LMU
LUDWIG-MAXIMILIANS-UNIVERSITÄT MÜNCHEN
INSTITUT FÜR STATISTIK

Gerhard Tutz & Micha Schneider

Mixture Models for Ordinal Responses with a Flexible Uncertainty Component

Technical Report Number 203, 2017
Department of Statistics
University of Munich

http://www.stat.uni-muenchen.de

Mixture Models for Ordinal Responses with a Flexible Uncertainty Component

Gerhard Tutz , Micha Schneider
[1] Ludwig-Maximilians-Universität München, Akademiestraße 1, 80799 München

Abstract

In classical mixture models for ordinal data with an uncertainty component the uniform distribution is used to model indecision. In the approach proposed here the discrete uniform distribution is replaced by a more flexible distribution, which is centered in the middle of the response categories. The resulting model allows to distinguish between a tendency to middle categories and a tendency to extreme categories. By linking these preferences to explanatory variables one can investigate which persons show a tendency to these response styles. It is demonstrated that severe bias might occur if inadvertently the uniform distribution is used to model uncertainty. An application to attitudes on the performance of health services illustrates the advantages of the more flexible model.

Keywords: Ordinal responses, response styles, rating scales, mixture models, CUP model, CUB model

1 Introduction

In recent years a class of mixture models for ordinal data has been introduced that considers the choice of a response category as resulting from a mixture of a deliberate choice and uncertainty. In the original CUB model (for Combination of discrete Uniform and shifted Binomial random variables), see D'Elia and Piccolo (2005), the deliberate choice is modelled by a binomial distribution and the uncertainty by a discrete uniform distribution. Various models with different specifications of the distributions of the deliberate choice and the uncertainty part have been proposed since then, see, for example, Iannario and Piccolo (2010), Iannario et al. (2012), Iannario and Piccolo (2012b), Iannario (2012a), Iannario (2012b), Manisera and Zuccolotto (2014), and Tutz et al. (2016). An introduction and overview on the modelling approaches was given by Iannario and Piccolo (2012a).

The basic assumption of most of these extensions is that uncertainty follows a discrete uniform distribution. Although the uniform distribution is the most simple conceivable model, the assumption that all categories, including middle and extreme categories, share the same degree of uncertainty is rather strong. In particular it excludes the preference of middle or extreme categories, which is a response style that is often found in applications. In the present paper we propose a more flexible uncertainty component which is able to capture response styles.

The presence of response styles has been found in many studies, see, for example, (Clarke, 2000; Van Herk et al., 2004), (Marin et al., 1992) and (Meisenberg and Williams, 2008). Several modelling approaches have been proposed for repeated measurements within the framework of item response models, see Bolt and Johnson (2009), Bolt and Newton (2011), Johnson (2003), Eid and Rauber (2000). More recently tree type approaches have been considered. They typically assume a nested structure where first a decision about the direction of the response and then about the strength is obtained, see, for example, De Boeck and Partchev (2012), Jeon and De Boeck (2015), and Böckenholt (2012). Mixture modelling of response styles by use of latent class models has been investigated by Moors (2004), Kankaraš and Moors (2009), Moors (2010), and Van Rosmalen et al. (2010).

The mixture considered here does not assume that responses on several items are available as is usually assumed in item response theory. We aim at separating the deliberate choice from the tendency to middle or extreme categories by using a mixture model in the tradition of CUB models. However, in contrast to these models we consider an uncertainty component that can account for response styles. By linking the uncertainty component to covariates, the model is able to uncover which person characteristics determine the response style. An alternative model for single items, which uses an explicit parametrization instead of a mixture, was proposed more recently by Tutz and Berger (2016).

The paper is organized as follows: in Section 2 we consider uncertainty as a relevant component quite often present in human choices. Thus CUB models and models with alternative parameterizations are briefly reviewed. Then the new class of models with more flexible uncertainty components is introduced. In Section 3 we investigate the consequences of fitting misspecified models in a simulation study. Section 4 gives the details of the fitting algorithm and in Section 5 the model is used to investigate the satisfaction with the Health Service in European Countries.

2 Mixture Models for Ordinal Responses

In the following we briefly consider an extended form of the CUB model. Then we consider alternative specifications of the uncertainty component.

2.1 Mixture Models for the Combination of Uncertainty and Preference

Let in a regression model the response of an individual R_i given explanatory variables take values from ordered categories $\{1,\dots,k\}$. The general mixture model we consider has the form

$$P(R_i = r|\boldsymbol{x}_i) = \pi_i P_M(Y_i = r|\boldsymbol{x}_i) + (1-\pi_i) P_U(U_i = r), \tag{1}$$

where R_i is the observed response, Y_i denotes the unobserved random variable that represents the deliberate choice, that is, the preference on the ordinal scale and U_i is the unobserved uncertainty component. Thus the observed response results from a discrete mixture of the preference and the uncertainty component, therefore the name CUP for Combination of Uncertainty and Preference. Both variables Y_i and U_i take values from $\{1,\dots,k\}$.

In model (1) the distribution of Y_i is determined by $P_M(Y_i = r|\boldsymbol{x}_i)$, which can be any ordinal model M. In CUB models and the extension considered by Tutz et al. (2016) the uncertainty component is specified by the uniform distribution, $P_U(U_i = r) = 1/k$. It has been argued that the uniform distribution is the most simple model that represent a totally random decision, for more motivation see also Iannario and Piccolo (2012a). The assumption of a more flexible distribution than the uniform distribution is the central issue here but postponed to the next section. Instead we consider briefly the ordinal models that can be used in the preference part.

In traditional CUB models the distribution of Y_i is specified as a shifted binomial distribution, that is,

$$P_M(Y_i = r|\boldsymbol{x}_i) = \binom{k-1}{r-1} \xi_i^{k-r}(1-\xi_i)^{r-1}, \quad r \in \{1,\dots,k\}.$$

In extended versions (Tutz et al. (2016)) more general models as the cumulative or the adjacent categories models are used. Cumulative models have the form

$$P(Y_i \le r|\boldsymbol{x}_i) = F(\gamma_{0r} + \boldsymbol{x}_i^T\boldsymbol{\gamma}), \quad r = 1,\dots,k-1,$$

where $F(.)$ is a cumulative distribution function and $-\infty = \gamma_{00} < \gamma_{01} < \dots < \gamma_{0k} = \infty$. The most widely used model from this class of models is the cumulative logit model, which uses the logistic distribution $F(.)$ It is also called *proportional odds model* and has the form

$$\log\left(\frac{P(Y_i \le r|\boldsymbol{x}_i)}{P(Y_i > r|\boldsymbol{x}_i)}\right) = \gamma_{0r} + \boldsymbol{x}_i^T\boldsymbol{\gamma}, \quad r = 1,\dots,k-1.$$

An alternative choice is the *adjacent categories model* given by

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

If the probability $P(Y_i = r|Y_i \geq r, \boldsymbol{x}_i)$ represents the probability of failure in (time) category r given category r is reached it can be seen as a discrete hazard. The specific model that uses the logistic distribution is the *adjacent categories logit model*

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

A general discussion of ordinal models is found in McCullagh (1980), Agresti (2010), Agresti (2013) and Tutz (2012).

2.2 Models with a Flexible Uncertainty Component

The uniform distribution as uncertainty component has the advantage of simplicity. However, it implies that uncertainty is uniformly distributed over the response categories. A more flexible concept allows that uncertainty may express itself in a stronger tendency toward middle or extreme categories. In particular persons who are undecided or have no strong opinion may have a tendency to choose middle categories and not choose at random from the whole spectrum of categories. Therefore, instead of the uniform distribution we use a specific version of the beta binomial distribution.

