Communications for Statistical Applications and Methods
2014, Vol. 21, No. 1, 45–60

DOI: http://dx.doi.org/10.5351/CSAM.2014.21.1.045
ISSN 2287-7843

Korean Welfare Panel Data: A Computational Bayesian Method for Ordered Probit Random Effects Models

Hyejin Lee[a], Minjung Kyung[1,a]

[a]Department of Statistics, Duksung Women's University, Korea

Abstract

We introduce a MCMC sampling for a generalized linear normal random effects model with the ordered probit link function based on latent variables from suitable truncated normal distribution. Such models have proven useful in practice and we have observed numerically reasonable results in the estimation of fixed effects when the random effect term is provided. Applications that utilize Korean Welfare Panel Study data can be difficult to model; subsequently, we find that an ordered probit model with the random effects leads to an improved analyses with more accurate and precise inferences.

Keywords: Ordered probit models, generalized linear mixed models, Gibbs sampling, hierarchical models.

1. Introduction

Generalized linear models (GLMs) have enjoyed considerable attention over the years and have provided a flexible framework to model discrete responses with a variety of error structures. If we have observations that are discrete or categorical, $\boldsymbol{y} = (y_1, \ldots, y_n)$, such data can often be assumed to be independent and from a distribution in the exponential family. The classic book by McCullagh and Nelder (1989) describes these models in detail; see also the more recent developments of Dey, *et al.*(2000) of Fahrmeir and Tutz (2001).

1.1. Generalized linear mixed models

A Generalized Linear Mixed model (GLMM) is an extension of a GLM that allows random effects, and probides flexibility to develop a more suitable model when the observations are correlated, or where there may be other underlying phenomena that contribute to the resulting variability. Thus, the GLMM can be specified to accomodate outcome variables that are conditional on mixtures of possible correlated random and fixed effects (Breslow and Clayton, 1993; Buonaccorsi, 1996; Wang *et al.*, 1998; Wolfinger and O'Connell, 1993). For example, assume a Bernoulli selection process where we observe Y_i according to

$$Y_i \sim \text{ Bernoulli}(p_i), \quad i = 1, \ldots, n,$$

where y_i is 1 or 0, thus $p_i = \mathrm{E(Y_i)}$ is the probability of a success for the i^{th} observation. Moreover, using a link function $g(\cdot)$, we can express the transformed mean as a linear function,

$$g(p_i) = \mathbf{X}_i\boldsymbol{\beta} + \psi_i,$$

This study is supported by Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education, Science and Technology (2012R1A1A1008444).

[1] Corresponding author: Professor, Department of Statistics, Duksung Women's University, 33, Samyangro 144-gil, Dobong Gu, Seoul 132-714, Korea. E-mail: mkyung@duksung.ac.kr

Published online 31 January 2014 / journal homepage: http://csam.or.kr

where $\mathbf{X}_i\boldsymbol{\beta}$ covariates associated with the i^{th} observation, $\boldsymbol{\beta}$ is the coefficient vector, and ψ_i is a random effect accounting for subject-specific deviation from the underlying model. The ψ_is are usually assumed to be distributed as $N(0, \sigma^2_\psi)$. Thus, this is a special case of a GLMM in which the exponential family is Bernoulli. Details of such models, covering both statistical inferences and computational methods, can be found in the recent texts by McCulloch and Searle (2001) and Jiang (2007).

There have been Markov chain Monte Carlo (MCMC) methods developed for the analysis of GLMMs with random effects modeled with normal distribution. The posteriors of parameters and the random effects are typically numerically intractable, especially when the dimension of the random effects is greater than one, however, there has been significant progress in the development of sampling schemes. For example, Damien *et al.* (1999) proposed a Gibbs sampler using auxiliary variables for sampling non-conjugate and hierarchical models. They mentioned that the assessments of convergence remains a major problem with the algorithm. However, their methods are slice sampling methods derived from the full conditional posterior distribution. Neal (2003) provided convergence properties of the posterior for slice sampling. Chib *et al.* (1998) and Chib and Winkelmann (2001) provided Metropolis-Hastings (M-H) algorithms for various kinds of GLMMs. They proposed a multivariate-t distribution as a candidate density in an M-H implementation, taking the mean equal to the posterior mode, and variance equal to the inverse of the Hessian evaluated at the posterior mode.

Gill and Casella (2009) proposed another GLMM variation of that considered latent clustering structure in random effects. Political science data was considered by using a GLMM with an ordered probit link, specifically modelling the stress, from public service, of Senate-confirmed political appointees as a reason for their shout tenure. For the analysis, a semi-parametric Bayesian approach was adopted that used the Dirichlet process instead of the normal distribution to model the random effect.

1.2. Summary

In this paper, we develop algorithms for a normal random effects model with probit link for the ordered categorical responses. We adapt the truncated normal sampling of Albert and Chib (1993) for the *"cutpoints"* between categories. In Section 2, we discuss the generalized linear mixed models for binary and ordinal outcomes. Section 3 describes a Gibbs sampler for the model parameters and the cutpoints between categories. Section 4 contains illustrative applications with a simulation study. In Section 5, a Gibbs sampler is applied to the Korea Welfare Panel Study data set. Section 6 summarizes the contributions and adds some perspective.

2. Generalized Linear Mixed Models with Probit Link Function

Albert and Chib (1993) introduced how truncated normal sampling could be used to implement the Gibbs sampler for a probit model for binary responses. They consider the latent variable such that

$$Z_i = X_i\beta + \psi_i + \epsilon_i, \quad \epsilon_i \sim N\left(0, \sigma^2\right), \tag{2.1}$$

and

$$\begin{cases} y_i = 1, & \text{if } Z_i > 0, \\ y_i = 0, & \text{if } Z_i \leq 0, \end{cases}$$

for $i = 1, \ldots, n$. Where X_i's are covariates associated with the i^{th} observation, $\boldsymbol{\beta}$ is the coefficient vector, and ψ_i denotes a random effect where $N(0, \tau^2)$. Y_i are independent Bernoulli random variables with the probability of success, $p_i = \Phi\left(\mathbf{X}_i\boldsymbol{\beta}/\sigma\right)$. This data augmentation approach provides the same simplification for censored regression models.

2.1. Generalized linear mixed model with probit link function for ordinal outcomes

Albert and Chip (1993), Gill and Casella (2009) showed how truncated normal sampling can be implemented in the Gibbs procedure for the ordinal outcomes with probit link. This model assumes that Y_i takes one of J ordered categories and we observed iid Y_i according to

$$Y_i \sim \text{Multinomial}\left(1, (p_1, p_2, \ldots, p_c)\right), \quad i = 1, \ldots, n, \tag{2.2}$$

where $p_j = P[Y_i = j]$, $j = 1, \ldots, c$, also, the p_j are ordered by a probit model for the random variable Z_i

$$p_j = P\left(\gamma_{j-1} \leq Z_i \leq \gamma_j\right), \tag{2.3}$$

where the "cutpoints" between categories have the property that

$$-\infty = \gamma_0 < \gamma_1 < \cdots < \gamma_c = \infty.$$

For the j^{th} order responses, $Y_i = j$ if $\gamma_{j-1} < Z_i \leq \gamma_j$ with the probability of success, p_j,

$$p_j = \Phi\left(\frac{\gamma_j - X_i\beta - \psi_i}{\sigma}\right) - \left(\frac{\gamma_{j-1} - X_i\beta - \psi_i}{\sigma}\right), \tag{2.4}$$

where Φ is the cumulative function of the standard normal. Without loss of generality, we set $\sigma^2 \equiv 1$. The ordered mulinomial model is written in the form of linking the cumulative response probabilities with the linear regression structure. This data augmentation method can be combined easily with the Gibbs sampling outlined in the next section.

