analysed with two different models: *Model 1* (no RS) is a simple model without response style effects where the linear predictor is specified as $\eta_{itr} = \gamma_{0tr} + \boldsymbol{x}_{it}^T\boldsymbol{\gamma} + b_i$, while *model 2* (with RS) corresponds to the proposed model containing response style effects with the linear predictor η_{itr} equal to the linear predictor in the DGP as described in (5.1) and (5.2), respectively.

5.2 Simulation results

We are mainly interested in the fixed effects which are contained in both models, namely the threshold parameters γ_{0tr} and the location covariate effects $\boldsymbol{\gamma}$. Figures 3 (for the adjacent categories model) and 4 (for the cumulative model) display box plots of the mean absolute deviation (MAD), separately for these two types of parameters and for the single simulation settings.

We distinguish between *model 1* (*no RS*) and *model 2* (*with RS*) by using different colours. With growing strength of the response style caused by the random effect and by the explanatory variables, the MAD of the parameters is growing when fitting *model 1* while it remains rather constant for *model 2*. This effect is much stronger for the threshold parameters than for the covariate effects. This behaviour can probably be explained by the fact that the threshold parameters cover both the general location of the different responses as well as the general (i.e., not individual-specific) response styles across the studied population, while the covariate effects only cover location effects connected to $\boldsymbol{x}_{it}$. It is seen that ignoring individual response style effects can lead to poor estimates of the parameters in the model.

In the supplementary materials accompanying this manuscript, box plots for the estimates for each of the 40 threshold parameters can be found, separately for *models 1* and 2, for the nine simulation setting, and for the adjacent categories model and the cumulative model. For the sake of simplicity, within this manuscript in Figure 5, we only present the respective results for the most extreme setting (with strong response style effects both for the random effects and the covariate effects) in the adjacent categories model. It can be seen that in *model 2*, the parameters are estimated unbiased, while the estimates in *model 1* are severely biased in many cases.

6 Application to pre-election data

The method is applied to data from the GLES (Roßteutscher et al., 2017). The GLES is a long-term study of the German electoral process. It collects pre- and post-election data for several federal elections. The data we are using originate from the pre-election survey for the German federal election in 2017. In this specific part of the study, the participants were asked about specific political fears. More precisely, the participants were asked: 'How afraid are you due to the ...'

1. refugee crisis?
2. global climate change?
3. international terrorism?
4. globalization?
5. political developments in Turkey?
6. use of nuclear energy?

The answers were measured on Likert scales from 1 (not afraid at all) to 7 (very afraid). As explanatory variables in the model we used:

Abitur High School Diploma (1: Abitur/A levels; 0: else),
Age Age of the participant,
EastWest (1: East Germany/former GDR; 0: West Germany/former FRG),
Gender (1: female; 0: male),
Unemployment (1: currently unemployed; 0: else).

The age of the participants ranges from 15 years to 94 years. The variable EastWest refers to the current place of residence where all Berlin residents are