A random variable U with support $\{1, \dots, k\}$ follows a beta-binomial distribution, $U \sim BetaBin(k, \alpha, \beta)$, if the mass function is given by

$$f(u) = \begin{cases} \binom{k-1}{u-1}\frac{B(\alpha+u-1,\beta+k-u+1)}{B(\alpha,\beta)} & u \in \{1, \dots, k\} \\ 0 & \text{otherwise,} \end{cases}$$

where $\alpha, \beta > 0$ and $B(\alpha, \beta)$ is the beta function defined as

$$B(\alpha, \beta) = \Gamma(\alpha)\Gamma(\beta)/\Gamma(\alpha+\beta) = \int_0^1 t^{\alpha-1}(1-t)^{\beta-1}dt.$$

With $\mu = \alpha/(\alpha+\beta)$ and $\delta = 1/(\alpha+\beta+1)$ one obtains

$$E(U) = (k-1)\mu + 1, \quad \text{var}(U) = (k-1)\mu(1-\mu)[1+(k-2)\delta].$$

As $\delta \to 0$, the beta-binomial distribution converges to the binomial distribution $B(k, \mu)$.

Since we aim at modelling a tendency to middle categories we choose a fixed value $\mu = 0.5$ and therefore $\alpha = \beta, \delta = 1/(2\alpha+1)$ to obtain

$$\text{E}(U) = (k+1)/2.$$

For the variance one obtains

$$\text{var}(U) = ((k-1)/4)\frac{2\alpha+k-1}{2\alpha+1}.$$

The restricted beta-binomial is determined by the parameters α and k. An interesting extreme cases is $\alpha = 0$, which yields

$$\mathrm{var}(U) = ((k-1)^2/4),$$

and corresponds to a two point distribution on 1 and k. If α tends to infinity one obtains

$$\mathrm{var}(U) = ((k-1)/4).$$

Therefore, the parameter α determines the concentration of the distribution in the middle, for small values the probability mass is concentrated in the end points, for $\alpha = 1$ one obtains the discrete uniform distribution and for $\alpha \to \infty$ one obtains a (shifted) binomial distribution, which is symmetric around its mean $(k-1)/2$.

Figure 1 shows the beta-binomial distribution for selected values of α. In the case of an odd number of categories the modus is at the middle category.

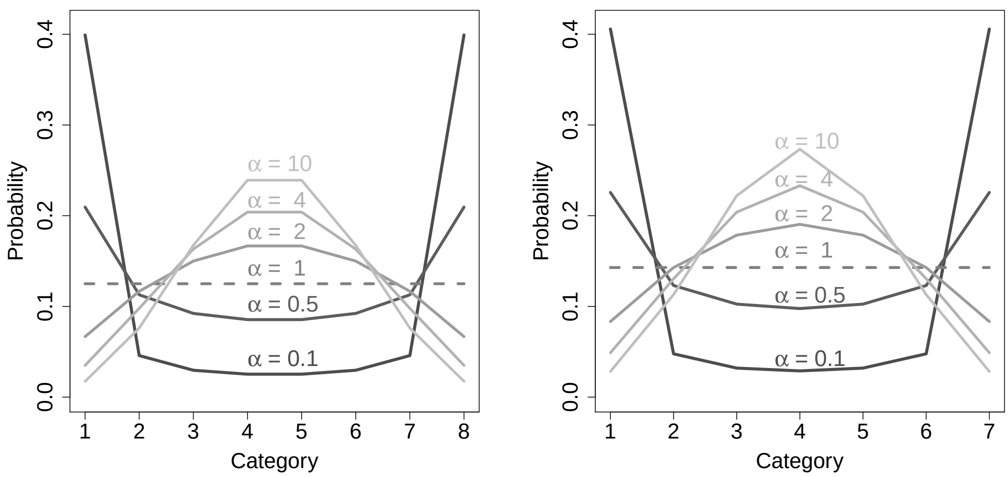

FIGURE 1: *Probability mass on categories for various values of α for 8 categories (left panel) and 7 categories (right panel).*

While mixture models in the tradition of CUB models use the uniform distribution, the beta-binomial distribution provides a wider concept of uncertainty in mixture models. An exception among CUB-type models is the model proposed by Gottard et al. (2016). It allows that the uncertainty is given by a parabolic or a triangular distribution. However, one has to choose the mode of the triangular distribution, therefore a priori information is needed. Moreover, the uncertainty distribution is not linked to explanatory variables as is done in the approach proposed here (see next section).

2.3 Parametrization

In the general mixture model (1) the preference for categories is determined by the covariates $\boldsymbol{x}_i$ within the ordinal model that is used in the preference part. However, also the strength of the tendency to middle or extreme categories may depend on covariates. Therefore, we let the parameter α depend on covariates $\boldsymbol{w}_i^T = (1, w_{i1}, \dots, w_{im})$, which can be different or identical to the covariates $\boldsymbol{x}_i$. A simple link is given by

$$\alpha = \exp(\boldsymbol{w}_i^T \boldsymbol{\alpha}) = \exp(\alpha_0) \exp(\alpha_1)^{w_{i1}} \ldots . \exp(\alpha_m)^{w_{im}},$$

where $\boldsymbol{\alpha}^T = (\alpha_0, \dots, \alpha_m)$. The parameter α_j contains the effect of the j-th covariate. The parameter α changes by the factor $\exp(\alpha_j)$ if w_{ij} increases by one unit. The parameters determine how a variable influences the tendency to middle or extreme categories. It should be noted that in the case without covariates one has the simple reparameterization $\alpha = \exp(\alpha_0)$.

The model (1) with a beta-binomial mixture component is called the BetaMix model. Although it is a generalization of CUP models the intention of the modelling approach is quite different. In CUP models the uncertainty is specified by a discrete uniform distribution. The underlying assumption is that a person is torn between his/her preference and uncertainty. The uncertainty is such that each category has the same probability. The BetaMix model is composed of a preference model and a model that represents a tendency to middle or extreme categories. It allows to model not only the preference as a function of covariates but also the tendency to middle or extreme categories as a function of covariates. One may see, for example, differences in the preference of middle or extreme categories induced by covariates like gender. Therefore, response patterns induced by explanatory variables can be identified.

The family of models considered here can be specified by Mix(structured part, uncertainty part). The structured part indicates which model is used to model the deliberate choice, and the uncertainty part indicates which distribution is used to model the uncertainty. Examples are

Mix(Binomial,Uniform) (or CUB), which means that the structured response follows binomial distribution and uncertainty is determined by the uniform distribution

Mix(Cumulative,Uniform) (or CUP), which means that the structured response is determined by a cumulative model, the uncertainty is the same as in the previous example

Mix(Cumulative, BetaBin), which means that the uncertainty is determined by the beta-binomial distribution

Mix(Binomial,BetaBin), which means that the structured response follows binomial distribution and uncertainty is determined by the beta-binomial distribution

The model Mix(Cumulative, BetaBin($\alpha = 1$)) is equivalent to the CUP cumulative model and Mix(Binomial, BetaBin($\alpha = 1$)) is equivalent to the CUB model.

3 Simulations

In the following we investigate the consequences of fitting misspecified models in a simulation study. In particular we compare the performance of the proposed model with the models that use a uniform distribution in the uncertainty part. First we compare the Mix(Cumulative,Uniform) (or cumulative CUP) and the Mix(Cumulative, BetaBin), and then CUB and Mix(Binomial,BetaBin).