3. Sampling Schemes for the Ordered Probit Random Effects Models

An overview of the general sampling scheme is as follows. We have group parameters

(i) $\mathbf{Z}$, the latent variables

(ii) $(\beta, \boldsymbol{\psi}, \tau^2.\gamma)$, the model parameters.

We iterate between these parameters until convergence:

1. Conditional on, $\beta, \boldsymbol{\psi}, \tau^2, \gamma$, generate $\mathbf{Z}$, the new latent variables.
2. Conditional on, $\beta, \boldsymbol{\psi}, \tau^2, \mathbf{Z}$, generate γ, the new cutpoints.
3. Conditional on, $\beta, \boldsymbol{\psi}, \gamma, \mathbf{Z}$, generate τ^2, the new parameters of random effects.
4. Conditional on, $\beta, \tau^2, \gamma, \mathbf{Z}$, generate $\boldsymbol{\psi}$, the new random effects.
5. Conditional on, $\boldsymbol{\psi}, \tau^2, \gamma, \mathbf{Z}$, generate β, the new coefficients.

For the model parameters, $\left(\beta, \boldsymbol{\psi}, \tau^2, \gamma\right)$, we add the priors

$$\begin{aligned} \beta|\sigma^2 &\sim N\left(\mathbf{0}, d\sigma^2 I\right) \\ \boldsymbol{\psi} &\sim N\left(\mathbf{0}, \tau^2 I\right) \\ \tau^2 &\sim \text{IG}(a, b) \end{aligned} \tag{3.1}$$

where $d > 1$ and $\mathrm{IG}(a, b)$ is inverted gamma distribution with parameters a and b. We will fix a value σ^2 and prior of γ assumed a diffuse prior.

For the ordinary responses with probit link, the likelihood function of model parameters and the latent variable in (2.1) is given by

$$L\left(\beta, \boldsymbol{\psi}, \tau^2, \gamma, \mathbf{Z}|\mathbf{y}\right) = \prod_{i=1}^{n}\left[\frac{1}{\sqrt{2\pi\sigma^2}}\exp\left\{-\frac{1}{2\sigma^2}\left(Z_i - X_i\beta - \psi_i\right)\right\}\right.$$
$$\left.\times\ \frac{1}{\sqrt{2\pi\tau^2}}\exp\left\{-\frac{1}{2\tau^2}\boldsymbol{\psi}^T\boldsymbol{\psi}\right\} \times \left\{\sum_{j=1}^{J} I(Y_i = j) I\left(\gamma_{j-1} < Z_i < \gamma_j\right)\right\}\right], \tag{3.2}$$

where $\mathbf{Z} = (Z_1, Z_2, \ldots, Z_n)$ is a vector of latent variables and $\boldsymbol{\psi} = (\psi_1, \psi_2, \ldots, \psi_n)$ is a vector of subject specific random effect variables.

With priors in (3.1), the full conditional distributions of $(\boldsymbol{\beta}, \boldsymbol{\psi}, \tau^2, \gamma, \mathbf{Z})$ are given by

$$\begin{aligned}
\boldsymbol{\beta}|\boldsymbol{\psi}, \tau^2, \gamma, \mathbf{Z}, \mathbf{y} &\sim N_p\left(\tilde{\boldsymbol{\beta}}, \sigma^2\Sigma_\beta\right)\\
\boldsymbol{\psi}|\boldsymbol{\beta}, \tau^2, \gamma, \mathbf{Z}, \mathbf{y} &\sim N_n\left(\tilde{\boldsymbol{\psi}}, \Sigma_\psi\right)\\
\tau^2|\boldsymbol{\beta}, \boldsymbol{\psi}, \gamma, \mathbf{Z}, \mathbf{y} &\sim \mathrm{IG}\left(\frac{n}{2} + a, \frac{1}{2}\boldsymbol{\psi}^T\boldsymbol{\psi} + b\right)\\
Z_i|\boldsymbol{\beta}, \boldsymbol{\psi}, \tau^2, \gamma, y_i = j &\sim N\left(\mathbf{X}_i\boldsymbol{\beta} + \psi_i, \sigma^2\right) \text{ truncated at the left and right by } \gamma_{j-1} \text{ and } \gamma_j,
\end{aligned} \tag{3.3}$$

where

$$\begin{aligned}
\Sigma_\beta^{-1} &= \frac{I}{d} + \mathbf{X}^T\mathbf{X}, & \tilde{\beta} &= \left(\frac{I}{d} + \mathbf{X}^T\mathbf{X}\right)^{-1}\mathbf{X}^T,\\
\Sigma_\psi^{-1} &= \left(I\frac{1}{\sigma^2} + \frac{1}{\tau^2}\right), & \tilde{\psi} &= \left\{\left(\frac{1}{\sigma^2} + \frac{1}{\tau^2}\right)\right\}^{-1}\frac{1}{\sigma^2}(\mathbf{Z} - \mathbf{X}\beta).
\end{aligned} \tag{3.4}$$

The full conditional density of γ_j given $\beta, \boldsymbol{\psi}, \tau^2, \mathbf{Z}, \mathbf{y}$ and $\{\gamma_k, k \neq j\}$ is given by

$$\pi\left(\gamma_j|\beta, \boldsymbol{\psi}, \tau^2, \mathbf{Z}, \mathbf{y}\right) \propto \prod_{i=1}^{n} 1(Y_i = j) 1\left(\gamma_{j-1} < Z_i < \gamma_j\right) + 1\left(Y_i = j + 1\right) 1\left(\gamma_j < Z_i < \gamma_{j+1}\right). \tag{3.5}$$

This conditional distribution can be seen to be uniform on the interval $[\max\{\max(Z_i : Y_i = j), \gamma_{j-1}\}, \min\min(Z_i : Y_i = j + 1), \gamma_{j+1}\}]$ (Albert and Chib, 1993). Details are in Appendix A.

4. Simulation Study

We conduct a simulation study to evaluate the procedure for the ordered categorical outcomes with random effects.

Using the GLMM with the ordered probit link function, we generated $n = 100$ samples with fixed parameters $\beta = (1, 2, 3)$, $\tau^2 = 2$ and cutpoints $\gamma = (-5, 0, 5, 10)$. We then generated X_1 and X_2 independently from $N(0, 1)$, and used the fixed design matrix to generate the categorical outcome Y. The Gibbs sampler was repeated 10000 times and saved 5000 draws for the posterior inference.

Table 1: Estimation of coefficients of the GLM, GLMM with ordered probit link and `MCMCoprobit`.

Parameter	True value	Mean(SD)			C.I.95%		
		GLM	GLMM	MCMCoprobit	GLM	GLMM	MCMCoprobit
β_0	1	-1.483 (1.856)	0.390 (1.870)	2.204 (0.251)	-4.106, 1.819	-3.397, 3.469	1.737, 2.595
β_1	2	1.570 (0.265)	2.451 (0.523)	1.030 (0.153)	1.107, 2.150	1.628, 3.693	0.736, 1.340
β_2	3	2.297 (0.326)	3.609 (0.759)	1.611 (0.149)	1.733, 3.056	2.498, 5.523	1.324, 1.907
γ_1	-5	-5.312 (1.869)	-5.665 (2.516)	-	-8.032, -1.604	-10.288, -1.761	-
γ_2	0	-1.990 (1.848)	-0.437 (1.939)	2.047 (0.370)	-4.260, 1.306	-4.260, 2.707	1.605, 2.369
γ_3	5	1.520 (1.995)	5.169 (1.984)	4.383 (0.254)	-1.520, 4.792	0.999, 8.374	4.033, 4.817
γ_4	10	5.574 (2.220)	11.453 (2.814)	9.944 (0.680)	1.911, 9.598	6.220, 17.229	8.900, 10.525
τ_2	2	-	1.622 (1.277)	-	-	0.376, 5.169	-

4.1. Ordered probit models

We compare the GLMM to the GLM of ordered probit link with the prior distribution of β from $\beta|\sigma^2 \sim N(0, d\sigma^2 I)$, where d is 5. $\pi(\tau^2) \sim \mathrm{IG}(3,2)$ and $\pi(\gamma) \propto c$, a flat prior for cutpoints. The starting points of β's are set to the maximum likelihood (ML) estimates, γ's are $(-10, 0, 5, 10)$, and τ^2 is 2. In addition, we compare the proposed sampling to the results from `MCMCoprobit` function in MCMCpackages of `R` (Andrew *et al.*, 2011). All summaries in the tables are posterior means and standard deviations calculated from the Gibbs sampler after burnin.

