Title	Bayesian Dynamic Panel-Ordered Probit Model and Its Application to Subjective Well-Being
Author(s)	Hasegawa, Hikaru
Citation	Communications in Statistics : Simulation and Computation, 38(6), 1321-1347 https://doi.org/10.1080/03610910902903133
Issue Date	2009-06
Doc URL	https://hdl.handle.net/2115/43089
Rights	This is an electronic version of an article published in Communications in Statistics : Simulation and Computation, 38(6) June 2009, 1321-1347. Communications in Statistics : Simulation and Computation is available online at: http://www.informaworld.com/openurl?genre=article&issn=0361-0918&volume=38&issue=6&spage=1321
Type	journal article
File Information	CSSC09.pdf

Bayesian Dynamic Panel Ordered Probit Model and Its Application to Subjective Well-Being

HIKARU HASEGAWA*

Graduate School of Economics and Business Administration

Hokkaido University

Abstract

This article proposes a Bayesian method for estimating the dynamic panel ordered probit model. We compare four alternative algorithms for the estimation of ordered probit models. Furthermore, this article presents the empirical results of the application of the dynamic ordered probit model to subjective well-being based on the micro-level survey data extracted from the Japanese Panel Survey of Consumers. The results of the application reveal that income and savings have positive effects on life satisfaction, whereas our posterior results show that marriage and labor force participation have negative effects on the same.

Keywords Generalized Gibbs sampler; Life satisfaction; Markov chain Monte Carlo (MCMC); Well-being function.

Mathematics Subject Classification 62F15; 62P20.

Proposed running head Bayesian Dynamic Panel Ordered Probit Model.

*Address correspondence to Hikaru Hasegawa, Graduate School of Economics and Business Administration, Hokkaido University, Kita 9, Nishi 7, Kita-ku, Sapporo 060-0809, Japan; E-mail: hasegawa@econ.hokudai.ac.jp.

1 Introduction

In a questionnaire survey concerning subjective outcomes such as life satisfaction, choices are often ordinally arranged. An ordered probit model can be used to statistically analyze the data on such ordinal choices. From the frequentist viewpoint, the ordered probit model can be estimated by using the maximum likelihood (ML) method.[1]

On the other hand, posterior distributions play a central role in the Bayesian analysis. Recent developments in the Markov chain Monte Carlo (MCMC) method enable us to sample unobservable parameters from their posterior distributions even if the models are complicated. Albert and Chib (1993) propose the Bayesian approach for the estimation of the ordered probit model by using the MCMC method. They use the latent variable representation for the estimation of ordered probit models. Cowles (1996) and Nandram and Chen (1996) propose sampling algorithms for the Bayesian ordered probit model that are more efficient than that of Albert and Chib (1993). Further, Jeliazkov et al. (2008) review the identification and modeling problems in the ordered probit models.

This article proposes a Bayesian dynamic panel ordered probit model. Although Müller and Czado (2005) consider the Bayesian estimation of the autoregressive ordered probit model by adapting Liu and Sabatti's (2000) method, they do not deal with the dynamic panel data model. Our model uses the latent variable representation for ordinal choices, and the model for latent variable is a dynamic version of the Hausman and Taylor (1981) model. The dynamic panel data models with unobserved effects encounter a problem with respect to the initial conditions (Hsiao, 2003, Section 7.5). Wooldridge (2005) proposes a general method for resolving this problem in dynamic panel models including the ordered probit model in the non-Bayesian framework. One of the contributions of this article is to provide the Bayesian estimation framework that formulates the initial conditions problem in dynamic panel models. According to Jeliazkov et al. (2008), there are several kind of estimation algorithms for ordered probit models. Using the simulated data and the real data, this article compares the performances of estimation algorithms within the framework of Bayesian dynamic panel ordered probit model.

We apply the Bayesian dynamic panel ordered probit model to subjective well-being as a real data application.[2] The number of studies on well-being continue to increase and form a large amount of the literature.[3] One of the important issues in well-being studies is the choice of the sources of well-being, that is, selecting the explanatory variable of the well-being function (Frey and Stutzer, 2005, pp.213–217). As we will describe in Section 5, based on the application of the dynamic ordered probit model to the data on young Japanese women, income and savings are shown to have positive effects on well-being, whereas marriage and labor force participation are shown to have negative effects.

The article proceeds as follows. In Section 2, we provide the Bayesian dynamic panel ordered probit model. Section 3 describes the alternative four

---

[1] See, for example, Greene (2003, Chapter 21).

[2] In this article, we use life satisfaction as the measurement of well-being.

[3] See, for example, Blanchflower and Oswald (2004), Clark and Oswald (2002), and Frey and Stutzer (2002a,b, 2005).

algorithms for the estimation of the Bayesian dynamic panel ordered probit model. The comparisons among the four algorithms are provided in Section 4. Section 5 presents the empirical results of the application of the dynamic panel ordered probit model to subjective well-being, based on the micro-level survey data extracted from the Japanese Panel Survey of Consumers (JPSC), which is conducted by the Institute for Research on Household Economics. In Section 6, we provide a brief conclusion.

2   The Dynamic Panel Ordered Probit Model

In this section, we describe the Bayesian dynamic panel ordered probit model. Let y_{it} denote the ordinal discrete response of individual i at time t for $t = 0, 1, \cdots, T$ and $i = 1, \cdots, n$; that is, $y_{it} = c$ for $c = 1, \cdots, C$, where $t = 0$ denotes an initial period. Further, let z_{it} denote the latent variable of individual i at time t such that

$$y_{it} = c \ \ \text{if} \ \ z_{it} \in (\gamma_{c-1}, \gamma_c], \ t = 0, 1, \cdots, T, \ i = 1, \cdots, n, \ c = 1, \cdots, C, \tag{1}$$

where γ_c is a cutoff point of ordinal response. We specify that

$$-\infty = \gamma_0 < \gamma_1 = 0 < \gamma_2 < \cdots < \gamma_{C-1} < \gamma_C = \infty,$$

where the condition $\gamma_1 = 0$ is required to establish the identifiability of the cutoff parameters.[4] The latent variable z_{it} is assumed to be determined by the following linear model: for $i = 1, \cdots, n$,

$$\begin{cases} z_{it} = \phi z_{i(t-1)} + \boldsymbol{x}'_{it}\boldsymbol{\beta} + \boldsymbol{w}'_i\boldsymbol{\delta} + \alpha_i + u_{it}, \ t = 1, \cdots, T \\ z_{i0} = \boldsymbol{x}'_{i0}\boldsymbol{\beta}_0 + \boldsymbol{w}'_i\boldsymbol{\delta}_0 + \alpha_i + u_{i0}, \ t = 0, \end{cases}$$

where $\boldsymbol{x}_{it} = (x_{it1}, \cdots, x_{itk})'$ is a vector of time-variant explanatory variables, $\boldsymbol{w}_i = (w_{i1}, \cdots, w_{ip})'$ is a vector of cross-sectional time-invariant explanatory variables, α_i is a random effect, $\boldsymbol{\beta} = (\beta_1, \cdots, \beta_k)'$, $\boldsymbol{\delta} = (\delta_1, \cdots, \delta_p)'$, $\boldsymbol{\beta}_0 = (\beta_{01}, \cdots, \beta_{0k})'$, and $\boldsymbol{\delta}_0 = (\delta_{01}, \cdots, \delta_{0p})'$.[5] We assume that u_{it} are independent and identically normally distributed with zero mean and unit variance, that is, $u_{it} \sim \text{N}(0, 1)$. Then, we have

$$\begin{cases} z_{it}|\cdots \sim \text{N}(\phi z_{i(t-1)} + \boldsymbol{x}'_{it}\boldsymbol{\beta} + \boldsymbol{w}'_i\boldsymbol{\delta} + \alpha_i, 1), \ t = 1, \cdots, T \\ z_{i0}|\cdots \sim \text{N}(\boldsymbol{x}'_{i0}\boldsymbol{\beta}_0 + \boldsymbol{w}'_i\boldsymbol{\delta}_0 + \alpha_i, 1), \ t = 0 \end{cases} \tag{2}$$

for $i = 1, \cdots, n$, where "$|\cdots$" denotes conditioning on the other unspecified variables in the equation.

To complete the Bayesian model, we introduce the prior distributions of the parameters $(\phi, \boldsymbol{\beta}, \boldsymbol{\delta}, \boldsymbol{\beta}_0, \boldsymbol{\delta}_0, \boldsymbol{\alpha}, \gamma)$, where $\boldsymbol{\alpha} = (\alpha_1, \cdots, \alpha_n)'$ and $\boldsymbol{\gamma} = (\gamma_2, \cdots, \gamma_{C-1})'$. We assume that the prior distribution of the vector of random effects $\boldsymbol{\alpha}$ has a hierarchical structure.[6] The prior distributions are specified as follows:

$$\begin{aligned} p(\phi, \boldsymbol{\beta}, \boldsymbol{\delta}, \boldsymbol{\beta}_0, \boldsymbol{\delta}_0, \boldsymbol{\alpha}, \mu, \tau, \boldsymbol{\gamma}) & \\ = p(\phi)p(\boldsymbol{\beta})p(\boldsymbol{\delta})p(\boldsymbol{\beta}_0)p(\boldsymbol{\delta}_0)p(\boldsymbol{\alpha}|\mu, \tau)p(\mu)p(\tau)p(\boldsymbol{\gamma}). \end{aligned} \tag{3}$$

---

[4]See, for example, Albert and Chib (1993, p.673) and Johnson and Albert (1999, p.131).
[5]This model is a dynamic version of the Hausman and Taylor (1981) model. See, for example, Baltagi (2001, p.139).
[6]See, for example, Koop *et al.* (2007, Example 14.4, p.210).

3

Further, from Bayes' theorem, the joint posterior distribution can be written as follows:

$$
\begin{aligned}
p(\phi, \boldsymbol{\beta}, \boldsymbol{\delta}, \boldsymbol{\beta}_0, \boldsymbol{\delta}_0, \boldsymbol{\alpha}, \mu, \tau, \boldsymbol{\gamma}, \boldsymbol{z}, \boldsymbol{z}_0 | \boldsymbol{y}, \boldsymbol{y}_0) &\propto p(\phi, \boldsymbol{\beta}, \boldsymbol{\delta}, \boldsymbol{\beta}_0, \boldsymbol{\delta}_0, \boldsymbol{\alpha}, \mu, \tau, \boldsymbol{\gamma}, \boldsymbol{z}, \boldsymbol{z}_0) \\
&\times p(\boldsymbol{y}, \boldsymbol{y}_0 | \phi, \boldsymbol{\beta}, \boldsymbol{\delta}, \boldsymbol{\beta}_0, \boldsymbol{\delta}_0, \boldsymbol{\alpha}, \mu, \tau, \boldsymbol{\gamma}, \boldsymbol{z}, \boldsymbol{z}_0) \\
&= p(\phi, \boldsymbol{\beta}, \boldsymbol{\delta}, \boldsymbol{\beta}_0, \boldsymbol{\delta}_0, \boldsymbol{\alpha}, \mu, \tau, \boldsymbol{\gamma}) \\
&\times \prod_{i=1}^{n} \Bigg[\Big\{ p(\boldsymbol{z}_i | \phi, \boldsymbol{\beta}, \boldsymbol{\delta}, \alpha_i, \mu, \tau, \boldsymbol{\gamma}, z_{i0}) p(\boldsymbol{y}_i | \phi, \boldsymbol{\beta}, \boldsymbol{\delta}, \alpha_i, \mu, \tau, \boldsymbol{\gamma}, \boldsymbol{z}_i, z_{i0}) \Big\} \\
&\times \Big\{ p(z_{i0} | \phi, \boldsymbol{\beta}_0, \boldsymbol{\delta}_0, \alpha_i, \mu, \tau, \boldsymbol{\gamma}) p(y_{i0} | \phi, \boldsymbol{\beta}_0, \boldsymbol{\delta}_0, \alpha_i, \mu, \tau, \boldsymbol{\gamma}, z_{i0}) \Big\} \Bigg],
\end{aligned}
\tag{4}
$$

where

$$
\begin{aligned}
&\boldsymbol{y}_i = (y_{i1}, y_{i2}, \cdots, y_{iT})', \ \ \boldsymbol{z}_i = (z_{i1}, z_{i2}, \cdots, z_{iT})', \ \ i = 1, \cdots, n \\
&\boldsymbol{y} = (\boldsymbol{y}_1', \boldsymbol{y}_2', \cdots, \boldsymbol{y}_n')', \ \ \boldsymbol{z} = (\boldsymbol{z}_1', \boldsymbol{z}_2', \cdots, \boldsymbol{z}_n')' \\
&\boldsymbol{y}_0 = (y_{10}, y_{20}, \cdots, y_{n0})', \ \ \boldsymbol{z}_0 = (z_{10}, z_{20}, \cdots, z_{n0})'.
\end{aligned}
$$

3 Algorithms for the Estimation of the Bayesian Dynamic Panel Ordered Probit Model

Jeliazkov et al. (2008) consider three algorithms for the univariate ordered probit model, that is to say, the algorithms based on Albert and Chib (1993), Albert and Chib (2001) and Chen and Dey (2000). It is known that the MCMC sample for the cutoff points by using the algorithm of Albert and Chib (1993) has high autocorrelation. To take care of this problem, Albert and Chib (2001) and Chen and Dey (2000) propose the algorithms for drawing the cutoff points by transforming them. Furthermore, Müller and Czado (2005) consider the Bayesian estimation of the autoregressive ordered probit model by adapting Liu and Sabatti's (2000) method. In this section, we extend these existing algorithms to the Bayesian dynamic ordered probit model.

