Self-Assessed Social Position and Poverty*

HIKARU HASEGAWA
Graduate School of Economics and Business Administration
Hokkaido University

KAZUHIRO UEDA†
Faculty of Economics
Nihon Fukushi University

February 9, 2016

Abstract

In this article, we used microlevel data extracted from the 2006 Japanese General Social Surveys to analyze the relationships between self-assessed social position and socioeconomic factors such as income and poverty. We provided the posterior results of the estimation of the Bayesian multivariate ordered probit model and proposed an inequality measure for self-assessed social position on the basis of the posterior results. We state the inequality measure "regret" and show that the distributions of regret differ for people above and below the poverty line.

KEYWORDS: Markov chain Monte Carlo (MCMC), multivariate ordered probit model, ordinal explanatory variables

JEL CLASSIFICATION : C11, C35, I3

* *Notes*: The authors appreciate the comments of Professor Klasen, the editor of this journal, and two anonymous referees, which improve the article greatly. The work of the first author was supported in part by a Grant-in-Aid for Scientific Research (No.22530203) from the Japan Society for the Promotion of Science (JSPS), while that of the first and second authors was supported in part by a Grant-in-Aid for Scientific Research (No.21530203) from the JSPS. Further, The Japanese General Social Surveys (JGSS) are designed and carried out by the JGSS Research Center at Osaka University of Commerce (Joint Usage / Research Center for Japanese General Social Surveys accredited by Minister of Education, Culture, Sports, Science and Technology), in collaboration with the Institute of Social Science at the University of Tokyo.

†Corresponding to: Kazuhiro Ueda, Faculty of Economics, Nihon Fukushi University, 4703295, Aichi, Japan (ueda@n-fukushi.ac.jp).

1 INTRODUCTION

As the standard of living improved because of the rapid postwar development of the Japanese economy, a middle-class consciousness prevailed in Japan. However, economic inequality and poverty have emerged as social issues, as the Japanese economy has been sluggish and the aging population has increased in the 20 years following the bursting of the bubble economy in the early 1990s. The factors that affect people's perception of social stratification may also be influenced by the response to these changes in socioeconomic conditions.[1]

People's perception of social stratification is related to their subjective well-being and life satisfaction, and a change therein has a visible effect on economic and public policy. For example, when people perceive that their social stratification lowers relative to others or when the distribution of their perception is biased, their frustration against society and the government increases. Thus, national and local political situations can be sensitive to these people's subjective frustration, which can lead to social uncertainty in some situations. In light of these problems, the method to investigate how people precisely perceive to which stratification they belong should be developed for adopting careful policies. Thus, analyzing the determinants of a person's perception of his or her social position, hereafter called "self-assessed social position," is important, as is the analysis of subjective well-being and life satisfaction.

While there have been many studies on self-assessed social position in sociology, there are few studies on the subject in economics. Therefore, the development of statistical methods for analyzing self-assessed social position is required for economics. This problem will be addressed by the methods used in the studies on subjective well-being, which have been under intense investigation in economics.

Subjective well-being is considered to be directly related to income. However, the seminal work of Easterlin (1974) shows that economic growth and subjective well-being are not always positively correlated; subsequently, many studies on issues pertaining to this relationship have been conducted (Blanchflower and Oswald, 2004; Layard, 2005; Di Tella and MacCulloch, 2006; Stevenson and Wolfers, 2008; Di Tella et al., 2010). Our article addresses this issue by focusing on people's perception of social position.

Many studies deal with Easterlin's paradox and the relationship between subjective well-being or life satisfaction and socioeconomic or demographic factors. These studies not only verify Easterlin's paradox but also address other issues such as the availability of subjective data on topics including subjective well-being and life satisfaction in economic analysis and the econometric analysis of ordinal data.

Using panel survey data in Russia, Ravallion and Lokshin (2001) narrow the broad concept of subjective welfare to economic welfare and analyze the relationship between subjective welfare and income. In addition, using the ordered probit method, they investigate how economic welfare is influenced by not only income but also by individual characteristics such as gender, age, marital status, education, and health. Ravallion and Lokshin (2002) analyze the relationship between subjective and objective economic welfare by using the ratio of

[1] See Ishida (2001) for a sociological analysis on class mobility in postwar Japan. See Shirahase (2010) regarding the change in Japanese class identification.

total household income to the poverty line as an indicator of objective economic welfare. They point out that people who perceive themselves as poor are not statistically classified as poor in Russia and that a discrepancy exists between subjective welfare and income class.

Blanchflower and Oswald (2004) define a reported well-being function that can be used to analyze subjective ordinal data. This function relates a person's self-reported well-being to the person's true well-being or utility function, which is a function of income and demographic and personal factors. They use an ordered logit model to estimate the function. Di Tella and MacCulloch (2006) estimate subjective well-being in Europe and the United States using an ordered probit model in which the explanatory variables are not only personal characteristics–including personal income position, employment status, gender, education, and marital status–but also macro variables–such as GDP per capita, life expectancy, and unemployment rate.

Although the adaptation of subjective well-being to a change in income has been proposed as an explanation of Easterlin's paradox, Di Tella et al. (2010) use panel data to estimate the adaptation of happiness not only to income but also to status; in this article, happiness is life satisfaction and status is the relative standing of one's job. They show that adaptation to income and status differs across sub-groups based on gender, political stance, and employment. Using panel data, Luttmer (2005) analyzes the relationship between a person's reported well-being and his or her neighbor's earnings and shows that subjective well-being is affected by the person's relative position.

Analyzing reported happiness data and availability, Di Tella and MacCulloch (2006) find that comparing the happiness of two persons is problematic because of individual differences in the perception of the amount of happiness obtained from consuming various goods. However, the problems of comparing happiness are substantially reduced when analyzing groups rather than individuals (Di Tella and MacCulloch, 2006, p.29).

Furthermore, Stutzer and Frey (2010) review recent advances of studies on subjective well-being in economics and directions of development. They indicate that many open issues for positive analysis exist in economics and that new insights from these studies can stimulate and expand the debate on happiness.