Table 1 provides the numerical summary of this process. The resulting estimates for βs are $(-1.483, 1.570, 2.297)$ with standard error $(1.856, 0.262, 0.326)$ from GLM and $(0.390, 2.451, 3.609)$ with standard error $(1.870, 0.523, 0.759)$ from GLMM. The resulting estimates for βs from GLMM are fairly closer to the true values, $(1, 2, 3)$, than those from GLM, although the standard errors for GLMM are slightly larger than that of GLM. As well as, the resulting estimates of `MCMCoprobit` are $(2.204, 1.030, 1.611)$ with standard error $(0.251, 0.153, 0.149)$. It seems that the `MCMCoprobit` output does not adequately estimate the true value. In addtion, the 95% Credible Intervals (CIs) of βs do not contain the true values, $(1, 2, 3)$. Note that our proposed sampling estimates the cutpoints γs well compared to other methods. The point estimates from the Gibbs sampler are the posterior means of γ's that is $(-5.312, -1.990, 1.520, 5.574)$ from GLM and $(-5.665, -0.437, 5.169, 11.453)$ from GLMM. It means that the resulting of γs from GLMM are numerically closed to the true value. Therefore, GLM is not enough to detect the true cutpoints, and the random effects models captures heterogeneous hidden structure. Especially, the 95% Credible Intervals (CIs) of γ_3 and γ_4 based on the GLM do not contain the true values, 5 and 10. `MCMCoprobit` provides two ways to sample the cutpoints, Cowles sampling, and Albert and Chib sampling. We compare our proposed GLMM sampling to the results from `MCMCoprobit` based on the Albert and Chib method. `MCMCoprobit` sampling does not consider the random effect terms and it does not provide the posterior estimates of γ_1 because this function conducts the first element is normalized to zero. The posterior means of γs are $(2.047, 4.383, 9.944)$; therefore, the cutpoints are unclear.

In this regard, as we expected, the our proposed sampling based on the GLMM is more reasonable compared to GLM performance and existing sampling methods, `MCMCoprobit`. To evaluate the convergence of components, we consider the autocorrelation function (ACF) plots given in Figure A.1 in

Table 2: Estimate of the coefficients from the GLM, GLMM and `MCMCoprobit` with ordered probit link by thinning.

Parameter	True value	Mean(SD)			C.I.95%		
		GLM	GLMM	MCMCoprobit	GLM	GLMM	MCMCoprobit
β_0	1	−1.510 (1.824)	0.370 (1.868)	2.212 (0.254)	−4.017, 1.734	−3.321, 3.469	1.736, 2.580
β_1	2	1.572 (0.239)	2.432 (0.534)	1.031 (0.138)	1.221, 2.026	1.670, 3.626	0.737, 1.255
β_2	3	2.293 (0.312)	3.543 (0.804)	1.601 (0.142)	1.844, 2.909	2.505, 5.651	1.327, 1.841
γ_1	−5	−5.349 (1.895)	−5.703 (2.424)	-	−7.896, −1.594	−10.197, −1.910	-
γ_2	0	−2.023 (1.798)	−0.454 (1.980)	2.055 (0.373)	−4.358, 1.239	−4.073, 2.738	1.605, 2.369
γ_3	5	1.501 (1.971)	5.100 (2.075)	4.383 (0.254)	−1.245, 4.760	1.069, 8.450	4.033, 4.817
γ_4	10	5.491 (2.257)	11.449 (2.823)	9.960 (0.684)	2.178, 9.428	6.732, 16.272	8.900, 10.525
τ_2	2	-	1.541 (1.027)	-	-	0.473, 3.830	-

Appendix B. We observe that the MCMC samples of βs and γs from a latent variable Gibbs sampler exhibits a strong autocorrelation; consequently, we used thinning to reduce autocorrelation.

4.2. Thinning step

We considered thinning by taking every 100th value the last 5000 draws. Figure A.2 in Appendix B is the ACF plot of βs and γs after the thinning process and we observe that the autocorrelations are reduced. Table 2 provides the numerical summary. The resulting estimates for βs are $(-1.510, 1.572, 2.293)$ from GLM, $(0.370, 2.432, 3.543)$ from GLMM and $(2.212, 1.031, 1.601)$ from `MCMCoprobit`.

Similar to the above outcome, the GLMM estimates of βs are numerically closer to the true values than the GLM and the `MCMCoprobit`. Likewise, in the case of γs, the GLMM estimates are closer to the true values such that the posterior means of s are $(-5.349, -2.023, 1.501, 5.491)$ from GLM, $(-5.703, -0.454, 5.100, 11.449)$ from GLMM and $(2.055, 4.383, 9.960)$ from `MCMCoprobit`. Considering 95% equal tail credible interval for γ_3 and γ_4 of the GLM, the credible interval does not contain the true value. Subsequently, the resulting estimates based on the GLMMare better than the GLM and `MCMCoprobit` even after the thinning step.

5. Data Analysis

We consider a real data for the performance of the GLMM with an ordered probit link function and compare the GLMM and the GLM with ordered probit link function. The Korea Welfare Panel Study (KWPS) since 2006 developed by the Korean Institute of Social and Health Affairs in conjunction with Social Welfare Research Institute of Seoul National University. It covers every Korean region and is researched by household type. KWPS provides information about household characteristic, household economic status, and the economic activities of household members. Also, the supply-demand situation and the supply-demand need of social welfare systems are also considered. The general focus is on household economy instability based on the deterioration of poverty and social polarization and the growth of irregular workers and youth unemployment. The goal of the study is contribute to policy and promote a policy-effect by grasping the dynamic change in household type, income, and employment of the poor and near-poor. In addtion, it is expected to contribute to policy-making and feed-back through understanding the dynamic conditions of living, the welfare supply-demand, welfare needs of each population group by income, economic activities, age and assessing policy effect.

The data used here is about the self-perceived health status of 20 64 age group based on the 2012 Korean Welfare Panel Study. Fehir (1988) defined that self-perceived health status is the absence of disease, a state of well-being and the capability to function in the face of changing circumstances. Thus the self-perceived health status has advantages in terms of health promotion. It covers the physical, mental, cognition and social spectrum as a tool for individual health status which cannot be measured by medical methods (Ware, 1987; Oh *et al.*, 2006; Lee and Kim, 2013); in addtion, self-perceived health status is that it affects the death rate (Hoeymans *et al.*, 1997; Scott *et al.*, 1997).

