On Polychoric and Polyserial Partial Correlation Coefficients: A Bayesian Approach

HIKARU HASEGAWA*
Graduate School of Economics and Business Administration
Hokkaido University

The first version: January 29, 2012
This version: February 13, 2016

Abstract

This article provides the estimation method for multivariate polychoric and polyserial correlation coefficients by using the simulation-based Bayesian method. It also shows that the *partial* version of the polychoric and polyserial correlation coefficients can be estimated using the corresponding estimates of the *simple* version. A simulation study illustrates the proposed method. Further, an application of the method to subjective well-being data is provided.
KEYWORDS: Gibbs sampler, Markov chain Monte Carlo (MCMC), Metropolis-Hastings (M-H) algorithm, well-being.

JEL CLASSIFICATION: **C11**, **C35**.

1 Introduction

In the social sciences, we often encounter a situation where correlations between ordinal discrete variables or between an ordinal discrete variable and a continuous variable need to investigated. For instance, in the happiness study field, it is important to investigate the relationships between happiness and other kinds of satisfaction or between happiness and income. The polychoric and the polyserial correlation coefficients are used for considering the relationships between two ordinal variables and between an ordinal discrete variable and a continuous variable, respectively.

To consider more than two ordinal discrete and/or continuous variables, we have to employ multivariate models. Poon and Lee (1987) estimate the multivariate polychoric and polyserial correlation coefficients by using the maximum likelihood method. This article employs a Bayesian method to estimate these two correlation coefficients. In a seminal work, Albert (1992) uses a latent bivariate normal distribution to estimate a polychoric correlation coefficient from

*Address for correspondence: 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)

the Bayesian point of view by using the Gibbs sampler. In this article, we extend his idea to the problem of estimating the multivariate polychoric and polyserial correlation coefficients using the Markov chain Monte Carlo (MCMC) method. This article employs the Bayesian method proposed by Chen and Dey (2000) that is used to estimate a multivariate ordered probit model.[1]

In a situation where more than two variables exist, it is important to consider the relationship between the two variables after excluding the other variables' effects. Therefore, we also estimate the *partial* version of the polychoric and polyserial correlation coefficients. In the estimation process of the multivariate ordinal data model described in the following sections, we can easily obtain the *simple* and *partial* versions of the polychoric and polyserial correlation coefficients and their standard deviations. This is one merit of using the MCMC method.

The article proceeds as follows. In Section 2, we describe the Bayesian model and its estimation procedure by using the Chen and Dey (2000) algorithm. Further, we provide an estimation procedure for the correlation coefficients. In Section 3, a simulation study illustrates the proposed method. In Section 4, we provide an application of the method to subjective well-being data. In Section 6 we provide the concluding remarks.

2 Bayesian Model

2.1 Estimated Model

For $i = 1, \cdots, n$, let y_{1ij} $(j = 1, \cdots, m)$ and y_{2il} $(l = 1, \cdots, q)$ denote the ordinal discrete and the continuous variables, respectively. The discrete variable y_{1ij} makes ordinal choices, that is, $y_{1ij} = c$ for $c = 1, \cdots, C_j$. We assume the following model for the ordinal discrete variable y_{1ij}:

$$y_{1ij} = c,\ z_{ij} \in (\gamma_{j(c-1)}, \gamma_{jc}],\ c = 1, \cdots, C_j;\ j = 1, \cdots, m;\ i = 1, \cdots, n, \tag{1}$$

where z_{ij} denotes the latent variable and γ_{jc} is a cutoff point for the jth ordinal response. Following Chen and Dey (2000), we specify that

$$\begin{aligned} &-\infty = \gamma_{j0} < \gamma_{j1} = 0 < \gamma_{j2} < \cdots < \gamma_{j(C_j-1)} = 1 < \gamma_{jC_j} = \infty, \\ &\quad j = 1, \cdots, m, \end{aligned} \tag{2}$$

where the conditions $\gamma_{j1} = 0$ and $\gamma_{j(C-1)} = 1$ are required to establish the identifiability of the cutoff parameters.[2] We define $\boldsymbol{\gamma}_j = (\gamma_{j2}, \cdots, \gamma_{j(C_j-2)})'$ $(j = 1, \cdots, m)$ and $\boldsymbol{\gamma} = (\boldsymbol{\gamma}_1', \cdots, \boldsymbol{\gamma}_m')'$.

[1] Kottas et al. (2005) also consider a nonparametric Bayesian model for multivariate ordinal data and take into account the estimation of the polychoric correlation coefficients in the multivariate situation. Further, Zhang et al. (2006) consider the multivariate probit model, and provide the estimation method for the tetrachoric or polychoric correlation using the PX-MH algorithm.

[2] Usually, $\boldsymbol{\Sigma}$ is assumed to be a correlation matrix for identification. See, for example, Chib and Greenberg (1998, p.348). In this direction, Zhang et al. (2006) consider the multivariate probit model, and provide the estimation method of a correlation matrix using the PX-MH algorithm. However, since this article employs the method provided in Chen and Dey (2000), we can specify the model using a covariance matrix instead of a correlation matrix. Further, as described in Section 2.2, we can estimate a correlation matrix after estimating the covariance matrix.

The latent variable z_{ij} is assumed to be determined by the following model:

$$z_{ij} = \beta_{1j} + u_{1ij},\ j = 1, \cdots, m;\ i = 1, \cdots, n. \tag{3}$$

Defining

$$\boldsymbol{z}_i = \begin{pmatrix} z_{i1} \\ \vdots \\ z_{im} \end{pmatrix},\ \boldsymbol{\beta}_2 = \begin{pmatrix} \beta_{11} \\ \vdots \\ \beta_{1m} \end{pmatrix},\ \boldsymbol{u}_{1i} = \begin{pmatrix} u_{1i1} \\ \vdots \\ u_{1im} \end{pmatrix},$$

(3) is rewritten as

$$\boldsymbol{z}_i = \boldsymbol{\beta}_1 + \boldsymbol{u}_{1i},\ i = 1, \cdots, n. \tag{4}$$

For the continuous variable y_{2il}, we consider the following model:

$$y_{2il} = \beta_{2l} + u_{2il},\ l = 1, \cdots, q;\ i = 1, \cdots, n. \tag{5}$$

Defining

$$\boldsymbol{y}_{2i} = \begin{pmatrix} y_{2i1} \\ \vdots \\ y_{2iq} \end{pmatrix},\ \boldsymbol{\beta}_2 = \begin{pmatrix} \beta_{21} \\ \vdots \\ \beta_{2q} \end{pmatrix},\ \boldsymbol{u}_{2i} = \begin{pmatrix} u_{2i1} \\ \vdots \\ u_{2iq} \end{pmatrix},$$

(5) is rewritten as

$$\boldsymbol{y}_{2i} = \boldsymbol{\beta}_2 + \boldsymbol{u}_{2i},\ i = 1, \cdots, n. \tag{6}$$

Now, defining $\boldsymbol{w}_i = (\boldsymbol{z}_i', \boldsymbol{y}_{2i})'$, (4) and (6) can be rewritten as

$$\boldsymbol{w}_i = \begin{pmatrix} \boldsymbol{z}_i \\ \boldsymbol{y}_{2i} \end{pmatrix} = \boldsymbol{\beta} + \boldsymbol{u}_i,$$

where

$$\boldsymbol{\beta} = \begin{pmatrix} \boldsymbol{\beta}_1 \\ \boldsymbol{\beta}_2 \end{pmatrix},\ \boldsymbol{u}_i = \begin{pmatrix} \boldsymbol{u}_{1i} \\ \boldsymbol{u}_{2i} \end{pmatrix}.$$

Assuming that $\boldsymbol{u}_i \sim \mathrm{N}(\boldsymbol{0}, \boldsymbol{\Sigma})$, we have

$$\boldsymbol{w}_i \sim \mathrm{N}(\boldsymbol{\beta}, \boldsymbol{\Sigma}),\ i = 1, \cdots, n, \tag{7}$$

where $\boldsymbol{\Sigma}$ is an $(m+q) \times (m+q)$ positive definite covariance matrix. Further, we collect up $\boldsymbol{y}_{1i} = (y_{1i1}, \cdots, y_{1im})', \boldsymbol{y}_{2i}, \boldsymbol{z}_i, \boldsymbol{w}_i$ for $i = 1, \cdots, n$ as follows:

$$\boldsymbol{y}_1 = \begin{pmatrix} \boldsymbol{y}_{11} \\ \vdots \\ \boldsymbol{y}_{1n} \end{pmatrix},\ \boldsymbol{y}_2 = \begin{pmatrix} \boldsymbol{y}_{21} \\ \vdots \\ \boldsymbol{y}_{2n} \end{pmatrix},\ \boldsymbol{z} = \begin{pmatrix} \boldsymbol{z}_1 \\ \vdots \\ \boldsymbol{z}_n \end{pmatrix},\ \boldsymbol{w} = \begin{pmatrix} \boldsymbol{w}_1 \\ \vdots \\ \boldsymbol{w}_n \end{pmatrix}.$$