3.1 Algorithm based on Albert and Chib (1993)

The prior distributions are specified as follows:

$$
\begin{cases}
\phi \sim \mathrm{N}(\tilde{\phi}, \tilde{\omega}) 1_{(\phi \in (-1,1))}, \ \boldsymbol{\beta} \sim \mathrm{N}(\tilde{\boldsymbol{\beta}}, \tilde{\boldsymbol{B}}), \ \boldsymbol{\delta} \sim \mathrm{N}(\tilde{\boldsymbol{\delta}}, \tilde{\boldsymbol{D}}) \\
\boldsymbol{\beta}_0 \sim \mathrm{N}(\tilde{\boldsymbol{\beta}}_0, \tilde{\boldsymbol{B}}_0), \ \boldsymbol{\delta}_0 \sim \mathrm{N}(\tilde{\boldsymbol{\delta}}_0, \tilde{\boldsymbol{D}}_0) \\
\alpha_i \sim \mathrm{N}(\mu, \tau), \ i = 1, \cdots, n, \ \mu \sim \mathrm{N}(\tilde{\mu}, \tilde{\kappa}), \ \tau^{-1} \sim \mathrm{Gam}(\tilde{a}, \tilde{b}) \\
p(\boldsymbol{\gamma}) \propto 1_{(\boldsymbol{\gamma} \in \mathcal{C})}, \ \mathcal{C} = \{\boldsymbol{\gamma} : 0 < \gamma_2 < \cdots < \gamma_{C-1} < \infty\},
\end{cases}
\tag{5}
$$

where $1_{(\cdot)}$ is an indicator function and $\mathrm{Gam}(a, b)$ denotes a gamma distribution with parameters a, b. The algorithm of drawing the parameters in the dynamic panel ordered probit model based on the algorithm of Albert and Chib (1993) is as follows:[7]

Algorithm 1:

---

[7]The details of this sampling algorithm are provided in Appendix A.

1. Sample $\phi, \boldsymbol{\beta}, \boldsymbol{\delta}, \boldsymbol{\beta}_0, \boldsymbol{\delta}_0, \boldsymbol{\alpha}, \mu, \tau$ from their full conditional distributions (FCDs).

2. Sample $\boldsymbol{z}, \boldsymbol{z}_0$ from their FCDs.

3. Sample $\boldsymbol{\gamma}$ from its FCD.

3.2 Algorithm using Liu and Sabatti's method (2000)

We can use the algorithm proposed in Liu and Sabatti (2000) for estimating our Bayesian dynamic ordered probit model. To apply Liu and Sabatti's (2000) method, we specify the hyperparameters in the prior distributions (5) as follows:

$$\tilde{\boldsymbol{\beta}} = \boldsymbol{0}, \ \tilde{\boldsymbol{\delta}} = \boldsymbol{0}, \ \tilde{\boldsymbol{\beta}}_0 = \boldsymbol{0}, \ \tilde{\boldsymbol{\delta}}_0 = \boldsymbol{0}, \ \tilde{\mu} = 0.$$

The algorithm of drawing the parameters in the dynamic panel ordered probit model as per the algorithm of Liu and Sabatti (2000) is as follows:[8]

Algorithm 2:

1. Sample $\phi, \boldsymbol{\beta}, \boldsymbol{\delta}, \boldsymbol{\beta}_0, \boldsymbol{\delta}_0, \boldsymbol{\alpha}, \mu, \tau$ from their FCDs.

2. Sample $\boldsymbol{z}, \boldsymbol{z}_0$ from their FCDs.

3. Sample $\boldsymbol{\gamma}$ from its FCD.

4. Implement Liu and Sabatti's generalized Gibbs sampler.

3.3 Algorithm based on Albert and Chib (2001)

Albert and Chib (2001) propose the algorithm for drawing the cutoff points by transforming them as follows:

$$\zeta_c = \log(\gamma_c - \gamma_{c-1}), \ c = 2, \cdots, C-1, \tag{6}$$

where $\boldsymbol{\zeta}(\boldsymbol{\gamma}) = (\zeta_2, \cdots, \zeta_{C-1})'$ is unrestricted. Thus, we specify the prior distribution of $\boldsymbol{\gamma}$, $p(\boldsymbol{\gamma}) = p(\boldsymbol{\zeta}(\boldsymbol{\gamma}))$, as $\boldsymbol{\zeta}(\boldsymbol{\gamma}) \sim \mathrm{N}(\tilde{\zeta}, \tilde{\boldsymbol{G}})$ based on the transformation (6). The algorithm of drawing the parameters in the dynamic panel ordered probit model based on the algorithm of Albert and Chib (2001) is as follows:[9]

Algorithm 3:

1. Sample $\phi, \boldsymbol{\beta}, \boldsymbol{\delta}, \boldsymbol{\beta}_0, \boldsymbol{\delta}_0, \boldsymbol{\alpha}, \mu, \tau$ from their FCDs.

2. Sample $\boldsymbol{z}, \boldsymbol{z}_0$ from their FCDs.

3. Sample $\boldsymbol{\zeta}$ from the Metropolis-Hastings (M-H) algorithm.

4. Calculate $\gamma_c = \sum\limits_{j=1}^{c} \exp \zeta_j, \ c = 2, \cdots, C-1.$

---

[8]The details of this sampling algorithm are provided in Appendix B. Further, to use the scale group defined in Appendix B, we have to specify the hyperparameters in the prior distributions $\tilde{\boldsymbol{\beta}} = \boldsymbol{0}, \ \tilde{\boldsymbol{\delta}} = \boldsymbol{0}, \ \tilde{\boldsymbol{\beta}}_0 = \boldsymbol{0}, \ \tilde{\boldsymbol{\delta}}_0 = \boldsymbol{0}$ and $\tilde{\mu} = 0$.

[9]The details of this sampling algorithm are provided in Appendix C.

3.4 Algorithm based on Chen and Dey (2000)

Although the unit variance of disturbance, that is, $\text{var}(u_{it}) = 1$, is standard identification restriction, there are another ways to identify the ordered probit model. Nandram and Chen (1996) and Chen and Dey (2000) provide an identification restriction based on the reparameterization of the cutoff points, that is, $\gamma_{C-1} = 1$. Under this identification condition, model (2) can be modified as

$$
\begin{cases}
z_{it}|\cdots \sim \text{N}(\phi z_{i(t-1)} + \boldsymbol{x}'_{it}\boldsymbol{\beta} + \boldsymbol{w}'_i\boldsymbol{\delta} + \alpha_i, \sigma^2), \ t = 1,\cdots,T \\
z_{i0}|\cdots \sim \text{N}(\boldsymbol{x}'_{i0}\boldsymbol{\beta}_0 + \boldsymbol{w}'_i\boldsymbol{\delta}_0 + \alpha_i, \sigma^2), \ t = 0
\end{cases}
\tag{7}
$$

for $i = 1,\cdots,n$. Further, Chen and Dey (2000) propose the following transformation of cutoff points:

$$
\zeta_c = \log\left(\frac{\gamma_c - \gamma_{c-1}}{1 - \gamma_c}\right), \ c = 2,\cdots,C-2,
\tag{8}
$$

where $\boldsymbol{\zeta}(\boldsymbol{\gamma}) = (\zeta_2,\cdots,\zeta_{C-2})'$ is unrestricted. The prior distributions are specified as follows:

$$
\begin{aligned}
& p(\phi,\boldsymbol{\beta},\boldsymbol{\delta},\boldsymbol{\beta}_0,\boldsymbol{\delta}_0,\boldsymbol{\alpha},\mu,\tau,\sigma^2,\boldsymbol{\gamma}) \\
& = p(\phi)p(\boldsymbol{\beta})p(\boldsymbol{\delta})p(\boldsymbol{\beta}_0)p(\boldsymbol{\delta}_0)p(\boldsymbol{\alpha}|\mu,\tau)p(\mu)p(\tau)p(\sigma^2)p(\boldsymbol{\gamma}).
\end{aligned}
\tag{9}
$$

We specify the same prior distributions of parameters as in (5) except for σ^2 and $\boldsymbol{\gamma}$. Their prior distributions are specified as follows:

$$
\sigma^{-2} \sim \text{Gam}(\tilde{c},\tilde{d}), \ \boldsymbol{\zeta}(\boldsymbol{\gamma}) \sim \text{N}(\tilde{\zeta},\tilde{\boldsymbol{G}}),
\tag{10}
$$

where the prior distribution of $\boldsymbol{\gamma}$, $p(\boldsymbol{\gamma}) = p(\boldsymbol{\zeta}(\boldsymbol{\gamma}))$, is based on the transformation (8). The algorithm of drawing the parameters in the dynamic panel ordered probit model based on the algorithm of Nandram and Chen (1996) and Chen and Dey (2000) is as follows:[10]

Algorithm 4:

1. Sample $\phi,\boldsymbol{\beta},\boldsymbol{\delta},\boldsymbol{\beta}_0,\boldsymbol{\delta}_0,\boldsymbol{\alpha},\mu,\tau,\sigma^2$ from their FCDs.
2. Sample $\boldsymbol{z},\boldsymbol{z}_0$ from their FCDs.
3. Sample $\boldsymbol{\zeta}$ from the M-H algorithm.
4. Calculate $\gamma_c = \dfrac{\gamma_{c-1} + \exp\zeta_c}{1 + \exp\zeta_c}, \ c = 2,\cdots,C-2.$

4 Numerical Example with Simulated Data

This section provides a numerical example with simulated data for comparing the performances of four algorithms. In the numerical example, we set $n = 200$ and $T = 9$, and the ordered response variables y_{it} ($i = 1,\cdots,200$, $t = 0,1,\cdots,9$) are assumed to take four values, that is, $y_{it} = 1,2,3,4$. We also assume that the latent variables are distributed as follows:

$$
\begin{cases}
z_{it}|\cdots \sim \text{N}(\phi z_{i(t-1)} + x_{it}\beta_1 + w_i\delta_1 + \alpha_i, 1), \ t = 1,\cdots,9 \\
z_{i0}|\cdots \sim \text{N}(x_{i0}\beta_{01} + w_i\delta_{01} + \alpha_i, 1), \ t = 0
\end{cases}
$$

---

[10]The details of this sampling algorithm are provided in Appendix D.

for $i = 1, \cdots, 200$, where the parameters and the variables are specified as:

$$\phi = 0.5, \ \beta_1 = \beta_{01} = 2.0, \ \delta_1 = \delta_{01} = 1.5, \ \gamma_2 = 5.0, \ \gamma_3 = 10.0$$
$$x_{i0} \sim N(1,3), \ x_{it} = 0.3x_{i(t-1)} + v_{it}, \ v_{it} \sim N(0,1)$$
$$w_i \sim N(2,4), \ \alpha_i \sim N(-3,4).$$

The summary statistics of these variables are provided in Tables 1 and 2.

The MCMC simulation was run for 12,000 iterations with the thinning interval equal to 5; the first 2,000 samples were discarded as the burn-in period.[11] The posterior results obtained thereafter are generated using Ox version 5.00 (Doornik, 2007). Table 3 presents the posterior results for the estimation of ϕ, β_1, δ_1, β_{01}, δ_{01}, and $\boldsymbol{\gamma}$.[12] In Table 3, "95%CI" denotes the 95% credible interval. "CD" denotes the convergence diagnostic statistic proposed by Geweke (1992).[13] The sixth column of Table 3 shows the p-values for CD.

According to CD, Algorithm 1 does not exhibit good performances. ϕ in Algorithm 3 does not converge. By contrast, Algorithms 2 and 4 perform well according to CD. In the next section, we use Algorithms 2 and 4 for estimating the dynamic ordered probit model by using the real data.

5 Application to the Analysis of Life Satisfaction

5.1 Data

Our empirical analysis uses the micro-level survey data extracted from the Japanese Panel Survey of Consumers (JPSC), which is conducted by the Institute for Research on Household Economics. Since 1993, the survey has been administered to 1,500 women representative of the entire female population in the age range of 24–34 years in 1993. The JPSC data provide several measures of satisfaction pertaining to young women in Japan. This article employs the 6th wave (1998) to the 11th wave (2003) of JPSC data. From the above panel data, we select the data pertaining to women who responded to the question on life satisfaction over a period of 6 years from 1998 through 2003. Finally, the total number of recorded responses stood at 338.

We use the following variables:

---

[11] We obtained 12,000 samples by running 60,000 iterations and sampling one observation at every five iterations. After obtaining 12,000 samples, we discarded the first 2,000 samples.

[12] The results of parameters in Algorithm 4 except for ϕ are those of the transformed parameters divided by σ.

[13] CD can be defined as follows: for the given sequence $\{g(j) \mid j = 1, 2, \cdots, n_s\}$, if the sequence is stationary,

$$CD = \frac{\bar{g}_A - \bar{g}_B}{\hat{S}_A(0)/n_A + \hat{S}_B(0)/n_B} \to N(0,1),$$

where

$$\bar{g}_A = \frac{1}{n_A} \sum_{j=1}^{n_A} g(j), \ \bar{g}_B = \frac{1}{n_B} \sum_{j=n_*}^{n_s} g(j) \ (n_* = n_s - n_B + 1),$$

and $\hat{S}_A(0)$ and $\hat{S}_B(0)$ denote consistent spectral density estimates. In this numerical example, we set $n_A = 1,000$ and $n_B = 5,000$, and calculate $\hat{S}_A(0)$ and $\hat{S}_B(0)$ using Parzen windows with bandwidths of 100 and 500, respectively.