The elderly have a self-perceived health status that usually recognizes that their health condition deteriorates with age by Luoh and Herzog (2002) and Lee and Kim (2013). However, Stoller (1984) and Oh *et al.* (2006) discussed that the elderly were more positive about the health condition as age increases. In addition, there is a different result not related to age self-perceived health status (Hoeymans *et al.*, 1997). For the sex-related opinion, there exist different points of view in that elderly women have a poorer objectively assessed health status for a given self-assessment of health than elderly men (Fillenbaum, 1979). Income and education level have a significant effect on self-perceived health status. The higher income or the higher education level, the better self-perceived health (von dem Knesebeck *et al.*, 2003; Luoh and Herzog, 2002; Otiniano *et al.*, 2003). Functional status measures the functional limitations or disabilities and is an important determinant for self-rated health in the elderly and affects self-rated health negatively (Hoeymans *et al.*, 1997).

The self-perceived health status study in elderly has been proceeded widely and frequently in the social sciences compared to the study in 20–64 age groups. This is because most studies are focused on the population aging and healthcare. In addition, not many research tracks have used KWPS data for self-perceived health status in Korea. Thus, we focus on the analyses of 20–64 age group based on some factors instead because the 20–64 age group is the economically productive population.

The KWPS data has 7690 cases in 20–64 age groups in the year of 2012. Our response variable is the answer to how they feel about their health status. This is measured on a five-point Likert scale: 1. excellent, 2. good, 3. so-so, 4. not so good, 5. bad. For the data analyses, we coded in reverse scale. In order to understand the implications of the questions above, many factors from this survey have been considered. For the supplemental variables, we considered `sex` (0 = female, 1 = male), `age` (20–64), educational level (`edu`), marital status (`marry`) (1 = marry, 2 = bereavement, 3 = divorce, 4 = separation, 5 = single, 6 = etc.), `religion` (0 = no, 1 = yes) and residential region (`region`) (1 = Seoul, 2 = Incheon/Gyeonggi, 3 = Busan/Gyeongsangnam/Ulsan, 4 = Daegu/Gyeongsangbuk, 5 = Daejeon/Chungcheongnam, 6 = Gangwon/Chungcheongbuk, 7 = Gwangju/Jeollanam/Jeollabuk/Jeju). As candidates of important factors considered for our response variable were `dis_c` for the disability (0 = non disability, 1 = disability), `trt_n` the number of outpatient medical examination and treatment during survey year, `cho_d` a dichotomous variable indicating that respondents have chronic disease or not (0 = no, 1 = yes), and indication of lower income (`income`) (0 = general, 1 = low income).

5.1. Analysis results

We ran the Markov chain for 50,000 iterations and saved the last 25,000 for Bayesian analysis. Table 3 provides estimates of the coefficients and cutpoints from two approaches: The GLM and The GLMM

Table 3: Ordered probit models for KWPS data.

Coefficient	GLM Probit			GLMM Probit		
	Mean	SD	95%CI	Mean	SD	95%CI
Intercept	7.807	0.118	7.576, 8.038	9.411	0.311	8.799, 10.020
sex	0.123	0.027	0.069, 0.175	0.322	0.071	0.181, 0.460
age	−0.018	0.002	−0.021, −0.015	−0.047	0.005	−0.056, −0.038
ede	0.054	0.012	0.030, 0.078	0.154	0.033	0.090, 0.219
dis_c	−0.558	0.051	−0.658, −0.457	−1.480	0.141	−1.761, −1.208
marry	−0.048	0.010	−0.067, −0.029	−0.123	0.025	−0.172, −0.074
religion	0.031	0.027	−0.020, 0.083	0.081	0.070	−0.058, 0.218
trt_n	−0.009	0.001	−0.010, −0.007	−0.023	0.002	−0.028, −0.019
cho_c	−0.260	0.012	−0.284, −0.236	−0.694	0.035	−0.766, −0.627
region	0.017	0.006	0.004, 0.030	0.047	0.018	0.013, 0.080
income	−0.408	0.037	−0.481, −0.337	−1.088	0.102	−1.291, −0.887
(bad)(not so good)	3.724	0.058	3.603, 3.834	−1.354	0.296	−1.942, −0.847
(not so good)(so so)	5.129	0.031	5.067, 5.178	2.383	0.143	2.107, 2.637
(so so)(good)	5.969	0.016	5.937, 5.999	4.649	0.073	4.520, 4.768
(good)(excellent)	7.938	0.014	7.907, 7.964	9.881	0.059	9.809, 10.038
τ^2	-	-	-	6.068	0.383	5.457, 6.899

with ordered probit link function from the Gibbs sampler.

We observe that standard errors of the coefficients from the GLMM tend to be numerically larger than GLM while there are no changes in signs; however, resulting cutpoints are interesting. Based on the posterior mean of beta, we predict the values of the latent variable Z in (2.3).

Then based on the posterior means of the cutpoints, we categorize Z's into 5 groups. The posterior means of cutpoints are $(3.724, 5.129, 5.969, 7.938)$ from GLM and $(-1.354, 2.383, 4.649, 9.881)$ from GLMM. The results indicate that the boundaries of cutpoints are ambiguous from GLM results, (not good)/(so) and (good)/(so), are especially vague and the difference between them is about 1. The results of cutpoints from GLM are a biased positive side. In this regard, from the simulation study, we notice that the GLMM is more accurate than the GLM for data analysis assessment. This is because there exist hidden aspects in the data detected by a random effect. Thus it provides a more accurate prediction for self-perceived health status compared to the model without a random effect term.

The convergence of coefficients of GLM and GLMM can be assessed form the ACF plots in Figure A.3 in Appendix B. The plots indicate that the samples of coefficients from GLMM exhibits a stronger autocorrelation than GLM and might be the reason for the estimation of the random effect terms. We thinned every 100th MCMC sample because of the strong autocorrelation; however, we use original 25000 samples after burnin because the estimates are not different from the output without thinning. Tables and Figures of thinned data are omitted. Table 3 shows that there are few changes in statistical reliability and no changes in the estimated coefficients (posterior means) sign among the two models; in addtion, the magnitudes of the effects are uniformly larger with the enhanced model GLMM compared to the standard GLM. This indicates that there exist extra information in the data that is detected by the random effect. Self-perceived health status is reliable for different regions. It seems that respondent who live in southern areas tend to positively answer questions on health status and this may be due to better circumstances in southern areas that improve health status. Also singles tend to recognize that their health status is not good compared to the married couples. Males seems to be more positive for the subjective health status than females in the 20–64 age group. In the cases of other coefficients, younger age, high educational level, non disability, the low frequency of outpatient medical, non-chronic disease, and the better financial status group tend to answer positively for the self-perceived health status. While the above variables are estimated to affect to the self-perceived health status, the religion seems unreliable, because 95% credible interval includes 0 values in it.

6. Discussion

This paper considered Bayesian Gibbs sampling schemes for the orderly categorized outcomes considering random effects. We studied the properties of Markov Chain Monte Carlo estimation tools for the GLMM with ordered probit link, noting that they incorporate latent information such that models fit better in the presence of unexplained heterogeneity in the data. This turns out to be ideal to analyze of survey data that has diverse hidden information.

The random effect models show that the cutpoint boundaries from GLMM are more accurate and precisely estimated than GLM based on simulation and data analysis. This is because that there exist the hidden aspects in the data detected by a random effect and it provides more accurate prediction for the self-perceived health status compared to the model without random effect term in 2012 KWPS data of 20–62 age group. However, the convergence of the GLMM parameters are slower. It is because that the model with random effect is a richer model and the random effect cannot be checked but includes hidden and structred information. The existence of a set of statistically reliable coefficient estimates in the presence of substantial heterogeneity in the actors studied shows that the GLMM model can uncover common patterns by modelling the unexplained variance as a latent information. This turns out to be ideal for the empirical analysis of KWPS data, which has diverse operators to generate heterogenous data.