We use a response with $k = 7$ categories and $n = 2000$ observations. The data were simulated from a mixture model with different values for π, α and γ. For the mixture weights π the values we used $0.5, 0.7$ and 0.8. The effect of the structure component γ was fixed at -1 and -2. The intercepts of the cumulative model were set to $-4, -3, -2, -1, 0, 1$, in the shifted binomial model we used 1. The range of the α-values was $\{0.01, 0.1, 0.25, 0.5, 1, 2, 4, 10, 100\}$ so that both the tendency to the middle categories with $\alpha > 1$ and the tendency to extreme categories with $\alpha < 1$ are covered. Also the special case $\alpha = 1$, in which the uncertainty components of CUP and BetaMix are identical, is included. For each parameter combination 500 data sets were simulated from the model with the beta-binomial-distribution. The beta-binomial model as well as the model with uniform distribution were fitted. Then the performance of the new proposed model is compared to the performance of the misspecified model with a uniform distribution.

Before given detailed tables for all used combinations of π, α and γ we show some exemplary box plots. Figure 2 displays the estimated parameters for different α-values for both models with π set to 0.7 and γ set to -1. Each boxplot consists of 500 samples. The results of the beta-binomial model are displayed on the left hand side and the results of the CUP-Model on the right hand side. The top row shows the π-values and the middle row the γ-estimates. For the beta-binomial model all the estimates are close to the true parameters regardless which response style is true. The model is able to capture both a strong tendency to the middle category as well as a strong tendency to extreme categories. On the right hand side the different response styles are neglected and it is always assumed that the uncertainty component follows a uniform distribution. It is seen that estimates are strongly biased if the model is unable to account for the response style. If the true α-value is far away form $\alpha = 1$, which is assumed by the CUP model, there is a large discrepancy between the true parameter values and the estimated parameters. For example, if $\alpha = 0.01$, which indicates a strong

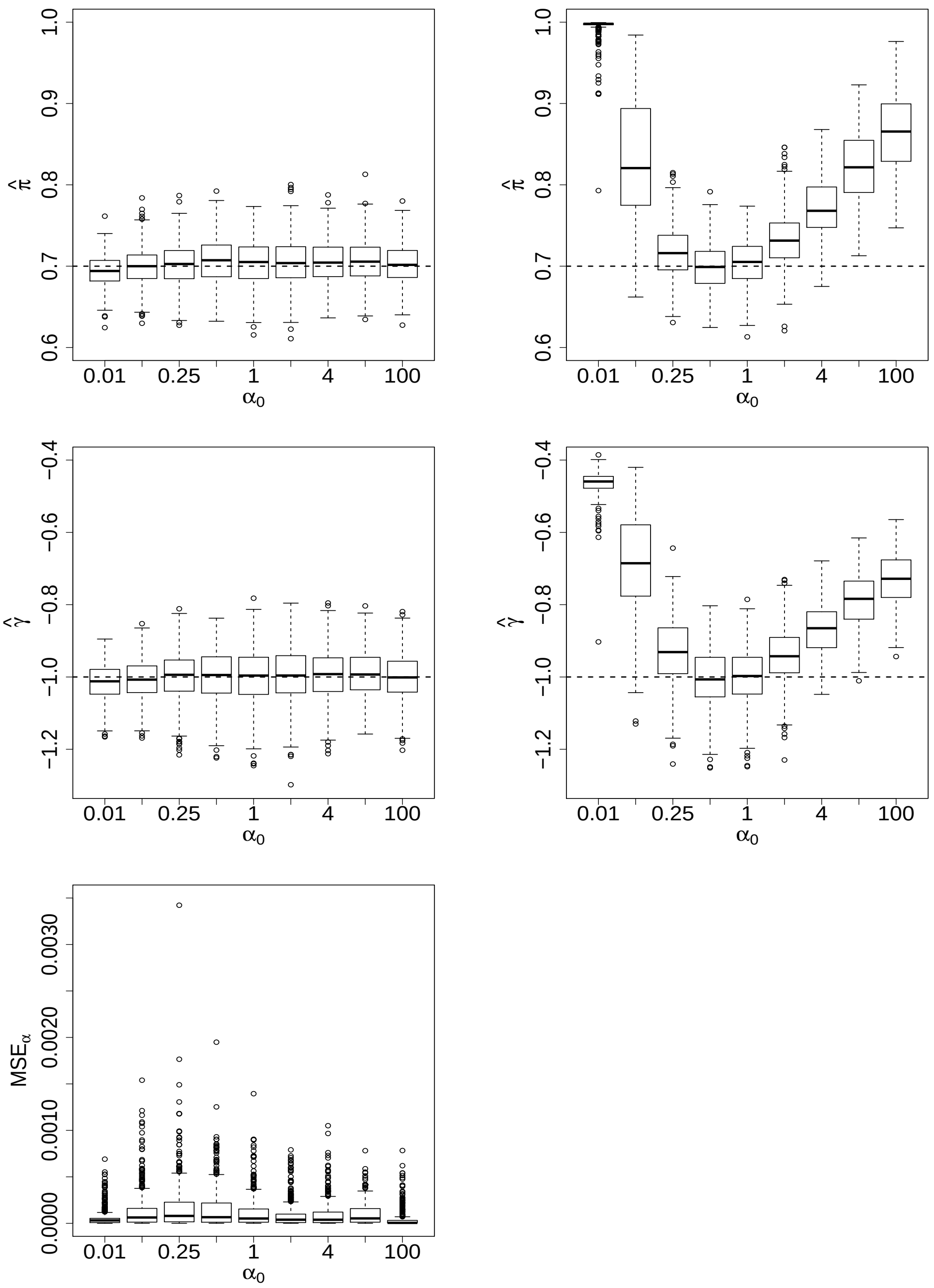

FIGURE 2: *Estimated parameters $\hat{\pi}$, $\hat{\gamma}$ for the Betamix model on the left and the CUP model on the right (true values are $\pi = 0.7$ and $\gamma = -1$). The true α-values are $\{$0.01, 0.1, 0.25, 0.5, 1, 2, 4, 10, 100$\}$. For the Betamix model also the MSEs are given.*

tendency to the extreme categories, the CUP-model estimates a π-value which is close to one. Thus, one would falsely infer that no uncertainty component is needed. At the same time the strength of the effect of the variable is underestimated. If there is a strong tendency to the middle categories the results are similar. So by using the uniform distribution as a possible response style not only the π-values but also the γ-values are strongly biased if the data generating model contains a specific response style.

To investigate the accuracy of estimates we consider the mean squared error. For the comparison we use the log proportions

$$\text{lp} = \frac{1}{S}\sum_{i=1}^{S} \log \frac{\text{MSE(uniform)}_i}{\text{MSE(beta-binomial)}_i},$$

where $\text{MSE(beta-binomial)}_i$ denotes the mean squared error in the ith sample if the beta-binomial model is fitted and MSE(uniform) the mean squared error if the uniform model is fitted. Positive values of lp indicate that the uniform model yields estimates that are worse than the estimates obtained by the beta-binomial model.

Table 1 and 2 show the log proportions for γ and π for several parameter combinations. In the case of $\alpha = 1$ the log proportions are close to zero so that both models fit equally. But there is a strong monotone increase when the true α-values are more and more away from $\alpha = 1$. For example, one obtains for $(\pi, \gamma, \alpha) = (0.5, -1, 4)$ $lp = 0.6509$, which means that the MSE of the uniform model is 1.92 times the MSE of the beta-binomial model, for $(\pi, \gamma, \alpha) = (0.5, -2, 4)$ one has $lp = 1.4235$ denoting that the MSE of the uniform model is 4.15 times the MSE of the beta-binomial model. It is also seen that for small values of π the proportions of γ-values are larger than for large values of π (close to 1), therefore for small values of π a wrong response style has stronger impact on the γ-parameters. For larger value of γ one obtains larger log proportions.