- ordinal dependent variables: `lifesat` (y, z)

- time-variant explanatory variables: `married` (x_1), `loginc` (x_2), `logsav` (x_3), `inwork` (x_4)

- cross-sectional time-invariant explanatory variables: `age` (w_1) and `educ` (w_2),

where `lifesat` denotes life satisfaction, for which the number of ordinal responses is $C = 4$;[14] `married` (x_1), a dummy variable that takes 1 if a respondent is married, and 0 otherwise; `loginc` (x_2), the logarithm of the previous year's per capita real annual household income (ten-thousand yen, base year = 2000); `logsav` (x_3), the logarithm of per capita real household savings (ten-thousand yen, base year = 2000); `inwork` (x_4), a dummy variable that takes 1 if a respondent is working, and 0 otherwise.[15] In addition, `age` (w_1) represents the age of the respondent in 1998; and `educ` (w_2), the years of education of the respondent. The summary statistics of these variables are provided in Tables 4 and 5.

The estimated well-being functions are as follows:

$$z_{it} = \phi z_{i(t-1)} + \beta_1 x_{it1} + \beta_2 x_{it2} + \beta_3 x_{it3} + \beta_4 x_{it4}$$
$$+\delta_1 w_{i1} + \delta_2 w_{i2} + \alpha_i + u_{it}, \ t = 1, \cdots, 5$$
$$z_{i0} = \beta_{01} x_{i01} + \beta_2 x_{i02} + \beta_3 x_{i03} + \beta_4 x_{i04} + \delta_{01} w_{i1} + \delta_{02} w_{i2} + \alpha_i + u_{i0},$$

where the initial period $(t = 0)$ is year 1998.

5.2 Posterior results

The MCMC simulation was run for 30,000 iterations with the thinning interval equal to 5; the first 10,000 samples were discarded as the burn-in period.[16] Table 6 presents the posterior results for the estimation of ϕ, $\boldsymbol{\beta}$, $\boldsymbol{\delta}$, $\boldsymbol{\beta}_0$, $\boldsymbol{\delta}_0$, and $\boldsymbol{\gamma}$.[17] According to the convergence diagnostic statistic (CD), parameters in Algorithm 4 except for ϕ, β_4, β_{01}, β_{04} and γ_3 do not converge. By contrast, all parameters in Algorithm 2 converge. Therefore, we use the posterior results in Algorithm 2. The following points can be observed from the posterior results in Algorithm 2.

- According to the 95% credible interval, the lagged latent variable $(z_{(-1)})$ has a significantly positive effect on life satisfaction.[18]

---

[14]The number of ordinal categories of life satisfaction is originally $C = 5$, including "strongly dissatisfied," "dissatisfied," "moderately satisfied," "satisfied," and "strongly satisfied." However, since the frequency of "strongly dissatisfied" is small, this article combines two categories "strongly dissatisfied" and "dissatisfied" into a single category, "strongly dissatisfied and dissatisfied."

[15]In the case of zero income or zero savings, we set `loginc` (x_2) or `logsav` (x_3) equal to 0. Further, we use the consumer price index as a deflator for constructing real variables.

[16]We obtained 30,000 samples by running 150,000 iterations and sampling one observation at every five iterations. After obtaining 30,000 samples, we discarded the first 10,000 samples.

[17]The results of parameters in Algorithm 4 except for ϕ are those of the transformed parameters divided by σ.

[18]Here, the term "significant" is used if the 95% credible interval for a parameter does not include zero. See also Koop (2003, p.124).

- Two dummy variables, the marriage (`married`) and labor force participation dummies (`inwork`), have significantly negative effects on life satisfaction for $t \geq 1$, but they are not significant at the initial period $t = 0$.

- Income (`loginc`) and savings (`logsav`) have significantly positive effects on life satisfaction for all the periods.

- The time-invariant explanatory variables (`age` and `educ`) are not significant for all the periods.

- Three cutoff points (γ_2, γ_3, and γ_4) are significantly estimated.

As in Wooldridge (2002, pp.505–506), we are more interested in the response probabilities $\Pr(y_{it} = c)$ than in the coefficient parameters themselves. Table 7 shows the posterior results for the partial effects of the explanatory variables on the response probabilities.[19] The following points can be observed from the posterior results in Table 7.

- The convergence of the posterior distributions can be confirmed from the p-values for CD.

- $z_{(-1)}$, `married`, `loginc`, `logsav`, and `inwork` have significantly partial effects on the response probabilities, but the partial effects of `age` and `educ` on the response probabilities are not significant.

- The lagged latent variable ($z_{(-1)}$) increases the probabilities $\Pr(y = 3)$ and $\Pr(y = 4)$ by 3.7 and 2.2 percent points, respectively; income (`loginc`) increases the probabilities $\Pr(y = 3)$ and $\Pr(y = 4)$ by 11.0 and 7.2 percent points, respectively; and savings (`logsav`) increases the probabilities $\Pr(y = 3)$ and $\Pr(y = 4)$ by 3.2 and 2.0 percent points, respectively. Therefore, income has the largest positive effect on life satisfaction in terms of the response probabilities.

- Per contra, the marriage dummy (`married`) decreases the probabilities $\Pr(y = 3)$ and $\Pr(y = 4)$ by 4.8 and 3.7 percent points, respectively. Further, the labor force participation dummy (`inwork`) decreases the probabilities $\Pr(y = 3)$ and $\Pr(y = 4)$ by 3.6 and 2.2 percent points, respectively. Therefore, marriage and labor force participation have negative effects on life satisfaction in terms of the response probabilities. The negative effect of marriage on life satisfaction contrasts with the findings of previous studies.[20] This may be attributable to the respondents of the sample. The respondents of the JPSC data are relatively young women, and most of the single women in the sample live with their parents. If young women do not live with their parents after marriage, the marriage would lead to a significant decline in single women's standard of living, which would decrease their life satisfaction.

---

[19]The details of partial effects of the continuous explanatory variables are provided in Appendix E.

[20]See, for example, Clark and Oswald (2002, p.1141).

6   Conclusion

This article proposed a Bayesian method for estimating the dynamic panel ordered probit model, which addressed the initial conditions problem, and compared the four alternative algorithms for the model. Our result of simulated data suggested that the algorithms using Liu and Sabatti's (2000) method and Chen and Dey (2000) perform well. In addition, this article provided the empirical results of the application of the dynamic ordered probit model to subjective well-being using the micro-level survey data obtained from the JPSC, which is conducted by the Institute for Research on Household Economics. The posterior results suggested that the algorithm using the Liu and Sabatti's method exhibits a better performance than that of Chen and Dey (2000). Furthermore, as in Wooldridge (2002, pp.505–506), we focused on the response probabilities of the ordered probit model rather than on the coefficient parameters themselves. The posterior results revealed that income and savings have positive effects on life satisfaction in terms of the response probabilities, whereas our posterior results showed that marriage and labor force participation have negative effects.

A   Derivation of Algorithm 1

A.1   FCD of ϕ

The full conditional distribution (FCD) of ϕ is

$$\phi|\cdots \sim \mathrm{N}(\hat{\phi}, \hat{\omega})1_{(\phi \in (-1,1))},$$

where

$$\hat{\omega} = \left(\frac{1}{\tilde{\omega}} + \sum_{i=1}^{n} \sum_{t=1}^{T} z_{i(t-1)}^2 \right)^{-1}$$

$$\hat{\phi} = \hat{\omega} \left[\frac{\tilde{\phi}}{\tilde{\omega}} + \sum_{i=1}^{n} \sum_{t=1}^{T} z_{i(t-1)}(z_{it} - \boldsymbol{x}_{it}'\boldsymbol{\beta} - \boldsymbol{w}_i'\boldsymbol{\delta} - \alpha_i) \right].$$

We can utilize Damien and Walker's (2001) algorithm for sampling from a truncated normal distribution.[21]

---

[21] Using Damien and Walker's algorithm, we can sample ϕ from the following steps.

1. Let $x = \dfrac{\phi - \hat{\phi}}{\sqrt{\hat{\omega}}}$, we have $\phi \in (-1, 1) \Leftrightarrow x \in \left(\dfrac{-1 - \hat{\phi}}{\sqrt{\hat{\omega}}}, \dfrac{1 - \hat{\phi}}{\sqrt{\hat{\omega}}} \right)$.

2. Introduce the latent variable y which has joint density with x given by

$$p(x, y) \propto 1_{(y \in (0, \exp(-x^2/2)))} 1_{(x \in (a,b))},$$

where $a = \dfrac{-1 - \hat{\phi}}{\sqrt{\hat{\omega}}}, b = \dfrac{1 - \hat{\phi}}{\sqrt{\hat{\omega}}}$.

3. We have the following FCDs:

$$y|x \sim \mathrm{U}\left(0, \exp\left(-\frac{x^2}{2} \right) \right)$$

$$x|y \sim \mathrm{U}\left(\max\left\{ \frac{-1 - \hat{\phi}}{\sqrt{\hat{\omega}}}, -\sqrt{-2\log y} \right\}, \min\left\{ \frac{1 - \hat{\phi}}{\sqrt{\hat{\omega}}}, \sqrt{-2\log y} \right\} \right).$$

A.2   FCD of $\boldsymbol{\beta}$

The FCD of $\boldsymbol{\beta}$ is

$$\boldsymbol{\beta}|\cdots \sim \mathrm{N}(\hat{\boldsymbol{\beta}}, \hat{\boldsymbol{B}}),$$

where

$$\hat{\boldsymbol{B}} = \left(\tilde{\boldsymbol{B}}^{-1} + \sum_{i=1}^{n} \sum_{t=1}^{T} \boldsymbol{x}_{it} \boldsymbol{x}_{it}' \right)^{-1}$$

$$\hat{\boldsymbol{\beta}} = \hat{\boldsymbol{B}} \left[\tilde{\boldsymbol{B}}^{-1} \tilde{\boldsymbol{\beta}} + \sum_{i=1}^{n} \sum_{t=1}^{T} \boldsymbol{x}_{it} (z_{it} - \phi z_{i(t-1)} - \boldsymbol{w}_i' \boldsymbol{\delta} - \alpha_i) \right].$$

A.3   FCD of $\boldsymbol{\delta}$

The FCD of $\boldsymbol{\delta}$ is

$$\boldsymbol{\delta}|\cdots \sim \mathrm{N}(\hat{\boldsymbol{\delta}}, \hat{\boldsymbol{D}}),$$

where

$$\hat{\boldsymbol{D}} = \left(\tilde{\boldsymbol{D}}^{-1} + T \sum_{i=1}^{n} \boldsymbol{w}_i \boldsymbol{w}_i' \right)^{-1}$$

$$\hat{\boldsymbol{\delta}} = \hat{\boldsymbol{D}} \left[\tilde{\boldsymbol{D}}^{-1} \tilde{\boldsymbol{\delta}} + \sum_{i=1}^{n} \sum_{t=1}^{T} \boldsymbol{w}_i (z_{it} - \phi z_{i(t-1)} - \boldsymbol{x}_{it}' \boldsymbol{\beta} - \alpha_i) \right].$$

A.4   FCD of $\boldsymbol{\beta}_0$

The FCD of $\boldsymbol{\beta}_0$ is

$$\boldsymbol{\beta}_0|\cdots \sim \mathrm{N}(\hat{\boldsymbol{\beta}}_0, \hat{\boldsymbol{B}}_0),$$

where

$$\hat{\boldsymbol{B}}_0 = \left(\tilde{\boldsymbol{B}}_0^{-1} + \sum_{i=1}^{n} \boldsymbol{x}_{i0} \boldsymbol{x}_{i0}' \right)^{-1}$$

$$\hat{\boldsymbol{\beta}}_0 = \hat{\boldsymbol{B}}_0 \left[\tilde{\boldsymbol{B}}_0^{-1} \tilde{\boldsymbol{\beta}}_0 + \sum_{i=1}^{n} \boldsymbol{x}_{i0} (z_{i0} - \boldsymbol{w}_i' \boldsymbol{\delta}_0 - \alpha_i) \right].$$

A.5   FCD of $\boldsymbol{\delta}_0$

The FCD of $\boldsymbol{\delta}_0$ is

$$\boldsymbol{\delta}_0|\cdots \sim \mathrm{N}(\hat{\boldsymbol{\delta}}_0, \hat{\boldsymbol{D}}_0),$$

---

4. Calculate $\phi = \hat{\phi} + \sqrt{\hat{\omega}} x$.

where

$$\hat{\boldsymbol{D}}_0 = \left(\tilde{\boldsymbol{D}}_0^{-1} + \sum_{i=1}^{n} \boldsymbol{w}_i \boldsymbol{w}_i' \right)^{-1}$$

$$\hat{\boldsymbol{\delta}}_0 = \hat{\boldsymbol{D}}_0 \left[\tilde{\boldsymbol{D}}_0^{-1} \tilde{\boldsymbol{\delta}}_0 + \sum_{i=1}^{n} \boldsymbol{w}_i (z_{i0} - \boldsymbol{x}_{i0}' \boldsymbol{\beta}_0 - \alpha_i) \right].$$