In simulation study, to compare our computation, we compared our Bayesian sampling with random effects to the Bayesian sampling based on `MCMCoprobit` in R. `MCMCoprobit` is a sampling based on the GLM without random effect and does not provide a first cutpoint because it is normalized to zero on the first element. The results show that researchers face difficulties to interpret exact cutpoints. The GLM is not enough to detect the heterogeneous hidden structure even though the GLMM with the ordered probit link is not totally new. The GLMM clearly detects the cutpoint compared to the regular GLM because of the subject specific random effect. In addition, there are not available Bayesian sampling tools for the GLM with random effects with detect the interpretable cutpoints are unabailable. Therefore our proposed sampling method provide more reasonable estimates for the model parameters and the cut points.

Finally, we note that although we have concentrated on ordered probit models, the results will apply directly to a wider class of generalized linear mixed models. There are implementation problems with the Gibbs sampler that arise with models such as the logit, where one needs to use either a slice sampler or a Metropolis-Hastings step. KWPS data has been collected since 2006 and it is reasonable to consider the full data in a longitudinal set up with random effects that capture the hidden structure to contribute to the variability.

Appendix: Model parameters

Appendix A. Generating the Model Parameters

With prior in (3.1) and likelihood function in (3.2), a full conditional Gibbs sampler of $(\boldsymbol{\beta}, \boldsymbol{\psi}, \tau^2, \gamma, \mathbf{Z})$ are following.

The conditional posterior distribution of $\boldsymbol{\beta}$ is

$$
\begin{aligned}
&\pi\left(\boldsymbol{\beta}|\boldsymbol{\psi}, \tau^2, \gamma, \mathbf{Z}, \mathbf{y}\right)\\
&\quad\propto \exp\left[-\frac{1}{2\sigma^2}\left\{(\mathbf{Z}-\mathbf{X}\boldsymbol{\beta}-\boldsymbol{\psi})^T(\mathbf{Z}-\mathbf{X}\boldsymbol{\beta}-\boldsymbol{\psi})+\frac{1}{d}\boldsymbol{\beta}^T\boldsymbol{\beta}\right\}\right]\\
&\quad\propto \exp\left[-\frac{1}{2\sigma^2}\left(\boldsymbol{\beta}-\tilde{\boldsymbol{\beta}}\right)^T\left(\frac{\mathbf{I}}{d}+\mathbf{X}^T\mathbf{X}\right)\left(\boldsymbol{\beta}-\tilde{\boldsymbol{\beta}}\right)\right]\\
&\qquad\times\exp\left[\frac{1}{2\sigma^2}\left\{\mathbf{X}^T(\mathbf{Z}-\boldsymbol{\psi})^T\right\}^T\left(\frac{\mathbf{I}}{d}+\mathbf{X}^T\mathbf{X}\right)^{-1}\left\{\mathbf{X}^T(\mathbf{Z}-\boldsymbol{\psi})^T\right\}-\frac{1}{2\sigma^2}\left\{(\mathbf{Z}-\psi)^T(\mathbf{Z}-\boldsymbol{\psi})\right\}\right].
\end{aligned}
$$

Here,

$$
\begin{aligned}
&\{(\mathbf{Z}-\boldsymbol{\psi})-\mathbf{X}\boldsymbol{\beta}\}^T\{(\mathbf{Z}-\boldsymbol{\psi})-\mathbf{X}\boldsymbol{\beta}\}+\frac{1}{d}\boldsymbol{\beta}^T\boldsymbol{\beta}\\
&=(\mathbf{Z}-\psi)^T(\mathbf{Z}-\boldsymbol{\psi})-2\boldsymbol{\beta}^T\mathbf{X}^T(\mathbf{Z}-\boldsymbol{\psi})+\boldsymbol{\beta}^T\mathbf{X}^T\mathbf{X}\boldsymbol{\beta}+\frac{1}{d}\boldsymbol{\beta}^T\boldsymbol{\beta}\\
&=\boldsymbol{\beta}^T\left(\frac{\mathbf{I}}{d}+\mathbf{X}^T\mathbf{X}\right)\boldsymbol{\beta}-2\boldsymbol{\beta}^T\left\{\mathbf{X}^T(\mathbf{Z}-\boldsymbol{\psi})\right\}+\left\{\mathbf{X}^T(\mathbf{Z}-\boldsymbol{\psi})\right\}\left(\frac{\mathbf{I}}{d}+\mathbf{X}^T\mathbf{X}\right)^{-1}\left(\frac{\mathbf{I}}{d}+\mathbf{X}^T\mathbf{X}\right)\\
&\quad\times\left(\frac{\mathbf{I}}{d}+\mathbf{X}^T\mathbf{X}\right)^{-1}\left\{\mathbf{X}^T(\mathbf{Z}-\boldsymbol{\psi})\right\}\\
&\quad-\left\{\mathbf{X}^T(\mathbf{Z}-\boldsymbol{\psi})\right\}\left(\frac{\mathbf{I}}{d}+\mathbf{X}^T\mathbf{X}\right)^{-1}\left(\frac{\mathbf{I}}{d}+\mathbf{X}^T\mathbf{X}\right)\times\left(\frac{\mathbf{I}}{d}+\mathbf{X}^T\mathbf{X}\right)^{-1}\left\{\mathbf{X}^T(\mathbf{Z}-\boldsymbol{\psi})\right\}+(\mathbf{Z}-\boldsymbol{\psi})^T(\mathbf{Z}-\boldsymbol{\psi})\\
&=\left(\boldsymbol{\beta}-\tilde{\boldsymbol{\beta}}\right)^T\left(\frac{\mathbf{I}}{d}+\mathbf{X}^T\mathbf{X}\right)\left(\boldsymbol{\beta}-\tilde{\boldsymbol{\beta}}\right)-\left\{\mathbf{X}^T(\mathbf{Z}-\boldsymbol{\psi})\right\}^T\left(\frac{\mathbf{I}}{d}+\mathbf{X}^T\mathbf{X}\right)^{-1}\left\{\mathbf{X}^T(\mathbf{Z}-\boldsymbol{\psi})\right\}+(\mathbf{Z}-\boldsymbol{\psi})^T(\mathbf{Z}-\boldsymbol{\psi}).
\end{aligned}
$$

Thus,

$$\therefore\ \boldsymbol{\beta}|\boldsymbol{\psi}, \tau^2, \gamma, \mathbf{Z}, \mathbf{y} \sim N\left(\tilde{\boldsymbol{\beta}}, \sigma^2\Sigma_\beta\right),$$

where $\Sigma_\beta^{-1} = \mathbf{I}/d + \mathbf{X}^T\mathbf{X}$, $\tilde{\boldsymbol{\beta}} = (\mathbf{I}/d + \mathbf{X}^T\mathbf{X})^{-1}\{\mathbf{X}^T(\mathbf{Z}-\boldsymbol{\psi})\}$.