For the accuracy of the estimated response style we do not use the mean squared errors of the α-values. The reason is the scaling of the parameter. For very large α-values the beta-binomial-distribution is close to the binomial distribution, which is obtained if α-values is infinitely large. Consequently very large α-values may be different in their absolute value but lead to nearly the same distribution function. Therefore, we use the mean squared errors of the estimated distributions

$$MSE_{\alpha} = \frac{1}{S}\sum_{i=1}^{S}\left(\frac{1}{k}\sum_{k=1}^{K}(Pr_i(U=k|\hat{\alpha}) - Pr_i(U=k|\alpha))^2\right).$$

As seen from Table 3 in all settings the MSE_{α} is less than 0.0004 and there is no structure visible. The last panel in Figure 2 shows the corresponding box plots, which are all close to zero. It is seen that the Betamix model is able to fit the true response style very well.

π	γ					α				
		0.01	0.1	0.25	0.5	1	2	4	10	100
0.5	-1	1.1120	1.0887	1.0061	0.2167	0.0007	0.2717	0.6509	0.8644	0.9398
0.7	-1	0.8909	0.5479	0.1319	-0.0138	0.0006	0.1043	0.2375	0.3690	0.4783
0.8	-1	0.2506	0.0925	-0.0080	-0.0334	-0.0008	0.0660	0.1343	0.2060	0.2711
0.5	-2	6.1968	5.6751	5.3883	1.0003	-0.0014	0.3862	1.4235	3.5169	4.8718
0.7	-2	6.3268	4.1285	1.2782	-0.0426	-0.0231	0.1646	0.2528	1.0761	1.8724
0.8	-2	2.4443	0.8960	0.2808	-0.0005	-0.0036	0.1269	0.3452	0.5010	0.7295

TABLE 1: *Log proportions of γ-values. Positive values indicate that γ estimates of the CUP model are farer away from the true γ-values than the estimates of the Betamix model.*

π	γ					α				
		0.01	0.1	0.25	0.5	1	2	4	10	100
0.5	-1	7.5542	7.0457	6.1898	0.9142	-0.0923	2.2108	4.9798	6.2124	6.7589
0.7	-1	6.6484	4.3927	0.7165	-0.1107	0.0312	1.0513	2.9954	4.1365	4.9370
0.8	-1	2.0368	0.8023	0.4547	-0.0873	-0.0320	0.5957	1.7226	2.6127	3.5027
0.5	-2	7.7898	7.6342	6.9940	2.2677	0.0124	0.4676	1.0724	3.3313	5.8254
0.7	-2	7.0814	5.0817	2.5478	0.4908	-0.0218	0.2590	0.4474	0.4942	0.8461
0.8	-2	3.6844	2.3575	0.9614	0.1400	-0.0077	0.1195	0.2181	0.5144	0.4748

TABLE 2: *Log proportions of π-values. Positive values indicate that π estimates of the CUP model are farer away from the true π-values than the estimates of the Betamix model.*

Similar results are obtained if the shifted binomial distribution and therefore the CUB model is used in the preference part. Now we compare Mix(Binomial,Uniform) (or CUB) with Mix(Binomial,Betabin). Figure 3 and 4 show the same setting as before, they compare the beta-binomial distribution with the uniform distribution in the uncertainty part, but now the shifted binomial distribution determines the preference component of both models. The figures show the results for $\gamma = -1$ as well as $\gamma = -2$. The well specified model can deal with different α and γ-values. But there are clear discrepancies in the misspecified models. For extreme α-values the estimates of γ and π in the misspecified models are poor. In the case of $\gamma = -1$ the π-values are underestimated for α-values smaller than one and overestimated for α-values greater than one. But for $\gamma = -2$ the opposite behaviour is observed. In both cases the γ estimates show the same trend. In Table 4 and 5 the results for all combinations are displayed. In general, there is clear discrepancy in the misspecified models but the direction (i.e. over or underestimation of the parameter) can vary. If the uniform distribution is the true uncertainty component the CUB-model seems

π	γ	α								
		0.01	0.1	0.25	0.5	1	2	4	10	100
0.5	-1	0.0000	0.0001	0.0001	0.0001	0.0001	0.0000	0.0000	0.0001	0.0000
0.7	-1	0.0001	0.0001	0.0002	0.0002	0.0001	0.0001	0.0001	0.0001	0.0000
0.8	-1	0.0001	0.0003	0.0003	0.0002	0.0002	0.0002	0.0002	0.0001	0.0001
0.5	-2	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
0.7	-2	0.0000	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0000
0.8	-2	0.0000	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0000

TABLE 3: *Mean squares errors that measure the discrepancy between the estimated and the true beta-binomial distribution.*

to be a bit closer to the true π-values than the model with the betabinomial-distribution. But the log proportions are close to zero so that the differences of the π-estimates in both models are very small. Moreover, in the Betabin model the uncertainty component has to be estimated which is more difficult than assuming that α is exactly fixed at 1 as in the CUB model. In all other cases the Betabin model clearly outperforms the CUB-model in terms of accuracy of the parameter estimates.

π	γ	α								
		0.01	0.1	0.25	0.5	1	2	4	10	100
0.5	-1	7.9908	6.9918	5.6090	3.8647	-0.2223	3.1404	4.3818	4.9940	5.4806
0.7	-1	6.3887	5.6636	4.4604	2.7079	-0.2538	2.0550	3.3342	4.0105	4.2475
0.8	-1	5.3687	4.7483	3.7107	2.0522	-0.0751	1.4637	2.4287	3.1594	3.3904
0.5	-2	6.7311	6.9559	6.2788	3.5601	-0.1111	1.8015	3.2645	3.2836	3.4449
0.7	-2	5.6132	4.9745	3.6663	1.7816	-0.0643	0.7877	1.5349	2.0155	2.2496
0.8	-2	4.9915	4.1609	3.0496	1.0264	-0.0259	0.6279	0.9362	1.3346	1.4698

TABLE 4: *Log proportions of γ-values. Positive values indicate that γ estimates of the CUB model are farer away from the true γ-values than the estimates of the Betabin model.*

π	γ	α								
		0.01	0.1	0.25	0.5	1	2	4	10	100
0.5	-1	2.0685	1.7739	0.9497	-0.0189	-0.0027	0.6860	1.8735	3.2835	4.0075
0.7	-1	1.2640	0.7599	0.7672	0.3249	0.0407	0.3471	0.8965	1.4374	1.9017
0.8	-1	1.6844	1.2142	0.7175	0.4247	-0.0611	0.0645	0.2233	0.8416	1.1060
0.5	-2	4.4434	5.5704	5.9981	3.1680	0.0910	1.0716	2.1484	2.7990	3.2827
0.7	-2	3.8922	3.2800	2.2376	0.6130	-0.1299	0.6281	1.4431	2.1902	2.7078
0.8	-2	3.3097	2.3723	1.4503	0.3256	0.1253	0.5703	1.1688	1.6480	1.9694

TABLE 5: *Log proportions of π-values. Positive values indicate that π estimates of the CUB model are farer away from the true π-values than the estimates of the Betabin model.*