A.6   FCD of $\boldsymbol{\alpha}$

The FCD of α_i is

$$\alpha_i | \cdots \sim \mathrm{N}(\hat{\alpha}_i, \hat{\tau}), \ i = 1, \cdots, n,$$

where

$$\hat{\tau} = \left(\frac{1}{\tau} + T + 1 \right)^{-1}$$

$$\hat{\alpha}_i = \hat{\tau} \left[\frac{\mu}{\tau} + \sum_{t=1}^{T} (z_{it} - \phi z_{i(t-1)} - \boldsymbol{x}_{it}' \boldsymbol{\beta} - \boldsymbol{w}_i' \boldsymbol{\delta}) + (z_{i0} - \boldsymbol{x}_{i0}' \boldsymbol{\beta}_0 - \boldsymbol{w}_i' \boldsymbol{\delta}_0) \right].$$

A.7   FCD of μ

The FCD of μ is

$$\mu | \cdots \sim \mathrm{N}(\hat{\mu}, \hat{\kappa}),$$

where

$$\hat{\kappa} = \left(\frac{1}{\tilde{\kappa}} + \frac{n}{\tau} \right)^{-1}, \ \hat{\mu} = \hat{\kappa} \left(\frac{\tilde{\mu}}{\tilde{\kappa}} + \frac{1}{\tau} \sum_{i=1}^{n} \alpha_i \right).$$

A.8   FCD of τ

The FCD of τ^{-1} is

$$\tau^{-1} | \cdots \sim \mathrm{Gam}(\hat{a}, \hat{b}),$$

where

$$\hat{a} = \tilde{a} + \frac{n}{2}, \ \hat{b} = \tilde{b} + \frac{1}{2} \sum_{i=1}^{n} (\alpha_i - \mu)^2.$$

A.9   FCD of $\boldsymbol{z}_0$

The FCD of z_{i0} is

$$z_{i0} | \cdots \sim \mathrm{N} \left(\hat{z}_{i0}, \frac{1}{1 + \phi^2} \right) 1_{(z_{i0} \in (\gamma_{c-1}, \gamma_c])}, \ i = 1, \cdots, n,$$

where

$$\hat{z}_{i0} = \frac{\phi(z_{i1} - \boldsymbol{x}_{i1}' \boldsymbol{\beta} - \boldsymbol{w}_i' \boldsymbol{\delta} - \alpha_i) + (\boldsymbol{x}_{i0}' \boldsymbol{\beta}_0 + \boldsymbol{w}_i' \boldsymbol{\delta}_0 + \alpha_i)}{1 + \phi^2}.$$

A.10 FCD of z

The FCD of z_{it} is

$$z_{it}| \cdots \sim \mathrm{N}\left(\hat{z}_{it}, \frac{1}{1+\phi^2}\right) 1_{(z_{it} \in (\gamma_{c-1}, \gamma_c])}, \ t = 1, \cdots, T-1, \ i = 1, \cdots, n,$$

where

$$\hat{z}_{it} = \frac{(\phi z_{i(t-1)} + \boldsymbol{x}'_{it}\boldsymbol{\beta} + \boldsymbol{w}'_i\boldsymbol{\delta} + \alpha_i) + \phi(z_{i(t+1)} - \boldsymbol{x}'_{i(t+1)}\boldsymbol{\beta} - \boldsymbol{w}'_i\boldsymbol{\delta} - \alpha_i)}{1+\phi^2}.$$

Further, the FCD of z_{iT} is

$$z_{iT}| \cdots \sim \mathrm{N}(\phi z_{i(T-1)} + \boldsymbol{x}'_{iT}\boldsymbol{\beta} + \boldsymbol{w}'_i\boldsymbol{\delta} + \alpha_i, 1)1_{(z_{iT} \in (\gamma_{c-1}, \gamma_c])}, \ i = 1, \cdots, n.$$

A.11 FCD of γ

The FCD of γ_c is

$$\gamma_c| \cdots \sim \mathrm{U}\left(\max\left\{\max_{i,t}(z_{it}|y_{it} = c), \gamma_{c-1}\right\}, \min\left\{\min_{i,t}(z_{it}|y_{it} = c+1), \gamma_{c+1}\right\}\right),$$
$$c = 2, \cdots, C-1.$$

B Derivation of Algorithm 2

The FCDs of parameters are the same as those with $\tilde{\boldsymbol{\beta}} = \boldsymbol{0}$, $\tilde{\boldsymbol{\delta}} = \boldsymbol{0}$, $\tilde{\boldsymbol{\beta}}_0 = \boldsymbol{0}$, $\tilde{\boldsymbol{\delta}}_0 = \boldsymbol{0}$, $\tilde{\mu} = 0$ in Appendix A.

Following Liu and Sabatti (2000), we consider the generalized Gibbs sampler using the following scale group:

$$\Gamma = \{g > 0 : g(\phi, \boldsymbol{\beta}, \boldsymbol{\delta}, \boldsymbol{\beta}_0, \boldsymbol{\delta}_0, \boldsymbol{\alpha}, \mu, \tau, \boldsymbol{\gamma}, \boldsymbol{z}, \boldsymbol{z}_0)$$
$$= (\phi, g\boldsymbol{\beta}, g\boldsymbol{\delta}, g\boldsymbol{\beta}_0, g\boldsymbol{\delta}_0, g\boldsymbol{\alpha}, g\mu, \tau, g\boldsymbol{\gamma}, g\boldsymbol{z}, g\boldsymbol{z}_0)\}.$$

We define the left-Haar measure for Γ such as $L(dg) = g^{-1}dg$. Then, we have

$$p(g(\phi, \boldsymbol{\beta}, \boldsymbol{\delta}, \boldsymbol{\beta}_0, \boldsymbol{\delta}_0, \boldsymbol{\alpha}, \mu, \tau, \boldsymbol{\gamma}, \boldsymbol{z}, \boldsymbol{z}_0)|\boldsymbol{y}, \boldsymbol{y}_0)$$

$$\propto \exp\left[-\frac{g^2}{2}\left\{\boldsymbol{\beta}'\tilde{\boldsymbol{B}}^{-1}\boldsymbol{\beta} + \boldsymbol{\delta}'\tilde{\boldsymbol{D}}^{-1}\boldsymbol{\delta} + \boldsymbol{\beta}'_0\tilde{\boldsymbol{B}}_0^{-1}\boldsymbol{\beta}_0 + \boldsymbol{\delta}'_0\tilde{\boldsymbol{D}}_0^{-1}\boldsymbol{\delta}_0\right.\right.$$

$$+ \frac{1}{\tau}\sum_{i=1}^n(\alpha_i - \mu)^2 + \frac{\mu^2}{\tilde{\kappa}} + \sum_{i=1}^n\sum_{t=1}^T(z_{it} - \phi z_{i(t-1)} - \boldsymbol{x}'_{it}\boldsymbol{\beta} - \boldsymbol{w}'_i\boldsymbol{\delta} - \alpha_i)^2$$

$$\left.\left.+ \sum_{i=1}^n(z_{i0} - \boldsymbol{x}'_{i0}\boldsymbol{\beta}_0 - \boldsymbol{w}'_i\boldsymbol{\delta}_0 - \alpha_i)^2\right\}\right] \exp\left[-\frac{1}{2\tilde{\omega}}(\phi - \tilde{\phi})^2\right] 1_{(\phi \in (0,1))}$$

$$\times 1_{(\boldsymbol{\gamma} \in \mathcal{C})}\left[\prod_{i=1}^n\prod_{t=1}^T\left(\sum_{c=1}^C 1_{(y_{it}=c)}1_{(z_{it} \in (\gamma_{c-1}, \gamma_c])}\right)\right]$$

$$\times \left[\prod_{i=1}^n\left(\sum_{c=1}^C 1_{(y_{i0}=c)}1_{(z_{i0} \in (\gamma_{c-1}, \gamma_c])}\right)\right].$$

Further, we have the following Jacobian:[22]

$$|J_g(\phi, \boldsymbol{\beta}, \boldsymbol{\delta}, \boldsymbol{\beta}_0, \boldsymbol{\delta}_0, \boldsymbol{\alpha}, \mu, \tau, \boldsymbol{\gamma}, \boldsymbol{z}, \boldsymbol{z}_0)| = g^{(T+2)n+2k+2p+C-1}.$$

Therefore the FCD of g is

$$
\begin{aligned}
p(g|\cdots) &\propto p(g(\phi, \boldsymbol{\beta}, \boldsymbol{\delta}, \boldsymbol{\beta}_0, \boldsymbol{\delta}_0, \boldsymbol{\alpha}, \mu, \tau, \boldsymbol{\gamma}, \boldsymbol{z}, \boldsymbol{z}_0)|\boldsymbol{y}, \boldsymbol{y}_0) \\
&\quad \times |J_g(\phi, \boldsymbol{\beta}, \boldsymbol{\delta}, \boldsymbol{\beta}_0, \boldsymbol{\delta}_0, \boldsymbol{\alpha}, \mu, \tau, \boldsymbol{\gamma}, \boldsymbol{z}, \boldsymbol{z}_0)|g^{-1} \\
&\propto \exp\left[-\frac{g^2}{2}\left\{\boldsymbol{\beta}'\tilde{\boldsymbol{B}}^{-1}\boldsymbol{\beta} + \boldsymbol{\delta}'\tilde{\boldsymbol{D}}^{-1}\boldsymbol{\delta} + \boldsymbol{\beta}_0'\tilde{\boldsymbol{B}}_0^{-1}\boldsymbol{\beta}_0 + \boldsymbol{\delta}_0'\tilde{\boldsymbol{D}}_0^{-1}\boldsymbol{\delta}_0\right.\right. \\
&\quad\left.\left. + \frac{1}{\tau}\sum_{i=1}^{n}(\alpha_i - \mu)^2 + \frac{\mu^2}{\tilde{\kappa}} + \sum_{i=1}^{n}\sum_{t=1}^{T}(z_{it} - \phi z_{i(t-1)} - \boldsymbol{x}_{it}'\boldsymbol{\beta} - \boldsymbol{w}_i'\boldsymbol{\delta} - \alpha_i)^2\right.\right. \\
&\quad\left.\left. + \sum_{i=1}^{n}(z_{i0} - \boldsymbol{x}_{i0}'\boldsymbol{\beta}_0 - \boldsymbol{w}_i'\boldsymbol{\delta}_0 - \alpha_i)^2\right\}\right]g^{(T+2)n+2k+2p+C-2}.
\end{aligned}
$$

Putting $g = \sqrt{h}$, we have

$$h|\cdots \sim \text{Gam}\left(\frac{(T+2)n + 2k + 2p + C - 1}{2}, \hat{s}\right),$$

where

$$
\begin{aligned}
\hat{s} = \frac{1}{2}\left[\boldsymbol{\beta}'\tilde{\boldsymbol{B}}^{-1}\boldsymbol{\beta} + \boldsymbol{\delta}'\tilde{\boldsymbol{D}}^{-1}\boldsymbol{\delta} + \boldsymbol{\beta}_0'\tilde{\boldsymbol{B}}_0^{-1}\boldsymbol{\beta}_0 + \boldsymbol{\delta}_0'\tilde{\boldsymbol{D}}_0^{-1}\boldsymbol{\delta}_0\right. \\
+ \frac{1}{\tau}\sum_{i=1}^{n}(\alpha_i - \mu)^2 + \frac{\mu^2}{\tilde{\kappa}} + \sum_{i=1}^{n}\sum_{t=1}^{T}(z_{it} - \phi z_{i(t-1)} - \boldsymbol{x}_{it}'\boldsymbol{\beta} - \boldsymbol{w}_i'\boldsymbol{\delta} - \alpha_i)^2 \\
\left. + \sum_{i=1}^{n}(z_{i0} - \boldsymbol{x}_{i0}'\boldsymbol{\beta}_0 - \boldsymbol{w}_i'\boldsymbol{\delta}_0 - \alpha_i)^2\right].
\end{aligned}
$$

Liu and Sabatti's algorithm

- Using usual Gibbs sampler, generate $\phi, \boldsymbol{\beta}, \boldsymbol{\delta}, \boldsymbol{\beta}_0, \boldsymbol{\delta}_0, \boldsymbol{\alpha}, \mu, \tau, \boldsymbol{\gamma}, \boldsymbol{z}, \boldsymbol{z}_0$.
- Generate h from $p(h|\cdots)$ and put $g = \sqrt{h}$. Using g, implement the following update: $\boldsymbol{\beta} \leftarrow g\boldsymbol{\beta}$, $\boldsymbol{\delta} \leftarrow g\boldsymbol{\delta}$, $\boldsymbol{\beta}_0 \leftarrow g\boldsymbol{\beta}_0$, $\boldsymbol{\delta}_0 \leftarrow g\boldsymbol{\delta}_0$, $\boldsymbol{\alpha} \leftarrow g\boldsymbol{\alpha}$, $\mu \leftarrow g\mu$, $\boldsymbol{\gamma} \leftarrow g\boldsymbol{\gamma}$, $\boldsymbol{z} \leftarrow g\boldsymbol{z}$, $\boldsymbol{z}_0 \leftarrow g\boldsymbol{z}_0$.

---

[22]Since $\boldsymbol{\beta}$ has k, $\boldsymbol{\delta}$ has p, $\boldsymbol{\beta}_0$ has k, $\boldsymbol{\delta}_0$ has p, $\boldsymbol{\alpha}$ has n, $\boldsymbol{\gamma}$ has $C - 2$, $\boldsymbol{z}$ has nT, $\boldsymbol{z}_0$ has n, μ has one g, the exponential part of g becomes $(T + 2)n + 2k + 2p + C - 1$.

