<table>
<thead>
<tr>
<th>Title</th>
<th>Bayesian Dynamic Panel-Ordered Probit Model and Its Application to Subjective Well-Being</th>
</tr>
</thead>
<tbody>
<tr>
<td>Author(s)</td>
<td>Hasegawa, Hikaru</td>
</tr>
<tr>
<td>Citation</td>
<td>Communications in Statistics : Simulation and Computation, 38(6): 1321-1347</td>
</tr>
<tr>
<td>Issue Date</td>
<td>2009-06</td>
</tr>
<tr>
<td>Doc URL</td>
<td><a href="http://hdl.handle.net/2115/43089">http://hdl.handle.net/2115/43089</a></td>
</tr>
<tr>
<td>Type</td>
<td>article (author version)</td>
</tr>
<tr>
<td>File Information</td>
<td>CSSC09.pdf</td>
</tr>
</tbody>
</table>

Hokkaido University Collection of Scholarly and Academic Papers : HUSCAP
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).
variables, \( u \) and identically normally distributed with zero mean and unit variance, that is, \( w \) at time \( t \). Further, let \( z_i \) denote the latent variable of individual \( i \) at time \( t \) such that

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

(1)

where \( \gamma_c \) is a cutoff point of ordinal response. We specify that

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

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 \( i = 1, \cdots, n, \)

\[
\begin{align*}
z_{it} &= \phi z_{i(t-1)} + x_{it}' \beta + w_i \delta + \alpha_i + u_{it}, \quad t = 1, \cdots, T \\
z_{i0} &= x_{i0}' \beta_0 + w_i \delta_0 + \alpha_i + u_{i0}, \quad t = 0
\end{align*}
\]

where \( x_{it} = (x_{it1}, \cdots, x_{itp})' \) is a vector of time-variant explanatory variables, \( w_i = (w_{i1}, \cdots, w_{ip})' \) is a vector of cross-sectional time-invariant explanatory variables, \( \alpha_i \) is a random effect, \( \beta = (\beta_1, \cdots, \beta_k)' \), \( \delta = (\delta_1, \cdots, \delta_p)' \), \( \beta_0 = (\beta_{01}, \cdots, \beta_{0k})' \), and \( \delta_0 = (\delta_{01}, \cdots, \delta_{0p})' \). \(^5\) We assume that \( u_{it} \) are independent and identically normally distributed with zero mean and unit variance, that is, \( u_{it} \sim N(0, 1) \). Then, we have

\[
\begin{align*}
\begin{cases}
\begin{align*}
z_{it} &\sim N(\phi z_{i(t-1)} + x_{it}' \beta + w_i \delta + \alpha_i, 1), \quad t = 1, \cdots, T \\
z_{i0} &\sim N(x_{i0}' \beta_0 + w_i \delta_0 + \alpha_i, 1), \quad t = 0
\end{align*}
\end{cases}
\end{align*}
\]

(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 \( (\phi, \beta, \delta, \beta_0, \delta_0, \alpha, \gamma) \), where \( \alpha = (\alpha_1, \cdots, \alpha_n)' \) and \( \gamma = (\gamma_2, \cdots, \gamma_{C-1})' \).

We assume that the prior distribution of the vector of random effects \( \alpha \) has a hierarchical structure. \(^6\) The prior distributions are specified as follows:

\[
p(\phi, \beta, \delta, \beta_0, \delta_0, \alpha, \mu, \tau, \gamma) = p(\phi) p(\beta) p(\delta) p(\beta_0) p(\delta_0) p(\alpha | \mu, \tau) p(\mu) p(\tau) p(\gamma).
\]

(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).
Further, from Bayes’ theorem, the joint posterior distribution can be written as follows:

\[
p(\phi, \beta, \delta, \beta_0, \delta_0, \alpha, \tau, \zeta, y, y_0) \propto p(\phi, \beta, \delta, \beta_0, \delta_0, \alpha, \tau, \zeta, z_0) \\
\times p(y, y_0|\phi, \beta, \delta, \beta_0, \delta_0, \alpha, \tau, \zeta, z_0) \\
= p(\phi, \beta, \delta, \beta_0, \delta_0, \alpha, \tau, \zeta) \\
\times \prod_{i=1}^{n} \left\{ p(z_i|\phi, \beta, \delta, \alpha_i, \alpha, \tau, \zeta, z_0) p(y_i|\phi, \beta, \delta, \alpha_i, \alpha, \tau, \zeta, z_i, z_0) \right\} \\
\times \left\{ p(z_0|\phi, \beta, \delta_0, \alpha_i, \alpha, \tau, \zeta, z_0) p(y_0|\phi, \beta, \delta_0, \alpha_i, \alpha, \tau, \zeta, z_0) \right\}, \tag{4}
\]

where

\[
y_i = (y_{i1}, y_{i2}, \ldots, y_{iT})', \quad z_i = (z_{i1}, z_{i2}, \ldots, z_{iT})', \quad i = 1, \ldots, n \\
y = (y_1, y_2, \ldots, y_n)', \quad z = (z_1, z_2, \ldots, z_n)' \\
y_0 = (y_{10}, y_{20}, \ldots, y_{n0})', \quad z_0 = (z_{10}, z_{20}, \ldots, z_{n0})'.
\]

### 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:

\[
\begin{align*}
\phi & \sim N(\phi, \tilde{\omega} I_{\phi \in (-1,1)}), \quad \beta \sim N(\beta, \tilde{B}), \quad \delta \sim N(\delta, \tilde{D}) \\
\beta_0 & \sim N(\beta_0, \tilde{B}_0), \quad \delta_0 \sim N(\delta_0, D_0) \\
\alpha_i & \sim N(\mu, \tau), \quad i = 1, \ldots, n, \quad \mu \sim N(\mu, \tilde{\kappa}), \quad \tau^{-1} \sim \text{Gam}(\tilde{a}, \tilde{b}) \\
p(\gamma) & \propto 1(\gamma \in \mathbb{C}), \quad \mathbb{C} = \{ \gamma : 0 < \gamma_2 < \cdots < \gamma_{C-1} < \infty \},
\end{align*}
\]

where \(1(\cdot)\) is an indicator function and \(\text{Gam}(a, b)\) 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:

**Algorithm 1:**

The details of this sampling algorithm are provided in Appendix A.
1. Sample \( \phi, \beta, \delta, \beta_0, \delta_0, \alpha, \mu, \tau \) from their full conditional distributions (FCDs).
2. Sample \( z, z_0 \) from their FCDs.
3. Sample \( \gamma \) 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:

\[
\begin{align*}
\beta &= 0, \quad \delta = 0, \quad \beta_0 = 0, \quad \delta_0 = 0, \quad \tilde{\mu} = 0.
\end{align*}
\]

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 \( \phi, \beta, \delta, \beta_0, \delta_0, \alpha, \mu, \tau \) from their FCDs.
2. Sample \( z, z_0 \) from their FCDs.
3. Sample \( \gamma \) from its FCD.