Using Japanese survey data, the present article aims to empirically investigate the effect of income and poverty on self-assessed social position and the relationship between self-assessed social position and individual characteristics such as gender, age, and education, as in previous studies. We use microlevel survey data extracted from the 2006 Japanese General Social Surveys (JGSS-2006), a nationwide survey conducted using two-stage stratified random sampling.

In questionnaire surveys concerning subjective outcomes such as self-assessed social position, choices are often arranged ordinally. The data on such ordinal choices can be statistically analyzed using an ordered probit model. The JGSS-2006 also contains several responses that indicate satisfaction in areas such as family life, as well as self-assessed social position; therefore, we estimate a multivariate ordered probit model for these variables. From the frequentist viewpoint, an ordered probit model can be estimated using the maximum likelihood method.[2] After the seminal work of Albert and Chib (1993), who utilize latent variable representation, a Bayesian analysis using the Markov chain Monte

[2]See, for example, Greene (2008, Chapter 23).

Carlo (MCMC) method has gained popularity for estimating the ordered probit model. Although the latent variables are unknown, their full conditional distributions (FCD) follow a truncated normal distribution. This makes the estimation of the ordered probit model very tractable in Bayesian analysis.[3] Further, we can apply Chen and Dey's (2000) Bayesian multivariate ordered probit model to two or more questionnaire items. Because the latent variables in the multivariate ordered probit model are correlated, we must consider this correlation while estimating the model. Restrictions must be imposed on the parameters in order to identify them. A sufficient condition for this identification problem is that the covariance matrix of the latent variables is defined as a form of the correlation matrix. However, this hinders the estimation of the model. Chen and Dey (2000) have successfully overcome this difficulty by using a joint reparameterization of the correlation matrix and cutoff points for the ordinal data.[4]

The model allows us to include a new concept, "regret," to measure self-assessed social position. Hasegawa and Ueda (2011) introduce regret to measure subjective well-being and define it as the probability with which a respondent of the survey who selects a choice in a multiple-choice question pertaining to social position does not choose any other option indicative of a better social position. Thus, regret is used to analyze inequality in self-assessed social position, whose data are given in ordinal variables, since individual regret, the average regret of a group of people, and some inequality measures of regret can be computed.[5]

This article proceeds as follows. In Section 2, following Chen and Dey (2000), we describe the Bayesian multivariate ordered probit model. Further, we describe a method of estimating the coefficients of ordinal explanatory variables on the basis of the posterior results of the multivariate ordered probit model. Section 3 provides the posterior results of the estimation of this model on self-assessed social position by using microlevel survey data extracted from JGSS-2006. In Section 4, using the values of the latent variables, we define the probability related to an individual's self-assessed social position and an inequality measure for self-assessed social position. Section 5 provides brief concluding remarks.

2 BAYESIAN MULTIVARIATE ORDERED PROBIT MODEL

2.1 *Chen and Dey's Model*

Following Chen and Dey (2000, pp.135–140), this section describes the Bayesian multivariate probit model. Let y_{ij} denote the ordinal discrete response of individual i to question j for $i = 1, \cdots, n$ and $j = 1, \cdots, m$; that is, $y_{ij} = c$ for $c = 1, \cdots, C_j$. Further, let z_{ij} denote the latent variable of individual i to

[3] See Albert and Chib (1993).

[4] Jeliazkov et al. (2009) discuss in depth the identification problems of univariate and multivariate ordered probit models.

[5] As Kalmijn and Veenhoven (2005) and Hasegawa and Ueda (2011) point out, differences in people's happiness are not measurable, and thus often ignored. Therefore, this comparability of a subjective index (social position in our article) is an advantage of using regret.

question j such that

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

where γ_{jc} is a cutoff point for the jth ordinal response. While in Chen and Dey (2000), the number of cutoff points in each equation is the same, (1) allows for a different number of cutoff points. We specify that

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

where condition $\gamma_{j1} = 0$ is required for establishing the identifiability of the cutoff parameters.[6] Latent variable z_{ij} is assumed to be determined by linear model

$$z_{ij} = \boldsymbol{x}_{ij}'\boldsymbol{\beta}_j + u_{ij}, \ i = 1, \cdots, n; \ j = 1, \cdots, m, \tag{3}$$

where

$$\boldsymbol{x}_{ij} = \begin{pmatrix} x_{ij1} \\ x_{ij2} \\ \vdots \\ x_{ijk_j} \end{pmatrix}, \ \boldsymbol{\beta}_j = \begin{pmatrix} \beta_{j1} \\ \beta_{j2} \\ \vdots \\ \beta_{jk_j} \end{pmatrix}, \ i = 1, \cdots, n; \ j = 1, \cdots, m.$$

Defining $\boldsymbol{\beta} = (\boldsymbol{\beta}_1', \boldsymbol{\beta}_2', \cdots, \boldsymbol{\beta}_m')'$ and

$$\boldsymbol{z}_i = \begin{pmatrix} z_{i1} \\ z_{i2} \\ \vdots \\ z_{im} \end{pmatrix}, \ \boldsymbol{X}_i = \begin{pmatrix} \boldsymbol{x}_{i1}' & & & \mathbf{O} \\ & \boldsymbol{x}_{i2}' & & \\ & & \ddots & \\ \mathbf{O} & & & \boldsymbol{x}_{im}' \end{pmatrix} = \operatorname{diag}(\boldsymbol{x}_{i1}', \boldsymbol{x}_{i2}', \cdots, \boldsymbol{x}_{im}'),$$

$$\boldsymbol{u}_i = \begin{pmatrix} u_{i1} \\ u_{i2} \\ \vdots \\ u_{im} \end{pmatrix}, \ i = 1, \cdots, n,$$

the linear model for the latent variables is rewritten as

$$\boldsymbol{z}_i = \boldsymbol{X}_i\boldsymbol{\beta} + \boldsymbol{u}_i, \ i = 1, \cdots, n.$$