To complete the Bayesian model, we introduce the prior distributions of the parameters $p(\boldsymbol{\beta}, \boldsymbol{\gamma}, \boldsymbol{\Sigma})$. On the basis of Bayes' theorem, the joint posterior distribution can be written as

$$\begin{aligned} p(\boldsymbol{\beta}, \boldsymbol{\gamma}, \boldsymbol{\Sigma}, \boldsymbol{z} | \boldsymbol{y}_1, \boldsymbol{y}_2) &\propto p(\boldsymbol{\beta}, \boldsymbol{\gamma}, \boldsymbol{\Sigma}, \boldsymbol{z}) p(\boldsymbol{y}_1, \boldsymbol{y}_2 | \boldsymbol{\beta}, \boldsymbol{\gamma}, \boldsymbol{\Sigma}, \boldsymbol{z}) \\ &= p(\boldsymbol{\beta}, \boldsymbol{\gamma}, \boldsymbol{\Sigma}) \left[\prod_{i=1}^{n} p(\boldsymbol{w}_i | \boldsymbol{\beta}, \boldsymbol{\gamma}, \boldsymbol{\Sigma}) p(\boldsymbol{y}_{1i} | \boldsymbol{\beta}, \boldsymbol{\gamma}, \boldsymbol{\Sigma}, \boldsymbol{z}_i) \right]. \end{aligned}$$

Further, defining

$$\mathcal{G}_{ij} = \left(\gamma_{j(c-1)}, \gamma_{jc}\right] \text{ if } y_{ij} = c,\ j = 1, \cdots, m;\ i = 1, \cdots, n$$
$$\mathcal{G}_i = \mathcal{G}_{i1} \times \cdots \times \mathcal{G}_{im},\ i = 1, \cdots, n,$$

we have

$$p(\boldsymbol{y}_{1i}|\boldsymbol{\beta}, \boldsymbol{\gamma}, \boldsymbol{\Sigma}, \boldsymbol{z}_i) = 1_{(\boldsymbol{z}_i \in \mathcal{G}_i)},\ i = 1, \cdots, n,$$

where $1_{(\cdot)}$ is an indicator function.[3] Now, we specify the prior distributions as follows:

$$p(\boldsymbol{\beta}, \boldsymbol{\gamma}, \boldsymbol{\Sigma}) = p(\boldsymbol{\beta})p(\boldsymbol{\gamma})p(\boldsymbol{\Sigma}) = p(\boldsymbol{\beta}) \left\{ \prod_{j=1}^{m} p(\boldsymbol{\gamma}_j) \right\} p(\boldsymbol{\Sigma}),$$

where

$$\boldsymbol{\beta} \sim \mathrm{N}(\boldsymbol{\beta}_0, \boldsymbol{B}_0),\ \boldsymbol{\Sigma}^{-1} \sim \mathrm{W}(\kappa_0, \boldsymbol{Q}_0^{-1}),$$

and $\mathrm{W}(\kappa_0, \boldsymbol{Q}_0^{-1})$ denotes a Wishart distribution with degrees of freedom κ_0 and scale matrix $\boldsymbol{Q}_0^{-1}$. Further, we introduce the prior distribution of $\boldsymbol{\gamma}_j$, $p(\boldsymbol{\gamma}_j) = p(\boldsymbol{\delta}_j(\boldsymbol{\gamma}_j))$, based on the following transformation for the cutoff points (Chen and Dey, 2000, p.140):

$$\delta_{jc} = \log\left(\frac{\gamma_{jc} - \gamma_{j(c-1)}}{1 - \gamma_{jc}}\right),\ c = 2, \cdots, C_j - 2,$$

where $\boldsymbol{\delta}_j(\boldsymbol{\gamma}_j) = (\delta_{j2}, \cdots, \delta_{j(C_j-2)})'$ $(j = 1, \cdots, m)$.[4] We specify $p(\boldsymbol{\delta}_j(\boldsymbol{\gamma}_j))$ as follows:

$$\boldsymbol{\delta}_j(\boldsymbol{\gamma}_j) \sim \mathrm{N}(\boldsymbol{\delta}_{j0}, \boldsymbol{D}_{j0}),\ j = 1, \cdots, m.$$

Thus, the joint posterior distribution can be written as

$$\begin{aligned} p(\boldsymbol{\beta}, \boldsymbol{\gamma}, \boldsymbol{\Sigma}, \boldsymbol{z}|\boldsymbol{y}_1, \boldsymbol{y}_2) &\propto p(\boldsymbol{\beta}, \boldsymbol{\gamma}, \boldsymbol{\Sigma}) \\ &\times \left\{ \prod_{i=1}^{n} 1_{(\boldsymbol{z}_i \in \mathcal{G}_i)} |\boldsymbol{\Sigma}|^{-\frac{1}{2}} \exp\left[-\frac{1}{2}(\boldsymbol{w}_i - \boldsymbol{\beta})'\boldsymbol{\Sigma}^{-1}(\boldsymbol{w}_i - \boldsymbol{\beta}) \right] \right\}. \end{aligned} \tag{8}$$

Using the sampling scheme of MCMC, we can sample the parameters $(\boldsymbol{\beta}, \boldsymbol{\gamma}, \boldsymbol{\Sigma}, \boldsymbol{z})$ from the joint posterior distribution (8). Appendix A provides the sampling algorithms' details.

2.2 Correlation Matrices

Using the sample of $\boldsymbol{\Sigma}$, we can calculate the *simple* version of the polychoric and polyserial correlation matrix at each iteration of MCMC:

$$\boldsymbol{R}_s = \{r_{ij}^s\} = \boldsymbol{D}_s \boldsymbol{\Sigma} \boldsymbol{D}_s, \tag{9}$$

[3] See Chib and Greenberg (1998, p.349).
[4] For more details on this transformation, see Jeliazkov et al. (2008, p.124).

where $\boldsymbol{D}_s = \mathrm{diag}(1/\sqrt{\sigma_{11}}, \cdots, 1/\sqrt{\sigma_{m+q,m+q}})$ and σ_{jj} is the jth diagonal element of $\boldsymbol{\Sigma}$. Further, we can calculate the *partial* version of the polychoric and polyserial correlation coefficients defined as the lower triangular part without the diagonal elements of the following matrix:

$$\boldsymbol{R}_p = \{r_{ij}^p\} = -\boldsymbol{D}_r \boldsymbol{R}_s^{-1} \boldsymbol{D}_r, \tag{10}$$

where $\boldsymbol{D}_r = \mathrm{diag}(1/\sqrt{r^{11}}, \cdots, 1/\sqrt{r^{m+q,m+q}})$ and r^{jj} is the jth diagonal element of $\boldsymbol{R}_s^{-1}$.

3 Numerical Example with Simulated Data

This section provides a numerical example with simulated data for investigating the performances of our approach. In the numerical example, we set $n = 100$, $m = 3$ and $q = 1$, and assumed that the ordered response variables y_{1ij} ($i = 1, \cdots, 100,\ j = 1, 2, 3$) take four values, that is, $y_{1ij} = 1, 2, 3, 4$. We also assume that $\boldsymbol{w}_i \sim \mathrm{N}(\boldsymbol{\beta}, \boldsymbol{\Sigma})$ in (7) are specified as follows:

$$\boldsymbol{\beta} = (0.5, 0.5, 0.5, 2.0)',\ \boldsymbol{\Sigma} = \mathrm{diag}(\boldsymbol{\sigma})\boldsymbol{R}\,\mathrm{diag}(\boldsymbol{\sigma}),$$

where $\boldsymbol{\sigma} = (0.5, 0.5, 0.5, 1.0)'$. Following Zhang et al. (2006, p.887), we use the correlation matrix $\boldsymbol{R} = (r_{ij})$ with four parameters $\boldsymbol{\rho} = (\rho_1, \rho_2, \rho_3, \rho_4)'$ and $r_{ij} = \prod_{l=i}^{j} \rho_l$ for $i < j$. In the numerical example, we set $\boldsymbol{\rho} = (0.7, 0.8, 0.4, 0.6)'$. After specifying these data generating structures, we obtain the values of $\boldsymbol{w}_i$ ($i = 1, \cdots, 100$). Further, specifying $\gamma_{j2} = 0.6$ for $j = 1, 2, 3$, we obtain the values y_{1ij} ($i = 1, \cdots, 100,\ j = 1, 2, 3$). The summary statistics of the data are provided in Table 1.