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:

\[
\zeta_c = \log(\gamma_c - \gamma_{c-1}), \quad c = 2, \cdots, C - 1,
\]

where \( \zeta(\gamma) = (\zeta_2, \cdots, \zeta_{C-1})' \) is unrestricted. Thus, we specify the prior distribution of \( \gamma, \ p(\gamma) = p(\zeta(\gamma)) \), as \( \zeta(\gamma) \sim N(\xi, \Sigma) \) 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 \( \phi, \beta, \delta, \beta_0, \delta_0, \alpha, \mu, \tau \) from their FCDs.
2. Sample \( z, z_0 \) from their FCDs.
3. Sample \( \zeta \) from the Metropolis-Hastings (M-H) algorithm.
4. Calculate \( \gamma_c = \sum_{j=1}^{c} \exp \zeta_j, \quad c = 2, \cdots, C - 1. \)

\(^8\)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 \( \beta = 0, \delta = 0, \beta_0 = 0, \delta_0 = 0 \) 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

\[
\begin{align*}
\{z_{it}|\cdots| & \sim N(\phi z_{i(t-1)} + x'_{it}\beta + w'\delta + \alpha_i, \sigma^2), \ t = 1, \cdots, T \\
\{z_{i0}|\cdots| & \sim N(x'_{i0}\beta_0 + w'\delta_0 + \alpha_i, \sigma^2), \ t = 0
\end{align*}
\]

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

\[
p(\phi, \beta, \delta, \beta_0, \delta_0, \alpha, \mu, \tau, \sigma^2, \gamma) = p(\phi)p(\beta)p(\delta)p(\beta_0)p(\delta_0)p(\alpha|\mu, \tau)p(\mu)p(\tau)p(\sigma^2)p(\gamma). \tag{9}
\]

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

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

where the prior distribution of \(\gamma, p(\gamma) = p(\zeta(\gamma))\), 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:

Algorithm 4:

1. Sample \(\phi, \beta, \delta, \beta_0, \delta_0, \alpha, \mu, \tau, \sigma^2\) from their FCDs.
2. Sample \(z, z_0\) from their FCDs.
3. Sample \(\zeta\) from the M-H algorithm.
4. Calculate \(\gamma_c = \frac{\gamma_{c-1} + \exp\zeta_c}{1 + \exp\zeta_c}, \ c = 2, \cdots, C - 2\).

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:

\[
\begin{align*}
\{z_{it}|\cdots| & \sim N(\phi z_{i(t-1)} + x_{it}\beta_1 + w_1\delta_1 + \alpha_i, 1), \ t = 1, \cdots, 9 \\
\{z_{i0}|\cdots| & \sim N(x_{i0}\beta_{01} + w_1\delta_{01} + \alpha_i, 1), \ t = 0
\end{align*}
\]

\[10^{\text{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:

$$\phi = 0.5, \quad \beta_1 = \beta_{01} = 2.0, \quad \delta_1 = \delta_{01} = 1.5, \quad \gamma_2 = 5.0, \quad \gamma_3 = 10.0$$

$$x_{i0} \sim N(1, 3), \quad x_{it} = 0.3x_{i(t-1)} + v_{it}, \quad v_{it} \sim N(0, 1)$$

$$w_i \sim N(2, 4), \quad \alpha_i \sim N(-3, 4).$$

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.\footnote{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.}

The posterior results obtained thereafter are generated using Ox version 5.00 (Doornik, 2007). Table 3 presents the posterior results for the estimation of $\phi$, $\beta_1$, $\delta_1$, $\beta_{01}$, $\delta_{01}$, and $\gamma$.\footnote{The results of parameters in Algorithm 4 except for $\phi$ are those of the transformed parameters divided by $\sigma$.} In Table 3, “95%CI” denotes the 95% credible interval. “CD” denotes the convergence diagnostic statistic proposed by Geweke (1992).\footnote{CD can be defined as follows: for the given sequence $\{g(j) \mid j = 1, 2, \cdots, n_s\}$, if the sequence is stationary,

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

where

$$\hat{\gamma}_A = \frac{1}{n_A} \sum_{j=1}^{n_A} g(j), \quad \hat{\gamma}_B = \frac{1}{n_B} \sum_{j=n_s}^{n_A} g(j) \quad (n_s = n_A - n_B + 1),$$

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.}

The posterior results obtained thereafter are generated using Ox version 5.00 (Doornik, 2007). Table 3 presents the posterior results for the estimation of $\phi$, $\beta_1$, $\delta_1$, $\beta_{01}$, $\delta_{01}$, and $\gamma$.\footnote{CD can be defined as follows: for the given sequence $\{g(j) \mid j = 1, 2, \cdots, n_s\}$, if the sequence is stationary,

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

where

$$\hat{\gamma}_A = \frac{1}{n_A} \sum_{j=1}^{n_A} g(j), \quad \hat{\gamma}_B = \frac{1}{n_B} \sum_{j=n_s}^{n_A} g(j) \quad (n_s = n_A - n_B + 1),$$

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.}

The sixth column of Table 3 shows the $p$-values for CD.

According to CD, Algorithm 1 does not exhibit good performances. $\phi$ 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:

$$w_i \sim N(2, 4), \quad \alpha_i \sim N(-3, 4).$$
• 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:

\[
\begin{align*}
  z_{it} &= \phi z_{i(t-1)} + \beta_1 x_{it1} + \beta_2 x_{it2} + \beta_3 x_{it3} + \beta_4 x_{it4} \\
  & \quad + \delta_1 w_{i1} + \delta_2 w_{i2} + \alpha_i + u_{it}, \quad t = 1, \ldots, 5 \\
  z_{i0} &= \beta_0 x_{i01} + \beta_2 x_{i02} + \beta_3 x_{i03} + \beta_4 x_{i04} + \delta w_{i1} + \delta_{i2} w_{i2} + \alpha_i + u_{i0},
\end{align*}
\]

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 \(\phi, \beta, \delta, \beta_0, \delta_0, \) and \(\gamma.\)\(^{17}\) According to the convergence diagnostic statistic (CD), parameters in Algorithm 4 except for \(\phi, \beta_4, \beta_{01}, \beta_{04} \) and \(\gamma_3\) 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 \(\phi\) are those of the transformed parameters divided by \(\sigma.\)

\(^{18}\)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 (\( \gamma_2, \gamma_3, \) and \( \gamma_4 \)) 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. Furthermore, 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 $\phi$

The full conditional distribution (FCD) of $\phi$ is

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

where

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

$$
\hat{\phi} = \hat{\omega} \left[ \frac{\phi}{\omega} + \sum_{i=1}^{n} \sum_{t=1}^{T} z_{it} (z_{it} \beta - w_i \delta - \alpha_i) \right].
$$

We can utilize Damien and Walker’s (2001) algorithm for sampling from a truncated normal distribution.\textsuperscript{21}

\textsuperscript{21}Using Damien and Walker’s algorithm, we can sample $\phi$ from the following steps.

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

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

3. We have the following FCDs:

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

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

The FCD of $\beta$ is

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

where

$$\hat{B} = \left( \tilde{B}^{-1} + \frac{n}{T} \sum_{i=1}^{n} x_{it} x_{it}' \right)^{-1}$$

$$\hat{\beta} = \hat{B} \left[ \tilde{B}^{-1} \tilde{\beta} + \frac{n}{T} \sum_{i=1}^{n} x_{it} (z_{it} - \phi z_{i(t-1)} - w_i \delta - \alpha_i) \right].$$

A.3 FCD of $\delta$

The FCD of $\delta$ is

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

where

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

$$\hat{\delta} = \hat{D} \left[ \tilde{D}^{-1} \tilde{\delta} + \frac{T}{n} \sum_{i=1}^{n} w_i (z_{it} - \phi z_{i(t-1)} - x_{it}' \beta - \alpha_i) \right].$$

A.4 FCD of $\beta_0$

The FCD of $\beta_0$ is

$$\beta_0 | \cdots \sim N(\hat{\beta}_0, \hat{B}_0),$$

where

$$\hat{B}_0 = \left( \tilde{B}_0^{-1} + \frac{n}{T} \sum_{i=1}^{n} x_{i0} x_{i0}' \right)^{-1}$$

$$\hat{\beta}_0 = \hat{B}_0 \left[ \tilde{B}_0^{-1} \tilde{\beta}_0 + \frac{n}{T} \sum_{i=1}^{n} x_{i0} (z_{i0} - w_i \delta_0 - \alpha_i) \right].$$

A.5 FCD of $\delta_0$

The FCD of $\delta_0$ is

$$\delta_0 | \cdots \sim N(\hat{\delta}_0, \hat{D}_0),$$

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

11
where
\[
\begin{align*}
\hat{D}_0 &= \left( \hat{D}_0^{-1} + \sum_{i=1}^{n} w_i w'_i \right)^{-1} \\
\delta_0 &= \hat{D}_0 \left[ \hat{D}_0^{-1} \delta_0 + \sum_{i=1}^{n} w_i (z_{i0} - x'_{i0} \beta_0 - \alpha_i) \right].
\end{align*}
\]