C Derivation of Algorithm 3

The FCDs of parameters are same as in Section A except for the FCD of $\boldsymbol{\gamma}$. The FCD of $\boldsymbol{\gamma}$ is

$$p(\boldsymbol{\gamma}|\cdots) \propto p(\boldsymbol{\zeta}(\boldsymbol{\gamma})) \prod_{i,t:y_{it}=2} [\Phi(\gamma_2 - \tilde{z}_{it}) - \Phi(-\tilde{z}_{it})]$$

$$\times \prod_{i,t:y_{it}=3} [\Phi(\gamma_3 - \tilde{z}_{it}) - \Phi(\gamma_2 - \tilde{z}_{it})]$$

$$\times \cdots \times \prod_{i,t:y_{it}=C-1} [\Phi(\gamma_{C-1} - \tilde{z}_{it}) - \Phi(\gamma_{C-2} - \tilde{z}_{it})]$$

$$\times \prod_{i,t:y_{it}=C} [1 - \Phi(\gamma_{C-1} - \tilde{z}_{it})],$$

where $\Phi(\cdot)$ is the distribution function of the standard normal distribution and

$$\tilde{z}_{it} = \begin{cases} \phi z_{i(t-1)} + \boldsymbol{x}_{it}'\boldsymbol{\beta} + \boldsymbol{w}_i'\boldsymbol{\delta} + \alpha_i & t = 1, \cdots, T \\ \boldsymbol{x}_{i0}'\boldsymbol{\beta}_0 + \boldsymbol{w}_i'\boldsymbol{\delta}_0 + \alpha_i & t = 0. \end{cases}$$

Since the Jacobian of the transformation of $\boldsymbol{\gamma}$ to $\boldsymbol{\zeta}$ is $\exp\left(\sum_{c=2}^{C-1} \zeta_c\right)$, we have

$$p(\boldsymbol{\zeta}|\cdots) = p(\boldsymbol{\gamma}|\cdots) \exp\left(\sum_{c=2}^{C-1} \zeta_c\right).$$

We use the following Metropolis-Hastings (M-H) algorithm for generating $\boldsymbol{\zeta}$. We use a multivariate t distribution, $\mathrm{Mt}(\boldsymbol{\zeta}|\hat{\boldsymbol{\zeta}}, \hat{\boldsymbol{\Sigma}}_\zeta, \nu)$, as a proposal distribution for generating $\boldsymbol{\zeta}$, where $\hat{\boldsymbol{\zeta}}$ is the mode of $p(\boldsymbol{\zeta}|\cdots)$ and

$$\hat{\boldsymbol{\Sigma}}_\zeta = \left\{\left[-\frac{\partial^2 \log p(\boldsymbol{\zeta}|\cdots)}{\partial \boldsymbol{\zeta} \partial \boldsymbol{\zeta}'}\right]_{\boldsymbol{\zeta}=\hat{\boldsymbol{\zeta}}}\right\}^{-1}.$$

The M-H algorithm for generating $\boldsymbol{\zeta}$ is as follows:

1. Let $\boldsymbol{\zeta}^{(l)}$ denote the value of $\boldsymbol{\zeta}$ at the lth iteration.

2. At the $(l+1)$th iteration, sample $\boldsymbol{\zeta}^p$ from $\mathrm{Mt}(\boldsymbol{\zeta}|\hat{\boldsymbol{\zeta}}, \hat{\boldsymbol{\Sigma}}_\zeta, \nu)$.

3. The transition probability from $\boldsymbol{\zeta}^{(l)}$ to $\boldsymbol{\zeta}^p$ is

$$\alpha = \min\left\{\frac{p(\boldsymbol{\zeta}^p|\cdots)\,\mathrm{Mt}(\boldsymbol{\zeta}^{(l)}|\hat{\boldsymbol{\zeta}}, \hat{\boldsymbol{\Sigma}}_\zeta, \nu)}{p(\boldsymbol{\zeta}^{(l)}|\cdots)\,\mathrm{Mt}(\boldsymbol{\zeta}^p|\hat{\boldsymbol{\zeta}}, \hat{\boldsymbol{\Sigma}}_\zeta, \nu)}, 1\right\}.$$

4. Generate $u \sim \mathrm{U}(0,1)$, the uniform distribution on $(0,1)$, and take

$$\boldsymbol{\zeta}^{(l+1)} = \begin{cases} \boldsymbol{\zeta}^p & \text{if } u < \alpha \\ \boldsymbol{\zeta}^{(l)} & \text{otherwise.} \end{cases}$$

We can obtain $\boldsymbol{\gamma}$ from $\boldsymbol{\zeta}$ using the equation

$$\gamma_c = \sum_{j=1}^c \exp \zeta_j, \ c = 2, \cdots, C-1.$$

D Derivation of Algorithm 4

D.1 FCD of ϕ

The FCD of ϕ is

$$\phi | \cdots \sim \mathrm{N}(\hat{\phi}, \hat{\omega}) 1_{(\phi \in (-1,1))},$$

where

$$\hat{\omega} = \left(\frac{1}{\tilde{\omega}} + \frac{1}{\sigma^2} \sum_{i=1}^{n} \sum_{t=1}^{T} z_{i(t-1)}^2 \right)^{-1}$$

$$\hat{\phi} = \hat{\omega} \left[\frac{\tilde{\phi}}{\tilde{\omega}} + \frac{1}{\sigma^2} \sum_{i=1}^{n} \sum_{t=1}^{T} z_{i(t-1)} (z_{it} - \boldsymbol{x}_{it}' \boldsymbol{\beta} - \boldsymbol{w}_i' \boldsymbol{\delta} - \alpha_i) \right].$$

D.2 FCD of $\boldsymbol{\beta}$

The FCD of $\boldsymbol{\beta}$ is

$$\boldsymbol{\beta} | \cdots \sim \mathrm{N}(\hat{\boldsymbol{\beta}}, \hat{\boldsymbol{B}}),$$

where

$$\hat{\boldsymbol{B}} = \left(\tilde{\boldsymbol{B}}^{-1} + \frac{1}{\sigma^2} \sum_{i=1}^{n} \sum_{t=1}^{T} \boldsymbol{x}_{it} \boldsymbol{x}_{it}' \right)^{-1}$$

$$\hat{\boldsymbol{\beta}} = \hat{\boldsymbol{B}} \left[\tilde{\boldsymbol{B}}^{-1} \tilde{\boldsymbol{\beta}} + \frac{1}{\sigma^2} \sum_{i=1}^{n} \sum_{t=1}^{T} \boldsymbol{x}_{it} (z_{it} - \phi z_{i(t-1)} - \boldsymbol{w}_i' \boldsymbol{\delta} - \alpha_i) \right].$$

D.3 FCD of $\boldsymbol{\delta}$

The FCD of $\boldsymbol{\delta}$ is

$$\boldsymbol{\delta} | \cdots \sim \mathrm{N}(\hat{\boldsymbol{\delta}}, \hat{\boldsymbol{D}}),$$

where

$$\hat{\boldsymbol{D}} = \left(\tilde{\boldsymbol{D}}^{-1} + \frac{T}{\sigma^2} \sum_{i=1}^{n} \boldsymbol{w}_i \boldsymbol{w}_i' \right)^{-1}$$

$$\hat{\boldsymbol{\delta}} = \hat{\boldsymbol{D}} \left[\tilde{\boldsymbol{D}}^{-1} \tilde{\boldsymbol{\delta}} + \frac{1}{\sigma^2} \sum_{i=1}^{n} \sum_{t=1}^{T} \boldsymbol{w}_i (z_{it} - \phi z_{i(t-1)} - \boldsymbol{x}_{it}' \boldsymbol{\beta} - \alpha_i) \right].$$

D.4 FCD of $\boldsymbol{\beta}_0$

The FCD of $\boldsymbol{\beta}_0$ is

$$\boldsymbol{\beta}_0 | \cdots \sim \mathrm{N}(\hat{\boldsymbol{\beta}}_0, \hat{\boldsymbol{B}}_0),$$

where

$$\hat{\boldsymbol{B}}_0 = \left(\tilde{\boldsymbol{B}}_0^{-1} + \frac{1}{\sigma^2} \sum_{i=1}^n \boldsymbol{x}_{i0} \boldsymbol{x}_{i0}' \right)^{-1}$$

$$\hat{\boldsymbol{\beta}}_0 = \hat{\boldsymbol{B}}_0 \left[\tilde{\boldsymbol{B}}_0^{-1} \tilde{\boldsymbol{\beta}}_0 + \frac{1}{\sigma^2} \sum_{i=1}^n \boldsymbol{x}_{i0} (z_{i0} - \boldsymbol{w}_i' \boldsymbol{\delta}_0 - \alpha_i) \right].$$

D.5   FCD of $\boldsymbol{\delta}_0$

The FCD of $\boldsymbol{\delta}_0$ is

$$\boldsymbol{\delta}_0 | \cdots \sim \mathrm{N}(\hat{\boldsymbol{\delta}}_0, \hat{\boldsymbol{D}}_0),$$

where

$$\hat{\boldsymbol{D}}_0 = \left(\tilde{\boldsymbol{D}}_0^{-1} + \frac{1}{\sigma^2} \sum_{i=1}^n \boldsymbol{w}_i \boldsymbol{w}_i' \right)^{-1}$$

$$\hat{\boldsymbol{\delta}}_0 = \hat{\boldsymbol{D}}_0 \left[\tilde{\boldsymbol{D}}_0^{-1} \tilde{\boldsymbol{\delta}}_0 + \frac{1}{\sigma^2} \sum_{i=1}^n \boldsymbol{w}_i (z_{i0} - \boldsymbol{x}_{i0}' \boldsymbol{\beta}_0 - \alpha_i) \right].$$

D.6   FCD of $\boldsymbol{\alpha}$

The FCD of α_i is

$$\alpha_i | \cdots \sim \mathrm{N}(\hat{\alpha}_i, \hat{\tau}), \ i = 1, \cdots, n,$$

where

$$\hat{\tau} = \left(\frac{1}{\tau} + \frac{T+1}{\sigma^2} \right)^{-1}$$

$$\hat{\alpha}_i = \hat{\tau} \left[\frac{\mu}{\tau} + \frac{1}{\sigma^2} \left\{ \sum_{t=1}^T (z_{it} - \phi z_{i(t-1)} - \boldsymbol{x}_{it}' \boldsymbol{\beta} - \boldsymbol{w}_i' \boldsymbol{\delta}) \right. \right.$$

$$\left. \left. + (z_{i0} - \boldsymbol{x}_{i0}' \boldsymbol{\beta}_0 - \boldsymbol{w}_i' \boldsymbol{\delta}_0) \right\} \right].$$

D.7   FCD of μ

The FCD of μ is

$$\mu | \cdots \sim \mathrm{N}(\hat{\mu}, \hat{\kappa}),$$

where

$$\hat{\kappa} = \left(\frac{1}{\tilde{\kappa}} + \frac{n}{\tau} \right)^{-1}, \ \hat{\mu} = \hat{\kappa} \left(\frac{\tilde{\mu}}{\tilde{\kappa}} + \frac{1}{\tau} \sum_{i=1}^n \alpha_i \right).$$

D.8 FCD of τ

The FCD of τ^{-1} is

$$\tau^{-1}|\cdots \sim \mathrm{Gam}(\hat{a}, \hat{b}),$$

where

$$\hat{a} = \tilde{a} + \frac{n}{2}, \ \hat{b} = \tilde{b} + \frac{1}{2}\sum_{i=1}^{n}(\alpha_i - \mu)^2.$$

D.9 FCD of σ^2

The FCD of σ^{-2} is

$$\sigma^{-2}|\cdots \sim \mathrm{Gam}(\hat{c}, \hat{d}),$$

where

$$\hat{c} = \tilde{c} + \frac{n}{2}(T+1)$$

$$\hat{d} = \tilde{d} + \frac{1}{2}\left[\sum_{i=1}^{n}\sum_{t=1}^{T}(z_{it} - \phi z_{i(t-1)} - \boldsymbol{x}_{it}'\boldsymbol{\beta} - \boldsymbol{w}_i'\boldsymbol{\delta} - \alpha_i)^2\right.$$

$$\left. + \sum_{i=1}^{n}(z_{i0} - \boldsymbol{x}_{i0}'\boldsymbol{\beta}_0 - \boldsymbol{w}_i'\boldsymbol{\delta}_0 - \alpha_i)^2\right]$$

D.10 FCD of z_0

The FCD of z_{i0} is

$$z_{i0}|\cdots \sim \mathrm{N}\left(\hat{z}_{i0}, \frac{\sigma^2}{1+\phi^2}\right)1_{(z_{i0}\in(\gamma_{c-1},\gamma_c])}, \ i = 1,\cdots,n,$$

where

$$\hat{z}_{i0} = \frac{\phi(z_{i1} - \boldsymbol{x}_{i1}'\boldsymbol{\beta} - \boldsymbol{w}_i'\boldsymbol{\delta} - \alpha_i) + (\boldsymbol{x}_{i0}'\boldsymbol{\beta}_0 + \boldsymbol{w}_i'\boldsymbol{\delta}_0 + \alpha_i)}{1+\phi^2}.$$

D.11 FCD of z

The FCD of z_{it} is

$$z_{it}|\cdots \sim \mathrm{N}\left(\hat{z}_{it}, \frac{\sigma^2}{1+\phi^2}\right)1_{(z_{it}\in(\gamma_{c-1},\gamma_c])}, \ t = 1,\cdots,T-1, \ i = 1,\cdots,n,$$

where

$$\hat{z}_{it} = \frac{(\phi z_{i(t-1)} + \boldsymbol{x}_{it}'\boldsymbol{\beta} + \boldsymbol{w}_i'\boldsymbol{\delta} + \alpha_i) + \phi(z_{i(t+1)} - \boldsymbol{x}_{i(t+1)}'\boldsymbol{\beta} - \boldsymbol{w}_i'\boldsymbol{\delta} - \alpha_i)}{1+\phi^2}.$$