Now, we assume that $\boldsymbol{u}_i \sim \mathrm{N}(\mathbf{0}, \boldsymbol{\Sigma})$; that is,

$$\boldsymbol{z}_i \sim \mathrm{N}(\boldsymbol{X}_i\boldsymbol{\beta}, \boldsymbol{\Sigma}), \ i = 1, \cdots, n, \tag{4}$$

where $\boldsymbol{\Sigma}$ is an $m \times m$ positive definite covariance matrix. To ensure the identification of parameters, some restriction must be imposed on $\boldsymbol{\Sigma}$. Here, following Chen and Dey (2000, p.136), in addition to $\gamma_{j1} = 0$, we assume that $\gamma_{j(C_j-1)} = 1$.[7] We define $\boldsymbol{\gamma}_j = (\gamma_{j2}, \cdots, \gamma_{j(C_j-2)})'$ $(j = 1, \cdots, m)$ and $\boldsymbol{\gamma} = (\boldsymbol{\gamma}_1', \cdots, \boldsymbol{\gamma}_m')'$.

[6] See, for example, Albert and Chib (1993, p.673), and Johnson and Albert (1999, p.131).

[7] Usually, $\boldsymbol{\Sigma}$ is assumed to be a correlation matrix for identification. See, for example, Chib and Greenberg (1998, p.348).

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

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

Further, defining

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

we have

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

where $1_{(\cdot)}$ is an indicator function.[8] Now, we specify the prior distributions as follows:

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

where

$$\boldsymbol{\beta}_j \sim \mathrm{N}(\boldsymbol{\beta}_{j0}, \boldsymbol{B}_{j0}),\ j = 1, \cdots, m,\ \boldsymbol{\Sigma}^{-1} \sim \mathrm{W}(\kappa_0, \boldsymbol{Q}_0^{-1}),$$

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 $\boldsymbol{\delta}_j = (\delta_{j2}, \cdots, \delta_{j(C_j-2)})'$ $(j = 1, \cdots, m)$. 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

$$\begin{aligned} &p(\boldsymbol{\beta}, \boldsymbol{\gamma}, \boldsymbol{\Sigma}, \boldsymbol{z}|\boldsymbol{y}) \propto p(\boldsymbol{\beta}, \boldsymbol{\gamma}, \boldsymbol{\Sigma}) \\ &\quad \times \left\{ \prod_{i=1}^{n} 1_{(\boldsymbol{z}_i \in \mathcal{G}_i)} |\boldsymbol{\Sigma}|^{-\frac{1}{2}} \exp\left[-\frac{1}{2} (\boldsymbol{z}_i - \boldsymbol{X}_i \boldsymbol{\beta})' \boldsymbol{\Sigma}^{-1} (\boldsymbol{z}_i - \boldsymbol{X}_i \boldsymbol{\beta}) \right] \right\}. \qquad (5) \end{aligned}$$

Using the MCMC sampling scheme, we can sample parameters $(\boldsymbol{\beta}, \boldsymbol{\gamma}, \boldsymbol{\Sigma}, \boldsymbol{z})$ from the joint posterior distribution (5). Appendix A provides details on the sampling algorithms.

[8] See Chib and Greenberg (1998, p.349).

2.2 *Relationship between One Ordinal Variable and Other Ordinal Variables*

We can use the multivariate ordered probit model (4) to investigate the relationship between one ordinal variable y_1 and the others $(y_2, \cdots, y_m)$. Dropping the suffix i in (4), we consider population regression $\boldsymbol{z} \sim \mathrm{N}(\boldsymbol{X\beta}, \boldsymbol{\Sigma})$, where $\boldsymbol{X} = \mathrm{diag}(\boldsymbol{x}'_1, \cdots, \boldsymbol{x}'_m)$, and divide $\boldsymbol{z}$ as $\boldsymbol{z} = \left(z_1, \boldsymbol{z}'_{(-1)}\right)'$. Suppose that z_1 is a latent variable associated with a dependent variable of interest y_1 and that $\boldsymbol{z}_{(-1)}$ is a vector of latent variables corresponding to the other variables $(y_2, \cdots, y_m)$. Multivariate normal model $\boldsymbol{z} \sim \mathrm{N}(\boldsymbol{X\beta}, \boldsymbol{\Sigma})$ can be used to predict z_1 given $\boldsymbol{z}_{(-1)}$.[9]

Then, we have

$$p(\boldsymbol{z}|\boldsymbol{X}, \cdots) = p\left(z_1, \boldsymbol{z}_{(-1)}|\boldsymbol{X}, \cdots\right) = p\left(z_1|\boldsymbol{z}_{(-1)}, \boldsymbol{X}, \cdots\right) p\left(\boldsymbol{z}_{(-1)}|\boldsymbol{X}, \cdots\right),$$

where "$| \cdots$" denotes the conditioning of the other unspecified variables in the equation. On the basis of the property of the multivariate normal distribution, we obtain

$$z_1|\boldsymbol{z}_{(-1)}, \boldsymbol{X}, \cdots \sim \mathrm{N}(\tilde{\mu}_1, \tilde{\sigma}_{11}), \tag{6}$$

where

$$\tilde{\mu}_1 = \boldsymbol{x}'_1\boldsymbol{\beta}_1 + \boldsymbol{\sigma}'_{(-1)}\boldsymbol{\Sigma}^{-1}_{(-1)}\left(\boldsymbol{z}_{(-1)} - \boldsymbol{X}_{(-1)}\boldsymbol{\beta}_{(-1)}\right) \tag{7}$$

$$\tilde{\sigma}_{11} = \sigma_{11} - \boldsymbol{\sigma}'_{(-1)}\boldsymbol{\Sigma}^{-1}_{(-1)}\boldsymbol{\sigma}_{(-1)} \tag{8}$$

$$\boldsymbol{X} = \begin{pmatrix} \boldsymbol{x}'_1 & \boldsymbol{0}' \\ \mathbf{O} & \boldsymbol{X}_{(-1)} \end{pmatrix}, \ \boldsymbol{\beta} = \begin{pmatrix} \boldsymbol{\beta}_1 \\ \boldsymbol{\beta}_{(-1)} \end{pmatrix}, \ \boldsymbol{\Sigma} = \begin{pmatrix} \sigma_{11} & \boldsymbol{\sigma}'_{(-1)} \\ \boldsymbol{\sigma}_{(-1)} & \boldsymbol{\Sigma}_{(-1)} \end{pmatrix}.$$