A.6 FCD of \( \alpha \)
The FCD of \( \alpha_i \) is
\[
\alpha_i | \cdots \sim N(\hat{\alpha}_i, \hat{\tau}), \; i = 1, \cdots, n,
\]
where
\[
\begin{align*}
\hat{\tau} &= \left( \frac{1}{T} + T + 1 \right)^{-1} \\
\hat{\alpha}_i &= \hat{\tau} \left[ \frac{\mu}{\tau} + \sum_{t=1}^{T} (z_{it} - \phi z_{i(t-1)} - x'_{i(t-1)} \beta - w'_i \delta) + (z_{i0} - x'_{i0} \beta_0 - w'_i \delta_0) \right].
\end{align*}
\]

A.7 FCD of \( \mu \)
The FCD of \( \mu \) is
\[
\mu | \cdots \sim N(\hat{\mu}, \hat{\kappa}),
\]
where
\[
\begin{align*}
\hat{\kappa} &= \left( \frac{1}{\kappa} + \frac{n}{\tau} \right)^{-1}, \; \hat{\mu} = \kappa \left( \frac{\hat{\mu}}{\kappa} + \frac{1}{\tau} \sum_{i=1}^{n} \alpha_i \right).
\end{align*}
\]

A.8 FCD of \( \tau \)
The FCD of \( \tau^{-1} \) is
\[
\tau^{-1} | \cdots \sim \text{Gam}(\hat{\alpha}, \hat{b}),
\]
where
\[
\begin{align*}
\hat{\alpha} &= \alpha + \frac{n}{2}, \; \hat{b} = \beta + \frac{n}{2} \sum_{i=1}^{n} (\alpha_i - \mu)^2.
\end{align*}
\]

A.9 FCD of \( z_{i0} \)
The FCD of \( z_{i0} \) is
\[
\begin{align*}
z_{i0} | \cdots \sim N \left( \hat{z}_{i0}, \frac{1}{1 + \phi^2} \right) 1_{(z_{i0} \in (\gamma_1, \gamma_2))}, \; i = 1, \cdots, n,
\end{align*}
\]
where
\[
\hat{z}_{i0} = \frac{\phi (z_{i1} - x'_{i1} \beta - w'_i \delta - \alpha_i) + (x'_{i0} \beta_0 + w'_i \delta_0 + \alpha_i)}{1 + \phi^2}.
\]

12
A.10 FCD of \( z \)

The FCD of \( z \) is

\[
\tilde{z}_{it} \sim N \left( \frac{1}{1 + \phi^2} \right) 1_{\{z_{it} \in (\gamma_{c-1}, \gamma_c)\}}, \quad t = 1, \ldots, T - 1, \quad i = 1, \ldots, n,
\]

where

\[
\tilde{z}_{it} = \frac{(\phi z_{i(t-1)} + x'_{it} + \alpha_i) + \phi(x'_{i(t+1)} - x'_{i(t+1)} - w'd - \alpha_i)}{1 + \phi^2}.
\]

Further, the FCD of \( z_{iT} \) is

\[
z_{iT} \sim N(\phi z_{iT-1} + x'_{iT} + \alpha_i, 1) 1_{\{z_{iT} \in (\gamma_{c-1}, \gamma_c)\}}, \quad i = 1, \ldots, n,
\]

A.11 FCD of \( \gamma \)

The FCD of \( \gamma_c \) is

\[
\gamma_c \sim U \left( \max_{i,t} \{\max_{i} \{z_{it} | y_{it} = c\}, \gamma_{c-1}\}, \min_{i,t} \{\min_{i} \{z_{it} | y_{it} = c + 1\}, \gamma_{c+1}\} \right),
\]

\( c = 2, \ldots, C - 1 \).

B Derivation of Algorithm 2

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

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

\[
\Gamma = \{g > 0 : g(\phi, \beta, \delta, \beta_0, \delta_0, \alpha, \mu, \tau, \gamma, z, z_0) \}
\]

\[
= \{\phi, g\beta, g\delta, g\beta_0, g\delta_0, g\alpha, g\mu, \tau, g\gamma, gz, gz_0)\}.
\]

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

\[
p(g(\phi, \beta, \delta, \beta_0, \delta_0, \alpha, \mu, \tau, \gamma, z, z_0) | y, y_0)
\]

\[
\propto \exp \left[ -\frac{g^2}{2} \left( \beta' \bar{\beta}^{-1} \beta + \delta' \bar{\delta}^{-1} \delta + \beta_0' \bar{\beta}_0^{-1} \beta_0 + \delta_0' \bar{\delta}_0^{-1} \delta_0 
\right.
\]

\[
+ \frac{1}{\tau} \sum_{i=1}^{n} (\alpha_i - \mu)^2 + \frac{\mu^2}{K} + \sum_{i=1}^{n} \sum_{t=1}^{T} (z_{it} - \phi z_{i(t-1)} - x'_{it} - w'_{i} - \alpha_i)^2
\]

\[
+ \frac{1}{\tau} \sum_{i=1}^{n} (z_{i0} - x'_{i0} - w'_{i0} - \alpha_i)^2 \right] \exp \left[ -\frac{g^2}{2} (\phi - \bar{\phi})^2 \right] \left( \phi \in (0, 1) \right)
\]

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

\[
\times \left[ \prod_{i=1}^{n} \sum_{c=1}^{C} 1_{\{y_{i0} = c\}} 1_{\{z_{i0} \in (\gamma_{c-1}, \gamma_c)\}} \right].
\]
Further, we have the following Jacobian:\(^{22}\)

\[
|J_g(\phi, \beta, \delta, \beta_0, \delta_0, \alpha, \mu, \tau, \gamma, z, z_0)| = g^{(T+2)n + 2k + 2p + C - 1}.
\]

Therefore the FCD of \(g\) is

\[
p(g|\cdots) = p(g(\phi|\beta, \delta, \beta_0, \delta_0, \alpha, \mu, \tau, \gamma, z, z_0)|y, y_0) \times |J_g(\phi, \beta, \delta, \beta_0, \delta_0, \alpha, \mu, \tau, \gamma, z, z_0)|g^{-1}
\]

\[
\propto \exp\left\{-\frac{g^2}{2}\left[\beta' \tilde{B}^{-1}\beta + \delta' \tilde{D}^{-1}\delta + \beta_0' \tilde{B}_0^{-1}\beta_0 + \delta_0' \tilde{D}_0^{-1}\delta_0
\right.
\right.
\]

\[
\left.+ \frac{1}{T} \sum_{i=1}^{n} (\alpha_i - \mu)^2 + \frac{\mu^2}{k} + \sum_{i=1}^{n} \sum_{t=1}^{T} (z_{it} - \phi z_{i(t-1)} - x_i' \beta - x_i' \delta - \alpha_i)^2
\right.
\]

\[
\left.+ \sum_{i=1}^{n} (z_{i0} - x_i' \beta_0 - x_i' \delta_0 - \alpha_i)^2\right]\}
\]

\[
g^{(T+2)n + 2k + 2p + C - 2}.
\]

Putting \(g = \sqrt{\lambda}\), we have

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

where

\[
\frac{\lambda}{\lambda} = \frac{1}{2}\left[\beta' \tilde{B}^{-1}\beta + \delta' \tilde{D}^{-1}\delta + \beta_0' \tilde{B}_0^{-1}\beta_0 + \delta_0' \tilde{D}_0^{-1}\delta_0
\right.
\]

\[
\left.+ \frac{1}{T} \sum_{i=1}^{n} (\alpha_i - \mu)^2 + \frac{\mu^2}{k} + \sum_{i=1}^{n} \sum_{t=1}^{T} (z_{it} - \phi z_{i(t-1)} - x_i' \beta - x_i' \delta - \alpha_i)^2
\right.
\]

\[
\left.+ \sum_{i=1}^{n} (z_{i0} - x_i' \beta_0 - x_i' \delta_0 - \alpha_i)^2\right].
\]