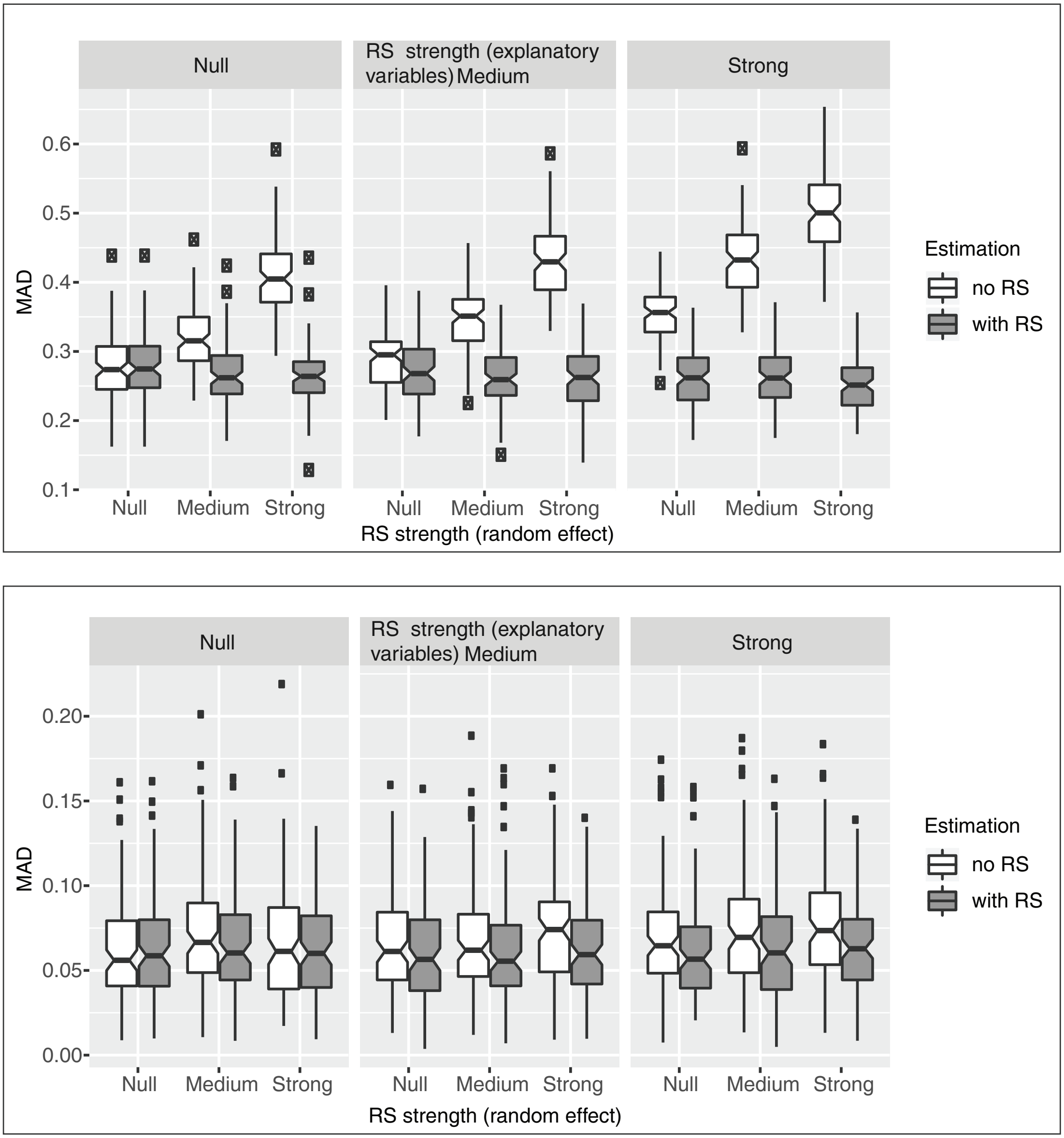
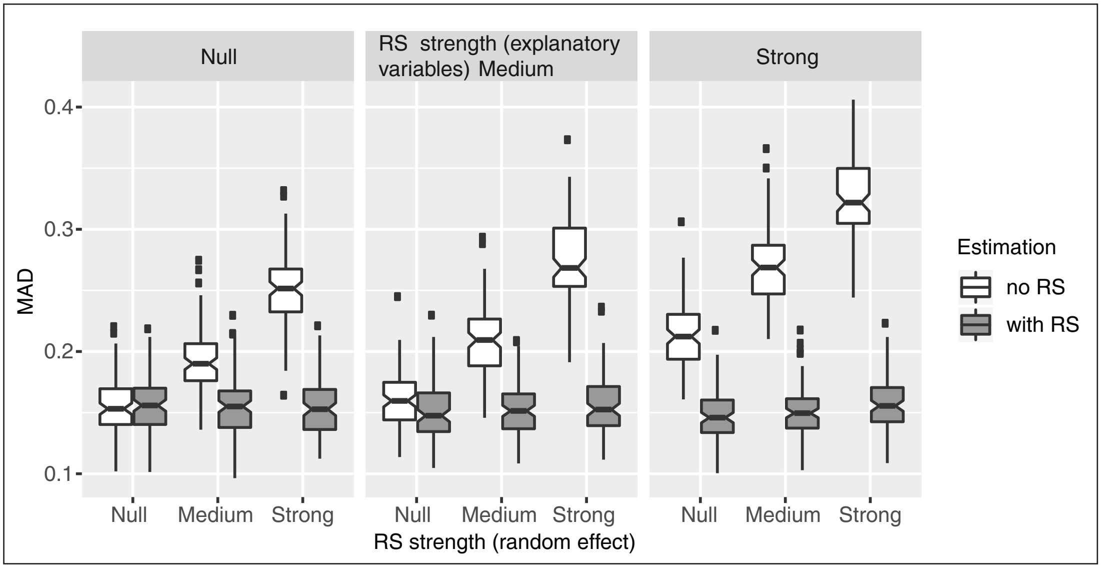


Figure 3 Mean absolute deviation (MAD) for threshold parameters γ_{0tr} (a) and for covariate parameters γ (b) depending on the response style strength of the random effect and of the explanatory variables, separately for estimation *with* (black) and *without* (grey) response style effects in the adjacent categories model.