$\boldsymbol{\alpha} = \boldsymbol{\Sigma}^{-1}_{(-1)}\boldsymbol{\sigma}_{(-1)} = (\alpha_2, \cdots, \alpha_m)'$ becomes a coefficient vector of $\boldsymbol{z}_{(-1)}$ in the regression of z_1 on $\boldsymbol{z}_{(-1)}$, given $\boldsymbol{X}$.[10] The coefficients of the explanatory variables in $\boldsymbol{X}$ can be calculated from $\boldsymbol{x}'_1\boldsymbol{\beta}_1 - \boldsymbol{\sigma}'_{(-1)}\boldsymbol{\Sigma}^{-1}_{(-1)}\boldsymbol{X}_{(-1)}\boldsymbol{\beta}_{(-1)}$.

3 EMPIRICAL ANALYSIS OF JAPANESE SELF-ASSESSED SOCIAL POSITION DATA

3.1 *Data*

Our empirical analysis uses microlevel survey data extracted from JGSS-2006. The survey population was comprised of men and women between 21–89 years of age and was stratified by six regional blocks, with each block stratified by cities/districts of four different population sizes. The total number of sampling locations was 526, and the number of sampling locations for each stratum was adjusted so that approximately 15 individuals were sampled in each location. Data were collected via face-to-face interviews and the placement (self-administered) method. The sample size of JGSS-2006 was $8,000$, and the

[9]See, for example, Hoff (2009, pp.118–122). Hoff (2009) uses the multivariate normal model to predict one or more variables given the others in the context of missing data problem.

[10]Hasegawa (2010) discusses this topic in more detail.

number of valid respondents was $4,254$. In the poverty rates announced by the Ministry of Health, Labour and Welfare (MHLW), for example, the poverty rate of households with an adult and a child/children is far higher than the overall poverty rate. Thus, for our analysis, we limited the households to those including two or more adults. A total of $2,220$ unit records were obtained after eliminating those that contained missing observations for the variables.

(Place Table 1 here.)

Table 1 presents the definitions of the variables used in the analysis.[11] The poverty line for equivalent income set by MHLW is based on disposable income; however, it is not applicable here because income in the JGSS data includes tax and social insurance premiums.[12] Therefore, a poverty line comparable to the income in the JGSS data must be calculated. According to Tanaka (2010), the burden ratio of tax and social insurance premiums in aggregate income is 0.155 for 2007. Assuming that the ratio does not vary significantly from 2006 to 2007, the ratio of aggregate income to disposable income is $1/(1-0.155)$. The nominal poverty line based on the disposable income in 2006 is 1,270 thousand yen, as announced by MHLM. Thus, the poverty line comparable to the income in the JGSS data is $1270/(1-0.155) = 1,503$ thousand yen.

(Place Tables 2(a), 2(b), 2(c), and 2(d) here.)

Tables 2 (a) to (d) present the summary statistics of the variables used in the analysis. Table 2 (a) shows the frequencies and percentages of the choices corresponding to the self-assessed social position (`SocPos10`), which includes seven ordinal choices from "less than or equal to 2" to "greater than or equal to 8."[13] Table 2 (b) shows satisfaction with family life (`SatFam`), with the household budget situation (`SatBudget`), and with health condition (`SatHealth`). These variables include five ordinal choices (from 1 to 5). Tables 2 (c) and (d) present the summary statistics of the explanatory variables. Statistics such as the mean and standard deviation of household income in Table 2 (d) are calculated from the numerical values explained in footnote 11.

[11] Variable `log(income)` is the logarithm of annual household income in ten thousand yen divided by the square root of the number of family members. In the case of zero income, we set `log(income)` equal to zero. The only income data provided in JGSS-2006 are those for the household income band. The number of income bands are 19, and they are 0 yen, less than 0.7 million yen, 0.7–1 million yen,$\cdots$, 18.5–23 million yen, and more than 23 million yen. Following Layard et al. (2008), we construct the numerical values of income on the basis of the income band. Layard et al. (2008, p.1850) construct the income data as follows: "In the cross-section surveys, only income bands are available and these we converted into numerical values using the mid point of each band. For respondents in the lowest income band, we assumed an income of two thirds of the upper limit of the band, and for respondents in the highest income band we assumed an income 1.5 times the lower income limit of the band."

[12] The poverty line announced by MHLW is available at the following URL of the Comprehensive Survey of Living Conditions:
`http://www.mhlw.go.jp/toukei/saikin/hw/k-tyosa/k-tyosa10/2-7.html`.

[13] See the definition of `SocPos10` in Table 1.