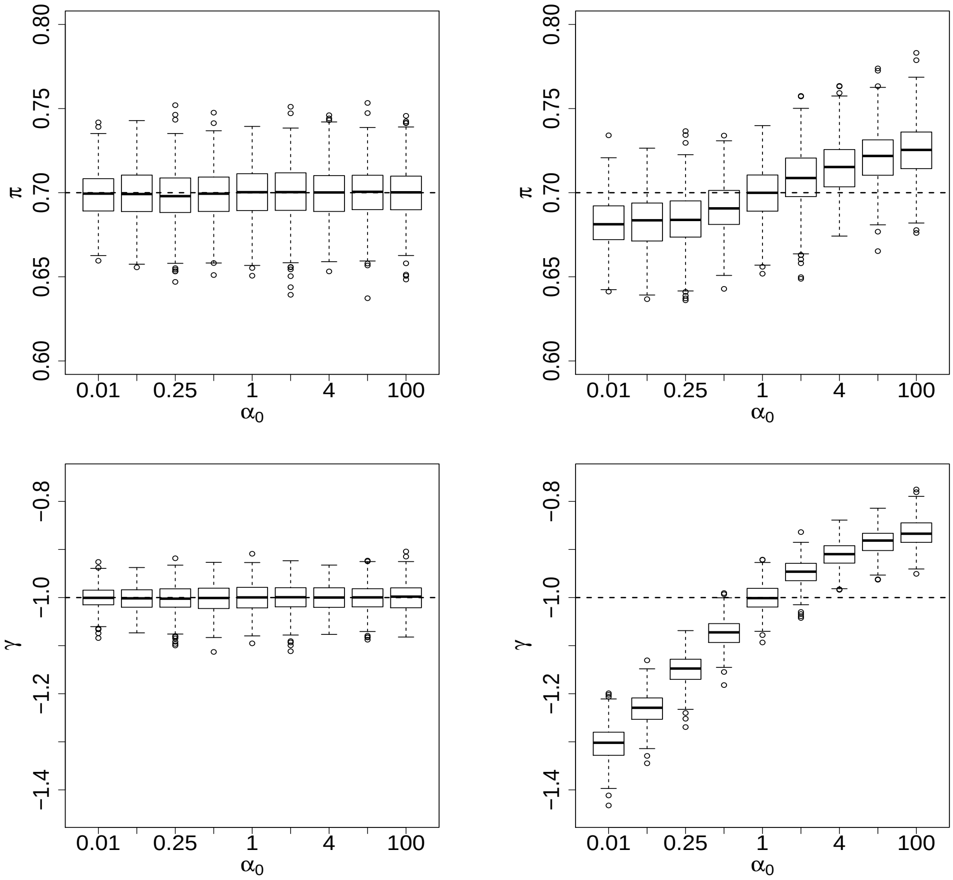

FIGURE 3: *Comparison of the estimated parameters $\hat{\pi}$, $\hat{\gamma}$ between the Betabin model on the left and the CUB model on the right for $\pi = 0.7$ and $\gamma = -1$. The true α-values are $\{0.01, 0.1, 0.25, 0.5, 1, 2, 4, 10, 100\}$. The MSE of α is only reasonable for the Betabin model.*

4 Estimation

The likelihood contribution of observation i when category y_i is observed is determined by

$$P(R_i = y_i|\boldsymbol{w}_i, \boldsymbol{x}_i) = \pi_i \, P_M(Y_i = y_i|\boldsymbol{x}_i) + (1 - \pi_i) \, P_U(U_i = y_i|\boldsymbol{w}_i) \qquad (2)$$

yielding the log-likelihood contribution

$$l_i(\boldsymbol{\gamma}, \boldsymbol{\alpha}) = \log(\pi_i \, P_M(Y_i = y_i|\boldsymbol{x}_i) + (1 - \pi_i) \, P_U(U_i = y_i|\boldsymbol{w}_i))$$

A way to obtain stable estimates is to consider it as a problem with incomplete data and use the EM algorithm Dempster et al. (1977). Therefore, let z_i^* denote the unknown mixture components that indicate whether y_i belongs to the first

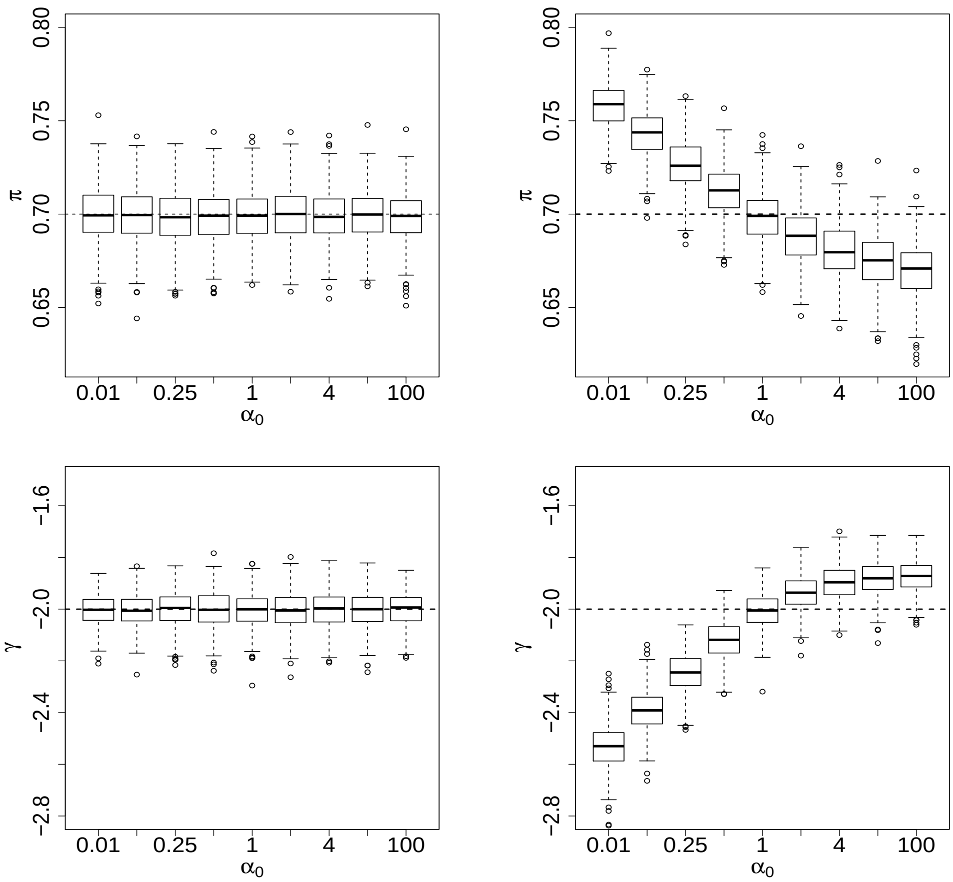

FIGURE 4: *Comparison of the estimated parameters $\hat{\pi}$, $\hat{\gamma}$ between the Betabin model on the left and the CUB model on the right for $\pi = 0.7$ and $\gamma = -2$. The true α-values are {0.01, 0.1, 0.25, 0.5, 1, 2, 4, 10, 100}. The MSE of α is only reasonable for the Betabin model.*

or second component of the mixture

$$z_i^* = \begin{cases} 1, \text{observation } y_i \text{ is from the first mixture component} \\ 0, \text{otherwise.} \end{cases}$$

The corresponding complete log-likelihood is given by

$$l_c(\boldsymbol{\gamma}, \boldsymbol{\alpha}) = \sum_{i=1}^{n} z_i^* \left\{\log(\pi_i) + \log(P_M(Y_i = y_i|\boldsymbol{x}_i))\right\} + (1 - z_i^*) \left\{\log(1 - \pi_i) + \log(P_U(U_i = y_i|\boldsymbol{w}_i))\right\}.$$