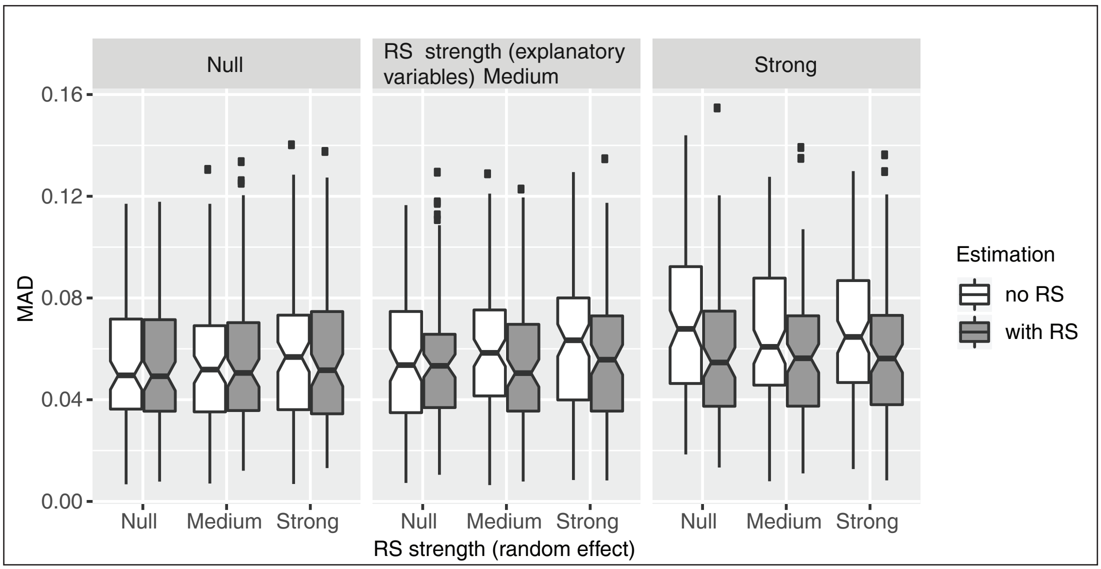


Figure 4 Mean absolute deviation (MAD) for threshold parameters γ_{0tr} (a) and for covariate parameters γ (b) depending on the response style strength of the random effect and of the explanatory variables, separately for estimation *with* (black) and *without* (grey) response style effects in the cumulative model.

assigned to East Germany. For our analysis, the original dataset consisting of 2 179 observations was reduced to the 2 036 complete observations. In order to make the effect sizes easier to compare all variables were standardized before the respective analyses.