Further, the FCD of z_{iT} is

$$z_{iT}|\cdots \sim \mathrm{N}\left(\phi z_{i(T-1)} + \boldsymbol{x}_{iT}'\boldsymbol{\beta} + \boldsymbol{w}_i'\boldsymbol{\delta} + \alpha_i, \sigma^2\right)1_{(z_{iT}\in(\gamma_{c-1},\gamma_c])},$$
$$i = 1,\cdots,n.$$

D.12 Sampling Algorithm of γ

The FCD of γ is

$$p(\boldsymbol{\gamma}|\cdots) \propto p(\boldsymbol{\zeta}(\boldsymbol{\gamma})) \prod_{i,t:y_{it}=2} \left[\Phi\left(\frac{\gamma_2 - \tilde{z}_{it}}{\sigma}\right) - \Phi\left(-\frac{\tilde{z}_{it}}{\sigma}\right) \right]$$

$$\times \prod_{i,t:y_{it}=3} \left[\Phi\left(\frac{\gamma_3 - \tilde{z}_{it}}{\sigma}\right) - \Phi\left(\frac{\gamma_2 - \tilde{z}_{it}}{\sigma}\right) \right]$$

$$\times \cdots \times \prod_{i,t:y_{it}=C-1} \left[\Phi\left(\frac{1 - \tilde{z}_{it}}{\sigma}\right) - \Phi\left(\frac{\gamma_{C-2} - \tilde{z}_{it}}{\sigma}\right) \right],$$

where

$$\tilde{z}_{it} = \begin{cases} \phi z_{i(t-1)} + \boldsymbol{x}'_{it}\boldsymbol{\beta} + \boldsymbol{w}'_i\boldsymbol{\delta} + \alpha_i & t = 1, \cdots, T \\ \boldsymbol{x}'_{i0}\boldsymbol{\beta}_0 + \boldsymbol{w}'_i\boldsymbol{\delta}_0 + \alpha_i & t = 0. \end{cases}$$

Since the Jacobian of the transformation of $\boldsymbol{\gamma}$ to $\boldsymbol{\zeta}$ is $\prod_{c=2}^{C-2} \frac{(1-\gamma_{c-1})\exp\zeta_c}{(1+\exp\zeta_c)^2}$, we have

$$p(\boldsymbol{\zeta}|\cdots) = p(\boldsymbol{\gamma}|\cdots) \prod_{c=2}^{C-2} \frac{(1-\gamma_{c-1})\exp\zeta_c}{(1+\exp\zeta_c)^2}.$$

We use the following Metropolis-Hastings (M-H) algorithm for generating $\boldsymbol{\zeta}$. We use a multivariate t distribution, $\mathrm{Mt}(\boldsymbol{\zeta}|\hat{\boldsymbol{\zeta}}, \hat{\boldsymbol{\Sigma}}_\zeta, \nu)$, as a proposal distribution for generating $\boldsymbol{\zeta}$, where $\hat{\boldsymbol{\zeta}}$ is the mode of $p(\boldsymbol{\zeta}|\cdots)$,

$$\hat{\boldsymbol{\Sigma}}_\zeta = \left\{ \left[-\frac{\partial^2 \log p(\boldsymbol{\zeta}|\cdots)}{\partial\boldsymbol{\zeta}\partial\boldsymbol{\zeta}'} \right]_{\boldsymbol{\zeta}=\hat{\boldsymbol{\zeta}}} \right\}^{-1}$$

and ν is the degrees of freedom. The M-H algorithm for generating $\boldsymbol{\zeta}$ is as follows:

1. Let $\boldsymbol{\zeta}^{(l)}$ denote the value of $\boldsymbol{\zeta}$ at the lth iteration.

2. At the $(l+1)$th iteration, sample $\boldsymbol{\zeta}^p$ from $\mathrm{Mt}(\boldsymbol{\zeta}|\hat{\boldsymbol{\zeta}}, \hat{\boldsymbol{\Sigma}}_\zeta, \nu)$.

3. The transition probability from $\boldsymbol{\zeta}^{(l)}$ to $\boldsymbol{\zeta}^p$ is

$$\alpha = \min\left\{ \frac{p(\boldsymbol{\zeta}^p|\cdots)\,\mathrm{Mt}(\boldsymbol{\zeta}^{(l)}|\hat{\boldsymbol{\zeta}}, \hat{\boldsymbol{\Sigma}}_\zeta, \nu)}{p(\boldsymbol{\zeta}^{(l)}|\cdots)\,\mathrm{Mt}(\boldsymbol{\zeta}^p|\hat{\boldsymbol{\zeta}}, \hat{\boldsymbol{\Sigma}}_\zeta, \nu)}, 1 \right\}.$$

4. Generate $u \sim \mathrm{U}(0,1)$, the uniform distribution on $(0,1)$, and take

$$\boldsymbol{\zeta}^{(l+1)} = \begin{cases} \boldsymbol{\zeta}^p & \text{if } u < \alpha \\ \boldsymbol{\zeta}^{(l)} & \text{otherwise.} \end{cases}$$

We can obtain $\boldsymbol{\gamma}$ from $\boldsymbol{\zeta}$ using the equation

$$\gamma_c = \frac{\gamma_{c-1} + \exp\zeta_c}{1 + \exp\zeta_c}, \quad c = 2, \cdots, C-2.$$

E The Partial Effects of the Explanatory Variables on the Response Probabilities

E.1 Algorithms 1–3

From (1) and (2), we have

$$
\begin{aligned}
\Pr(y_{it} = c) &= \Pr(\gamma_{c-1} < z_{it} \leq \gamma_c) \\
&= \begin{cases}
F\big(-\operatorname{E}(z_{it})\big) & c = 1 \\
F\big(\gamma_c - \operatorname{E}(z_{it})\big) - F\big(\gamma_{c-1} - \operatorname{E}(z_{it})\big) & 2 \leq c \leq C - 1 \\
1 - F\big(\gamma_{C-1} - \operatorname{E}(z_{it})\big) & c = C
\end{cases}
\end{aligned}
\tag{11}
$$
$$
t = 0, 1, \cdots, T, \ i = 1, \cdots, n,
$$

where $F(\cdot)$ denotes the distribution function of the standard normal distribution. If an explanatory variable x (such as $z_{i(t-1)}, x_{itj}, w_{ij}$) is continuous, we obtain the following partial effect of this explanatory variable on the response probabilities by differentiating (11): for $t = 1, \cdots, T, \ i = 1, \cdots, n,$

$$
\begin{aligned}
&\frac{\partial \Pr(y_{it} = c)}{\partial x} \\
&= \begin{cases}
-\beta_x f\big(-(\phi z_{i(t-1)} + \boldsymbol{x}_{it}'\boldsymbol{\beta} + \boldsymbol{w}_i'\boldsymbol{\delta} + \alpha_i)\big) & c = 1 \\
\beta_x \Big[f\big(\gamma_{c-1} - (\phi z_{i(t-1)} + \boldsymbol{x}_{it}'\boldsymbol{\beta} + \boldsymbol{w}_i'\boldsymbol{\delta} + \alpha_i)\big) \\
\quad - f\big(\gamma_c - (\phi z_{i(t-1)} + \boldsymbol{x}_{it}'\boldsymbol{\beta} + \boldsymbol{w}_i'\boldsymbol{\delta} + \alpha_i)\big)\Big] & 2 \leq c \leq C - 1 \\
\beta_x f\big(\gamma_{C-1} - (\phi z_{i(t-1)} + \boldsymbol{x}_{it}'\boldsymbol{\beta} + \boldsymbol{w}_i'\boldsymbol{\delta} + \alpha_i)\big) & c = C
\end{cases}
\end{aligned}
\tag{12}
$$

and for $t = 0, \ i = 1, \cdots, n,$

$$
\begin{aligned}
&\frac{\partial \Pr(y_{i0} = c)}{\partial x} \\
&= \begin{cases}
-\beta_x f\big(-(\boldsymbol{x}_{i0}'\boldsymbol{\beta}_0 + \boldsymbol{w}_i'\boldsymbol{\delta}_0 + \alpha_i)\big) & c = 1 \\
\beta_x \Big[f\big(\gamma_{c-1} - (\boldsymbol{x}_{i0}'\boldsymbol{\beta}_0 + \boldsymbol{w}_i'\boldsymbol{\delta}_0 + \alpha_i)\big) \\
\quad - f\big(\gamma_c - (\boldsymbol{x}_{i0}'\boldsymbol{\beta}_0 + \boldsymbol{w}_i'\boldsymbol{\delta}_0 + \alpha_i)\big)\Big] & 2 \leq c \leq C - 1 \\
\beta_x f\big(\gamma_{C-1} - (\boldsymbol{x}_{i0}'\boldsymbol{\beta}_0 + \boldsymbol{w}_i'\boldsymbol{\delta}_0 + \alpha_i)\big) & c = C,
\end{cases}
\end{aligned}
\tag{13}
$$

where β_x is a coefficient of the explanatory variable x. From the posterior results of the Bayesian analysis, we can calculate the individual partial effects of the explanatory variables on response probabilities (12) and (13). Thereafter, we can average these individual effects. For example, the partial effects of $z_{(-1)}$, x_j, and w_j on the response probabilities $\Pr(y = c)$ of $2 \leq c \leq C - 1$ are as

follows:

$$\frac{\partial \Pr(y=c)}{\partial z_{(-1)}} = \frac{\phi}{nT} \sum_{i=1}^{n} \sum_{t=1}^{T} \Big[f\big(\gamma_{c-1} - (\phi z_{i(t-1)} + \boldsymbol{x}_{it}'\boldsymbol{\beta} + \boldsymbol{w}_i'\boldsymbol{\delta} + \alpha_i)\big)$$

$$- f\big(\gamma_c - (\phi z_{i(t-1)} + \boldsymbol{x}_{it}'\boldsymbol{\beta} + \boldsymbol{w}_i'\boldsymbol{\delta} + \alpha_i)\big)\Big]$$

$$\frac{\partial \Pr(y=c)}{\partial x_j} = \frac{\beta_j}{nT} \sum_{i=1}^{n} \sum_{t=1}^{T} \Big[f\big(\gamma_{c-1} - (\phi z_{i(t-1)} + \boldsymbol{x}_{it}'\boldsymbol{\beta} + \boldsymbol{w}_i'\boldsymbol{\delta} + \alpha_i)\big)$$

$$- f\big(\gamma_c - (\phi z_{i(t-1)} + \boldsymbol{x}_{it}'\boldsymbol{\beta} + \boldsymbol{w}_i'\boldsymbol{\delta} + \alpha_i)\big)\Big]$$

$$+ \frac{\beta_{0j}}{n} \sum_{i=1}^{n} \Big[f\big(\gamma_{c-1} - (\boldsymbol{x}_{i0}'\boldsymbol{\beta}_0 + \boldsymbol{w}_i'\boldsymbol{\delta}_0 + \alpha_i)\big)$$

$$- f\big(\gamma_c - (\boldsymbol{x}_{i0}'\boldsymbol{\beta}_0 + \boldsymbol{w}_i'\boldsymbol{\delta}_0 + \alpha_i)\big)\Big]$$

$$\frac{\partial \Pr(y=c)}{\partial w_j} = \frac{\delta_j}{nT} \sum_{i=1}^{n} \sum_{t=1}^{T} \Big[f\big(\gamma_{c-1} - (\phi z_{i(t-1)} + \boldsymbol{x}_{it}'\boldsymbol{\beta} + \boldsymbol{w}_i'\boldsymbol{\delta} + \alpha_i)\big)$$

$$- f\big(\gamma_c - (\phi z_{i(t-1)} + \boldsymbol{x}_{it}'\boldsymbol{\beta} + \boldsymbol{w}_i'\boldsymbol{\delta} + \alpha_i)\big)\Big]$$

$$+ \frac{\delta_{0j}}{n} \sum_{i=1}^{n} \Big[f\big(\gamma_{c-1} - (\boldsymbol{x}_{i0}'\boldsymbol{\beta}_0 + \boldsymbol{w}_i'\boldsymbol{\delta}_0 + \alpha_i)\big)$$

$$- f\big(\gamma_c - (\boldsymbol{x}_{i0}'\boldsymbol{\beta}_0 + \boldsymbol{w}_i'\boldsymbol{\delta}_0 + \alpha_i)\big)\Big].$$