The EM algorithm treats z_i^* as missing data and maximizes the log-likelihood iteratively by using an expectation and a maximization step. During the E-step the conditional expectation of the complete log-likelihood given the observed data $\boldsymbol{y}^T = (y_1, \ldots, y_n)$ and the current estimate $\boldsymbol{\theta}^{(s)} = (\boldsymbol{\gamma}^{(s)}, \boldsymbol{\alpha}^{(s)})$,

$$M(\boldsymbol{\theta}|\boldsymbol{\theta}^{(s)}) = \mathrm{E}(l_c(\boldsymbol{\theta})|\boldsymbol{y}, \boldsymbol{\theta}^{(s)})$$

has to be computed. Because $l_c(\boldsymbol{\theta})$ is linear in the unobservable data z_i^*, it is only necessary to estimate the current conditional expectation of z_i^*. From Bayes's theorem follows

$$\begin{aligned}
E(z_i^*|\boldsymbol{y}, \boldsymbol{\theta}) &= P(z_i^* = 1|y_i, \boldsymbol{x}_i, \boldsymbol{w}_i, \boldsymbol{\theta}) = P(R = y_i|z_i^* = 1, \boldsymbol{x}_i, \boldsymbol{w}_i, \boldsymbol{\theta})/P(R = y_i|\boldsymbol{x}_i, \boldsymbol{w}_i, \boldsymbol{\theta}) \\
&= \pi_i\, P_M(Y_i = y_i|\boldsymbol{x}_i, \boldsymbol{\theta})/(\pi_i\, P_M(Y_i = y_i|\boldsymbol{x}_i) + (1 - \pi_i)\, P_U(U_i = y_i|\boldsymbol{w}_i)) \\
&= \hat{z}_i^* = \hat{z}^*.
\end{aligned}$$

This is the posterior probability that the observation y_i belongs to the first component of the mixture. Because there are no individual covariates determining the propensity to the structure component $\hat{z}_i^*$ the expectation $E(z_i^*|\boldsymbol{y}, \boldsymbol{\theta})$ is the same for all observations. For the s-th iteration one obtains

$$\begin{aligned}
M(\boldsymbol{\theta}|\boldsymbol{\theta}^{(s)}) &= \sum_{i=1}^{n} \hat{z}^* \left\{\log(\pi) + \log(P_M(Y_i = y_i|\boldsymbol{x}_i))\right\} \\
&\quad + (1 - \hat{z}^*) \left\{\log(1 - \pi) + \log(P_U(U_i = y_i|\boldsymbol{w}_i))\right\} \\
&= \underbrace{\sum_{i=1}^{n} \hat{z}^* \log(\pi) + (1 - \hat{z}^*) \log(1 - \pi)}_{M_1} \\
&\quad + \underbrace{\sum_{i=1}^{n} (1 - \hat{z}^*) \log(P_U(U_i = y_i|\boldsymbol{w}_i))}_{M_2} \\
&\quad + \underbrace{\sum_{i=1}^{n} \hat{z}^* \log(P_M(Y_i = y_i|\boldsymbol{x}_i))}_{M_3}.
\end{aligned}$$

The maximization in the M-Step uses the decomposition into M_1, M_2 and M_3. M_2 corresponds to the uncertainty component and M_3 to the structure component. M_1, M_2 and M_3 can be maximized separately with traditional software. For M_1 and the shifted binomial distribution (M_3 in CUB-models) we use the R-package MRSP by Poessnecker (2015). For the beta-binomial distribution (M_2) and the cumulative model(M_3 in CUP-models) we use the R-package VGAM by Yee (2016). In the s-th EM iteration M_1, M_2 and M_3 are not maximised until convergence is reached but only a few iterations in the sense of the generalized EM-Algorithm. So for given $\boldsymbol{\theta}^{(s)}$ one computes in the E-step the weights $\hat{z}^{*(s)}$ and in the M-step maximizes $M(\boldsymbol{\theta}|\boldsymbol{\theta}^{(s)})$, which yields the new estimates.

5 Application: Satisfaction with the Health Service in European Countries

To illustrate the new model we use the European Social Survey which measures the behaviour, attitudes and beliefs of populations in various European countries. We use the data of the 7th round in 2014, which is available at `http://www.europeansocialsurvey.org`. We focus on the attitude concerning the state of the health services measured on a Likert Scale from 0 "extremely bad" to 10 "extremely good". The covariates are gender (1: female), the age in decades (centered at 50), citizenship, the area of living (1: "big city" as reference, 2: "suburbs or outskirts of a big city", 3: "town or small city", 4: "country village", 5: "farm or home in the countryside"), the smoke behaviour (1: "I smoke daily", 2: "I smoke but not every day", 3: "I don't smoke now but I used to", 4: "I have only smoked a few times", 5: "I have never smoked" as reference) and if the person is handicapped in its daily activities in any way by any longstanding illness, disability, infirmity or mental health problem (1: "yes a lot", 2: "yes to some extent", 3: "no" as reference).

An identical model with the same covariates is fitted separately for several countries. We give detailed results for Germany and compare the estimated uncertainty propensity and gender effects across countries.

	estimate	BS.sd	BS.2.5	BS.97.5	
female	0.2778	0.0751	0.1486	0.4385	
age	0.0677	0.0239	0.0237	0.1181	
age^2	-0.1009	0.0122	-0.1283	-0.0798	
German citizen: No	-1.3709	0.2270	-1.8828	-0.9374	
domicil: suburb	0.1442	0.1405	-0.1212	0.4177	
domicil: town	0.2566	0.1082	0.0574	0.4792	
domicil: village	0.2402	0.1106	0.0366	0.4747	
domicil: countryside	0.0925	0.2153	-0.3162	0.5384	γ
handicapped: a lot	0.4302	0.1752	0.1254	0.7786	
handicapped: to some extent	0.4212	0.0999	0.2319	0.6397	
smoke: daily	0.3879	0.1175	0.1900	0.6403	
smoke: not every day	0.3936	0.2157	-0.0041	0.8214	
smoke: no, but used to	0.1042	0.0994	-0.0715	0.3067	
smoke: only a few times	-0.2471	0.1279	-0.4953	0.0035	
(Intercept)	3.8184	1.5363	1.8803	8.1662	
female	-2.3892	1.1699	-5.1968	-0.7173	
age	-0.6522	0.4172	-1.9058	-0.1083	
age^2	0.2528	0.1510	-0.0707	0.5546	α
handicapped: a lot	-3.5315	1.5560	-6.6147	-1.0599	
handicapped: to some extent	-1.8433	1.2455	-3.7856	0.2417	
$1-\pi$	0.1177	0.0349	0.0995	0.2123	

TABLE 6: *State of health services in Germany*

Table 6 shows the estimates of the BetaMix model for Germany with a cumulative model in the structure part. In the upper panel the effects on the preference part are displayed. Positive values indicate less satisfaction with the health services. It is seen that females are less satisfied with the health services in Germany than men. Persons who are not German citizen are happier with the health services than German citizens. It is often discussed if there is a difference between urban and rural health service supply. According to the model responders living in a town or in a village are significantly less happy with the health services than people living in a big city. For people living in the countryside or suburbs the difference to people living in a big city is non-significant. Also handicapped persons are less satisfied with the health services than non-handicapped persons. In the lower part the response style effects are displayed. Positive values indicate a tendency to the middle, negative values indicate a tendency to extreme categories. This follows from the parametrization of the α-values of the beta-binomial distribution, because for positive estimates one obtains exp(estimate) > 1 and therefore α increases. It is seen that females tend to choose more extreme categories than men. Handicapped persons also prefer more extreme categories than non-handicapped persons.