Liu and Sabatti’s algorithm

- Using usual Gibbs sampler, generate \(\phi, \beta, \delta, \beta_0, \delta_0, \alpha, \mu, \tau, \gamma, z, z_0\).
- Generate \(h\) from \(p(h|\cdots)\) and put \(g = \sqrt{\lambda}\). Using \(g\), implement the following update: \(\beta \leftarrow g\beta, \delta \leftarrow g\delta, \beta_0 \leftarrow g\beta_0, \delta_0 \leftarrow g\delta_0, \alpha \leftarrow g\alpha, \mu \leftarrow g\mu, \gamma \leftarrow g\gamma, z \leftarrow g\gamma, z_0 \leftarrow g\gamma_0.\)

\(^{22}\)Since \(\beta\) has \(k\), \(\delta\) has \(p\), \(\beta_0\) has \(k\), \(\delta_0\) has \(p\), \(\alpha\) has \(n\), \(\gamma\) has \(C - 2\), \(z\) has \(nT\), \(z_0\) has \(n\), \(\mu\) 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 $\gamma$. The FCD of $\gamma$ is

$$p(\gamma|\cdots) \propto p(\zeta(\gamma)) \prod_{i,t,y_{it}=2} \Phi(\gamma_2 - \bar{z}_{it}) - \Phi(-\bar{z}_{it}) \times \prod_{i,t,y_{it}=3} \Phi(\gamma_3 - \bar{z}_{it}) - \Phi(\gamma_2 - \bar{z}_{it}) \times \cdots \times \prod_{i,t,y_{it}=C-1} \Phi(\gamma_{C-1} - \bar{z}_{it}) - \Phi(\gamma_{C-2} - \bar{z}_{it}) \times \prod_{i,t,y_{it}=C} [1 - \Phi(\gamma_{C-1} - \bar{z}_{it})],$$

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

$$\bar{z}_{it} = \begin{cases} \phi z_{it(t-1)} + x'_{it}\beta + w'_{i}\delta + \alpha_i & t = 1, \cdots, T \\ x'_{t0}\beta_0 + w'_{i0}\delta_0 + \alpha_i & t = 0. \end{cases}$$

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

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

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

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

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

1. Let $\zeta^{(l)}$ denote the value of $\zeta$ at the $l$th iteration.
2. At the $(l+1)$th iteration, sample $\zeta^p$ from $Mt(\zeta|\hat{\zeta}, \hat{\Sigma}_\zeta, \nu)$.
3. The transition probability from $\zeta^{(l)}$ to $\zeta^p$ is

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

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

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

We can obtain $\gamma$ from $\zeta$ 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 $\phi$

The FCD of $\phi$ is

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

where

$$
\hat{\omega} = \left( \frac{1}{\hat{\omega}} + \frac{1}{\sigma^2} \sum_{i=1}^{n} \sum_{t=1}^{T} z_{i(t-1)}^2 \right)^{-1},
$$
$$
\hat{\phi} = \hat{\omega} \left[ \frac{\phi}{\hat{\omega}} + \frac{1}{\sigma^2} \sum_{i=1}^{n} \sum_{t=1}^{T} z_{i(t-1)}(z_{it} - x_{it}'\hat{\beta} - w_{it}'\hat{\delta} - \alpha_i) \right].
$$

D.2 FCD of $\beta$

The FCD of $\beta$ is

$$
\beta | \ldots \sim N(\hat{\beta}, \hat{\beta}),
$$

where

$$
\hat{\beta} = B^{-1} \hat{\beta} + \frac{1}{\sigma^2} \sum_{i=1}^{n} \sum_{t=1}^{T} x_{it}(z_{it} - x_{it}'\hat{\beta} - w_{it}'\hat{\delta} - \alpha_i).
$$

D.3 FCD of $\delta$

The FCD of $\delta$ is

$$
\delta | \ldots \sim N(\hat{\delta}, \hat{D}),
$$

where

$$
\hat{D} = \left( \hat{D}^{-1} + \frac{T}{\sigma^2} \sum_{i=1}^{n} w_{i}w_{i}' \right)^{-1},
$$
$$
\hat{\delta} = \hat{D} \left[ \hat{D}^{-1} \hat{\delta} + \frac{1}{\sigma^2} \sum_{i=1}^{n} \sum_{t=1}^{T} w_{i}(z_{it} - x_{it}'\hat{\beta} - w_{it}'\hat{\delta} - \alpha_i) \right].
$$

D.4 FCD of $\beta_0$

The FCD of $\beta_0$ is

$$
\beta_0 | \ldots \sim N(\hat{\beta}_0, \hat{B}_0),
$$
where

\[
\hat{B}_0 = \left( \tilde{B}_0^{-1} + \frac{1}{\sigma^2} \sum_{i=1}^{n} x_{i0} x_{i0}' \right)^{-1}
\]

\[
\hat{\beta}_0 = \hat{B}_0 \left[ \tilde{B}_0^{-1} \beta_0 + \frac{1}{\sigma^2} \sum_{i=1}^{n} x_{i0} (z_{i0} - w_i' \delta_0 - \alpha_i) \right].
\]

### D.5 FCD of $\delta_0$

The FCD of $\delta_0$ is

\[
\delta_0 | \cdots \sim N(\hat{\delta}_0, \check{D}_0),
\]

where

\[
\check{D}_0 = \left( D_0^{-1} + \frac{1}{\sigma^2} \sum_{i=1}^{n} w_i w_i' \right)^{-1}
\]

\[
\check{\delta}_0 = D_0 \left[ \tilde{D}_0^{-1} \tilde{\delta}_0 + \frac{1}{\sigma^2} \sum_{i=1}^{n} w_i (z_{i0} - x_{i0}' \beta_0 - \alpha_i) \right].
\]

### D.6 FCD of $\alpha$

The FCD of $\alpha_i$ is

\[
\alpha_i | \cdots \sim N(\hat{\alpha}_i, \check{\tau}), \ i = 1, \cdots, n,
\]

where

\[
\check{\tau} = \left( \frac{1}{\tau} + \frac{T + 1}{\sigma^2} \right)^{-1}
\]

\[
\check{\alpha}_i = \check{\tau} \left\{ \frac{\mu}{\tau} + \frac{1}{\sigma^2} \left[ \sum_{t=1}^{T} (z_{it} - \phi z_{i(t-1)} - x_{i0}' \beta - w_i' \delta) \right. \right.
\]

\[
+ (z_{i0} - x_{i0}' \beta_0 - w_i' \delta_0) \left. \left. \right]\right\}.
\]

### D.7 FCD of $\mu$

The FCD of $\mu$ is

\[
\mu | \cdots \sim N(\hat{\mu}, \check{\kappa}),
\]

where

\[
\check{\kappa} = \left( \frac{1}{\kappa} + \frac{n}{\tau} \right)^{-1}, \ \check{\mu} = \check{\kappa} \left( \frac{\mu}{\kappa} + \frac{1}{\tau} \sum_{i=1}^{n} \alpha_i \right).
\]
D.8 FCD of $\tau$

The FCD of $\tau^{-1}$ is

$$\tau^{-1} \cdot \cdot \cdot \sim \text{Gam}(\hat{a}, \hat{b}),$$

where

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

D.9 FCD of $\sigma^2$

The FCD of $\sigma^{-2}$ is

$$\sigma^{-2} \cdot \cdot \cdot \sim \text{Gam}(\tilde{c}, \tilde{d}),$$

where

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

$$\tilde{d} = \hat{d} + \frac{1}{2} \left[ \sum_{i=1}^{n} \sum_{t=1}^{T} (z_{it} - \phi z_{i(t-1)} - x_{it}'\beta - w_i'\delta - \alpha_i)^2 
+ \sum_{i=1}^{n} (z_{i0} - x_{i0}'\beta_0 - w_i'\delta_0 - \alpha_i)^2 \right].$$