E.2 Algorithm 4

From (1) and (7), we have

$$\Pr(y_{it} = c) = \Pr(\gamma_{c-1} < z_{it} \leq \gamma_c)$$

$$= \Pr\left(\frac{\gamma_{c-1} - \mathrm{E}(z_{it})}{\sigma} < \frac{z_{it} - \mathrm{E}(z_{it})}{\sigma} \leq \frac{\gamma_c - \mathrm{E}(z_{it})}{\sigma} \right)$$

$$= \begin{cases} F\left(\dfrac{-\mathrm{E}(z_{it})}{\sigma} \right) & c = 1 \\ F\left(\dfrac{\gamma_c - \mathrm{E}(z_{it})}{\sigma} \right) - F\left(\dfrac{\gamma_{c-1} - \mathrm{E}(z_{it})}{\sigma} \right) & 2 \leq c \leq C - 1 \\ 1 - F\left(\dfrac{1 - \mathrm{E}(z_{it})}{\sigma} \right) & c = C \end{cases} \quad (14)$$

$$t = 0, 1, \cdots, T, \ i = 1, \cdots, n.$$

If an explanatory variable x (such as $z_{i(t-1)}, x_{itj}, w_{ij}$) is continuous, we obtain the following partial effect of this explanatory variable on the response proba-

bilities by differentiating (14): for $t = 1, \cdots, T, \ i = 1, \cdots, n,$

$$
\frac{\partial \Pr(y_{it} = c)}{\partial x}
$$
$$
=
\begin{cases}
-\dfrac{\beta_x}{\sigma} f\left(-\dfrac{1}{\sigma}(\phi z_{i(t-1)} + \boldsymbol{x}'_{it}\boldsymbol{\beta} + \boldsymbol{w}'_i\boldsymbol{\delta} + \alpha_i)\right) & c = 1 \\[2ex]
\dfrac{\beta_x}{\sigma}\left[f\left(\dfrac{1}{\sigma}\left\{\gamma_{c-1} - (\phi z_{i(t-1)} + \boldsymbol{x}'_{it}\boldsymbol{\beta} + \boldsymbol{w}'_i\boldsymbol{\delta} + \alpha_i)\right\}\right)\right. \\[2ex]
\qquad \left. -f\left(\dfrac{1}{\sigma}\left\{\gamma_c - (\phi z_{i(t-1)} + \boldsymbol{x}'_{it}\boldsymbol{\beta} + \boldsymbol{w}'_i\boldsymbol{\delta} + \alpha_i)\right\}\right)\right] & 2 \le c \le C-1 \\[2ex]
\dfrac{\beta_x}{\sigma} f\left(\dfrac{1}{\sigma}\left\{1 - (\phi z_{i(t-1)} + \boldsymbol{x}'_{it}\boldsymbol{\beta} + \boldsymbol{w}'_i\boldsymbol{\delta} + \alpha_i)\right\}\right) & c = C
\end{cases}
$$

$$(15)$$

and for $t = 0, \ i = 1, \cdots, n,$

$$
\frac{\partial \Pr(y_{i0} = c)}{\partial x}
$$
$$
=
\begin{cases}
-\dfrac{\beta_x}{\sigma} f\left(-\dfrac{1}{\sigma}(\boldsymbol{x}'_{i0}\boldsymbol{\beta}_0 + \boldsymbol{w}'_i\boldsymbol{\delta}_0 + \alpha_i)\right) & c = 1 \\[2ex]
\dfrac{\beta_x}{\sigma}\left[f\left(\dfrac{1}{\sigma}\left\{\gamma_{c-1} - (\boldsymbol{x}'_{i0}\boldsymbol{\beta}_0 + \boldsymbol{w}'_i\boldsymbol{\delta}_0 + \alpha_i)\right\}\right)\right. \\[2ex]
\qquad \left. -f\left(\dfrac{1}{\sigma}\left\{\gamma_c - (\boldsymbol{x}'_{i0}\boldsymbol{\beta}_0 + \boldsymbol{w}'_i\boldsymbol{\delta}_0 + \alpha_i)\right\}\right)\right] & 2 \le c \le C-1 \\[2ex]
\dfrac{\beta_x}{\sigma} f\left(\dfrac{1}{\sigma}\left\{1 - (\boldsymbol{x}'_{i0}\boldsymbol{\beta}_0 + \boldsymbol{w}'_i\boldsymbol{\delta}_0 + \alpha_i)\right\}\right) & c = C,
\end{cases}
$$

$$(16)$$

where β_x is a coefficient of the explanatory variable x. From the posterior results of the Bayesian analysis, we can calculate the individual partial effects of the explanatory variables on response probabilities (15) and (16). Thereafter, we can average these individual effects. For example, the partial effects of $z_{(-1)}$, x_j, and w_j on the response probabilities $\Pr(y = c)$ of $2 \le c \le C-1$ are as

follows:

$$
\frac{\partial \Pr(y=c)}{\partial z_{(-1)}} = \frac{\phi}{nT\sigma} \sum_{i=1}^{n} \sum_{t=1}^{T} \left[f\left(\frac{1}{\sigma} \left\{ \gamma_{c-1} - (\phi z_{i(t-1)} + \boldsymbol{x}'_{it}\boldsymbol{\beta} + \boldsymbol{w}'_i\boldsymbol{\delta} + \alpha_i) \right\} \right) \right.
$$
$$
\left. - f\left(\frac{1}{\sigma} \left\{ \gamma_c - (\phi z_{i(t-1)} + \boldsymbol{x}'_{it}\boldsymbol{\beta} + \boldsymbol{w}'_i\boldsymbol{\delta} + \alpha_i) \right\} \right) \right]
$$

$$
\frac{\partial \Pr(y=c)}{\partial x_j} = \frac{\beta_j}{nT\sigma} \sum_{i=1}^{n} \sum_{t=1}^{T} \left[f\left(\frac{1}{\sigma} \left\{ \gamma_{c-1} - (\phi z_{i(t-1)} + \boldsymbol{x}'_{it}\boldsymbol{\beta} + \boldsymbol{w}'_i\boldsymbol{\delta} + \alpha_i) \right\} \right) \right.
$$
$$
\left. - f\left(\frac{1}{\sigma} \left\{ \gamma_c - (\phi z_{i(t-1)} + \boldsymbol{x}'_{it}\boldsymbol{\beta} + \boldsymbol{w}'_i\boldsymbol{\delta} + \alpha_i) \right\} \right) \right]
$$
$$
+ \frac{\beta_{0j}}{n\sigma} \sum_{i=1}^{n} \left[f\left(\frac{1}{\sigma} \left\{ \gamma_{c-1} - (\boldsymbol{x}'_{i0}\boldsymbol{\beta}_0 + \boldsymbol{w}'_i\boldsymbol{\delta}_0 + \alpha_i) \right\} \right) \right.
$$
$$
\left. - f\left(\frac{1}{\sigma} \left\{ \gamma_c - (\boldsymbol{x}'_{i0}\boldsymbol{\beta}_0 + \boldsymbol{w}'_i\boldsymbol{\delta}_0 + \alpha_i) \right\} \right) \right]
$$

$$
\frac{\partial \Pr(y=c)}{\partial w_j} = \frac{\delta_j}{nT\sigma} \sum_{i=1}^{n} \sum_{t=1}^{T} \left[f\left(\frac{1}{\sigma} \left\{ \gamma_{c-1} - (\phi z_{i(t-1)} + \boldsymbol{x}'_{it}\boldsymbol{\beta} + \boldsymbol{w}'_i\boldsymbol{\delta} + \alpha_i) \right\} \right) \right.
$$
$$
\left. - f\left(\frac{1}{\sigma} \left\{ \gamma_c - (\phi z_{i(t-1)} + \boldsymbol{x}'_{it}\boldsymbol{\beta} + \boldsymbol{w}'_i\boldsymbol{\delta} + \alpha_i) \right\} \right) \right]
$$
$$
+ \frac{\delta_{0j}}{n\sigma} \sum_{i=1}^{n} \left[f\left(\frac{1}{\sigma} \left\{ \gamma_{c-1} - (\boldsymbol{x}'_{i0}\boldsymbol{\beta}_0 + \boldsymbol{w}'_i\boldsymbol{\delta}_0 + \alpha_i) \right\} \right) \right.
$$
$$
\left. - f\left(\frac{1}{\sigma} \left\{ \gamma_c - (\boldsymbol{x}'_{i0}\boldsymbol{\beta}_0 + \boldsymbol{w}'_i\boldsymbol{\delta}_0 + \alpha_i) \right\} \right) \right].
$$

E.3   For Discrete Variables

If an explanatory variable x is discrete, we can obtain the partial effect of x on the response probability $\Pr(y=c)$ by calculating the average of $\Delta \Pr(y=c) = \Pr(y=c|x=x_0+1) - \Pr(y=c|x=x_0)$.

Acknowledgments

References

Albert, J. H. and S. Chib (1993). Bayesian analysis of binary and polychotomous response data. *Journal of the American Statistical Association 88*(422), 669–

679.

Albert, J. H. and S. Chib (2001). Sequential ordinal modeling with applications to survival data. *Biometrics 57*(3), 829–836.

Baltagi, B. H. (2001). *Econometric Analysis of Panel Data, 2nd ed.* Chichester: Wiley.

Blanchflower, D. G. and A. J. Oswald (2004). Well-being over time in Britain and the USA. *Journal of Public Economics 88*(7–8), 1359–1386.

Chen, M.-H. and D. K. Dey (2000). Bayesian analysis for correlated ordinal data models. In D. K. Dey, S. K. Ghosh, and B. K. Mallick (Eds.), *Generalized Linear Models: A Bayesian Perspective*, pp. 133–157. New York: Marcel Dekker.

Clark, A. E. and A. J. Oswald (2002). A simple statistical method for measuring how life events affect happiness. *International Journal of Epidemiology 31*(6), 1139–1144.

Cowles, M. K. (1996). Accelerating Monte Carlo Markov chain convergence for cumulative-link generalized linear models. *Statistics and Computing 6*(2), 101–111.

Damien, P. and S. G. Walker (2001). Sampling truncated normal, beta, and gamma densities. *Journal of Computational and Graphical Statistics 10*(2), 206–215.

Doornik, J. A. (2007). *An Object-oriented Matrix Programming Language Ox^{TM} 5.* London: Timberlake Consultants Ltd.

Frey, B. S. and A. Stutzer (2002a). *Happiness and Economics.* Princeton: Princeton University Press.

Frey, B. S. and A. Stutzer (2002b). What can economists learn from happiness research? *Journal of Economic Literature 40*(2), 402–435.

Frey, B. S. and A. Stutzer (2005). Happiness research: State and prospects. *Review of Social Economy 63*(2), 207–228.

Geweke, J. (1992). Evaluating the accuracy of sampling-based approaches to the calculation of posterior moments. In J. M. Bernardo, J. O. Berger, A. P. Dawid, and A. F. M. Smith (Eds.), *Bayesian Statistics 4*, pp. 169–193. Oxford: Oxford University Press.

Hausman, J. A. and W. E. Taylor (1981). Panel data and unobservable individual effects. *Econometrica 49*(6), 1377–1398.

Hsiao, C. (2003). *Analysis of Panel Data, 2nd ed.* New York: Cambridge University Press.

Jeliazkov, I., J. Graves, and M. Kutzbach (2008). Fitting and comparison of models for multivariate ordinal outcomes. In S. Chib, G. Koop, W. Griffiths, and D. Terrell (Eds.), *Advances in Econometrics: Bayesian Econometrics*, Volume 23, pp. 115–156. Bingley: Emerald.

Johnson, V. E. and J. H. Albert (1999). *Ordinal Data Modeling.* New York: Springer.

Koop, G. (2003). *Bayesian Econometrics.* Chichester: Wiley.

Liu, J. S. and C. Sabatti (2000). Generalized Gibbs sampler and multigrid Monte Carlo for Bayesian computation. *Biometrika 87*(2), 353–369.

Müller, G. and C. Czado (2005). An autoregressive ordered probit model with application to high-frequency financial data. *Journal of Computational and Graphical Statistics 14*(2), 320–338.

Nandram, B. and M.-H. Chen (1996). Reparameterizing the generalized linear model to accelerate Gibbs sampler convergence. *Journal of Statistical Computation and Simulation 54*(1–3), 129–144.

Wooldridge, J. M. (2002). *Econometric Analysis of Cross Section and Panel Data.* Cambridge: MIT Press.

Wooldridge, J. M. (2005). Simple solutions to the initial conditions problem in dynamic, nonlinear panel data models with unobserved heterogeneity. *Journal of Applied Econometrics 20*(1), 39–54.

Table 1: Numbers in categories of y

	1	2	3	4
period 0	24	80	67	29
period 1	28	67	60	45
period 2	30	59	61	50
period 3	31	56	62	51
period 4	30	53	70	47
period 5	36	51	57	56
period 6	39	47	60	54
period 7	41	51	54	54
period 8	35	58	57	50
period 9	38	43	67	52

Table 2: Summary statistics of explanatory variables

x					
	mean[a]	sd[a]	min[a]	50%[a]	max[a]
period 0	1.0021	1.6490	-3.3643	0.9026	8.1696
period 1	0.2128	1.1573	-2.5351	0.1995	4.2354
period 2	0.0348	1.1499	-2.7772	-0.0147	2.9722
period 3	0.0863	1.0062	-2.8848	0.1469	2.5879
period 4	-0.0053	1.0518	-2.9270	0.0018	2.8444
period 5	0.0149	1.1228	-2.4023	-0.1354	3.3439
period 6	-0.0502	1.0602	-3.4296	0.0204	2.5537
period 7	-0.0246	1.0200	-2.7011	-0.0001	2.7871
period 8	0.0384	1.1257	-2.5074	-0.0068	3.3792
period 9	-0.0037	1.0616	-2.8790	-0.1062	3.0002

w				
mean	sd	min	50%	max
1.9586	1.9500	-3.4154	1.8541	7.5332

a: 'mean' and 'sd' denote the sample mean and sample standard deviation. 'min,' '50%' and 'max' denote the minimum value, median and maximum value.