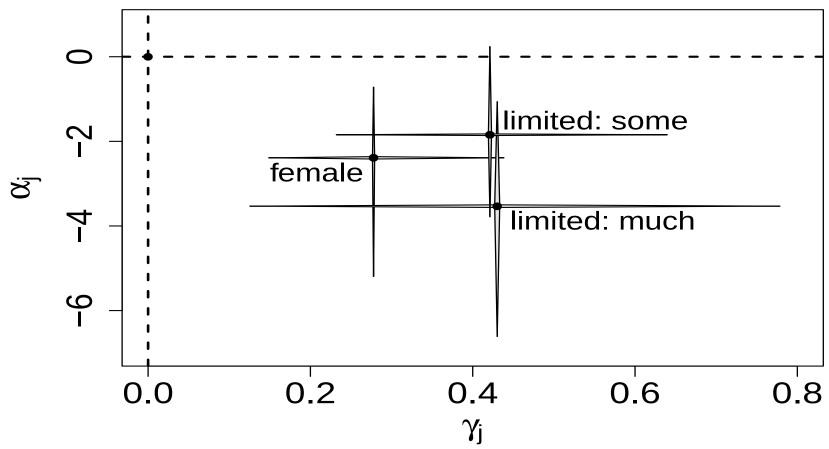

FIGURE 5: *State of health services in Germany: Gender and Handicap Effects*

In addition to giving estimates we use visualization tools to make the found effects easily accessible. In particular we use two-dimensional plots of the effects found in the preference part and the uncertainty part of the model. In the latter we use the response style parameters. More concrete, we plot the α and γ values together with the confidence intervals obtained by bootstrap to obtain a star for each binary variable and several stars for multi-categorical variables. Figure 5 shows the estimated effects (γ, α) of gender and being handicapped. Positive values in the γ-dimension indicate a tendency to negative statements concerning

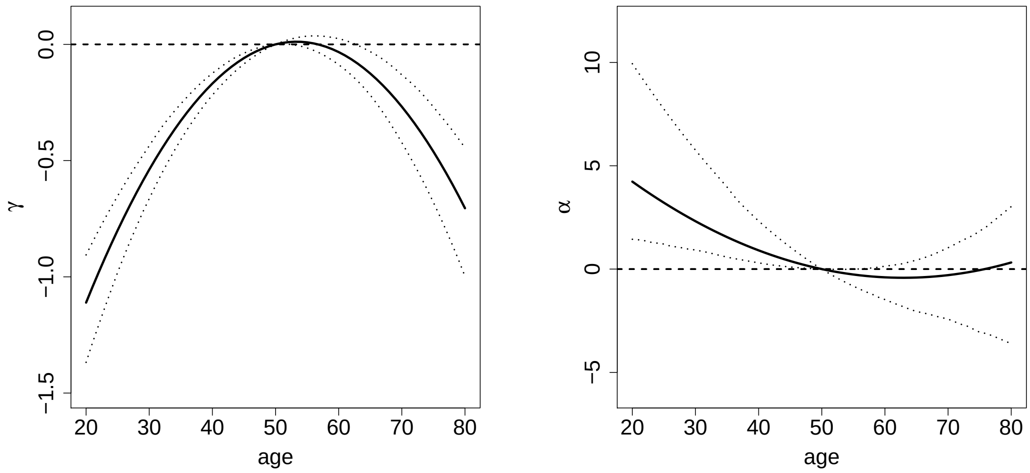

FIGURE 6: *State of health services in Germany: Age Effects*

the state of the health services, positive values in the α-dimension indicate a tendency to middle categories. It is seen that females tend to see the health services more sceptically and tend to choose more extreme categories. The effect of being handicapped is stronger than the gender effect in terms of a preference to categories indicating scepticism. The effects of being handicapped are almost the same in the preference part but differ in the uncertainty part. If a person is more handicapped it tends to choose more extreme categories. The effects are all significant except of "handicapped: to some extent" in the uncertainty component α. We used the 2.5% and 97.5% quantiles of the bootstrap samples instead of the bootstrap standard errors, because the distribution of the bootstrap standard errors may be skewed.

The effect of age is displayed in Figure 6. The dotted lines correspond to point-wise 95% bootstrap confidence intervals. They are constructed in such a way that in every bootstrap sample the age curve is calculated. Then the point-wise 2.5% and 97.5% quantiles are used to draw the dotted lines. On the left hand side the effect of age on the satisfaction of the health services is shown. It is seen that younger and older persons are more satisfied with the health services than persons in their 50s. The response style shows a different picture. Young persons below 50 years of age show a significant tendency to middle categories whereas for persons older than 50 years of age no significant tendency to middle or extreme categories can be detected.

For the comparison of countries we consider the performance of the BetaMix model, the estimates $1-\hat{\pi}$ and the effect of gender across countries. The countries considered are Austria (AT), Germany (DE), Denmark (DK), Spain (ES), Finland

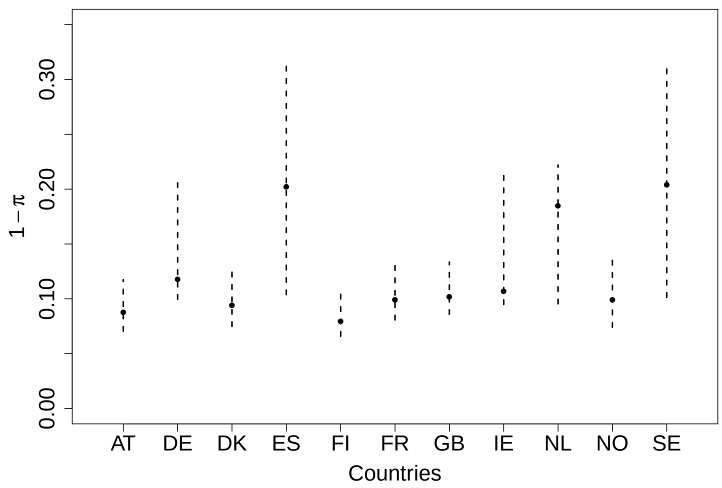

Figure 7: *State of health services: Influence of response style in different countries*

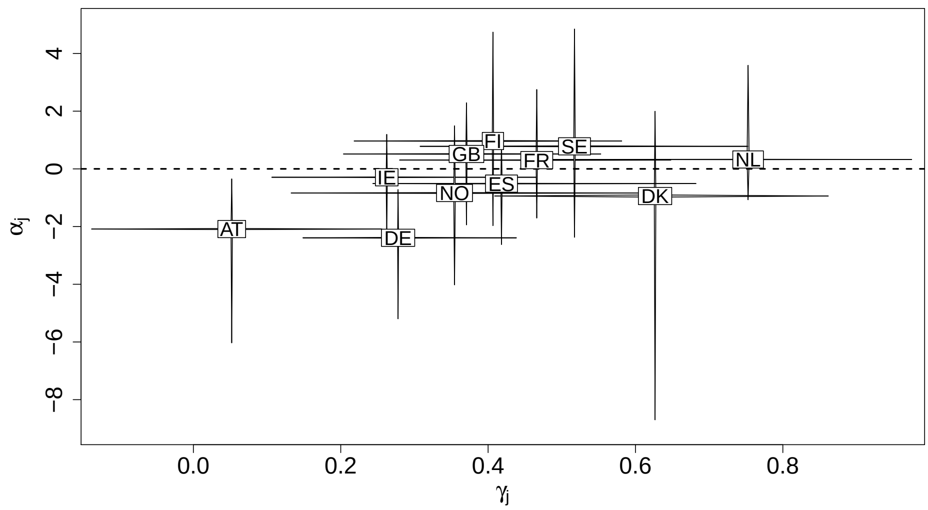

Figure 8: *State of health services: Influence of gender in different countries*

(FI), France (FR), Great Britain (GB), Ireland (IE), Netherlands (NL), Norway (NO) and Sweden (SE).