3.2 *Posterior Results of Estimated Equations*

The estimated equations are as follows:

$$
\begin{aligned}
z_{ij} = {} & \beta_{j1} + \beta_{j2}\,\texttt{DPov}_i + \beta_{j3}\,\texttt{Male}_i + \beta_{j4}\log\left(\frac{\texttt{Income}_i}{\texttt{PovLine}}\right) \\
& + \beta_{j5}\,\texttt{DPov}_i \times \left[\log\left(\frac{\texttt{Income}_i}{\texttt{PovLine}}\right) - \overline{\log\left(\frac{\texttt{Income}}{\texttt{PovLine}}\right)}\right] \\
& + \beta_{j6}\,\texttt{Age}_i + \beta_{j7}\,\texttt{Educ}_i + u_{ij},\ j = 1, \cdots, 4;\ i = 1, \cdots, n,
\end{aligned} \tag{9}
$$

where $j = 1$ corresponds to `SocPos10` and $j = 2, 3$, and 4 correspond to `SatFam`, `SatBudget`, and `SatHealth`, respectively. In the interaction term, $\overline{\log\left(\frac{\texttt{Income}}{\texttt{PovLine}}\right)}$ is the average value of $\log\left(\frac{\texttt{Income}}{\texttt{PovLine}}\right)$, and the former is subtracted from the latter. The coefficient of that term is estimated at the average value of $\log\left(\frac{\texttt{Income}}{\texttt{PovLine}}\right)$ according to Wooldridge (2008, pp.241–242).[14]

The MCMC simulation was run for 30,000 iterations, and the first 10,000 samples were discarded as the burn-in period. The posterior results obtained thereafter were generated using Ox version 6.3 (Doornik, 2009). We set the prior distributions as follows:

$$
\begin{aligned}
& \boldsymbol{\beta}_j \sim \mathrm{N}(\mathbf{0}, 100\boldsymbol{I}_7) \\
& \boldsymbol{\delta}_1(\gamma_1) \sim \mathrm{N}(\mathbf{0}, 100\boldsymbol{I}_4),\ \boldsymbol{\delta}_j(\boldsymbol{\gamma}_j) \sim \mathrm{N}(\mathbf{0}, 100\boldsymbol{I}_2),\ j = 2, \cdots, 4 \\
& \boldsymbol{\Sigma}^{-1} \sim \mathrm{W}(8, 50\boldsymbol{I}_4).
\end{aligned}
$$

Tables 3 and 4 present the posterior results of the estimation of (9). Table 5 provides the posterior results of the model with ordinal explanatory variables using equation (6). In these tables, "Mean," "SD," and "Median" denote the posterior mean, posterior standard deviation, and posterior median, respectively.[15] Further, "P-value" denotes the p-value for the convergence diagnostic statistic (CD) proposed by Geweke (1992).

(Place Table 3 here.)

From Table 3, we can infer the following:

- Income positively affects self-assessed social position (`SocPos10`) because the posterior mean of the coefficient of the logarithm of the income-deflated poverty line (β_{14}) is positive. However, the posterior means of the coefficient of the poverty dummy (β_{12}) and the coefficient of the interaction of the poverty dummy and the logarithm of the income-deflated poverty line (β_{15}) are negative. Therefore, poverty negatively affects `SocPos10`.

[14]In Layard et al. (2008, p.1853), the model includes estimations for not only log(`Income`) but also [log(`Income`)]^2. However, the correlation coefficient between those two variables is high, which makes the interpretation of their coefficients difficult. Thus, the model that includes only log(`income/PovLine`) is estimated in this article.

[15]The robustness of the results in Tables 3 and 5 is also checked by estimating the coefficients of (9) for different values of the poverty line (PL) in the supplementary material (Tables B.1 and B.2).

- The gender dummy (`Male`) positively affects `SatFam` because the sign of the posterior mean of β_{23} is positive. However, `Male` does not notably affect `SocPos10`, `SatBudget`, and `SatHealth` since the 95% credible intervals (CIs) of β_{13}, β_{33}, and β_{43} include zero.
- `Age` positively affects `SocPos10` and `SatBudget` because the signs of the posterior means of β_{16} and β_{36} are positive. However, `Age` does not notably affect `SatFam` and `SatHealth` because the 95% CIs of β_{26} and β_{46} include zero. However, `Age` negatively affects `SatHealth` with a 90% CI.
- Educational level (`Educ`) positively affects `SocPos10` because the sign of the posterior mean of β_{17} is positive. However, `Educ` does not notably affect the three kinds of satisfaction because the 95% CIs of β_{27}, β_{37}, and β_{47} include zero.

Furthermore, Table 4 shows that the 95% CIs do not include zero for all elements of $\mathbf{\Sigma}$.

(Place Table 4 here.)

Table 5 provides the posterior results of the model with ordinal explanatory variables. From this table, we can infer the following:

- `SatBudget` positively affects `SocPos10`, while `SatFam` and `SatHealth` do not notably affect it.
- Poverty negatively affects `SocPos10`, but income and educational level positively affect it.

(Place Table 5 here.)

4 INEQUALITY OF SELF-ASSESSED SOCIAL POSITION

4.1 *Probability Associated with Self-Assessed Social Position*

In this section, we consider the inequality of self-assessed social position. An advantage of Bayesian analysis is that the value of latent variable z_{i1} can be directly obtained from the posterior results. However, the value of z_{i1} may change with the identification restrictions on the parameters. Therefore, instead of latent variable z_{i1}, we consider the probability related to individual i's self-assessed social position, which is defined as

$$p_i = \Phi\left(\frac{z_{i1} - \tilde{\mu}_1}{\sqrt{\tilde{\sigma}_{11}}}\right), \; i = 1, \cdots, n, \tag{10}$$

where $\Phi(\cdot)$ is the distribution function of the standard normal distribution. $\tilde{\mu}_1$ and $\tilde{\sigma}_{11}$ are defined in (7) and (8), respectively. Using the probability related to self-assessed social position, we can calculate the inequality indices for self-assessed social position.

(Place Figure 1 here.)

Figure 1 contains two boxplots for each $y_1 = c$ $(c = 1, \cdots, 7)$; the left one shows the distributions of $\boldsymbol{p} = (p_1, \cdots, p_n)'$ of the respondents whose income is above the poverty line and the right one shows the same for the respondents whose income is below the poverty line. A comparison of these boxplots shows that the medians of the right boxplots are higher than those of the left for each $y_1 = c$, $(c = 1, \cdots, 7)$. One explanation for this finding may be that respondents above and below the poverty line differ in their perceptions of social position. The respondents below the poverty line are likely to perceive their social positions as higher because they may be satisfied with something other than income. The respondents above the poverty line, however, do not perceive their social position as higher than those who are below the poverty line because income may be an important factor for their satisfaction. Table 6 shows the number of respondents choosing $y_1 = 1, \cdots, 7$.