D.10 FCD of $z_0$

The FCD of $z_{i0}$ is

$$z_{i0} \cdot \cdot \cdot \sim N \left( \bar{z}_{i0}, \frac{\sigma^2}{1 + \phi^2} \right) 1_{(z_{i0} \in (\gamma_{c-1}, \gamma_c))}, \quad i = 1, \ldots, n,$$

where

$$\bar{z}_{i0} = \frac{\phi(z_i - x_{i1}'\beta - w_i'\delta - \alpha_i) + (x_{i0}'\beta_0 + w_i'\delta_0 + \alpha_i)}{1 + \phi^2}.$$

D.11 FCD of $z$

The FCD of $z_{it}$ is

$$z_{it} \cdot \cdot \cdot \sim N \left( \bar{z}_{it}, \frac{\sigma^2}{1 + \phi^2} \right) 1_{(z_{it} \in (\gamma_{c-1}, \gamma_c))}, \quad t = 1, \ldots, T - 1, \quad i = 1, \ldots, n,$$

where

$$\bar{z}_{it} = \frac{(\phi z_{i(t-1)} + x_{it}'\beta + w_i'\delta + \alpha_i) + \phi(z_{i(t+1)} - x_{i(t+1)}'\beta - w_i'\delta - \alpha_i)}{1 + \phi^2}.$$

Further, the FCD of $z_{iT}$ is

$$z_{iT} \cdot \cdot \cdot \sim N \left( \phi z_{i(T-1)} + x_{iT}'\beta + w_i'\delta + \alpha_i, \sigma^2 \right) 1_{(z_{iT} \in (\gamma_{c-1}, \gamma_c))}, \quad i = 1, \ldots, n.$$
D.12 Sampling Algorithm of $\gamma$

The FCD of $\gamma$ is

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

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

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

where

$$\bar{z}_{it} = \left\{ \begin{array}{ll}
\phi z_{it}(t-1) + x_{it}^t \beta + w_{it}^t \delta + \alpha_i & t = 1, \cdots, T \\
x_{it}^t \beta_0 + w_{it}^t \delta_0 + \alpha_i & t = 0.
\end{array} \right.$$ 

Since the Jacobian of the transformation of $\gamma$ to $\zeta$ is

$$\prod_{c=2}^{C-2} \left( 1 - \gamma_{c-1} \right) \exp \zeta_c \left( 1 + \exp \zeta_c \right)^2$$

we have

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

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

$$\hat{\zeta} = \left\{ - \frac{\partial^2 \log p(\zeta; \cdots)}{\partial \zeta \partial \zeta'} \right\}_{\zeta = \hat{\zeta}}^{-1}$$

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

1. Let $\zeta^{(l)}$ denote the value of $\zeta$ at the $l$th iteration.
2. At the $(l+1)$th iteration, sample $\zeta^p$ from $Mt(\zeta; \hat{\zeta}, \nu)$.
3. The transition probability from $\zeta^{(l)}$ to $\zeta^p$ is

$$\alpha = \min \left\{ \frac{p(\zeta^p; \cdots) Mt(\zeta^p; \hat{\zeta}, \nu)}{p(\zeta^{(l)}; \cdots) Mt(\zeta^{(l)}; \hat{\zeta}, \nu)}, 1 \right\}.$$ 

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

$$\zeta^{(l+1)} = \left\{ \begin{array}{ll}
\zeta^p & \text{if } u < \alpha \\
\zeta^{(l)} & \text{otherwise}.
\end{array} \right.$$ 

We can obtain $\gamma$ from $\zeta$ using the equation

$$\gamma_c = \frac{\gamma_{c-1} + \exp \zeta_c}{1 + \exp \zeta_c}, \quad c = 2, \cdots, C - 2.$$
The Partial Effects of the Explanatory Variables on the Response Probabilities

E.1 Algorithms 1–3

From (1) and (2), we have

\[
\Pr(y_{it} = c) = \Pr(\gamma_{c-1} < z_{it} \leq \gamma_c) = \begin{cases} 
F(-E(z_{it})) & c = 1 \\
F(\gamma_c - E(z_{it})) - F(\gamma_{c-1} - E(z_{it})) & 2 \leq c \leq C - 1 \\
1 - F(\gamma_{c-1} - E(z_{it})) & c = C 
\end{cases}
\tag{11}
\]

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 \(t = 1, \cdots, T, i = 1, \cdots, n,\)

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

and for \(t = 0, i = 1, \cdots, n,\)

\[
\frac{\partial \Pr(y_{i0} = c)}{\partial x} = \begin{cases} 
-\beta_x f(- (x'_{i0} \beta_0 + w'_{i0} \delta_0 + \alpha_i)) & c = 1 \\
\beta_x f(\gamma_{c-1} - (x'_{i0} \beta_0 + w'_{i0} \delta_0 + \alpha_i)) \\
- f(\gamma_c - (x'_{i0} \beta_0 + w'_{i0} \delta_0 + \alpha_i)) & 2 \leq c \leq C - 1 \\
\beta_x f(\gamma_{c-1} - (x'_{i0} \beta_0 + w'_{i0} \delta_0 + \alpha_i)) & c = C, 
\end{cases}
\tag{13}
\]

where \(\beta_x\) 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 \(z_{i(-1)}, x_j,\) and \(w_j\) on the response probabilities \(\Pr(y = c)\) of \(2 \leq c \leq C - 1\) are as
follows:
\[
\frac{\partial \Pr(y = c)}{\partial z(-1)} = \frac{\phi}{nT} \sum_{i=1}^{n} \sum_{t=1}^{T} \left[ f\left(\gamma_{c-1} - (\phi z_{i(t-1)} + x'_{it} \theta + w'_i \delta + \alpha_i)\right) - f\left(\gamma_{c} - (\phi z_{i(t-1)} + x'_{it} \theta + w'_i \delta + \alpha_i)\right) \right] \\
\frac{\partial \Pr(y = c)}{\partial x_j} = \frac{\beta_j}{nT} \sum_{i=1}^{n} \sum_{t=1}^{T} \left[ f\left(\gamma_{c-1} - (\phi z_{i(t-1)} + x'_{it} \theta + w'_i \delta + \alpha_i)\right) - f\left(\gamma_{c} - (x'_{it} \theta_0 + w'_i \delta_0 + \alpha_i)\right) \right] \\
+ \frac{\beta_{0j}}{n} \sum_{i=1}^{n} \left[ f\left(\gamma_{c-1} - (x'_{it} \theta_0 + w'_i \delta_0 + \alpha_i)\right) - f\left(\gamma_{c} - (x'_{it} \theta_0 + w'_i \delta_0 + \alpha_i)\right) \right] \\
\frac{\partial \Pr(y = c)}{\partial w_j} = \frac{\delta_j}{nT} \sum_{i=1}^{n} \sum_{t=1}^{T} \left[ f\left(\gamma_{c-1} - (\phi z_{i(t-1)} + x'_{it} \theta + w'_i \delta + \alpha_i)\right) - f\left(\gamma_{c} - (x'_{it} \theta_0 + w'_i \delta_0 + \alpha_i)\right) \right] \\
+ \frac{\delta_{0j}}{n} \sum_{i=1}^{n} \left[ f\left(\gamma_{c-1} - (x'_{it} \theta_0 + w'_i \delta_0 + \alpha_i)\right) - f\left(\gamma_{c} - (x'_{it} \theta_0 + w'_i \delta_0 + \alpha_i)\right) \right].
\]

E.2 Algorithm 4

From (1) and (7), we have
\[
\Pr(y_{it} = c) = \Pr(\gamma_{c-1} < z_{it} \leq \gamma_{c}) = \Pr\left(\gamma_{c-1} - \frac{E(z_{it})}{\sigma} \leq \frac{\gamma_{c} - E(z_{it})}{\sigma}\right)
= \begin{cases} 
F\left(\frac{-E(z_{it})}{\sigma}\right) & c = 1 \\
F\left(\frac{\gamma_{c} - E(z_{it})}{\sigma}\right) - F\left(\frac{\gamma_{c-1} - E(z_{it})}{\sigma}\right) & 2 \leq c \leq C - 1 \\
1 - F\left(\frac{1 - E(z_{it})}{\sigma}\right) & c = C
\end{cases}
\]
\[t = 0, 1, \ldots, T, \quad i = 1, \ldots, n.\]