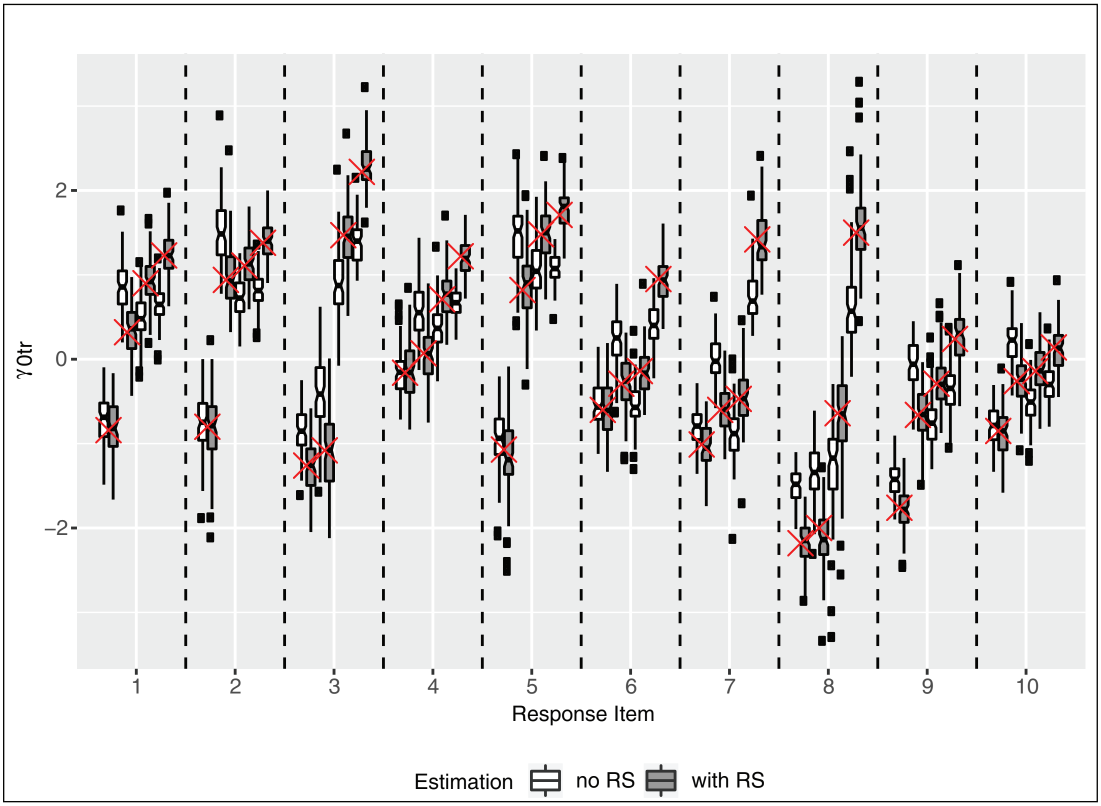


Figure 5 Parameter estimates separately for all threshold parameters in the simulation setting with strong response style effects both for the random effects and the covariate effects in the adjacent categories model. True values are marked with crosses.

Table 1 (Standardized) parameter estimates (and *p*-values) for all covariate effects from *models 1* and *2* (separately for the adjacent categories model and the cumulative model).

Adjacent categories model

	Age	Gender	Unemployment	EastWest	Abitur
Model 1: γ	0.129 (0.000)	0.123 (0.000)	−0.005 (0.686)	−0.013 (0.273)	−0.107 (0.000)
Model 2: γ	0.115 (0.000)	0.120 (0.000)	−0.008 (0.484)	−0.015 (0.196)	−0.100 (0.000)
Model 2: α	−0.066 (0.000)	−0.010 (0.303)	−0.013 (0.152)	−0.009 (0.330)	0.021 (0.026)

Cumulative model

	Age	Gender	Unemployment	EastWest	Abitur
Model 1: γ	−0.351 (0.000)	−0.318 (0.000)	0.015 (0.628)	0.029 (0.345)	0.286 (0.000)
Model 2: γ	−0.317 (0.000)	−0.333 (0.000)	0.026 (0.422)	0.048 (0.137)	0.278 (0.000)
Model 2: α	−0.066 (0.000)	0.010 (0.414)	−0.019 (0.126)	−0.021 (0.078)	0.013 (0.287)

We fitted two different models, in *model 1* possible response style effects (both the random effects and the covariate effects) are ignored, whereas they are included in *model 2*. For both models we used the adjacent categories approach as well as the cumulative modelling approach with logistic link. Table 1 shows the estimates of all covariate effects $\boldsymbol{\gamma}$ and $\boldsymbol{\alpha}$ (separately for the adjacent categories model and the cumulative model) for *models 1* and 2 together with their corresponding *p*–values. A striking difference between the adjacent categories model and the cumulative model is that the location parameters $\boldsymbol{\gamma}$ have reversed signs, which is simply an effect of the parameterization. While a positive parameter in the cumulative model means that the tendency to low response categories is increased (for increasing values of the variable), it means a tendency to higher response categories in the adjacent categories model. Otherwise, the two model types yield similar results.

Figure 6 gives a visualization of the estimated effects and confidence intervals (a similar visualization tool was proposed by Tutz and Berger (2016)). It shows the exponentials of the covariate effects both for the location effects γ (abscissa) and the response style effects α (ordinate) given in Table 1, together with the respective 95% confidence intervals, which determine the size of the stars. If the stars cross the no effects lines $\exp(\gamma) = 1$ and $\exp(\alpha) = 1$ the corresponding effects cannot be considered as significant.