The conditional posterior distribution of $\boldsymbol{\psi}$ is

$$
\begin{aligned}
&\pi\left(\boldsymbol{\psi}|\boldsymbol{\beta}, \tau^2, \gamma, \mathbf{Z}, \mathbf{y}\right)\\
&\quad\propto \exp\left[-\frac{1}{2}\left\{\frac{1}{\sigma^2}((\mathbf{Z}-\mathbf{X}\boldsymbol{\beta})\boldsymbol{\psi})^T((\mathbf{Z}-\mathbf{X}\boldsymbol{\beta})\boldsymbol{\psi})+\frac{1}{\tau^2}\boldsymbol{\psi}^T\boldsymbol{\psi}\right\}\right]\\
&\quad\propto \exp\left[-\frac{1}{2}\left\{\left(\boldsymbol{\psi}-\tilde{\boldsymbol{\psi}}\right)^T\left[\left(\frac{1}{\sigma^2}+\frac{1}{\tau^2}\right)\mathbf{I}\right]\left(\boldsymbol{\psi}-\tilde{\boldsymbol{\psi}}\right)\right\}\right]\\
&\qquad\times\exp\left[\frac{1}{2}\left\{\frac{1}{\sigma^2}\mathbf{I}(\mathbf{Z}-\mathbf{X}\boldsymbol{\beta})\right\}^T\left(\frac{1}{\sigma^2}+\frac{1}{\tau^2}\right)^{-1}\mathbf{I}\left\{\frac{1}{\sigma^2}\mathbf{I}(\mathbf{Z}-\mathbf{X}\boldsymbol{\beta})\right\}-\frac{1}{2\sigma^2}(\mathbf{Z}-\mathbf{X}\boldsymbol{\beta})^T(\mathbf{Z}-\mathbf{X}\boldsymbol{\beta})\right].
\end{aligned}
$$

Here,

$$
\begin{aligned}
&\frac{1}{\sigma^2}\{(\mathbf{Z}-\mathbf{X}\boldsymbol{\beta})-\boldsymbol{\psi}\}^T\{(\mathbf{Z}-\mathbf{X}\boldsymbol{\beta})-\boldsymbol{\psi}\}+\frac{1}{\tau^2}\boldsymbol{\psi}^T\boldsymbol{\psi}\\
&=\frac{1}{\sigma^2}\left\{(\mathbf{Z}-\mathbf{X}\boldsymbol{\beta})^T(\mathbf{Z}-\mathbf{X}\boldsymbol{\beta})-2\boldsymbol{\psi}^T(\mathbf{y}-\mathbf{X}\boldsymbol{\beta})+\boldsymbol{\psi}^T\boldsymbol{\psi}\right\}+\frac{1}{\tau^2}\boldsymbol{\psi}^T\boldsymbol{\psi}\\
&=\boldsymbol{\psi}^T\left\{\left(\frac{1}{\sigma^2}+\frac{1}{\tau^2}\right)\mathbf{I}\right\}\boldsymbol{\psi}-2\boldsymbol{\psi}^T\frac{1}{\sigma^2}(\mathbf{Z}-\mathbf{X}\boldsymbol{\beta})\\
&\quad+\left\{\frac{1}{\sigma^2}(\mathbf{Z}-\mathbf{X}\boldsymbol{\beta})\right\}^T\left\{\left(\frac{1}{\sigma^2}+\frac{1}{\tau^2}\right)\mathbf{I}\right\}^{-1}\left\{\left(\frac{1}{\sigma^2}+\frac{1}{\tau^2}\right)\mathbf{I}\right\}\left\{\left(\frac{1}{\sigma^2}+\frac{1}{\tau^2}\right)\mathbf{I}\right\}^{-1}\left\{\frac{1}{\sigma^2}(\mathbf{Z}-\mathbf{X}\boldsymbol{\beta})\right\}\\
&\quad+\frac{1}{\sigma^2}(\mathbf{Z}-\mathbf{X}\boldsymbol{\beta})^T(\mathbf{Z}-\mathbf{X}\boldsymbol{\beta})\\
&=\left(\boldsymbol{\psi}-\tilde{\boldsymbol{\psi}}\right)^T\left\{\left(\frac{1}{\sigma^2}+\frac{1}{\tau^2}\right)\mathbf{I}\right\}\left(\boldsymbol{\psi}-\tilde{\boldsymbol{\psi}}\right)-\left\{\frac{1}{\sigma^2}(\mathbf{Z}-\mathbf{X}\boldsymbol{\beta})\right\}^T\left\{\left(\frac{1}{\sigma^2}+\frac{1}{\tau^2}\right)\mathbf{I}\right\}^{-1}\left\{\frac{1}{\sigma^2}(\mathbf{Z}-\mathbf{X}\boldsymbol{\beta})\right\}\\
&\quad+\frac{1}{\sigma^2}(\mathbf{y}-\mathbf{X}\boldsymbol{\beta})^T(\mathbf{y}-\mathbf{X}\boldsymbol{\beta}).
\end{aligned}
$$

Thus,

$$\therefore\ \boldsymbol{\psi}|\boldsymbol{\beta},\tau^2,\gamma,\mathbf{Z},\mathbf{y}\sim N\left(\tilde{\boldsymbol{\psi}},\Sigma_\psi\right),$$

where $\Sigma_\psi^{-1}=(1/\sigma^2+1/\tau^2)\mathbf{I}$, $\tilde{\psi}=\{(1/\sigma^2+1/\tau^2)\mathbf{I}\}^{-1}(1/\sigma^2)(\mathbf{Z}-\mathbf{X}\boldsymbol{\beta})$.

The conditional posterior distribution of $\mathbf{Z}$ is

$$
\begin{aligned}
\pi\left(Z_i|\boldsymbol{\beta},\boldsymbol{\psi},\tau^2,\gamma,y_i=j\right)\propto\prod_{i=1}^{n}&\left\{1\left(Y_i=j\right)1\left(\gamma_{j-1}<Z_i<\gamma_j\right)+1\left(Y_i=j+1\right)1\left(\gamma_j<Z_i<\gamma_{j+1}\right)\right\}\\
&\times\exp\left[-\frac{1}{2\sigma^2}\left(Z_i-\mathbf{X}_i\boldsymbol{\beta}-\psi_i\right)^2\right].
\end{aligned}
$$

Thus,

$$Z_i|\boldsymbol{\beta},\boldsymbol{\psi},\tau^2,\gamma,y_j=j\sim N\left(\mathbf{X}_i\boldsymbol{\beta}+\psi_i,\sigma^2\right)\ \text{ truncated at the left (right) by } \gamma_j-1(\gamma_j).$$

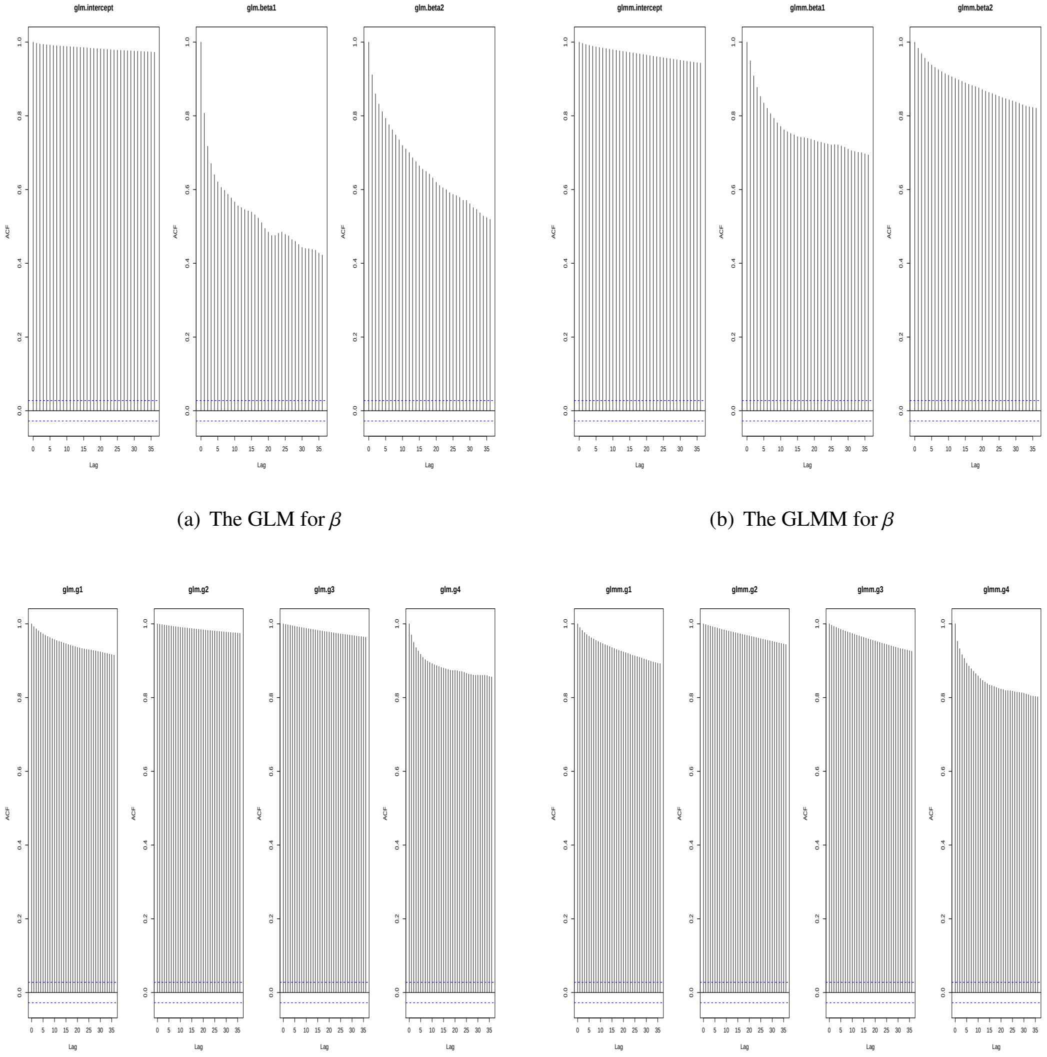
Finally, the fully conditional density of γ_j is