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 \( t = 1, \cdots, T, \ i = 1, \cdots, n, \)
\[
\frac{\partial \Pr(y_{it} = c)}{\partial x} = \begin{cases} 
\frac{\beta_x}{\sigma} f \left( \frac{1 - (\phi z_{i(t-1)} + x'_{it} \beta + w'_{i} \delta + \alpha_i)}{\sigma} \right) & c = 1 \\
\frac{\beta_x}{\sigma} \left[ f \left( \frac{1}{\sigma} \{ \gamma_{c-1} - (\phi z_{i(t-1)} + x'_{it} \beta + w'_{i} \delta + \alpha_i) \} \right) \\
- f \left( \frac{1}{\sigma} \{ \gamma_{c} - (\phi z_{i(t-1)} + x'_{it} \beta + w'_{i} \delta + \alpha_i) \} \right) \right] & 2 \leq c \leq C - 1 \\
\frac{\beta_x}{\sigma} f \left( \frac{1}{\sigma} \{ 1 - (\phi z_{i(t-1)} + x'_{it} \beta + w'_{i} \delta + \alpha_i) \} \right) & c = C
\end{cases}
\] (15)

and for \( t = 0, \ i = 1, \cdots, n, \)
\[
\frac{\partial \Pr(y_{i0} = c)}{\partial x} = \begin{cases} 
\frac{\beta_x}{\sigma} f \left( \frac{1 - (x'_{i0} \beta_0 + w'_{i0} \delta_0 + \alpha_i)}{\sigma} \right) & c = 1 \\
\frac{\beta_x}{\sigma} \left[ f \left( \frac{1}{\sigma} \{ \gamma_{c-1} - (x'_{i0} \beta_0 + w'_{i0} \delta_0 + \alpha_i) \} \right) \\
- f \left( \frac{1}{\sigma} \{ \gamma_{c} - (x'_{i0} \beta_0 + w'_{i0} \delta_0 + \alpha_i) \} \right) \right] & 2 \leq c \leq C - 1 \\
\frac{\beta_x}{\sigma} f \left( \frac{1}{\sigma} \{ 1 - (x'_{i0} \beta_0 + w'_{i0} \delta_0 + \alpha_i) \} \right) & c = C,
\end{cases}
\] (16)

where \( \beta_x \) 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 \( z_{i(-1)}, x_j, \) and \( w_j \) on the response probabilities \( \Pr(y = c) \) of \( 2 \leq c \leq C - 1 \) are as
follows:

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

\[
\frac{\partial \Pr(y = c)}{\partial x_j} = \frac{\beta_j}{nT\sigma} \sum_{i=1}^{n} \sum_{t=1}^{T} \left[ f \left( \frac{1}{\sigma} \left\{ \gamma_{c-1} - (\phi z_{i(t-1)} + x'_{it}\beta + w'_i\delta + \alpha_i) \right\} \right) 
- f \left( \frac{1}{\sigma} \left\{ \gamma_{c} - (x'_{i0}\beta_j + w'_i\delta_0 + \alpha_i) \right\} \right) \right]
+ \frac{\delta_{ij}}{nT\sigma} \sum_{i=1}^{n} \left[ f \left( \frac{1}{\sigma} \left\{ \gamma_{c-1} - (x'_{i0}\beta_j + w'_i\delta_0 + \alpha_i) \right\} \right) 
- f \left( \frac{1}{\sigma} \left\{ \gamma_{c} - (x'_{i0}\beta_j + w'_i\delta_0 + \alpha_i) \right\} \right) \right]
\]

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

The author appreciates the comments of two anonymous referees, which improve the article greatly. He is also grateful to the Institute for Research on Household Economics for providing the micro data of the Japanese Panel Survey of Consumers. Further, this work was supported in part by a Grant-in-Aid for Scientific Research (No.19530177) from the JSPS.

References


Table 1: Numbers in categories of $y$

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

Table 2: Summary statistics of explanatory variables

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

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

<table>
<thead>
<tr>
<th></th>
<th>mean</th>
<th>sd</th>
<th>min</th>
<th>50%</th>
<th>max</th>
</tr>
</thead>
<tbody>
<tr>
<td>period 0</td>
<td>1.9586</td>
<td>1.9500</td>
<td>-3.4154</td>
<td>1.8541</td>
<td>7.5332</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>mean</th>
<th>sd</th>
<th>min</th>
<th>50%</th>
<th>max</th>
</tr>
</thead>
<tbody>
<tr>
<td>period 0</td>
<td>1.9586</td>
<td>1.9500</td>
<td>-3.4154</td>
<td>1.8541</td>
<td>7.5332</td>
</tr>
</tbody>
</table>
Table 3: Posterior results for parameters (simulated data)

<table>
<thead>
<tr>
<th></th>
<th>Algorithm 1</th>
<th></th>
<th>Algorithm 2</th>
<th></th>
<th>Algorithm 3</th>
<th></th>
<th>Algorithm 4</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>mean(^a)</td>
<td>sd(^a)</td>
<td>mean/sd</td>
<td>95%CI(^a)</td>
<td>CD(^a)</td>
<td>p-val(^a)</td>
<td>mean(^a)</td>
</tr>
<tr>
<td>(\phi)</td>
<td>0.515</td>
<td>0.012</td>
<td>43.146</td>
<td>(0.491, 0.538)</td>
<td>–0.052</td>
<td>0.958</td>
<td>0.514</td>
</tr>
<tr>
<td>(\beta_{1})</td>
<td>2.023</td>
<td>0.070</td>
<td>31.546</td>
<td>(2.066, 2.338)</td>
<td>–4.559</td>
<td>0.000</td>
<td>2.079</td>
</tr>
<tr>
<td>(\delta_{1})</td>
<td>1.643</td>
<td>0.054</td>
<td>30.384</td>
<td>(1.540, 1.747)</td>
<td>–4.241</td>
<td>0.000</td>
<td>1.651</td>
</tr>
<tr>
<td>(\beta_{01})</td>
<td>2.143</td>
<td>0.083</td>
<td>25.731</td>
<td>(1.981, 2.308)</td>
<td>–4.138</td>
<td>0.000</td>
<td>2.079</td>
</tr>
<tr>
<td>(\delta_{01})</td>
<td>1.704</td>
<td>0.063</td>
<td>26.956</td>
<td>(1.585, 1.829)</td>
<td>–4.223</td>
<td>0.000</td>
<td>1.704</td>
</tr>
<tr>
<td>(\gamma_{2})</td>
<td>5.558</td>
<td>0.169</td>
<td>32.902</td>
<td>(5.275, 5.915)</td>
<td>–4.430</td>
<td>0.000</td>
<td>5.558</td>
</tr>
<tr>
<td>(\gamma_{3})</td>
<td>11.081</td>
<td>0.282</td>
<td>39.308</td>
<td>(10.568, 11.609)</td>
<td>–4.226</td>
<td>0.000</td>
<td>11.081</td>
</tr>
</tbody>
</table>

\(^a\): ‘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 \(\phi\) are transformed ones divided by \(\sigma\).
Table 4 (a): Life satisfaction (frequency)

<table>
<thead>
<tr>
<th>wave</th>
<th>strongly dissatisfied and dissatisfied</th>
<th>moderately satisfied</th>
<th>satisfied</th>
<th>strongly satisfied</th>
</tr>
</thead>
<tbody>
<tr>
<td>6</td>
<td>60</td>
<td>86</td>
<td>160</td>
<td>32</td>
</tr>
<tr>
<td>7</td>
<td>48</td>
<td>87</td>
<td>173</td>
<td>30</td>
</tr>
<tr>
<td>8</td>
<td>66</td>
<td>89</td>
<td>156</td>
<td>27</td>
</tr>
<tr>
<td>9</td>
<td>49</td>
<td>96</td>
<td>165</td>
<td>28</td>
</tr>
<tr>
<td>10</td>
<td>62</td>
<td>86</td>
<td>169</td>
<td>21</td>
</tr>
<tr>
<td>11</td>
<td>71</td>
<td>91</td>
<td>153</td>
<td>23</td>
</tr>
</tbody>
</table>