(Place Table 6 here.)

4.2 *Regret over Self-Assessed Position*

In this section, we use the inequality measure for self-assessed social position proposed by Hasegawa and Ueda (2011).[16] We calculate probability $r(z_{i1}|\gamma_{1c}, \boldsymbol{y})$, which denotes the difference between cutoff points γ_{1c} and z_{i1}:

$$\begin{aligned} r(z_{i1}|\gamma_{1c}, \boldsymbol{y}_i) &= \Pr(z_{i1} < z < \gamma_{1c}|\boldsymbol{y}_i) \\ &= \Pr\left(\frac{z_{i1} - \tilde{\mu}_1}{\sqrt{\tilde{\sigma}_{11}}} < \frac{z - \tilde{\mu}_1}{\sqrt{\tilde{\sigma}_{11}}} < \frac{\gamma_{1c} - \tilde{\mu}_1}{\sqrt{\tilde{\sigma}_{11}}} \;\middle|\; \boldsymbol{y}_i \right) \\ &= \begin{cases} \Phi\left(\frac{\gamma_{1c} - \tilde{\mu}_1}{\sqrt{\tilde{\sigma}_{11}}}\right) - \Phi\left(\frac{z_{i1} - \tilde{\mu}_1}{\sqrt{\tilde{\sigma}_{11}}}\right) & \text{if } z_{i1} < \gamma_{1c} \\ 0 & \text{if } z_{i1} \geq \gamma_{1c}, \end{cases} \\ & i = 1, \cdots, n, \; c = 1, \cdots, 7. \end{aligned} \tag{11}$$

We call $r(z_{i1}|\gamma_c, \boldsymbol{y}_i)$ the regretted self-assessed social position of individual i in category c.[17] Further, on the basis of the definition of regretted self-assessed

[16] Using the World Values Survey data, Hasegawa and Ueda (2011) propose the regret function for subjective well-being and provide posterior analyses on the inequality of subjective well-being.

[17] If a loss function is defined as $L(z_{i1}, \gamma_{1c}) = 1_{\left(z \in (z_{i1}, \gamma_{1c})\right)}$, the posterior risk function can be written as

$$r(z_{i1}|\gamma_{1c}, \boldsymbol{y}_i) = \mathrm{E}[L(z_{i1}, \gamma_{1c})|\cdots] = \int 1_{\left(z \in (z_{i1}, \gamma_{1c})\right)} p(z|\cdots) dz,$$

social position, we have $r\left(z_{i1}|\gamma_{1C_1}, \boldsymbol{y}_i\right) = 1-p_i$. When latent variable z_{i1}, which represents an evaluation of person i's social position, is greater than cutoff point γ_{1c}, from (1), he/she is supposed to perceive himself or herself to be in social position c. Thus, regret denotes the probability that person i does not choose a higher social position c as his or her social position. It can also be regarded as the value representing the degree of disappointment that the person has upon failing to attain a higher social position. As regret denotes probability and computation of aggregation and comparison is possible, an inequality of self-assessed social position can be measured by investigating its distribution. Therefore, regret is a useful method for economic and social analysis using survey data.

(Place Table 7 here.)

Table 7 provides the posterior results of regretted self-assessed social position. `regret1` to `regret7` denote the regretted self-assessed social position for categories $c = 1$ to $c = 7$, respectively. "Mean," "SD," and "Median" denote the sample mean, sample standard deviation, and sample median of individual posterior regretted self-assessed social position, respectively; for $j \leq c$,

$$\bar{r}_{cj} = \frac{1}{n_j} \sum_{i \in S_j} r\left(z_{i1}|\gamma_{1c}, \boldsymbol{y}_i\right), \ s_{cj} = \sqrt{\frac{1}{n_j - 1} \sum_{i \in S_j} \left(r\left(z_{i1}|\gamma_{1c}, \boldsymbol{y}_i\right) - \bar{r}_{cj}\right)^2}$$

$$\bar{r}^*_c = \frac{1}{n^*_c} \sum_{i \in S^*_c} r\left(z_{i1}|\gamma_{1c}, \boldsymbol{y}_i\right), \ s^*_c = \sqrt{\frac{1}{n^*_c - 1} \sum_{i \in S^*_c} \left(r\left(z_{i1}|\gamma_{1c}, \boldsymbol{y}_i\right) - \bar{r}_c\right)^2},$$

where $S_j = \{i : y_{i1} = j\}$, $n_j = \#S_j$ (the number of observations in S_j) and where $S^*_c = \{i : y_{i1} \leq c\}$, $n^*_c = \#S^*_c$ (the number of observations in S^*_c).[18]

which is the regret function defined in (11). For the Bayesian decision theory, see, for example, Press (2003, Chapter 11).

[18]For example, for `regret2` in Table 7, the Mean and SD of $y_1 = 1$ denote

$$\bar{r}_{21} = \frac{1}{n_1} \sum_{i \in S_1} r\left(z_{i1}|\gamma_{12}, \boldsymbol{y}_i\right), \ s_{21} = \sqrt{\frac{1}{n_1 - 1} \sum_{i \in S_1} \left(r\left(z_{i1}|\gamma_{12}, \boldsymbol{y}_i\right) - \bar{r}_{21}\right)^2},$$

and the Mean and SD of $y_1 = 2$ denote

$$\bar{r}_{22} = \frac{1}{n_1} \sum_{i \in S_2} r\left(z_{i1}|\gamma_{12}, \boldsymbol{y}_i\right), \ s_{22} = \sqrt{\frac{1}{n_1 - 1} \sum_{i \in S_2} \left(r\left(z_{i1}|\gamma_{12}, \boldsymbol{y}_i\right) - \bar{r}_{22}\right)^2},$$

where $S_1 = \{i : y_{i1} = 1\}$, $n_1 = \#S_1$, $S_2 = \{i : y_{i1} = 2\}$, and $n_2 = \#S_2$. Further, the Mean and SD of $y_1 \leq 2$ denote

$$\bar{r}^*_2 = \frac{1}{n^*_2} \sum_{i \in S^*_2} r\left(z_{i1}|\gamma_{12}, \boldsymbol{y}_i\right), \ s^*_2 = \sqrt{\frac{1}{n^*_2 - 1} \sum_{i \in S^*_2} \left(r\left(z_{i1}|\gamma_{12}, \boldsymbol{y}_i\right) - \bar{r}_2\right)^2},$$

where $S^*_2 = \{i : y_{i1} \leq 2\}$ and $n^*_2 = \#S^*_2$.