The MCMC simulation was run for 20,000 iterations with the thinning interval being 10; the first 5,000 samples were discarded as the burn-in period.[5] The posterior results were generated thereafter using Ox version 6.30 (Doornik, 2009), and the output analysis and convergence diagnosis of the MCMC samples were implemented by the R package `coda`. We set the prior distributions as follows:

$$\boldsymbol{\beta} \sim \mathrm{N}(\boldsymbol{0}, 100\boldsymbol{I}_4),\ \delta_{j2}(\gamma_{j2}) \sim \mathrm{N}(0, 100),\ j = 1, 2, 3,\ \boldsymbol{\Sigma}^{-1} \sim \mathrm{W}(6, \boldsymbol{Q}_0^{-1}), \tag{11}$$

where $\boldsymbol{Q}_0 = 0.02\boldsymbol{I}_4$. Table 2 presents the posterior results. In Table 2, "Mean," "SD," and "Median" denote the posterior mean, posterior standard deviation, and posterior median, respectively. The "$*$" in Median denotes that zero is not included in the 95% credible interval. "Geweke" and "HW" denote the p-values for the convergence diagnostic statistics proposed by Geweke (1992) and Heidelberger and Welch (1983), respectively. "GR" denotes the convergence diagnostic statistic proposed by Gelman and Rubin (1992) and modified by Brooks and Gelman (1998).[6] According to the convergence diagnostic statistics, we verify that the MCMC samples converge.

[5] We obtained 20,000 samples by running 200,000 iterations and sampling one observation at every ten iterations. After obtaining 20,000 samples, we discarded the first 5,000 samples.

[6] For calculating "GR", we implemented two additional chains generated from the same procedure as the first chain.

The prior distributions specified in (11) are less informative specifications. However, in particular, the specification of prior distribution of $\mathbf{\Sigma}$ might be explained in more detail. The full conditional distribution (FCD) of $\mathbf{\Sigma}^{-1}$ is

$$\mathbf{\Sigma}^{-1}|\cdots \sim \mathrm{W}(\tilde{\kappa}, \tilde{\boldsymbol{Q}}^{-1}),$$

where "$|\cdots$" denotes the conditioning of the other unspecified variables in the model, $\tilde{\kappa} = \kappa_0 + n$, and

$$\tilde{\boldsymbol{Q}} = \boldsymbol{Q}_0 + \sum_{i=1}^{n}(\boldsymbol{z}_i - \boldsymbol{X}_i\boldsymbol{\beta})(\boldsymbol{z}_i - \boldsymbol{X}_i\boldsymbol{\beta})'. \tag{12}$$

In (11), we set $\boldsymbol{Q}_0 = 0.02\boldsymbol{I}$. Thus, $\boldsymbol{Q}_0$ does not have a significant effect on the calculation of $\tilde{\boldsymbol{Q}}$ in (12). Table 3 provides the posterior results of the model for the different values of $\boldsymbol{Q}_0$. The case of $\boldsymbol{Q}_0 = \boldsymbol{I}$ is often used in the existing literature.[7] The case of $\boldsymbol{Q}_0 = 10\boldsymbol{I}$ is more informative than the others. From Table 3, the cases of $\boldsymbol{Q}_0 = 0.02$ and $\boldsymbol{Q}_0 = \boldsymbol{I}$ provide similar posterior results. The cases of $\boldsymbol{Q}_0 = 5\boldsymbol{I}$ and $\boldsymbol{Q}_0 = 10\boldsymbol{I}$ provide slightly different results from those of $\boldsymbol{Q}_0 = 0.02$ and $\boldsymbol{Q}_0 = \boldsymbol{I}$.

4 Application to the Analysis of Well-Being Data

4.1 Data

Our empirical application example 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. The JPSC data provide several measures of satisfaction pertaining to young women in Japan. This article employs the 14th wave (2006) of JPSC data. In this example, we select the data on single women who responded to all the questions on satisfaction. Our selection contains 515 recorded responses.

We use the following variables:

- Ordinal variables ($m = 6$):
  `Happiness` (y_{11}, z_1), `Life_Sat` (y_{12}, z_2), `Liv_Std` (y_{13}, z_3),
  `Inc_Sat` (y_{14}, z_4), `Cons_Sat` (y_{15}, z_5), `Health_Cond` (y_{16}, z_6)
- Continuous variable ($q = 1$):
  `Log_Inc` (y_{21}).

Here `Happiness` denotes the degree of happiness; `Life_Sat`, life satisfaction; `Liv_Std`, the subjective standard of living; `Inc_Sat`, satisfaction regarding revenue; `Cons_Sat`, satisfaction regarding consumption; and `Health_Cond`, the health condition of the respondent. In the literature of happiness, these variables are called the subjective well-being data. The responses numbers for the ordinal variables are $C_1 = C_2 = C_6 = 5$ and $C_3 = C_4 = C_5 = 4$. Furthermore, `Log_Inc` implies the logarithm of the annual earnings (in $10,000$ yen). The summary statistics of these variables are provided in Table 4.

[7] See, for example, Gelman and Hill (2007, p.377) and Jackman (2009, p.242).

4.2 Posterior results

The MCMC simulation was run for 20,000 iterations with the thinning interval being 10; the first 5,000 samples were discarded as the burn-in period. We set the prior distributions as follows:

$$\begin{aligned}
&\boldsymbol{\beta} \sim \mathrm{N}(\mathbf{0}, 100\boldsymbol{I}_7) \\
&\boldsymbol{\delta}_j(\gamma_j) \sim \mathrm{N}(\mathbf{0}, 100\boldsymbol{I}_2),\ j = 1, 2, 6,\ \delta_{j2}(\gamma_{j2}) \sim \mathrm{N}(0, 100),\ j = 3, 4, 5 \\
&\boldsymbol{\Sigma}^{-1} \sim \mathrm{W}(8, \boldsymbol{Q}_0^{-1}),\ \boldsymbol{Q}_0 = 0.02\boldsymbol{I}_7.
\end{aligned}$$

Table 5 presents the parameters' posterior results. Tables 6 and 7 provide the posterior results of the polyserial and polychoric (simple) correlation and the partial correlation coefficients, respectively. According to the convergence diagnostic statistics in these tables, we verify that the MCMC samples converge. From these tables, the following observations can be made:

- From Table 5, we observe that the Bayesian model is well estimated.
- All ordinal discrete variables are correlated in Table 6. However, from Table 7, after excluding other variables' effects, some ordinal discrete variables appear to be uncorrelated, *e.g.*, r^p_{31} (correlation between `Liv_Std` and `Happiness`).
- `Health_Cond` and `Inc_Sat` are positively correlated in Table 6 (r^s_{64}), but negatively in Table 7 (r^p_{64}).
- The continuous variable `Log_Inc` is positively correlated with `Life_Sat`, `Liv_Std`, and `Health_Con` in Table 6. However, in Table 7, after excluding other variables' effects, `Log_Inc` is correlated only with `Health_Con`.

Table 8 provides the posterior results of the model for the different values of $\boldsymbol{Q}_0$. From this table, the cases of $\boldsymbol{Q}_0 = 0.02$ and $\boldsymbol{Q}_0 = \boldsymbol{I}$ provide similar posterior results. The cases of $\boldsymbol{Q}_0 = 5\boldsymbol{I}$ and $\boldsymbol{Q}_0 = 10\boldsymbol{I}$ provide slightly different results from those of $\boldsymbol{Q}_0 = 0.02$ and $\boldsymbol{Q}_0 = \boldsymbol{I}$.

5 Concluding Remarks

In this article, we presented the estimation method for multivariate polychoric and polyserial correlation coefficients by using the simulation-based Bayesian method. Further, using the estimates of the *simple* version of polychoric and polyserial correlation coefficients, we provided the estimation procedure of the corresponding *partial* version.