$$\pi\left(\gamma_j|\boldsymbol{\beta},\boldsymbol{\psi},\tau^2,\mathbf{Z},\mathbf{y}\right)\propto\prod_{i=1}^{n}\left\{1(Y_i=j)1\left(\gamma_{j-1}<Z_i<\gamma_j\right)+1(Y_i=j+1)1\left(\gamma_j<Z_i<\gamma_{j+1}\right)\right\}.$$

This conditional distribution can be seen to be uniform on the interval

$$\left[\max\left\{\max(Z_i:Y_i=j),\gamma_{j-1}\right\},\min\left\{\min(Z_i:Y_i=j+1),\gamma_{j+1}\right\}\right].$$

Appendix B. Autocorrelation Function (ACF) Plots



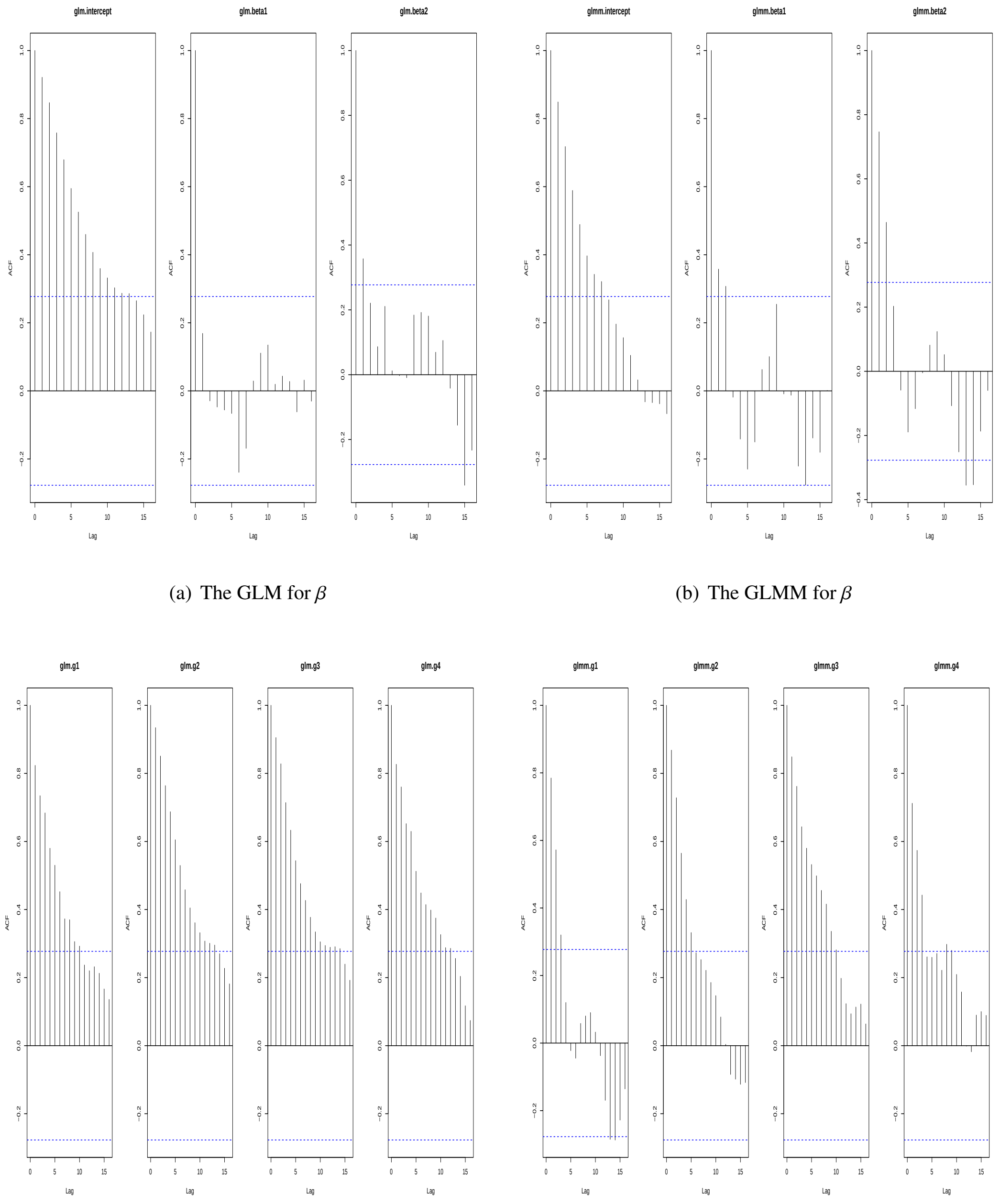
(a) The GLM for β

(b) The GLMM for β

(c) The GLM for cutpoints

(d) The GLMM for cutpoints

Figure A.1: *ACF Plots for* $(\beta_0, \beta_1, \beta_2)$ *and* $(\gamma_1, \gamma_2, \gamma_3, \gamma_4)$ *for the GLM and GLMM with ordered probit link.*