Table 3: Posterior results for parameters (simulated data)

			Algorithm 1			
	mean[a]	sd[a]	mean/sd	95%CI[a]	CD[a]	p-val[a]
ϕ	0.515	0.012	43.461	$(0.491, 0.538)$	-0.052	0.958
β_1	2.203	0.070	31.546	$(2.066, 2.338)$	-4.559	0.000
δ_1	1.643	0.054	30.384	$(1.540, 1.747)$	-4.241	0.000
β_{01}	2.143	0.083	25.731	$(1.981, 2.308)$	-4.138	0.000
δ_{01}	1.704	0.063	26.956	$(1.585, 1.829)$	-4.223	0.000
γ_2	5.558	0.169	32.902	$(5.275, 5.915)$	-4.430	0.000
γ_3	11.081	0.282	39.308	$(10.568, 11.609)$	-4.226	0.000
			Algorithm 2			
	mean[a]	sd[a]	mean/sd	95%CI[a]	CD[a]	p-val[a]
ϕ	0.513	0.012	41.858	$(0.489, 0.536)$	0.359	0.720
β_1	2.139	0.091	23.380	$(1.961, 2.321)$	0.045	0.964
δ_1	1.600	0.068	23.368	$(1.470, 1.736)$	0.140	0.889
β_{01}	2.079	0.103	20.088	$(1.877, 2.284)$	0.068	0.945
δ_{01}	1.651	0.080	20.615	$(1.496, 1.812)$	-0.246	0.805
γ_2	5.385	0.242	22.274	$(4.919, 5.864)$	-0.763	0.445
γ_3	10.735	0.433	24.770	$(9.901, 11.612)$	-0.028	0.978
			Algorithm 3			
	mean[a]	sd[a]	mean/sd	95%CI[a]	CD[a]	p-val[a]
ϕ	0.514	0.012	43.143	$(0.491, 0.537)$	2.187	0.029
β_1	2.210	0.089	24.841	$(2.044, 2.388)$	1.516	0.129
δ_1	1.651	0.068	24.357	$(1.520, 1.787)$	1.335	0.182
β_{01}	2.147	0.101	21.257	$(1.955, 2.351)$	1.712	0.087
δ_{01}	1.709	0.078	21.774	$(1.562, 1.869)$	1.921	0.055
γ_2	5.562	0.233	23.856	$(5.136, 6.045)$	1.805	0.071
γ_3	11.112	0.419	26.527	$(10.351, 11.965)$	1.814	0.070
			Algorithm 4[b]			
	mean[a]	sd[a]	mean/sd	95%CI[a]	CD[a]	p-val[a]
ϕ	0.496	0.012	40.296	$(0.472, 0.520)$	-1.757	0.079
β_1	2.287	0.108	21.201	$(2.085, 2.508)$	-0.041	0.967
δ_1	1.751	0.085	20.539	$(1.590, 1.927)$	1.118	0.264
β_{01}	2.204	0.117	18.859	$(1.986, 2.442)$	-0.344	0.731
δ_{01}	1.739	0.093	18.706	$(1.560, 1.924)$	-0.026	0.979
γ_2	5.665	0.273	20.724	$(5.147, 6.214)$	-0.030	0.976
γ_3	11.308	0.506	22.349	$(10.355, 12.339)$	-0.153	0.878

a: 'mean' and 'sd' denote the posterior mean and posterior standard deviation. '95%CI' denotes the 95% credible interval. 'CD' and 'p-val' denote the convergence diagnostic statistic of MCMC proposed by Geweke (1992) and its p value.

b: The values of parameters except for ϕ are transformed ones devided by σ.

Table 4 (a): Life satisfaction (frequency)

wave	strongly dissatisfied and dissatisfied 1	moderately satisfied 2	satisfied 3	strongly satisfied 4
6	60	86	160	32
7	48	87	173	30
8	66	89	156	27
9	49	96	165	28
10	62	86	169	21
11	71	91	153	23

Table 4 (b): Life satisfaction (%)

wave	strongly dissatisfied and dissatisfied 1	moderately satisfied 2	satisfied 3	strongly satisfied 4
6	17.8	25.4	47.3	9.5
7	14.2	25.7	51.2	8.9
8	19.5	26.3	46.2	8.0
9	14.5	28.4	48.8	8.3
10	18.3	25.4	50.0	6.2
11	21.0	26.9	45.3	6.8

Table 5 (a): Summary statistics 1

wave	married no(0)	married yes(1)	inwork no(0)	inwork yes(1)
6	77	261	159	179
7	72	266	147	191
8	62	276	153	185
9	85	253	125	213
10	76	262	131	207
11	84	254	126	212

Table 5 (b): Summary statistics 2

wave	mean[a]	sd[a]	min[a]	50%[a]	max[a]
			loginc		
6	5.7619	0.5324	2.2986	5.7979	7.0778
7	5.7682	0.5388	2.1284	5.8029	7.6070
8	5.7370	0.6343	0.0000	5.7429	7.5264
9	5.7553	0.7120	0.0000	5.8476	7.9722
10	5.7594	0.7173	0.0000	5.8259	7.9516
11	5.7689	0.5377	3.1072	5.7991	7.2181

wave	mean	sd	min	50%	max
			logsav		
6	3.1013	2.4569	0.0000	3.9021	8.5118
7	3.1273	2.4993	0.0000	3.9050	8.1432
8	3.2913	2.4653	0.0000	4.0559	8.1502
9	2.8816	2.6231	0.0000	3.4508	8.1776
10	3.1855	2.6601	0.0000	3.9282	8.4608
11	2.9902	2.6571	0.0000	3.5965	8.3235

explanatory variable	mean	sd	min	50%	max
age[b]	33.75	3.24	25	34	39
educ[b]	13.28	1.56	9	14	18

a: 'mean' and 'sd' denote the sample mean and sample standard deviation. 'min,' '50%' and 'max' denote the minimum value, median and maximum value.

b: The data of 'age' and 'educ' are obtained from Panel 6.

Table 6: Posterior results for parameters (JPSC data)

	mean[a]	sd[a]	mean/sd	95%CI[a]	CD[a]	p-val[a]
Algorithm 2						
ϕ ($z_{(-1)}$)	0.1562	0.0301	5.1953	$(0.0959, 0.2139)$	0.3487	0.7273
β_1 (married)	-0.2809	0.0873	-3.2193	$(-0.4498, -0.1106)$	0.0132	0.9895
β_2 (loginc)	0.2008	0.0455	4.4151	$(0.1111, 0.2899)$	-0.7148	0.4748
β_3 (logsav)	0.0665	0.0144	4.6148	$(0.0383, 0.0946)$	0.1413	0.8876
β_4 (inwork)	-0.1781	0.0572	-3.1130	$(-0.2894, -0.0658)$	-0.1227	0.9024
δ_1 (age)	0.0107	0.0084	1.2852	$(-0.0054, 0.0270)$	1.2050	0.2282
δ_2 (educ)	0.0052	0.0177	0.2916	$(-0.0294, 0.0399)$	1.3350	0.1819
β_{01} (married)	0.0033	0.2118	0.0154	$(-0.4088, 0.4172)$	-0.3903	0.6963
β_{02} (loginc)	0.2970	0.0941	3.1560	$(0.1114, 0.4812)$	-0.6232	0.5331
β_{03} (logsav)	0.0725	0.0345	2.0983	$(0.0057, 0.1403)$	-0.2439	0.8073
β_{04} (inwork)	-0.0442	0.1338	-0.3302	$(-0.3082, 0.2154)$	-1.5468	0.1219
δ_{01} (age)	-0.0302	0.0157	-1.9195	$(-0.0613, 0.0006)$	1.6928	0.0905
δ_{02} (educ)	0.0581	0.0343	1.6953	$(-0.0095, 0.1256)$	1.4569	0.1452
γ_2	0.8239	0.0319	25.8420	$(0.7619, 0.8868)$	1.9339	0.0531
γ_3	2.4624	0.0522	47.1700	$(2.3620, 2.5663)$	1.3778	0.1683
Algorithm 4						
ϕ ($z_{(-1)}$)	0.1601	0.0603	2.6577	$(0.0433, 0.2782)$	-0.9979	0.3183
β_1 (married)	-0.2475	0.1886	-1.3120	$(-0.6170, 0.1202)$	-4.8318	0.0000
β_2 (loginc)	0.2163	0.1075	2.0128	$(-0.0039, 0.4203)$	-3.4980	0.0005
β_3 (logsav)	0.0596	0.0307	1.9389	$(-0.0008, 0.1195)$	6.3138	0.0000
β_4 (inwork)	-0.1686	0.1237	-1.3626	$(-0.4066, 0.0730)$	1.6420	0.1006
δ_1 (age)	0.0152	0.0181	0.8376	$(-0.0200, 0.0501)$	-7.4842	0.0000
δ_2 (educ)	0.0122	0.0387	0.3158	$(-0.0651, 0.0872)$	-6.6902	0.0000
β_{01} (married)	0.0346	0.4613	0.0750	$(-0.8732, 0.9355)$	-1.3203	0.1868
β_{02} (loginc)	0.3221	0.2233	1.4423	$(-0.1254, 0.7518)$	-2.2819	0.0225
β_{03} (logsav)	0.0641	0.0736	0.8715	$(-0.0808, 0.2067)$	3.7356	0.0002
β_{04} (inwork)	-0.0434	0.2925	-0.1485	$(-0.6105, 0.5335)$	1.7447	0.0810
δ_{01} (age)	-0.0257	0.0356	-0.7214	$(-0.0951, 0.0436)$	-6.2490	0.0000
δ_{02} (educ)	0.0611	0.0765	0.7982	$(-0.0890, 0.2117)$	-6.5632	0.0000
γ_2	0.7791	0.0375	20.7760	$(0.7067, 0.8533)$	2.2615	0.0237
γ_3	2.3166	0.0533	43.5000	$(2.2130, 2.4213)$	0.7597	0.4474

a: 'mean' and 'sd' denote the posterior mean and posterior standard deviation. 95%CI denotes the 95% credible interval. 'CD' and 'p-val' denote the convergence diagnostic statistic of MCMC proposed by Geweke (1992) and its p value.

b: The values of parameters except for ϕ are transformed ones devided by σ.

Table 7: Posterior results for partial effects on response probabilities

explanatory variables		partial effects on response probabilities					
		mean[a]	sd[a]	mean/sd	95%CI[a]	CD[a]	p-val[a]
$z_{(-1)}$	$y=1$	−0.0385	0.0075	−5.1213	(−0.0529, −0.0234)	0.4257	0.6703
	$y=2$	−0.0200	0.0038	−5.2098	(−0.0273, −0.0123)	−1.3173	0.1877
	$y=3$	0.0366	0.0069	5.2790	(0.0226, 0.0499)	−0.1135	0.9097
	$y=4$	0.0219	0.0045	4.8807	(0.0131, 0.0307)	0.8923	0.3722
x_1	$y=1$	0.0529	0.0179	2.9577	(0.0168, 0.0869)	−0.1125	0.9104
married	$y=2$	0.0319	0.0113	2.8314	(0.0102, 0.0545)	0.5584	0.5766
	$y=3$	−0.0475	0.0154	−3.0740	(−0.0757, −0.0155)	−0.0010	0.9992
	$y=4$	−0.0373	0.0139	−2.6769	(−0.0660, −0.0114)	−0.3345	0.7380
x_2	$y=1$	−0.1196	0.0266	−4.5032	(−0.1715, −0.0672)	0.8817	0.3779
loginc	$y=2$	−0.0654	0.0149	−4.3958	(−0.0946, −0.0362)	0.3746	0.7079
	$y=3$	0.1126	0.0247	4.5582	(0.0634, 0.1609)	−0.7608	0.4468
	$y=4$	0.0724	0.0172	4.2127	(0.0394, 0.1070)	−0.6157	0.5381
x_3	$y=1$	−0.0335	0.0090	−3.7062	(−0.0514, −0.0162)	0.2893	0.7724
logsav	$y=2$	−0.0181	0.0049	−3.6672	(−0.0276, −0.0084)	−0.5238	0.6005
	$y=3$	0.0315	0.0083	3.7918	(0.0154, 0.0478)	−0.0378	0.9699
	$y=4$	0.0201	0.0058	3.4871	(0.0091, 0.0318)	0.0839	0.9331
x_4	$y=1$	0.0377	0.0126	2.9814	(0.0129, 0.0624)	0.0077	0.9939
inwork	$y=2$	0.0206	0.0072	2.8775	(0.0068, 0.0349)	1.4888	0.1366
	$y=3$	−0.0359	0.0120	−2.9928	(−0.0594, −0.0123)	−0.2444	0.8070
	$y=4$	−0.0224	0.0078	−2.8490	(−0.0381, −0.0072)	−1.0815	0.2795
w_1	$y=1$	−0.0010	0.0019	−0.5075	(−0.0047, 0.0028)	−1.3179	0.1876
age	$y=2$	−0.0005	0.0010	−0.4946	(−0.0025, 0.0015)	−1.4062	0.1597
	$y=3$	0.0010	0.0018	0.5219	(−0.0026, 0.0045)	1.3253	0.1851
	$y=4$	0.0005	0.0011	0.4713	(−0.0016, 0.0028)	1.3862	0.1657
w_2	$y=1$	−0.0032	0.0040	−0.8133	(−0.0110, 0.0046)	−1.3842	0.1663
educ	$y=2$	−0.0019	0.0021	−0.8934	(−0.0062, 0.0022)	−1.5455	0.1222
	$y=3$	0.0030	0.0038	0.8008	(−0.0044, 0.0103)	1.4429	0.1491
	$y=4$	0.0021	0.0024	0.9022	(−0.0024, 0.0069)	1.4382	0.1504

a: 'mean' and 'sd' denote the posterior mean and posterior standard deviation.

b: 'CD' denotes the convergence diagnostic statistic of MCMC proposed by Geweke (1992).