There are some differences in the estimates of $1 - \hat{\pi}$, which is a measure of the importance of the uncertainty component. Large values indicate the presence of response styles in the survey. Figure 7 shows the proportions of the response styles. The dotted lines correspond to the 2.5% and 97.5% bootstrap quantiles. In Germany (DE) the tendency to response styles is in the middle range. In Spain (ES) and Sweden (SE) the model estimates show higher proportions of the response style. The lowest estimated proportions are found for Austria (AT) and Finland (FI), with values 0.0876 and 0.0793, respectively.

Figure 8 displays the effect of gender across the different countries. As in Figure 5 the x-axis corresponds to the effect on the preference structure and the y-axsis to the effect of the response style. The confidence intervals are again obtained by bootstrap samples. For all countries the γ-parameters are positive which indicates that women are less satisfied with the health services of their country than men. The strongest effect can be found for the Netherlands and Denmark and the smallest for Austria. The effects are significant for all countries with the exception of Austria, for which the 95% bootstrap confidence interval contains zero.

In contrast, the gender effect in the response style is not homogeneous across countries. Positive α-parameters for Great Britain (GB), Finland (FI), France (FR), Netherlands (NL) and Sweden (SE) indicate that women show a weak tendency to the middle category. In the other countries the estimated α-parameters are negative. However, except for Austria and Germany the effects are not significant.

Table 7 compares the performances of the proposed BetaMix model and the simple CUP model when fitting the models with all covariates included for each country. It is seen that for all countries the deviance for the BetaMix model is smaller than for the CUP model. Also, for all countries except for Denmark the AIC values are smaller when fitting the BetaMix model. The largest reduction can be found for Germany (reduction by 42 in the deviance and 30 in the AIC).

Countries	Deviance Uniform	Deviance BetaMix	AIC Uniform	AIC BetaMix
AT	7358	7342	7408	7404
DE	12864	12822	12914	12884
DK	6078	6070	6128	6132
ES	8553	8532	8603	8594
FI	8126	8112	8176	8174
FR	7797	7778	7847	7840
GB	9684	9665	9734	9727
IE	10354	10336	10404	10398
NL	7611	7594	7661	7656
NO	5677	5657	5727	5719
SE	7421	7393	7471	7455

Table 7: *Comparison of CUP and BetaMix models*

6 Concluding Remarks

It has been shown that the modelling of the uncertainty component by a betabinomial distribution yields a more flexible model than traditional mixture models. The shape of the response style is allowed to depend on personal attributes and leads to a better understanding of the concept of uncertainty. The inclusion of covariate effects on the uncertainty also increases the interpretability of the model parameters. It has been demonstrated that ignoring the response style yields biased estimates. The applications demonstrate that the more flexible model outperforms the traditional model in most cases in terms of goodness-of-fit and AIC.

References

Agresti, A. (2010). *Analysis of Ordinal Categorical Data, 2nd Edition*. New York: Wiley.

Agresti, A. (2013). *Categorical Data Analysis, 3d Edition*. New York: Wiley.

Böckenholt, U. (2012). Modeling multiple response processes in judgment and choice. *Psychological Methods*, 665–678.

Bolt, D. M. and T. R. Johnson (2009). Addressing score bias and differential item functioning due to individual differences in response style. *Applied Psychological Measurement*.

Bolt, D. M. and J. R. Newton (2011). Multiscale measurement of extreme response style. *Educational and Psychological Measurement 71*(5), 814–833.

Clarke, I. (2000). Extreme response style in cross-cultural research: An empirical investigation. *Journal of Social Behavior & Personality*.

De Boeck, P. and I. Partchev (2012). Irtrees: Tree-based item response models of the glmm family. *Journal of Statistical Software 48*(1), 1–28.

D'Elia, A. and D. Piccolo (2005). A mixture model for preference data analysis. *Computational Statistics & Data Analysis 49*, 917–934.

Dempster, A. P., N. M. Laird, and D. B. Rubin (1977). Maximum likelihood from incomplete data via the EM algorithm. *Journal of the Royal Statistical Society B 39*, 1–38.

Eid, M. and M. Rauber (2000). Detecting measurement invariance in organizational surveys. *European Journal of Psychological Assessment 16*(1), 20.

Gottard, A., M. Iannario, and D. Piccolo (2016). Varying uncertainty in CUB. *Advances in Data Analysis and Classification 10*(2), 225–244.

Iannario, M. (2012a). Hierarchical cub models for ordinal variables. *Communications in Statistics-Theory and Methods 41*, 3110–3125.

Iannario, M. (2012b). Modelling *shelter* choices in a class of mixture models for ordinal responses. *Statistical Methods and Applications 21*, 1–22.

Iannario, M., M. Manisera, D. Piccolo, and P. Zuccolotto (2012). Sensory analysis in the food industry as a tool for marketing decisions. *Advances in Data Analysis and Classification 6*, 303–321.

Iannario, M. and D. Piccolo (2010). Statistical modelling of subjective survival probabilities. *Genus 66*, 17–42.

Iannario, M. and D. Piccolo (2012a). CUB models: Statistical methods and empirical evidence. In R. Kennett and S. Salini (Eds.), *Modern Analysis of Customer Surveys: with applications using R*, pp. 231–258. New York: Wiley.

Iannario, M. and D. Piccolo (2012b). Investigating and modelling the perception of economic security in the survey of household income and wealth. In M. S. C. Perna (Ed.), *Mathematical and Statistical Methods for Actuarial Sciences and Finance*, pp. 237–244. Berlin: Springer.

Jeon, M. and P. De Boeck (2015). A generalized item response tree model for psychological assessments. *Behavior Research Methods*, 1–16.

Johnson, T. R. (2003). On the use of heterogeneous thresholds ordinal regression models to account for individual differences in response style. *Psychometrika 68*(4), 563–583.

Kankaraš, M. and G. Moors (2009). Measurement equivalence in solidarity attitudes in europe insights from a multiple-group latent-class factor approach. *International Sociology 24*(4), 557–579.

Manisera, M. and P. Zuccolotto (2014). Modeling rating data with nonlinear CUB models. *Computational Statistics & Data Analysis 78*, 100–118.

Marin, G., R. J. Gamba, and B. V. Marin (1992). Extreme response style and acquiescence among hispanics the role of acculturation and education. *Journal of Cross-Cultural Psychology 23*(4), 498–509.

McCullagh, P. (1980). Regression model for ordinal data (with discussion). *Journal of the Royal Statistical Society B 42*, 109–127.

Meisenberg, G. and A. Williams (2008). Are acquiescent and extreme response styles related to low intelligence and education? *Personality and Individual Differences 44*(7), 1539–1550.

Moors, G. (2004). Facts and artefacts in the comparison of attitudes among ethnic minorities. a multigroup latent class structure model with adjustment for response style behavior. *European Sociological Review 20*(4), 303–320.

Moors, G. (2010). Ranking the ratings: A latent-class regression model to control for overall agreement in opinion research. *International Journal of Public Opinion Research 22*, 93–119.

Poessnecker, W. (2015). *MRSP: Multinomial Response Models with Structured Penalties*. R package version 0.6.11.

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

Tutz, G. and M. Berger (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*, 239–268.

Tutz, G., M. Schneider, M. Iannario, and D. Piccolo (2016). Mixture models for ordinal responses to account for uncertainty of choice. *Advances in Data Analysis and Classification, published online*.

Van Herk, H., Y. H. Poortinga, and T. M. Verhallen (2004). Response styles in rating scales evidence of method bias in data from six eu countries. *Journal of Cross-Cultural Psychology 35*(3), 346–360.

Van Rosmalen, J., H. Van Herk, and P. Groenen (2010). Identifying response styles: A latent-class bilinear multinomial logit model. *Journal of Marketing Research 47*, 157–172.

Yee, T. W. (2016). *VGAM: Vector Generalized Linear and Additive Models*. R package version 1.0-3.