The Bayesian model used in this article can be extended to a model with regressors, that is, a model including a multivariate ordered probit model and a multivariate linear regression model. Adding the vectors of regressors, the linear model for the latent variables (3) and the continuous variables (5) can be extended to

$$\begin{aligned}
&z_{ij} = \boldsymbol{x}'_{1ij}\boldsymbol{\beta}_{1j} + u_{1ij},\ j = 1, \cdots, m \\
&y_{2il} = \boldsymbol{x}'_{2il}\boldsymbol{\beta}_{2l} + u_{2il},\ l = 1, \cdots, q,
\end{aligned}$$

where

$$\boldsymbol{x}_{1ij} = \begin{pmatrix} x_{1ij1} \\ x_{1ij2} \\ \vdots \\ x_{1ijk_{1j}} \end{pmatrix}, \ \boldsymbol{\beta}_{1j} = \begin{pmatrix} \beta_{1j1} \\ \beta_{1j2} \\ \vdots \\ \beta_{1jk_{1j}} \end{pmatrix}, \ j = 1, \cdots, m$$

$$\boldsymbol{x}_{2il} = \begin{pmatrix} x_{2il1} \\ x_{2il2} \\ \vdots \\ x_{2ilk_{2l}} \end{pmatrix}, \ \boldsymbol{\beta}_{2l} = \begin{pmatrix} \beta_{2l1} \\ \beta_{2l2} \\ \vdots \\ \beta_{2lk_{2l}} \end{pmatrix}, \ l = 1, \cdots, q.$$

Defining $\boldsymbol{\beta}_1 = (\boldsymbol{\beta}'_{11}, \boldsymbol{\beta}'_{12}, \cdots, \boldsymbol{\beta}'_{1m})'$, $\boldsymbol{\beta}_2 = (\boldsymbol{\beta}'_{21}, \boldsymbol{\beta}'_{22}, \cdots, \boldsymbol{\beta}'_{2q})'$, and

$$\boldsymbol{z}_i = \begin{pmatrix} z_{i1} \\ z_{i2} \\ \vdots \\ z_{im} \end{pmatrix}, \ \boldsymbol{X}_{1i} = \operatorname{diag}(\boldsymbol{x}'_{1i1}, \boldsymbol{x}'_{1i2}, \cdots, \boldsymbol{x}'_{1im}), \ \boldsymbol{u}_i = \begin{pmatrix} u_{1i1} \\ u_{1i2} \\ \vdots \\ u_{1im} \end{pmatrix}$$

$$\boldsymbol{y}_{2i} = \begin{pmatrix} y_{2i1} \\ y_{2i2} \\ \vdots \\ y_{2iq} \end{pmatrix}, \ \boldsymbol{X}_{2i} = \operatorname{diag}(\boldsymbol{x}'_{2i1}, \boldsymbol{x}'_{2i2}, \cdots, \boldsymbol{x}'_{2iq}), \ \boldsymbol{u}_{2i} = \begin{pmatrix} u_{2i1} \\ u_{2i2} \\ \vdots \\ u_{2iq} \end{pmatrix},$$

(4) and (6) can be extended to

$$\begin{aligned} \boldsymbol{z}_i &= \boldsymbol{X}_{1i}\boldsymbol{\beta}_1 + \boldsymbol{u}_{1i}, \ i = 1, \cdots, n \\ \boldsymbol{y}_{2i} &= \boldsymbol{X}_{2i}\boldsymbol{\beta}_2 + \boldsymbol{u}_{2i}, \ i = 1, \cdots, n. \end{aligned}$$

Therefore, assuming that $\boldsymbol{u}_i \sim \mathrm{N}(\mathbf{0}, \boldsymbol{\Sigma})$, (7) can be replaced by

$$\boldsymbol{w}_i \sim \mathrm{N}(\boldsymbol{X}_i\boldsymbol{\beta}, \boldsymbol{\Sigma}), \ i = 1, \cdots, n, \tag{13}$$

where

$$\boldsymbol{w}_i = \begin{pmatrix} \boldsymbol{z}_i \\ \boldsymbol{y}_{2i} \end{pmatrix}, \ \boldsymbol{X}_i = \begin{pmatrix} \boldsymbol{X}_{1i} & \mathbf{O} \\ \mathbf{O} & \boldsymbol{X}_{2i} \end{pmatrix}, \ \boldsymbol{\beta} = \begin{pmatrix} \boldsymbol{\beta}_1 \\ \boldsymbol{\beta}_2 \end{pmatrix}, \ \boldsymbol{u}_i = \begin{pmatrix} \boldsymbol{u}_{1i} \\ \boldsymbol{u}_{2i} \end{pmatrix}.$$

Combining a prior distribution of $\boldsymbol{\beta}$, $\boldsymbol{\delta}$, and $\boldsymbol{\Sigma}$ with (13), we can obtain the posterior distribution of the parameters, and apply a similar algorithm for generating the parameters used in Section 2 to the posterior distribution. Further, the *simple* and *partial* versions of the polychoric and polyserial correlation coefficients can be estimated from the posterior result of $\boldsymbol{\Sigma}$ as described in Section 2.2.

A Sampling Algorithms

A.1 Sampling of $\boldsymbol{\beta}$ and $\boldsymbol{\Sigma}^{-1}$

We have the following full conditional distributions (FCDs) of $\boldsymbol{\beta}$ and $\boldsymbol{\Sigma}^{-1}$.

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

$$\boldsymbol{\beta}|\cdots \sim \mathrm{N}(\tilde{\boldsymbol{\beta}}, \tilde{\boldsymbol{B}}), \tag{14}$$

where "$|\cdots$" denotes the conditioning of the other unspecified variables in the model, and

$$\tilde{\boldsymbol{B}} = \left(\boldsymbol{B}_0^{-1} + n\boldsymbol{\Sigma}^{-1}\right)^{-1},\ \tilde{\boldsymbol{\beta}} = \tilde{\boldsymbol{B}}\left(\boldsymbol{B}_0^{-1}\boldsymbol{\beta}_0 + \boldsymbol{\Sigma}^{-1}\sum_{i=1}^{n}\boldsymbol{w}_i\right).$$

- The FCD of $\boldsymbol{\Sigma}^{-1}$ is

$$\boldsymbol{\Sigma}^{-1}|\cdots \sim \mathrm{W}(\tilde{\kappa}, \tilde{\boldsymbol{Q}}^{-1}), \tag{15}$$

where

$$\tilde{\kappa} = \kappa_0 + n,\ \tilde{\boldsymbol{Q}}^{-1} = \boldsymbol{Q}_0 + \sum_{i=1}^{n}(\boldsymbol{w}_i - \boldsymbol{\beta})(\boldsymbol{w}_i - \boldsymbol{\beta})'.$$

Applying Gibbs sampling to the FCDs of (14) and (15), we can generate $\boldsymbol{\beta}$ and $\boldsymbol{\Sigma}^{-1}$.

A.2 Sampling of $\boldsymbol{z}$ and $\boldsymbol{\gamma}$

Let $\boldsymbol{z}_{(j)} = (z_{1j}, z_{2j}, \cdots, z_{nj})'$ denote the vector of the jth element z_{ij} from $\boldsymbol{z}_i$ $(i = 1, \cdots, n)$. Further, let $\boldsymbol{z}_{(-j)}$ and $\boldsymbol{w}_{(-j)}$ denote the vector obtained by removing $\boldsymbol{z}_{(j)}$ from $\boldsymbol{z}$ and $\boldsymbol{w}$, respectively, and let $\boldsymbol{z}_{i(-j)}$ and $\boldsymbol{w}_{i(-j)}$ denote the vector obtained by removing z_{ij} from $\boldsymbol{z}_i$ and $\boldsymbol{w}_i$, respectively, where

$$\boldsymbol{w}_{(-j)} = \begin{pmatrix} \boldsymbol{z}_{(-j)} \\ \boldsymbol{y}_2 \end{pmatrix},\ \boldsymbol{w}_{i(-j)} = \begin{pmatrix} \boldsymbol{z}_{i(-j)} \\ \boldsymbol{y}_{2i} \end{pmatrix}.$$

We generate $\boldsymbol{\gamma}_j$ and $\boldsymbol{z}_{(j)}$ from the joint conditional distribution $p(\boldsymbol{\gamma}_j, \boldsymbol{z}_{(j)}|\boldsymbol{\beta}, \boldsymbol{\Sigma}, \boldsymbol{w}_{(-j)}, \boldsymbol{y}_1)$ $(j = 1, \cdots, m)$. The joint conditional distribution $p(\boldsymbol{\gamma}_j, \boldsymbol{z}_{(j)}|\boldsymbol{\beta}, \boldsymbol{\Sigma}, \boldsymbol{w}_{(-j)}, \boldsymbol{y}_1)$ can be written as

$$p(\boldsymbol{\gamma}_j, \boldsymbol{z}_{(j)}|\boldsymbol{\beta}, \boldsymbol{\Sigma}, \boldsymbol{w}_{(-j)}, \boldsymbol{y}_1) = p(\boldsymbol{\gamma}_j|\boldsymbol{\beta}, \boldsymbol{\Sigma}, \boldsymbol{w}_{(-j)}, \boldsymbol{y})p(\boldsymbol{z}_{(j)}|\boldsymbol{\gamma}_j, \boldsymbol{\beta}, \boldsymbol{\Sigma}, \boldsymbol{w}_{(-j)}, \boldsymbol{y}_1),$$
$$j = 1, \cdots, m.$$