It is immediately seen that location effects for Abitur, Age and Gender and the response style effect for Age are significant. In contrast to the cumulative model, in the adjacent categories model also Abitur has a significant response style effect. According to these estimates, the overall level of political fears is increased with increasing age and for women in comparison to men, while people with Abitur tend to have a lower level of fears than other respondents. On the other hand, with growing age people have an increasing tendency towards extreme categories, while people with Abitur show a tendency towards middle categories.

The random effects components are seen from the estimated (co-)variances for each of the models. For the adjacent categories models, we obtained $\hat{\sigma}_b^2 = 0.184$ (standard error $se(\hat{\sigma}_b^2) = 0.0110$) for *model 1* and

$$\hat{\boldsymbol{\Sigma}} = \begin{pmatrix} 0.166 & -0.002 \\ -0.002 & 0.077 \end{pmatrix}$$

for *model 2* with standard errors $se(\hat{\sigma}_b^2) = 0.0111$ and $se(\hat{\sigma}_a^2) = 0.00628$. For the cumulative models, we obtained $\hat{\sigma}_b^2 = 1.337$ (standard error $se(\hat{\sigma}_b^2) = 0.0694$) for *model 1* and

$$\hat{\boldsymbol{\Sigma}} = \begin{pmatrix} 1.276 & -0.185 \\ -0.185 & 0.159 \end{pmatrix}$$

for *model 2* with standard errors $se(\hat{\sigma}_b^2) = 0.0717$ and $se(\hat{\sigma}_a^2) = 0.0225$. In both cases, in *model 2* the variance of the random location effects is estimated to be slightly smaller than in *model 1*. There appears to be no strong correlation between the