4.3 *Measuring the Inequality of Self-Assessed Social Position*

In survey questionnaires such as JGSS, social position is investigated by asking respondents their perceptions of their own social positions. The choices are given in an ordinal scale. Dissatisfaction with their position differs even when respondents choose the same position. Regret can represent such dissatisfaction, and can be used to analyze the inequality of self-assessed social position since it is the probability that people do not choose a higher social position.

(Place Table 8 here.)

Table 8 shows the Gini coefficients of regretted self-social position. They are computed from (11) by using a person's regret for $c = 1, \cdots, 7$. The 95% CI does not include zero. The Gini coefficient decreases as c of `regret`c increases from $c = 1$ to 7. `regret1` denotes the degree to which a person whose self-assessed social position is the lowest feels dissatisfied about not attaining a higher social position. Table 6 shows that with regard to `regret1`, 106 samples have positive regret, while far more samples have zero regret. Thus, the inequality of regret in all samples of the posterior results is large, and the value of the Gini coefficient is high, near 1. The Gini coefficient is inversely related to c because more samples have positive regret and fewer have zero regret. In our results, the Gini coefficient decreases greatly when c increases from 4 to 5, since, as Table 6 shows, the number of samples of positive regret increases by 855.

Although this analysis uses single-year data, the regret of each year can be computed if cross-sectional data or panel data across years are available. Using these data to compute the Gini coefficient allows us to analyze changes in the inequality of regret; that is, we can compare the degree of dissatisfaction among persons who do not attain a higher social position.

Figure 2 shows the boxplots of regretted self-assessed social position and the details of its distribution. The figure titled as "$y_1 = 1$" denotes the distributions of regret for the respondents belonging to S_1 that fail to reach γ_{1c} ($c = 1, \cdots, 7$). The figures titled as "$y_1 \leq j$" ($j = 2, \cdots, 7$) denote the distributions of regret that respondents belonging to S_j^* fail to reach γ_{1c} ($c = 1, \cdots, 7$). The vertical axis represents regret and the horizontal axis represents c ($c = 1, \cdots, 7$), which denotes the choices of self-assessed social position. Two boxplots are drawn for each c ($c = 1, \cdots, 7$) on the horizontal axis in the figures. The left boxplot shows the distribution of regret for respondents above the poverty line, while the right shows that for respondents below the poverty line.

(Place Figure 2 here.)

From Figure 2, we infer the following:

- The interquartile range (IQR) of the boxplots in the figures for respondents below the poverty line is wider than that for respondents above the poverty line when $y_1 \leq 2, \cdots, y_1 \leq 7$.

- Comparing the boxplots of (A) and (B) at the label c of the horizontal axis in the figure titled as "$y_1 \leq c$ ($c = 1, \cdots, 7$)", the distributions of the respondents above the poverty line skew to the right except in the cases of $y_1 \leq 2$ and $y_1 \leq 4$. However, the distributions of respondents below the poverty line do not skew when $y_1 \leq 6$ and $y_1 \leq 7$, although their distributions skew to the left when $y_1 \leq 2$, $y_1 \leq 3$, $y_1 \leq 4$, and $y_1 \leq 5$. These findings suggest the following three points. Respondents above the poverty line are satisfied with their social position. Respondents below the poverty line who see themselves at a lower social position are not satisfied with their present situation. Among respondents below the poverty line, persons who see themselves at a higher social position hardly display a biased perception about being satisfied with their present social position.

- Similarly, the median of the regretted self-assessed social position of respondents below the poverty line is higher than that of respondents above the poverty line for each c ($c = 1, \cdots 7$) except in the case of $y_1 \leq 7$.

5 Concluding Remarks

In this article, using the microlevel data of JGSS-2006, we analyzed the relationships between self-assessed social position and socioeconomic factors such as income and poverty. We estimated the Bayesian multivariate ordered probit model because ordinal data are used for some explanatory variables that represent people's satisfaction with their family life, household economic situation, and health. We proposed an inequality measure on self-assessed social position, "regret," which allows us to investigate the effects of poverty on self-assessed social position through a change in its distribution, as it is derived from the probability that a person cannot choose a higher social position.

The main results of our empirical analysis are as follows.

- Income positively affects self-assessed social position, but poverty negatively affects it (see Table 3).

- Satisfaction with the household budget positively affects self-assessed social position, but satisfaction with family life and health do not notably affect it (see Table 5).

- When the ordinal explanatory variables—satisfaction with family life, with the household budget, and with health—are not controlled, `Age` positively affects self-assessed social position (see Table 3). However, when they are controlled (see Table 5), `Age` does not notably affect it.

- The median of the regretted self-assessed social position of respondents below the poverty line is higher than that of respondents above the poverty line in most cases (see Figure 2).

- When the dispersion of regretted self-assessed social position is calculated by the IQR of a boxplot, the dispersion of respondents below the poverty line is greater than that of respondents above the poverty line (see Figure 2).

- The distributions of the regretted self-assessed social position of respondents above the poverty line skew to the right at each label on the horizontal axis, c ($c = 1, \cdots, 7$), in most cases (see Figure 2). However, the distributions of regretted self-assessed social position have two different shapes among those who are below the poverty line. The distributions of the regret of persons whose self-assessed social position is lower skew to the left, but this phenomenon is not observed for persons whose self-assessed social position is higher. This suggests that persons above the poverty line tend to be satisfied with their self-assessed social position. Furthermore, persons below the poverty line whose self-assessed social position is lower tend to be dissatisfied with their current social position, but persons whose self-assessed social position is higher do not exhibit distinct dissatisfaction.