For convenience of expression, we replace the jth factor of $\boldsymbol{w}_i, \boldsymbol{\beta}, \boldsymbol{\Sigma}$ and $\boldsymbol{\Sigma}^{-1}$ as the first factor, that is,

$$\boldsymbol{w}_i = \begin{pmatrix} z_{ij} \\ \boldsymbol{w}_{i(-j)} \end{pmatrix},\ \boldsymbol{\beta} = \begin{pmatrix} \beta_{1j} \\ \boldsymbol{\beta}_{(-1j)} \end{pmatrix}$$
$$\boldsymbol{\Sigma} = \begin{pmatrix} \sigma_{1j} & \boldsymbol{\sigma}'_{(-1j)} \\ \boldsymbol{\sigma}_{(-1j)} & \boldsymbol{\Sigma}_{(-1j)} \end{pmatrix},\ \boldsymbol{\Sigma}^{-1} = \begin{pmatrix} \sigma^{1j} & \boldsymbol{\sigma}^{(-1j)\prime} \\ \boldsymbol{\sigma}^{(-1j)} & \boldsymbol{\Sigma}^{(-1j)}. \end{pmatrix}$$

Since $\boldsymbol{w}_i|\boldsymbol{\beta}, \boldsymbol{\Sigma}, \boldsymbol{\gamma} \sim \mathrm{N}(\boldsymbol{\beta}, \boldsymbol{\Sigma})$, from the property of the multivariate normal distribution we have

$$z_{ij}|\boldsymbol{\gamma}_j, \boldsymbol{\beta}, \boldsymbol{\Sigma}, \boldsymbol{w}_{(-j)}, \boldsymbol{y}_1 \sim \mathrm{N}(\tilde{\mu}_{ij}, \tilde{\sigma}_j)1_{(z_{ij}\in\mathcal{G}_{ij})},\ i = 1, \cdots, n;\ j = 1, \cdots, m, \tag{16}$$

where

$$\tilde{\mu}_{ij} = \beta_{1j} + \boldsymbol{\sigma}'_{(-1j)}\boldsymbol{\Sigma}^{-1}_{(-1j)}\left(\boldsymbol{w}_{i(-j)} - \boldsymbol{\beta}_{(-1j)}\right)$$

$$\tilde{\sigma}_j = \sigma_{1j} - \boldsymbol{\sigma}'_{(-1j)}\boldsymbol{\Sigma}^{-1}_{(-1j)}\boldsymbol{\sigma}_{(-1j)}.$$

The distribution of z_{ij} is a truncated normal distribution. We can utilize the method for sampling truncated normal variables proposed by Damien and Walker (2001).

Since $z_{1j}, z_{2j}, \cdots, z_{nj}$ are independent given $\boldsymbol{\gamma}_j, \boldsymbol{\beta}, \boldsymbol{\Sigma}$, we have

$$\begin{aligned}p(\boldsymbol{\gamma}_j|\boldsymbol{\beta}, \boldsymbol{\Sigma}, \boldsymbol{w}_{(-j)}, \boldsymbol{y}_1) \propto p(\boldsymbol{\delta}_j(\boldsymbol{\gamma}_j)) &\prod_{i:y_{ij}=2}\left[\Phi\left(\frac{\gamma_{j2}-\tilde{\mu}_{ij}}{\tilde{\sigma}_j}\right) - \Phi\left(-\frac{\tilde{\mu}_{ij}}{\tilde{\sigma}_j}\right)\right] \\ \times \prod_{i:y_{ij}=3}&\left[\Phi\left(\frac{\gamma_{j3}-\tilde{\mu}_{ij}}{\tilde{\sigma}_j}\right) - \Phi\left(\frac{\gamma_{j2}-\tilde{\mu}_{ij}}{\tilde{\sigma}_j}\right)\right] \\ \times \cdots \times \prod_{i:y_{ij}=C_j-1}&\left[\Phi\left(\frac{1-\tilde{\mu}_{ij}}{\tilde{\sigma}_j}\right) - \Phi\left(\frac{\gamma_{j(C_j-2)}-\tilde{\mu}_{ij}}{\tilde{\sigma}_j}\right)\right],\end{aligned}$$

where $\Phi(\cdot)$ is the distribution function of the standard normal distribution. Thus, the conditional distribution of $\boldsymbol{\delta}_j$ is

$$p(\boldsymbol{\delta}_j|\boldsymbol{\beta}, \boldsymbol{\Sigma}, \boldsymbol{w}_{(-j)}, \boldsymbol{y}_1) \propto p(\boldsymbol{\gamma}_j|\boldsymbol{\beta}, \boldsymbol{\Sigma}, \boldsymbol{w}_{(-j)}, \boldsymbol{y}_1)\prod_{c=2}^{C_j-2}\frac{\left(1-\gamma_{j(c-1)}\right)\exp(\zeta_{jc})}{\left(1+\exp(\zeta_{jc})\right)^2}. \tag{17}$$

We use a multivariate t distribution, $\mathrm{Mt}(\boldsymbol{\delta}_j|\tilde{\boldsymbol{\delta}}_j, \tilde{\boldsymbol{\Sigma}}_{\boldsymbol{\delta}_j}, \nu)$, as a proposal distribution for generating $\boldsymbol{\delta}_j$, where $\tilde{\boldsymbol{\delta}}_j$ is the mode of (17),

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

and ν is the degrees of freedom. The Metropolis-Hastings (M-H) algorithm for generating $\boldsymbol{\delta}_j$ is as follows:

1. Let $\boldsymbol{\delta}_j^{(t)}$ denote the value of $\boldsymbol{\delta}_j$ at the tth iteration.
2. At the $(t+1)$th iteration, sample $\boldsymbol{\delta}_j^p$ from $\mathrm{Mt}(\boldsymbol{\delta}_j|\tilde{\boldsymbol{\delta}}_j, \tilde{\boldsymbol{\Sigma}}_{\boldsymbol{\delta}_j}, \nu)$.
3. The transition probability from $\boldsymbol{\delta}_j^{(t)}$ to $\boldsymbol{\delta}_j^p$ is
$$\alpha = \min\left\{\frac{p(\boldsymbol{\delta}_j^p|\cdots)\,\mathrm{Mt}(\boldsymbol{\delta}_j^{(t)}|\tilde{\boldsymbol{\delta}}_j, \tilde{\boldsymbol{\Sigma}}_{\boldsymbol{\delta}_j}, \nu)}{p(\boldsymbol{\delta}_j^{(t)}|\cdots)\,\mathrm{Mt}(\boldsymbol{\delta}_j^p|\tilde{\boldsymbol{\delta}}_j, \tilde{\boldsymbol{\Sigma}}_{\boldsymbol{\delta}_j}, \nu)}, 1\right\}.$$
4. Generate $u \sim \mathrm{U}(0,1)$, the uniform distribution on $(0,1)$, and take
$$\boldsymbol{\delta}_j^{(t+1)} = \begin{cases}\boldsymbol{\delta}_j^p & \text{if } u < \alpha \\ \boldsymbol{\delta}_j^{(t)} & \text{otherwise.}\end{cases}$$

We can obtain $\boldsymbol{\gamma}_j$ from $\boldsymbol{\delta}_j$ by using the equation

$$\gamma_{jc} = \frac{\gamma_{j(c-1)} + \exp(\delta_{jc})}{1+\exp(\delta_{jc})},\ c = 2, \cdots, C_j - 2.$$

Acknowledgments

The author appreciates the comments of an anonymous referee and an anonymous associate editor, which improve the article greatly. Further, the author is grateful to the Institute for Research on Household Economics for providing the micro-data of the Japanese Panel Survey of Consumers. This work was supported in part by Grants-in-Aid for Scientific Research (No.22530203) from the JSPS.

References

Albert, J. H. (1992). Bayesian estimation of the polychoric correlation coefficient. *Journal of Statistical Computation and Simulation 44*(1–2), 47–61.

Brooks, S. P. and A. Gelman (1998). General methods for monitoring convergence of iterative simulations. *Journal of Computational and Graphical Statistics 7*(4), 434–455.

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.

Chib, S. and E. Greenberg (1998). Analysis of multivariate probit models. *Biometrika 85*(2), 347–361.

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. (2009). *An Object-oriented Matrix Programming Language Ox^{TM} 6*. London: Timberlake Consultants Ltd.

Gelman, A. and J. Hill (2007). *Data Analysis Using Regression and Multilevel/Hierarchical Models*. New York: Cambridge University Press.

Gelman, A. and D. B. Rubin (1992). Inference from iterative simulation using multiple sequences. *Statistical Science 7*(4), 457–472.