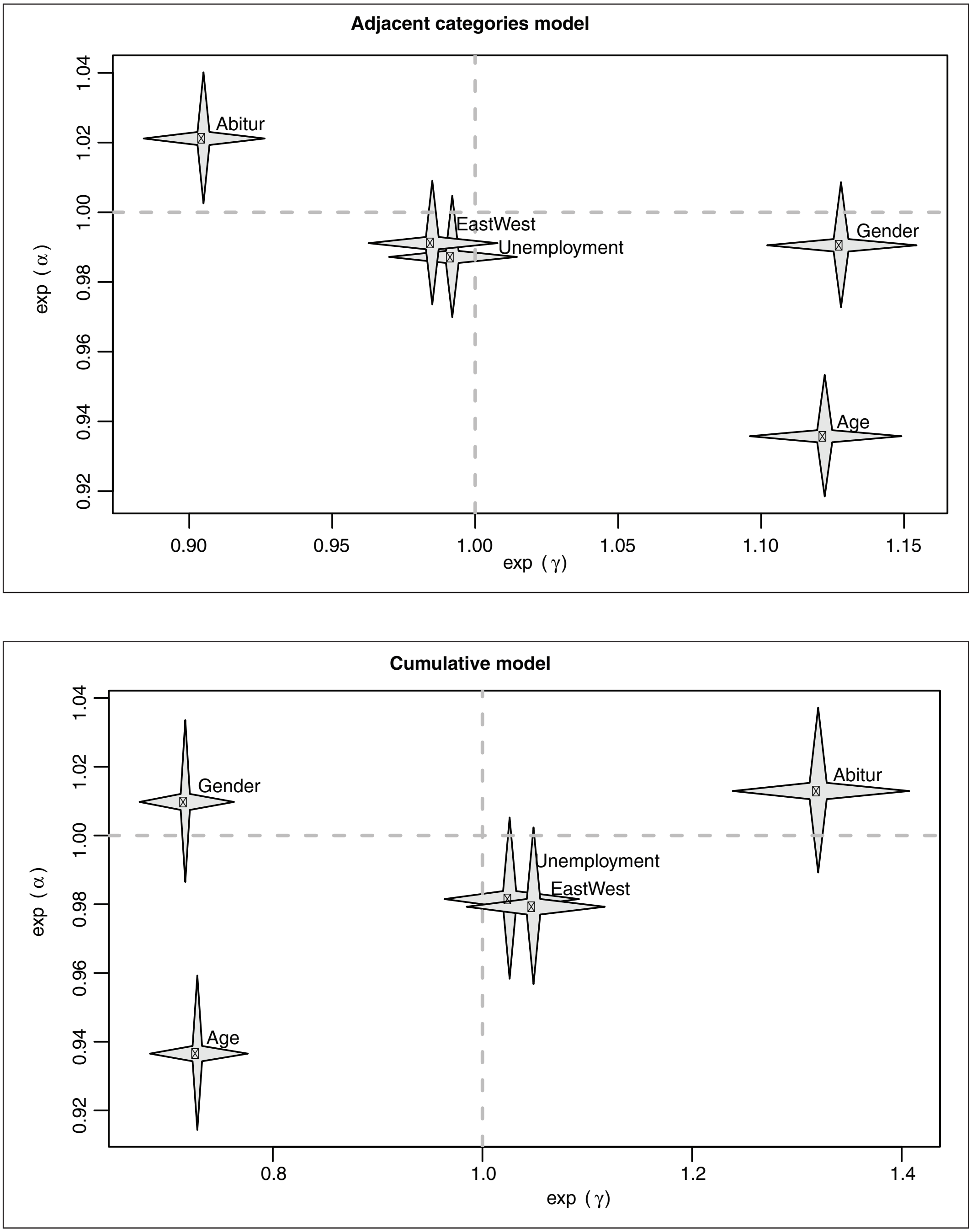


Figure 6 (Exponential) effects of explanatory variables in GLES data together with 95% confidence intervals both for location effects γ and response style effects α (separately for the adjacent categories model and the cumulative model).

random location effects and the random response style effects. It is obvious that the estimates of variances and covariances are quite different for the adjacent categories models and the cumulative models. However, this is not surprising given that a very different response function is used and that the random effects for the response styles have exponential form in the case of the cumulative model.

7 Concluding remarks

We considered response styles for multivariate ordinal responses. Response styles may be present whenever individuals use rating scales, that can be an assessment of their own feelings or ratings that evaluate the performance of others. As has been demonstrated, ignoring response styles may yield inferior estimates of location effects. Therefore, one should check if they can be ignored or have to be included. Beyond the effect on the accuracy of estimates, response styles are of interest by themselves, in particular if explanatory variables are included. The modelling of the dependence of response styles on covariates shows which groups of respondents tend to more or less extreme responses.

The proposed models are implemented in an add-on package for `R` (R Core Team, 2018), it is available from `https://github.com/Schaubert/MultOrdRS` and is supposed to be available from Comprehensive R Archive Network (CRAN) soon. The package also contains the data and the code from our application to the GLES (Roßteutscher et al., 2017) in Section 6 in order to make our results easily reproducible. For the (bivariate) normal distribution of the random effect parameters, we use (two-dimensional) Gauss–Hermite integration where the number of sample points can be specified by the user. The marginal likelihoods (and score functions) are implemented in `C++` and integrated into `R` via the packages `Rcpp` (Eddelbuettel et al., 2011) and `RcppArmadillo` (Eddelbuettel and Sanderson, 2014). The marginal likelihood is optimized numerically, either using the functions `optim()` or `nlminb()` in `R`.

Supplementary materials

Supplementary materials for this article including more detailed results of the simulation studies are available from `http://www.statmod.org/smij/archive.html`.

Declaration of conflicting interests

The authors declared no potential conflicts of interest with respect to the research, authorship and/or publication of this article.

Funding

The authors received no financial support for the research, authorship and/or publication of this article.

References

Agresti A (2009) *Analysis of Ordinal Categorical Data, 2nd Edition*. New York, NY: Wiley.

Anderson DA and Aitkin M (1985) Variance component models with binary response: Interviewer variability. *Journal of the Royal Statistical Society, Series B*, **47**, 203–10.

Anderson JA (1984) Regression and ordered categorical variables. *Journal of the Royal Statistical Society, B*, **46**, 1–30.

Baumgartner H and Steenkamp J-BE (2001) Response styles in marketing research: A cross-national investigation. *Journal of Marketing Research*, **38**, 143–56.

Böckenholt U (2017) Measuring response styles in Likert items. *Psychological Methods*, **22**, 69–83.

Bolt DM and Newton JR (2011) Multiscale measurement of extreme response style. *Educational and Psychological Measurement*, **71**, 814–33.

Booth JG and Hobert JP (1999) Maximizing generalized linear mixed model likelihoods with an automated Monte Carlo EM algorithm. *Journal of the Royal Statistical Society, B*, **61**, 265–85.

Breslow NE and Clayton DG (1993) Approximate inference in generalized linear mixed model. *Journal of the American Statistical Association*, **88**, 9–25.

Eddelbuettel D, Fran–ois R, Allaire J, Chambers J, Bates D and Ushey K (2011) Rcpp: Seamless R and C++ integration. *Journal of Statistical Software*, **40**, 1–18.