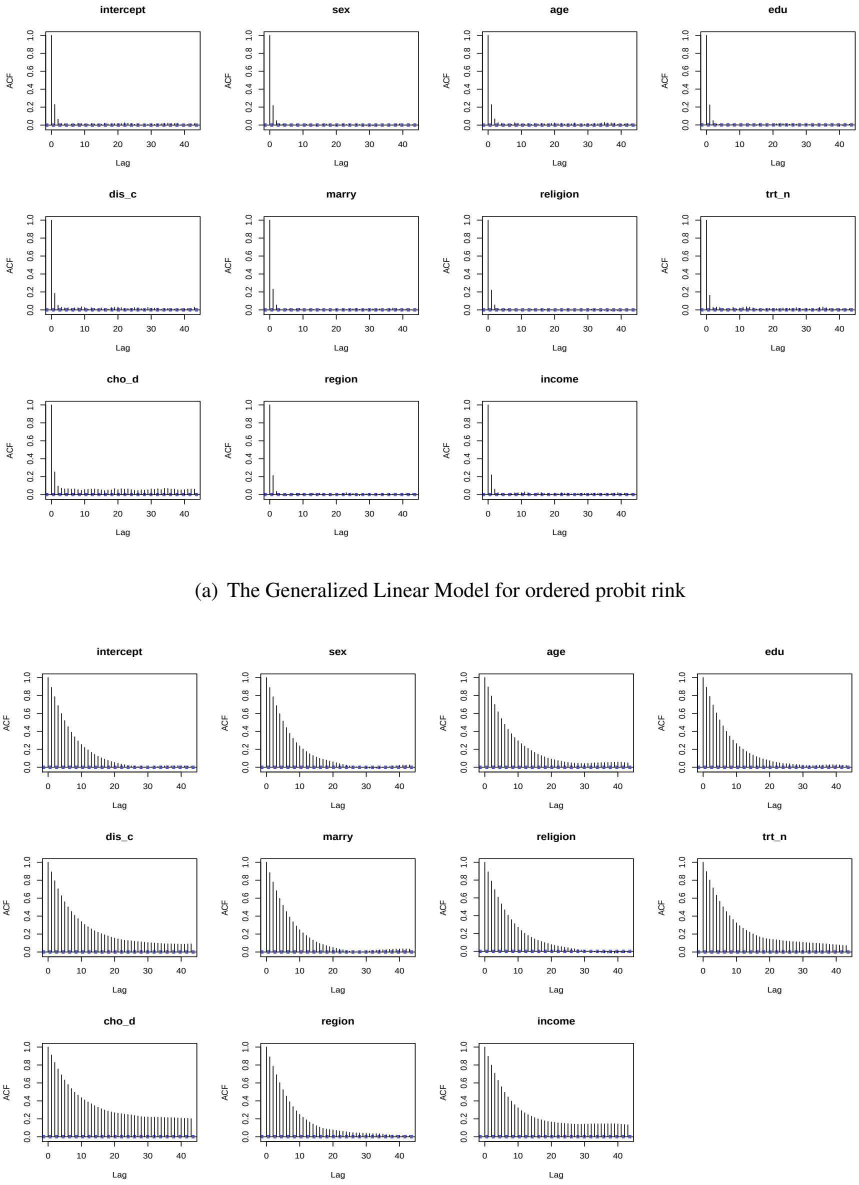
(a) The GLM for β

(b) The GLMM for β

(c) The GLM for cutpoints

(d) The GLMM for cutpoints

Figure A.2: *ACF Plots for thinning* $(\beta_0, \beta_1, \beta_2)$ *and* $(\gamma_1, \gamma_2, \gamma_3, \gamma_4)$ *for the GLM and GLMM with ordered probit link.*



(a) The Generalized Linear Model for ordered probit rink

(b) The Generalized Linear Mixed Model for ordered probit rink

Figure A.3: *ACF Plots for the coefficients of Ordered Probit Models on KWPS data.*

Acknowledgement

The authors would like to thank the referees and the associate editor for the informative feedback that has improved this version form a previous draft.

References

Albert, J. H. and Chib, S. (1993). Bayesian analysis of binary and polychotomous response data, *Journal of the American Statistical Association*, **88**, 669–679.

Andrew, D. Martin, Quinn, K. M. and Park, J. H. (2011). MCMCpack: Markov Chain Monte Carlo in R, *Journal of Statistical Software*, **42**, 1–21.

Breslow, N. E. and Clayton, D. G. (1993). Approximate inference in generalized linear mixed models, *Journal of the American Statistical Association*, **88**, 9–25.

Buonaccorsi, J. P. (1996). Measurement error in the response in the general linear model, *Journal of the American Statistical Association*, **91**, 633–642.

Chib, S., Greenberg, E. and Chen, Y. (1998). MCMC methods for fitting and comparing multinomial response models, Technical Report, Economics Working Paper Archive, Washington University at St. Louis.

Chib, S. and Winkelmann, R. (2001). Markov Chain Monte Carlo Analysis of Correlated Count data, *Journal of Business and Economic Statistics*, **19**, 428–435.

Damien, P., Wakefield, J. and Walker, S. (1999). Gibbs sampling for Bayesian non-conjugate and hierarchical models by using auxiliary variables, *Journal of the Royal Statistical Society, Series B*, **61**, 331–344.

Dey, D. K., Ghosh, S. K. and Mallick, B. K. (2000). *Generalized Linear Models: A Bayesian Perspective*, Marcel Dekker, New York.

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

Fehir, J. S. (1988). Self-rated health status, self-efficacy, motivation, and selected demographics as determinants of health-promoting lifestyle behavior in men 35 to 64 years old: A nursing investigation, Doctoral Dissertation, The University of Texas at Austin.

Fillenbaum, G. G. (1979). Social context and self-assessments of health among the elderly, *Journal of Health and Social Behavior*, **20**, 45–51.

Gill, J. and Casella, G. (2009). Nonparametric priors for ordinal Bayesian social science models: Specification and estimation, *Journal of the American Statistical Association*, **104**, 453–454.

Hoeymans, N., Feskens, E. J., Kromhout, D. and van den Bos, G. A. (1997). Ageing and the relationship between functional status and self-rated health in elderly man, *Social science and Medicine*, **45**, 1527–1536.

Jiang, J. (2007). *Linear and Generalized Linear Mixed Models and Their Applications*, Springer-Verlag, New York.

Lee, M. and Kim, D. (2013). Predictors of Korean Elderly People's self-rated Health Status and Moderating Effects of Socio-Economic Position. *The Korean Journal of Community Living Science*, **24**, 37–49.

Luoh, M. and Herzog, A. R. (2002). Individual consequences of volunteer and paid work in old age: Health and mortality, *Journal of Health and Social Behavior*, **43**, 490–509.

McCullagh, P. and Nelder, J. A. (1989). *Generalized Linear Models*, Second Edition, Chapman & Hall, New York.

McCulloch, C. E. and Searle, S. R. (2001). *Generalized, Linear, and Mixed Models*, John Wiley &

Sons, New York.

Neal, R. M. (2003). Slice sampling, *Annals of Statistics*, **31**, 705–741.

Oh, Y. H., Bae, H. O. and Kim, Y. S. (2006). A study on physical and mental function affecting self-perceived health of older persons in Korea. *Journal of the Korean Gerontological Society*, **26**, 461–476.

Otiniano, M. E., Cu, X. L., Ottenbacher, K. and Markides, K. S. (2003). The effect of diabetes combined with stroke on disability, self-rated health, and mortality in older Mexican Americans, *Academy of Physical Medicines and Rehabilitation*, **84**, 725–730.

Scott, W. K., Macera, C. A., Cornman, C. B. and Sharpe, P. A. (1997). Functional health status as a predictor of mortality in men and women over 65. *Journal of Clinical Epidemiology*, **50**, 291–296.

Stoller, E. P. (1984). Self-Assessments of health by the elderly: The impact of informal assistance, *Journal of Health and Social Behavior*, **25**, 260–270.

von dem Knesebeck, O., Luschen, G., Cocherham, W. C. and Siegrist, J. (2003). Socioeconomic status and health among the aged in the United States and Germany: A comparative cross-sectional study, *Social Science and Medicine*, **57**, 1643–1652.

Wang, N., Lin, X., Gutierrez, R. G. and Carroll, R. J. (1998). Bias analysis and SIMEX approach in generalized linear mixed measurement error models, *Journal of the American Statistical Association*, **93**, 249–261.

Ware, J. E. Jr. (1987). Standards for validating health measures: Definition and content, *Journal of Chronic Diseases*, **40**, 473–480.

Wolfinger, R. and O'Connell, M. (1993). Generalized linear mixed models: A Pseudo-likelihood approach, *Journal of Statistical Computation and Simulation*, **48**, 233–243.

Received October 24, 2013; Revised December 5, 2013; Accepted December 12, 2013