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.

Heidelberger, P. and P. D. Welch (1983). Run length control in the presence of an initial transient. *Operations Research 31*(6), 1109–1144.

Jackman, S. (2009). *Bayesian Analysis for the Social Sciences*. Chichester: Wiley.

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.

Kottas, A., P. Müller, and F. Quintana (2005). Nonparametric Bayesian modeling for multivariate ordinal data. *Journal of Computational and Graphical Statistics 14*(3), 610–625.

Poon, W.-Y. and S.-Y. Lee (1987). Maximum likelihood estimation of multivariate polyserial and polychoric correlation coefficients. *Psychometrika 52*(3), 409–430.

Zhang, X., W. J. Boscardin, and T. R. Belin (2006). Sampling correlation matrices in Bayesian models with correlated latent variables. *Journal of Computational and Graphical Statistics 15*(4), 880–896.

Table 1 (a): Summary statistics (ordinal discrete variables)

Variables\Choices	1	2	3	4
$\boldsymbol{y}_{11}$	10	45	26	19
$\boldsymbol{y}_{12}$	14	40	30	16
$\boldsymbol{y}_{13}$	20	39	22	19

Table 1 (b): Summary statistics (continuous variable)[a]

Varible	Mean	SD	Min	25%	50%	75%	Max
$\boldsymbol{y}_2$	1.882	1.043	-1.063	1.328	1.835	2.513	4.177

a: "Mean" and "SD" denote the sample mean and sample standard deviation, respectively. "Min," "25%," "50%," "75%," and "Max" denote the minimum value, 25%, 50%, 75% quantile values and maximum value, respectively.

Table 2: Posterior results (simulated data)[a]

	Mean	SD	Median	Geweke	HW	GR
β_{11}	0.5899	0.0527	0.5901*	0.210	0.197	1.00002
β_{12}	0.5180	0.0541	0.5178*	0.691	0.251	1.00000
β_{13}	0.4881	0.0640	0.4880*	0.900	0.832	1.00010
β_{14}	1.8806	0.1046	1.8796*	0.706	0.618	1.00001
γ_{12}	0.6426	0.0507	0.6435*	0.897	0.543	1.00012
γ_{22}	0.5631	0.0535	0.5629*	0.523	0.846	0.99998
γ_{32}	0.6172	0.0574	0.6181*	0.540	0.795	0.99998
σ_{11}	0.2102	0.0402	0.2057*	0.660	0.526	1.00001
σ_{21}	0.0806	0.0264	0.0787*	0.104	0.497	1.00004
σ_{31}	0.0756	0.0313	0.0740*	0.122	0.316	1.00017
σ_{41}	0.1136	0.0530	0.1109*	0.751	0.173	1.00007
σ_{22}	0.2245	0.0423	0.2195*	0.342	0.474	1.00004
σ_{32}	0.1176	0.0337	0.1146*	0.064	0.277	1.00000
σ_{42}	0.1382	0.0556	0.1349*	0.222	0.212	1.00008
σ_{33}	0.3254	0.0659	0.3178*	0.081	0.344	1.00008
σ_{43}	0.2856	0.0738	0.2794*	0.087	0.233	1.00006
σ_{44}	1.0891	0.1553	1.0749*	0.244	0.289	1.00010
r^s_{21}	0.3710	0.0972	0.3747*	0.183	0.761	1.00004
r^s_{31}	0.2893	0.1039	0.2930*	0.240	0.677	1.00013
r^s_{41}	0.2365	0.0995	0.2386*	0.830	0.173	1.00003
r^s_{32}	0.4356	0.0924	0.4395*	0.153	0.364	1.00003
r^s_{42}	0.2787	0.0980	0.2810*	0.352	0.190	1.00001
r^s_{43}	0.4787	0.0837	0.4828*	0.269	0.644	1.00002
r^p_{21}	0.2773	0.1065	0.2798*	0.395	0.613	1.00005
r^p_{31}	0.1007	0.1177	0.1029	0.603	0.519	1.00020
r^p_{41}	0.0952	0.1079	0.0967	0.571	0.083	1.00008
r^p_{32}	0.3151	0.1052	0.3191*	0.690	0.635	1.00002
r^p_{42}	0.0566	0.1102	0.0581	0.822	0.287	0.99998
r^p_{43}	0.4001	0.0927	0.4033*	0.558	0.617	1.00002

a: "Mean," "SD," and "Median" denote the posterior mean, posterior standard deviation, and posterior median, respectively. The "$*$" in Median denotes that zero is not included in the 95% credible interval. "Geweke" and "HW" denote the p-values for the convergence diagnostic statistics proposed by Geweke (1992) and Heidelberger and Welch (1983), respectively. "GR" denotes the convergence diagnostic statistic proposed by Gelman and Rubin (1992) and modified by Brooks and Gelman (1998).

Table 3: Posterior results (medians of coefficient parameters, simulated data)^a

	True	$\boldsymbol{Q}_0 = 0.02\boldsymbol{I}$	$\boldsymbol{Q}_0 = \boldsymbol{I}$	$\boldsymbol{Q}_0 = 5\boldsymbol{I}$	$\boldsymbol{Q}_0 = 10\boldsymbol{I}$
β_{11}	0.5000	0.5901*	0.5902*	0.5921*	0.5953*
β_{12}	0.5000	0.5178*	0.5180*	0.5187*	0.5202*
β_{13}	0.5000	0.4880*	0.4878*	0.4869*	0.4866*
β_{14}	2.0000	1.8796*	1.8797*	1.8799*	1.8798*
γ_{12}	0.6000	0.6435*	0.6429*	0.6416*	0.6406*
γ_{22}	0.6000	0.5629*	0.5635*	0.5652*	0.5664*
γ_{32}	0.6000	0.6181*	0.6184*	0.6206*	0.6230*
σ_{11}	0.2500	0.2057*	0.2239*	0.2941*	0.3772*
σ_{21}	0.1400	0.0787*	0.0798*	0.0850*	0.0915*
σ_{31}	0.0560	0.0740*	0.0756*	0.0823*	0.0908
σ_{41}	0.0672	0.1109*	0.1126*	0.1188*	0.1250
σ_{22}	0.2500	0.2195*	0.2383*	0.3097*	0.3946*
σ_{32}	0.0800	0.1146*	0.1164*	0.1248*	0.1349*
σ_{42}	0.0960	0.1349*	0.1367*	0.1421*	0.1483*
σ_{33}	0.2500	0.3178*	0.3417*	0.4332*	0.5397*
σ_{43}	0.1200	0.2794*	0.2841*	0.3012*	0.3187*
σ_{44}	1.0000	1.0749*	1.0844*	1.1239*	1.1735*
r^s_{21}	0.5600	0.3747*	0.3501*	0.2854*	0.2407*
r^s_{31}	0.2240	0.2930*	0.2769*	0.2341*	0.2041
r^s_{41}	0.1344	0.2386*	0.2312*	0.2094*	0.1899
r^s_{32}	0.3200	0.4395*	0.4139*	0.3449*	0.2957*
r^s_{42}	0.1920	0.2810*	0.2720*	0.2440*	0.2202*
r^s_{43}	0.2400	0.4828*	0.4716*	0.4359*	0.4045*
r^p_{21}	0.5242	0.2798*	0.2613*	0.2142*	0.1819
r^p_{31}	0.0517	0.1029	0.1046	0.1029	0.0989
r^p_{41}	0.0226	0.0967	0.0982	0.1004	0.1000
r^p_{32}	0.2174	0.3191*	0.2979*	0.2451*	0.2098*
r^p_{42}	0.0948	0.0581	0.0676	0.0872	0.0959
r^p_{43}	0.1906	0.4033*	0.3955*	0.3726*	0.3507*

a: "True" denotes the true values. The "$*$" denotes that zero is not included in the 95% credible interval.

Table 4 (a): Summary statistics (ordinal discrete variables)[a]

	1 (Very Poor)	2 (Poor)	3 (Average)	4 (Good)	5 (Very Good)
Happiness	9	45	118	274	69
(y_1)	(1.75)	(8.74)	(22.91)	(53.20)	(13.40)
Life_Sat	25	69	164	216	41
(y_2)	(4.85)	(13.40)	(31.84)	(41.94)	(7.96)
Liv_Std	46	153	248	68	
(y_3)	(8.93)	(29.71)	(48.16)	(13.20)	
Health_Cond	3	86	188	180	58
(y_6)	(0.58)	(16.70)	(36.50)	(34.95)	(11.26)
	1 (Very Poor)	2 (Poor)	3 (Good)	4 (Very Good)	
Inc_Sat	91	224	186	14	
(y_4)	(17.67)	(43.50)	(36.12)	(2.72)	
Cons_Sat	49	208	249	9	
(y_5)	(9.51)	(40.39)	(48.35)	(1.75)	

a: Values in parentheses denote percentage.