Eddelbuettel D and Sanderson C (2014) Rcpparmadillo: Accelerating r with high performance C++ linear algebra. *Computational Statistics and Data Analysis*, **71**, 1054–63. URL http://dx.doi.org/10.1016/j.csda.2013.02.005 (last accessed 16 December 2020).

Eid M and Rauber M (2000) Detecting measurement invariance in organizational surveys. *European Journal of Psychological Assessment*, **16**, 20.

Fahrmeir L and Tutz G (2001) *Multivariate Statistical Modelling Based on Generalized Linear Models*. New York, NY: Springer.

Hartzel J, Agresti A and Caffo B (2001). Multinomial logit random effects models. *Statistical Modelling*, **1**, 81–102.

Harville DA and Mee RW (1984) A mixed-model procedure for analyzing ordered categorical data. *Biometrics*, **40**, 393–408.

Hedeker D and Gibbons RB (1994) A random-effects ordinal regression model for multilevel analysis. *Biometrics*, **50**, 933–44.

Hinde J (1982). Compound Poisson regression models. In *GLIM 1982 International Conference on Generalized Linear Models*, edited by R. Gilchrist, pages 109–121. New York, NY: Springer-Verlag.

Jansen J (1990) On the statistical analysis of ordinal data when extravariation is present. *Applied Statistics*, **39**, 74–85.

Johnson TR (2003) On the use of heterogeneous thresholds ordinal regression models to account for individual differences in response style. *Psychometrika*, **68**, 563–83.

Liu Q and Pierce DA (1994) A note on Gauss–Hermite quadrature. *Biometrika*, **81**, 624–29.

Marin G, Gamba RJ and Marin BV (1992) Extreme response style and acquiescence among Hispanics the role of acculturation and education. *Journal of Cross-Cultural Psychology*, **23**, 498–509.

McCulloch C and Searle S (2001) *Generalized, Linear, and Mixed Models*. New York, NY: Wiley.

Meisenberg G and Williams A (2008) Are acquiescent and extreme response styles related to low intelligence and education? *Personality and Individual Differences*, **44**, 1539–50.

Messick S (1991) Psychology and methodology of response styles. In Improving Inquiry in Social Science: *A Volume in Honor of Lee J. Cronbach*, edited by RE Snow and DE Wiley, pages 161–200. Hillsdale, NJ: Lawrence Erlbaum Associates.

Peyhardi J, Trottier C and Guedon Y (2015) A new specification of generalized linear models for categorical responses. *Biometrika*, **102**, 889–906. doi: 10.1093/biomet/asv042. URL https://doi.org/10.1093/biomet/asv042 (last accessed 16 December 2020).

Pinheiro JC and Bates DM (1995) Approximations to the log-likelihood function in the nonlinear mixed-effects model. *Journal of Computational and Graphical Statistics*, **4**, 12–35.

R Core Team (2018) *R: A Language and Environment for Statistical Computing*. Vienna: R Foundation for Statistical Computing. URL https://www.R-project.org/ (last accessed 16 December 2020).

Roßteutscher S, Schmitt-Beck R, Schoen H, Weßels B, Wolf C, Bieber I, Stövsand L-C, Dietz M and Scherer P (2017) Pre-election cross section (GLES 2017). GESIS Data Archive, Cologne, ZA6800 Data le Version 2.0.0. URL https://doi.org/10.4232/1.12927 (last accessed 16 December 2020).

Tutz G (2012) *Regression for Categorical Data*. Cambridge: Cambridge University Press.

Tutz G and Berger M (2016) Response styles in rating scales - simultaneous modelling of content-related effects and the tendency to middle or extreme categories. *Journal of Educational and Behavioral Statistics*, **41**, 239–68.

——— (2017) Separating location and dispersion in ordinal regression models. *Econometrics and Statistics*, **2**, 131–48.

Tutz G and Hennevogl W (1996). Random effects in ordinal regression models. *Computational Statistics and Data Analysis*, **22**, 537–57.

Tutz G, Schauberger G and Berger M (2018) Response styles in the partial credit model. *Applied Psychological Measurement*, **42**, 407–27.

Van Vaerenbergh Y and Thomas TD (2013) Response styles in survey research: A literature review of antecedents, consequences, and remedies. *International Journal of Public Opinion Research*, **25**, 195–217.