Table 4 (b): Life satisfaction (%)

<table>
<thead>
<tr>
<th>wave</th>
<th>strongly dissatisfied and dissatisfied</th>
<th>moderately satisfied</th>
<th>satisfied</th>
<th>strongly satisfied</th>
</tr>
</thead>
<tbody>
<tr>
<td>6</td>
<td>17.8</td>
<td>25.4</td>
<td>47.3</td>
<td>9.5</td>
</tr>
<tr>
<td>7</td>
<td>14.2</td>
<td>25.7</td>
<td>51.2</td>
<td>8.9</td>
</tr>
<tr>
<td>8</td>
<td>19.5</td>
<td>26.3</td>
<td>46.2</td>
<td>8.0</td>
</tr>
<tr>
<td>9</td>
<td>14.5</td>
<td>28.4</td>
<td>48.8</td>
<td>8.3</td>
</tr>
<tr>
<td>10</td>
<td>18.3</td>
<td>25.4</td>
<td>50.0</td>
<td>6.2</td>
</tr>
<tr>
<td>11</td>
<td>21.0</td>
<td>26.9</td>
<td>45.3</td>
<td>6.8</td>
</tr>
</tbody>
</table>

Table 5 (a): Summary statistics 1

<table>
<thead>
<tr>
<th>wave</th>
<th>married</th>
<th>inwork</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>no(0)</td>
<td>yes(1)</td>
</tr>
<tr>
<td></td>
<td>no(0)</td>
<td>yes(1)</td>
</tr>
<tr>
<td>6</td>
<td>77</td>
<td>261</td>
</tr>
<tr>
<td>7</td>
<td>72</td>
<td>266</td>
</tr>
<tr>
<td>8</td>
<td>62</td>
<td>276</td>
</tr>
<tr>
<td>9</td>
<td>85</td>
<td>253</td>
</tr>
<tr>
<td>10</td>
<td>76</td>
<td>262</td>
</tr>
<tr>
<td>11</td>
<td>84</td>
<td>254</td>
</tr>
</tbody>
</table>
### Table 5 (b): Summary statistics 2

<table>
<thead>
<tr>
<th>wave</th>
<th>loginc</th>
<th>logsav</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>mean&lt;sup&gt;a&lt;/sup&gt;</td>
<td>sd&lt;sup&gt;a&lt;/sup&gt;</td>
</tr>
<tr>
<td>6</td>
<td>5.7619</td>
<td>0.5324</td>
</tr>
<tr>
<td>7</td>
<td>5.7682</td>
<td>0.5388</td>
</tr>
<tr>
<td>8</td>
<td>5.7370</td>
<td>0.6343</td>
</tr>
<tr>
<td>9</td>
<td>5.7553</td>
<td>0.7120</td>
</tr>
<tr>
<td>10</td>
<td>5.7594</td>
<td>0.7173</td>
</tr>
<tr>
<td>11</td>
<td>5.7689</td>
<td>0.5377</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>wave</th>
<th>mean</th>
<th>sd</th>
<th>min</th>
<th>50%</th>
<th>max</th>
</tr>
</thead>
<tbody>
<tr>
<td>6</td>
<td>3.1013</td>
<td>2.4569</td>
<td>0.0000</td>
<td>3.9021</td>
<td>8.5118</td>
</tr>
<tr>
<td>7</td>
<td>3.1273</td>
<td>2.4993</td>
<td>0.0000</td>
<td>3.9050</td>
<td>8.1432</td>
</tr>
<tr>
<td>8</td>
<td>3.2913</td>
<td>2.4653</td>
<td>0.0000</td>
<td>4.0559</td>
<td>8.1502</td>
</tr>
<tr>
<td>9</td>
<td>2.8816</td>
<td>2.6231</td>
<td>0.0000</td>
<td>3.4508</td>
<td>8.1776</td>
</tr>
<tr>
<td>10</td>
<td>3.1855</td>
<td>2.6601</td>
<td>0.0000</td>
<td>3.9282</td>
<td>8.4608</td>
</tr>
<tr>
<td>11</td>
<td>2.9902</td>
<td>2.6571</td>
<td>0.0000</td>
<td>3.5965</td>
<td>8.3235</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>explanatory variable</th>
<th>mean</th>
<th>sd</th>
<th>min</th>
<th>50%</th>
<th>max</th>
</tr>
</thead>
<tbody>
<tr>
<td>age&lt;sup&gt;b&lt;/sup&gt;</td>
<td>33.75</td>
<td>3.24</td>
<td>25</td>
<td>34</td>
<td>39</td>
</tr>
<tr>
<td>educ&lt;sup&gt;b&lt;/sup&gt;</td>
<td>13.28</td>
<td>1.56</td>
<td>9</td>
<td>14</td>
<td>18</td>
</tr>
</tbody>
</table>

<sup>a</sup> ‘mean’ and ‘sd’ denote the sample mean and sample standard deviation. ‘min,’ ‘50%’ and ‘max’ denote the minimum value, median and maximum value.

<sup>b</sup> The data of ‘age’ and ‘educ’ are obtained from Panel 6.
Table 6: Posterior results for parameters (JPSC data)

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Algorithm 2</th>
<th>Algorithm 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \phi(z_{(-1)}) )</td>
<td>mean</td>
<td>sd</td>
</tr>
<tr>
<td>( \beta_1 ) (married)</td>
<td>-0.2809</td>
<td>0.0873</td>
</tr>
<tr>
<td>( \beta_2 ) (loginc)</td>
<td>0.2008</td>
<td>0.0455</td>
</tr>
<tr>
<td>( \beta_3 ) (logsav)</td>
<td>0.0665</td>
<td>0.0144</td>
</tr>
<tr>
<td>( \beta_4 ) (inwork)</td>
<td>-0.1781</td>
<td>0.0572</td>
</tr>
<tr>
<td>( \delta_2 ) (age)</td>
<td>0.0107</td>
<td>0.0084</td>
</tr>
<tr>
<td>( \delta_3 ) (educ)</td>
<td>0.0052</td>
<td>0.0177</td>
</tr>
<tr>
<td>( \beta_{01} ) (married)</td>
<td>0.0033</td>
<td>0.2118</td>
</tr>
<tr>
<td>( \beta_{02} ) (loginc)</td>
<td>0.2970</td>
<td>0.0941</td>
</tr>
<tr>
<td>( \beta_{03} ) (logsav)</td>
<td>0.0725</td>
<td>0.0345</td>
</tr>
<tr>
<td>( \beta_{04} ) (inwork)</td>
<td>-0.0442</td>
<td>0.1338</td>
</tr>
<tr>
<td>( \delta_{01} ) (age)</td>
<td>-0.0302</td>
<td>0.0157</td>
</tr>
<tr>
<td>( \delta_{02} ) (educ)</td>
<td>0.0581</td>
<td>0.0343</td>
</tr>
<tr>
<td>( \gamma_2 )</td>
<td>0.8239</td>
<td>0.0319</td>
</tr>
<tr>
<td>( \gamma_3 )</td>
<td>2.4624</td>
<td>0.0522</td>
</tr>
<tr>
<td>( \phi(z_{(-1)}) )</td>
<td>mean a</td>
<td>sd a</td>
</tr>
<tr>
<td>( \beta_1 ) (married)</td>
<td>-0.2475</td>
<td>0.1886</td>
</tr>
<tr>
<td>( \beta_2 ) (loginc)</td>
<td>0.2163</td>
<td>0.1075</td>
</tr>
<tr>
<td>( \beta_3 ) (logsav)</td>
<td>0.0596</td>
<td>0.0307</td>
</tr>
<tr>
<td>( \beta_4 ) (inwork)</td>
<td>-0.1686</td>
<td>0.1237</td>
</tr>
<tr>
<td>( \delta_1 ) (age)</td>
<td>0.0152</td>
<td>0.0181</td>
</tr>
<tr>
<td>( \delta_2 ) (educ)</td>
<td>0.0122</td>
<td>0.0387</td>
</tr>
<tr>
<td>( \beta_{01} ) (married)</td>
<td>0.0346</td>
<td>0.4613</td>
</tr>
<tr>
<td>( \beta_{02} ) (loginc)</td>
<td>0.3221</td>
<td>0.2233</td>
</tr>
<tr>
<td>( \beta_{03} ) (logsav)</td>
<td>0.0641</td>
<td>0.0736</td>
</tr>
<tr>
<td>( \beta_{04} ) (inwork)</td>
<td>-0.0434</td>
<td>0.2925</td>
</tr>
<tr>
<td>( \delta_{01} ) (age)</td>
<td>-0.0257</td>
<td>0.0356</td>
</tr>
<tr>
<td>( \delta_{02} ) (educ)</td>
<td>0.0611</td>
<td>0.0765</td>
</tr>
<tr>
<td>( \gamma_2 )</td>
<td>0.7791</td>
<td>0.0375</td>
</tr>
<tr>
<td>( \gamma_3 )</td>
<td>2.3166</td>
<td>0.0533</td>
</tr>
</tbody>
</table>