Table 4 (b): Summary statistics (annual earnings, 10,000 yen)[a]

Mean	SD	Min	25%	50%	75%	Max
264.55	189.03	0	158	240	348.5	2800

a: "Mean" and "SD" denote the sample mean and sample standard deviation, respectively. "Min," "25%," "50%," "75%," and "Max" denote the minimum value, 25%, 50%, 75% quantile values and maximum value, respectively.

Table 5: Posterior results of parameters[a]

	Mean	SD	Median	Geweke	HW	GR
β_{11}	0.6562	0.0183	0.6563*	0.678	0.839	1.00002
β_{12}	0.5468	0.0182	0.5469*	0.857	0.959	1.00000
β_{13}	0.5465	0.0210	0.5463*	0.114	0.788	0.99998
β_{14}	0.3348	0.0195	0.3347*	0.964	0.854	1.00000
β_{15}	0.3947	0.0185	0.3947*	0.846	0.665	0.99999
β_{16}	0.6740	0.0189	0.6740*	0.523	0.772	0.99997
β_{21}	5.1987	0.0575	5.1978*	0.853	0.222	0.99999
γ_{12}	0.2601	0.0270	0.2601*	0.757	0.429	1.00001
γ_{13}	0.5194	0.0217	0.5193*	0.524	0.799	1.00001
γ_{22}	0.2482	0.0208	0.2480*	0.161	0.535	1.00000
γ_{23}	0.5494	0.0188	0.5497*	0.751	0.905	0.99999
γ_{32}	0.4307	0.0221	0.4307*	0.105	0.219	1.00001
γ_{42}	0.4449	0.0212	0.4448*	0.956	0.782	0.99999
γ_{52}	0.3997	0.0199	0.3994*	0.235	0.614	0.99999
γ_{62}	0.4205	0.0284	0.4207*	0.444	0.933	0.99999
γ_{63}	0.6983	0.0194	0.6985*	0.421	0.718	1.00001

a: "Mean," "SD," and "Median" denote the posterior mean, posterior standard deviation, and posterior median, respectively. The "$*$" in Median denotes that zero is not included in the 95% credible interval. "Geweke" and "HW" denote the p-values for the convergence diagnostic statistics proposed by Geweke (1992) and Heidelberger and Welch (1983), respectively. "GR" denotes the convergence diagnostic statistic proposed by Gelman and Rubin (1992) and modified by Brooks and Gelman (1998).

Table 5: Continued

	Mean	SD	Median	Geweke	HW	GR
σ_{11}	0.0996	0.0085	0.0992*	0.246	0.779	0.99998
σ_{21}	0.0718	0.0064	0.0715*	0.939	0.928	0.99997
σ_{31}	0.0513	0.0070	0.0511*	0.745	0.192	1.00002
σ_{41}	0.0510	0.0065	0.0509*	0.847	0.497	1.00000
σ_{51}	0.0401	0.0054	0.0399*	0.981	0.521	1.00001
σ_{61}	0.0428	0.0051	0.0426*	0.855	0.498	1.00000
σ_{71}	0.0264	0.0199	0.0262	0.953	0.132	1.00000
σ_{22}	0.1094	0.0084	0.1091*	0.125	0.355	0.99999
σ_{32}	0.0696	0.0076	0.0693*	0.178	0.778	1.00009
σ_{42}	0.0670	0.0070	0.0667*	0.113	0.262	1.00001
σ_{52}	0.0500	0.0057	0.0498*	0.189	0.165	1.00000
σ_{62}	0.0523	0.0054	0.0521*	0.379	0.692	1.00007
σ_{72}	0.0554	0.0204	0.0552*	0.865	0.380	1.00004
σ_{33}	0.1644	0.0129	0.1638*	0.820	0.907	1.00000
σ_{43}	0.0835	0.0086	0.0831*	0.126	0.636	1.00000
σ_{53}	0.0579	0.0069	0.0576*	0.609	0.981	1.00007
σ_{63}	0.0390	0.0059	0.0389*	0.308	0.990	1.00005
σ_{73}	0.1196	0.0258	0.1192*	0.190	0.529	0.99997
σ_{44}	0.1319	0.0112	0.1315*	0.308	0.879	0.99998
σ_{54}	0.0660	0.0067	0.0657*	0.456	0.182	1.00000
σ_{64}	0.0283	0.0052	0.0281*	0.639	0.850	0.99998
σ_{74}	0.0282	0.0227	0.0282	0.780	0.913	0.99999
σ_{55}	0.0890	0.0076	0.0886*	0.846	0.214	0.99999
σ_{65}	0.0273	0.0044	0.0272*	0.303	0.977	1.00002
σ_{75}	0.0300	0.0191	0.0299	0.742	0.866	0.99998
σ_{66}	0.0724	0.0072	0.0721*	0.246	0.474	1.00005
σ_{76}	0.0512	0.0167	0.0508*	0.235	0.172	1.00011
σ_{77}	1.6815	0.1056	1.6759*	0.998	0.574	0.99999

Table 6: Posterior results of polyserial and polychoric (simple) correlation coefficients[a]

	Mean	SD	Median	Geweke	HW	GR
r^s_{21}	0.6876	0.0281	0.6888*	0.988	0.142	0.99998
r^s_{31}	0.4008	0.0434	0.4015*	0.815	0.081	1.00004
r^s_{41}	0.4448	0.0421	0.4461*	0.676	0.219	1.00000
r^s_{51}	0.4256	0.0443	0.4262*	0.487	0.411	1.00003
r^s_{61}	0.5040	0.0393	0.5050*	0.999	0.786	1.00002
r^s_{71}	0.0646	0.0482	0.0647	0.996	0.160	1.00000
r^s_{32}	0.5186	0.0380	0.5194*	0.207	0.906	1.00011
r^s_{42}	0.5574	0.0363	0.5585*	0.317	0.258	1.00000
r^s_{52}	0.5069	0.0394	0.5082*	0.249	0.350	0.99999
r^s_{62}	0.5882	0.0338	0.5888*	0.487	0.733	1.00003
r^s_{72}	0.1292	0.0463	0.1296*	0.756	0.382	1.00002
r^s_{43}	0.5666	0.0362	0.5675*	0.092	0.228	1.00003
r^s_{53}	0.4790	0.0413	0.4796*	0.616	0.778	1.00011
r^s_{63}	0.3576	0.0442	0.3584*	0.066	0.737	1.00003
r^s_{73}	0.2272	0.0452	0.2278*	0.166	0.462	0.99998
r^s_{54}	0.6092	0.0352	0.6100*	0.575	0.137	0.99998
r^s_{64}	0.2892	0.0465	0.2894*	0.220	0.738	1.00000
r^s_{74}	0.0598	0.0477	0.0602	0.864	0.881	0.99999
r^s_{65}	0.3398	0.0460	0.3407*	0.072	0.466	1.00001
r^s_{75}	0.0776	0.0487	0.0779	0.726	0.835	0.99999
r^s_{76}	0.1467	0.0461	0.1469*	0.108	0.088	1.00010

a: "Mean," "SD," and "Median" denote the posterior mean, posterior standard deviation, and posterior median, respectively. The "$*$" in Median denotes that zero is not included in the 95% credible interval. "Geweke" and "HW" denote the p-values for the convergence diagnostic statistics proposed by Geweke (1992) and Heidelberger and Welch (1983), respectively. "GR" denotes the convergence diagnostic statistic proposed by Gelman and Rubin (1992) and modified by Brooks and Gelman (1998).

Table 7: Posterior results of polyserial and polychoric partial correlation coefficients[a]

	Mean	SD	Median	Geweke	HW	GR
r^p_{21}	0.4515	0.0450	0.4531*	0.395	0.189	1.00003
r^p_{31}	0.0166	0.0576	0.0164	0.567	0.560	1.00000
r^p_{41}	0.0589	0.0602	0.0587	0.870	0.984	1.00000
r^p_{51}	0.0685	0.0603	0.0686	0.716	0.710	1.00002
r^p_{61}	0.1714	0.0552	0.1721*	0.531	0.115	0.99998
r^p_{71}	−0.0526	0.0536	−0.0534	0.524	0.077	1.00005
r^p_{32}	0.1463	0.0570	0.1466*	0.137	0.483	1.00000
r^p_{42}	0.2239	0.0571	0.2250*	0.733	0.966	0.99998
r^p_{52}	0.0936	0.0589	0.0937	0.134	0.379	0.99998
r^p_{62}	0.3461	0.0498	0.3464*	0.710	0.837	1.00002
r^p_{72}	0.0336	0.0541	0.0332	0.943	0.605	1.00004
r^p_{43}	0.3175	0.0524	0.3179*	0.182	0.593	0.99998
r^p_{53}	0.1290	0.0573	0.1292*	0.093	0.293	1.00016
r^p_{63}	0.0749	0.0551	0.0753	0.212	0.896	1.00000
r^p_{73}	0.2048	0.0487	0.2056*	0.244	0.583	1.00000
r^p_{54}	0.3870	0.0506	0.3885*	0.863	0.382	1.00004
r^p_{64}	−0.1231	0.0569	−0.1237*	0.833	0.957	1.00008
r^p_{74}	−0.0765	0.0541	−0.0766	0.633	0.545	0.99998
r^p_{65}	0.0706	0.0576	0.0707	0.063	0.397	1.00004
r^p_{75}	−0.0083	0.0541	−0.0082	0.869	0.954	1.00000
r^p_{76}	0.0754	0.0500	0.0755	0.218	0.534	1.00005

a: "Mean," "SD," and "Median" denote the posterior mean, posterior standard deviation, and posterior median, respectively. The "$*$" in Median denotes that zero is not included in the 95% credible interval. "Geweke" and "HW" denote the p-values for the convergence diagnostic statistics proposed by Geweke (1992) and Heidelberger and Welch (1983), respectively. "GR" denotes the convergence diagnostic statistic proposed by Gelman and Rubin (1992) and modified by Brooks and Gelman (1998).

Table 8: Posterior results (medians of coefficient parameters)[a]

	$\boldsymbol{Q}_0 = 0.02\boldsymbol{I}$	$\boldsymbol{Q}_0 = \boldsymbol{I}$	$\boldsymbol{Q}_0 = 5\boldsymbol{I}$	$\boldsymbol{Q}_0 = 10\boldsymbol{I}$
β_{11}	0.6563*	0.6521*	0.6421*	0.6358*
β_{12}	0.5469*	0.5456*	0.5427*	0.5406*
β_{13}	0.5463*	0.5464*	0.5463*	0.5461*
β_{14}	0.3347*	0.3363*	0.3406*	0.3432*
β_{15}	0.3947*	0.3984*	0.4068*	0.4121*
β_{16}	0.6740*	0.6625*	0.6415*	0.6302*
β_{21}	5.1978*	5.1981*	5.1980*	5.1980*
γ_{12}	0.2601*	0.2474*	0.2172*	0.1967*
γ_{13}	0.5193*	0.5120*	0.4929*	0.4788*
γ_{22}	0.2480*	0.2419*	0.2271*	0.2161*
γ_{23}	0.5497*	0.5482*	0.5455*	0.5431*
γ_{32}	0.4307*	0.4300*	0.4271*	0.4239*
γ_{42}	0.4448*	0.4484*	0.4593*	0.4680*
γ_{52}	0.3994*	0.4036*	0.4137*	0.4205*
γ_{62}	0.4207*	0.3969*	0.3490*	0.3202*
γ_{63}	0.6985*	0.6873*	0.6671*	0.6559*
σ_{11}	0.0992*	0.1044*	0.1217*	0.1395*
σ_{21}	0.0715*	0.0725*	0.0752*	0.0778*
σ_{31}	0.0511*	0.0517*	0.0535*	0.0549*
σ_{41}	0.0509*	0.0516*	0.0538*	0.0557*
σ_{51}	0.0399*	0.0406*	0.0423*	0.0436*
σ_{61}	0.0426*	0.0445*	0.0486*	0.0513*
σ_{71}	0.0262	0.0268	0.0273	0.0278
σ_{22}	0.1091*	0.1129*	0.1273*	0.1433*
σ_{32}	0.0693*	0.0695*	0.0708*	0.0721*
σ_{42}	0.0667*	0.0672*	0.0691*	0.0710*
σ_{52}	0.0498*	0.0504*	0.0518*	0.0531*
σ_{62}	0.0521*	0.0539*	0.0579*	0.0608*
σ_{72}	0.0552*	0.0553*	0.0561*	0.0566*
σ_{33}	0.1638*	0.1674*	0.1817*	0.1986*
σ_{43}	0.0831*	0.0836*	0.0851*	0.0868*
σ_{53}	0.0576*	0.0582*	0.0593*	0.0604*
σ_{63}	0.0389*	0.0401*	0.0427*	0.0444*
σ_{73}	0.1192*	0.1195*	0.1207*	0.1223*
σ_{44}	0.1315*	0.1360*	0.1525*	0.1709*
σ_{54}	0.0657*	0.0663*	0.0682*	0.0702*
σ_{64}	0.0281*	0.0292*	0.0317*	0.0334*
σ_{74}	0.0282	0.0282	0.0293	0.0303
σ_{55}	0.0886*	0.0932*	0.1088*	0.1254*
σ_{65}	0.0272*	0.0283*	0.0305*	0.0319*
σ_{75}	0.0299	0.0304	0.0316	0.0323
σ_{66}	0.0721*	0.0796*	0.0999*	0.1183*
σ_{76}	0.0508*	0.0522*	0.0546*	0.0558*
σ_{77}	1.6759*	1.6805*	1.6883*	1.6980*

a: The "$*$" denotes that zero is not included in the 95% credible interval.

Table 8: Continued

	$Q_0 = 0.02I$	$Q_0 = I$	$Q_0 = 5I$	$Q_0 = 10I$
r^s_{21}	0.6888*	0.6682*	0.6056*	0.5515*
r^s_{31}	0.4015*	0.3921*	0.3606*	0.3306*
r^s_{41}	0.4461*	0.4340*	0.3963*	0.3621*
r^s_{51}	0.4262*	0.4126*	0.3688*	0.3311*
r^s_{61}	0.5050*	0.4888*	0.4416*	0.4007*
r^s_{71}	0.0647	0.0642	0.0607	0.0573
r^s_{32}	0.5194*	0.5071*	0.4666*	0.4287*
r^s_{42}	0.5585*	0.5442*	0.4972*	0.4550*
r^s_{52}	0.5082*	0.4925*	0.4416*	0.3973*
r^s_{62}	0.5888*	0.5696*	0.5144*	0.4676*
r^s_{72}	0.1296*	0.1275*	0.1213*	0.1150*
r^s_{43}	0.5675*	0.5551*	0.5120*	0.4720*
r^s_{53}	0.4796*	0.4667*	0.4227*	0.3836*
r^s_{63}	0.3584*	0.3483*	0.3177*	0.2902*
r^s_{73}	0.2278*	0.2261*	0.2187*	0.2111*
r^s_{54}	0.6100*	0.5908*	0.5311*	0.4808*
r^s_{64}	0.2894*	0.2820*	0.2578*	0.2352*
r^s_{74}	0.0602	0.0590	0.0578	0.0564
r^s_{65}	0.3407*	0.3289*	0.2930*	0.2627*
r^s_{75}	0.0779	0.0772	0.0740	0.0703
r^s_{76}	0.1469*	0.1435*	0.1335*	0.1250*
r^p_{21}	0.4531*	0.4346*	0.3857*	0.3496*
r^p_{31}	0.0164	0.0238	0.0425	0.0541
r^p_{41}	0.0587	0.0667	0.0807	0.0881
r^p_{51}	0.0686	0.0698	0.0739	0.0759
r^p_{61}	0.1721*	0.1771*	0.1829*	0.1805*
r^p_{71}	−0.0534	−0.0490	−0.0419	−0.0356
r^p_{32}	0.1466*	0.1475*	0.1490*	0.1479*
r^p_{42}	0.2250*	0.2172*	0.2004*	0.1894*
r^p_{52}	0.0937	0.0983	0.1053*	0.1073*
r^p_{62}	0.3464*	0.3325*	0.2984*	0.2732*
r^p_{72}	0.0332	0.0324	0.0315	0.0310
r^p_{43}	0.3179*	0.3140*	0.2961*	0.2789*
r^p_{53}	0.1292*	0.1323*	0.1362*	0.1361*
r^p_{63}	0.0753	0.0737	0.0725	0.0720
r^p_{73}	0.2056*	0.2030*	0.1930*	0.1840*
r^p_{54}	0.3885*	0.3717*	0.3293*	0.2986*
r^p_{64}	−0.1237*	−0.1087	−0.0703	−0.0448
r^p_{74}	−0.0766	−0.0731	−0.0621	−0.0526
r^p_{65}	0.0707	0.0688	0.0633	0.0610
r^p_{75}	−0.0082	−0.0084	−0.0062	−0.0040
r^p_{76}	0.0755	0.0747	0.0710	0.0677