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 \( \phi \) are transformed ones devied by \( \sigma \).
Table 7: Posterior results for partial effects on response probabilities

<table>
<thead>
<tr>
<th>explanatory variables</th>
<th>mean(^a)</th>
<th>partial effects on response probabilities</th>
<th>95%CI(^a)</th>
<th>CD(^a)</th>
<th>p-val(^a)</th>
</tr>
</thead>
<tbody>
<tr>
<td>z(-1)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>y = 1</td>
<td>-0.0385</td>
<td>-5.1213 (-0.0529, -0.0244)</td>
<td>0.4257</td>
<td>0.6703</td>
<td></td>
</tr>
<tr>
<td>y = 2</td>
<td>-0.0200</td>
<td>-5.2098 (-0.0273, -0.0123)</td>
<td>-1.3173</td>
<td>0.1877</td>
<td></td>
</tr>
<tr>
<td>y = 3</td>
<td>0.0366</td>
<td>5.2790 (0.0226, 0.0499)</td>
<td>-0.1135</td>
<td>0.9097</td>
<td></td>
</tr>
<tr>
<td>y = 4</td>
<td>0.0219</td>
<td>4.8807 (0.0131, 0.0307)</td>
<td>0.8923</td>
<td>0.3722</td>
<td></td>
</tr>
<tr>
<td>x1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>married</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>y = 1</td>
<td>0.0529</td>
<td>2.9577 (0.0168, 0.0869)</td>
<td>-0.1125</td>
<td>0.9104</td>
<td></td>
</tr>
<tr>
<td>y = 2</td>
<td>0.0319</td>
<td>2.8314 (0.0102, 0.0545)</td>
<td>0.5584</td>
<td>0.5766</td>
<td></td>
</tr>
<tr>
<td>y = 3</td>
<td>-0.0475</td>
<td>-3.0740 (-0.0757, -0.0155)</td>
<td>-0.0010</td>
<td>0.9992</td>
<td></td>
</tr>
<tr>
<td>y = 4</td>
<td>-0.0373</td>
<td>-2.6769 (-0.0660, -0.0114)</td>
<td>-0.3345</td>
<td>0.7380</td>
<td></td>
</tr>
<tr>
<td>x2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>loginc</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>y = 1</td>
<td>-0.1196</td>
<td>-4.5042 (-0.1715, -0.0672)</td>
<td>0.8817</td>
<td>0.3779</td>
<td></td>
</tr>
<tr>
<td>y = 2</td>
<td>-0.0654</td>
<td>-4.3958 (-0.0946, -0.0302)</td>
<td>0.3746</td>
<td>0.7079</td>
<td></td>
</tr>
<tr>
<td>y = 3</td>
<td>0.1126</td>
<td>4.5582 (0.0634, 0.1699)</td>
<td>-0.7608</td>
<td>0.4468</td>
<td></td>
</tr>
<tr>
<td>y = 4</td>
<td>0.0724</td>
<td>4.2127 (0.0394, 0.1070)</td>
<td>-0.6157</td>
<td>0.5381</td>
<td></td>
</tr>
<tr>
<td>x3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>logsav</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>y = 1</td>
<td>-0.0345</td>
<td>-3.7962 (-0.0514, -0.0162)</td>
<td>0.2893</td>
<td>0.7724</td>
<td></td>
</tr>
<tr>
<td>y = 2</td>
<td>-0.0181</td>
<td>-3.6672 (-0.0276, -0.0084)</td>
<td>-0.5238</td>
<td>0.6005</td>
<td></td>
</tr>
<tr>
<td>y = 3</td>
<td>0.0315</td>
<td>3.7918 (0.0154, 0.0478)</td>
<td>-0.0378</td>
<td>0.9699</td>
<td></td>
</tr>
<tr>
<td>y = 4</td>
<td>0.0201</td>
<td>3.4871 (0.0091, 0.0318)</td>
<td>0.0839</td>
<td>0.9331</td>
<td></td>
</tr>
<tr>
<td>x4</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>inwork</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>y = 1</td>
<td>0.0377</td>
<td>2.9814 (0.0129, 0.0624)</td>
<td>0.0077</td>
<td>0.9939</td>
<td></td>
</tr>
<tr>
<td>y = 2</td>
<td>0.0206</td>
<td>2.8775 (0.0068, 0.0349)</td>
<td>1.4888</td>
<td>0.1366</td>
<td></td>
</tr>
<tr>
<td>y = 3</td>
<td>-0.0359</td>
<td>-2.9028 (-0.0594, -0.0123)</td>
<td>-0.2444</td>
<td>0.8070</td>
<td></td>
</tr>
<tr>
<td>y = 4</td>
<td>-0.0224</td>
<td>-2.8490 (-0.0381, -0.0072)</td>
<td>-1.0815</td>
<td>0.2795</td>
<td></td>
</tr>
<tr>
<td>u1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>age</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>y = 1</td>
<td>-0.0005</td>
<td>-0.5075 (-0.0047, 0.0028)</td>
<td>-1.3179</td>
<td>0.1876</td>
<td></td>
</tr>
<tr>
<td>y = 2</td>
<td>-0.0005</td>
<td>-0.4946 (-0.0025, 0.0015)</td>
<td>-1.4062</td>
<td>0.1597</td>
<td></td>
</tr>
<tr>
<td>y = 3</td>
<td>0.0010</td>
<td>0.5219 (0.0026, 0.0043)</td>
<td>1.3253</td>
<td>0.1851</td>
<td></td>
</tr>
<tr>
<td>y = 4</td>
<td>0.0005</td>
<td>0.4713 (0.0016, 0.0028)</td>
<td>1.3862</td>
<td>0.1657</td>
<td></td>
</tr>
<tr>
<td>w2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>educ</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>y = 1</td>
<td>-0.0019</td>
<td>-0.8934 (-0.0062, 0.0022)</td>
<td>-1.5455</td>
<td>0.1222</td>
<td></td>
</tr>
<tr>
<td>y = 2</td>
<td>0.0030</td>
<td>0.8008 (-0.0044, 0.0103)</td>
<td>1.4429</td>
<td>0.1491</td>
<td></td>
</tr>
<tr>
<td>y = 3</td>
<td>0.0021</td>
<td>0.9022 (-0.0024, 0.0069)</td>
<td>1.4382</td>
<td>0.1504</td>
<td></td>
</tr>
</tbody>
</table>

\(a\): ‘mean’ and ‘sd’ denote the posterior mean and posterior standard deviation.
\(b\): ‘CD’ denotes the convergence diagnostic statistic of MCMC proposed by Geweke (1992).