Interestingly, the distributions of regret differ among those below the poverty line. Further investigation is required to determine the reason for the skewness of the distributions. For example, such people may give up trying to achieve higher social positions, or non-economic factors may decrease their regret.

The effects of income and poverty on self-assessed social position may depend on the changes in a variety of socioeconomic circumstances. Our future analysis will be extended across years and will use the multi-year JGSS data.

A SAMPLING ALGORITHMS

Following Chen and Dey (2000, pp.135–140), this appendix describes the sampling algorithm of Bayesian multivariate probit model.

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

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

$$\boldsymbol{z}_i = \begin{pmatrix} z_{ij} \\ \boldsymbol{z}_{i(-j)} \end{pmatrix}, \ \boldsymbol{X}_i = \begin{pmatrix} \boldsymbol{x}'_{ij} & \boldsymbol{0}' \\ \boldsymbol{0} & \boldsymbol{X}_{i(-j)} \end{pmatrix}, \ \boldsymbol{\beta} = \begin{pmatrix} \boldsymbol{\beta}_j \\ \boldsymbol{\beta}_{(-j)} \end{pmatrix}$$
$$\boldsymbol{\Sigma} = \begin{pmatrix} \sigma_{jj} & \boldsymbol{\sigma}'_{(-j)} \\ \boldsymbol{\sigma}_{(-j)} & \boldsymbol{\Sigma}_{(-j)} \end{pmatrix}, \ \boldsymbol{\Sigma}^{-1} = \begin{pmatrix} \sigma^{jj} & \boldsymbol{\sigma}^{(-j)\prime} \\ \boldsymbol{\sigma}^{(-j)} & \boldsymbol{\Sigma}^{(-j)} \end{pmatrix}.$$

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

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

$$\boldsymbol{\beta}_j | \cdots \sim \mathrm{N}(\tilde{\boldsymbol{\beta}}_j, \tilde{\boldsymbol{B}}_j), \ j = 1, \cdots, m, \tag{12}$$

where

$$\tilde{\boldsymbol{B}}_j = \left(\boldsymbol{B}_{j0}^{-1} + \sigma^{jj} \sum_{i=1}^{n} \boldsymbol{x}_{ij} \boldsymbol{x}'_{ij} \right)^{-1}$$
$$\tilde{\boldsymbol{\beta}}_j = \tilde{\boldsymbol{B}}_j \left[\boldsymbol{B}_{j0}^{-1} \boldsymbol{\beta}_{j0} + \sigma^{jj} \sum_{i=1}^{n} \boldsymbol{x}_{ij} z_{ij} + \sum_{i=1}^{n} \boldsymbol{x}_{ij} \boldsymbol{\sigma}^{(-j)\prime} \left(\boldsymbol{z}_{i(-j)} - \boldsymbol{X}_{i(-j)} \boldsymbol{\beta}_{(-j)} \right) \right].$$

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

$$\boldsymbol{\Sigma}^{-1}|\cdots \sim \mathrm{W}(\tilde{\kappa}, \tilde{\boldsymbol{Q}}^{-1}) \tag{13}$$

where

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

Applying Gibbs sampling to the FCDs of (12) and (13), we can generate $\boldsymbol{\beta}_j$ 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 $\boldsymbol{z}_{(-j)}$ denote the vector obtained by removing $\boldsymbol{z}_{(j)}$ from $\boldsymbol{z}$, and let $\boldsymbol{z}_{i(-j)}$ denote the vector of removing z_{ij} from $\boldsymbol{z}_i$. We generate $\boldsymbol{\gamma}_j$ and $\boldsymbol{z}_{(j)}$ from the joint conditional distribution $p(\boldsymbol{\gamma}_j, \boldsymbol{z}_{(j)}|\boldsymbol{\beta}, \boldsymbol{\Sigma}, \boldsymbol{z}_{(-j)}, \boldsymbol{y})$ $(j = 1, \cdots, m)$. The joint conditional distribution $p(\boldsymbol{\gamma}_j, \boldsymbol{z}_{(j)}|\boldsymbol{\beta}, \boldsymbol{\Sigma}, \boldsymbol{z}_{(-j)}, \boldsymbol{y})$ can be written as

$$\begin{aligned}
&p(\boldsymbol{\gamma}_j, \boldsymbol{z}_{(j)}|\boldsymbol{\beta}, \boldsymbol{\Sigma}, \boldsymbol{z}_{(-j)}, \boldsymbol{y}) \\
&\quad = p(\boldsymbol{\gamma}_j|\boldsymbol{\beta}, \boldsymbol{\Sigma}, \boldsymbol{z}_{(-j)}, \boldsymbol{y}) p(\boldsymbol{z}_{(j)}|\boldsymbol{\gamma}_j, \boldsymbol{\beta}, \boldsymbol{\Sigma}, \boldsymbol{z}_{(-j)}, \boldsymbol{y}),\ j = 1, \cdots, m.
\end{aligned}$$

Similar to the sampling of $\boldsymbol{\beta}_j$, for the sake of convenience of expression, we replace the jth factor as the first factor. Since $\boldsymbol{z}_i|\boldsymbol{\beta}, \boldsymbol{\Sigma}, \boldsymbol{\gamma} \sim \mathrm{N}(\boldsymbol{X}_i\boldsymbol{\beta}, \boldsymbol{\Sigma})$, from the property of the multivariate normal distribution we have

$$z_{ij}|\boldsymbol{\gamma}_j, \boldsymbol{\beta}, \boldsymbol{\Sigma}, \boldsymbol{z}_{(-j)}, \boldsymbol{y} \sim \mathrm{N}(\tilde{\mu}_{ij}, \tilde{\sigma}_{jj}) 1_{(z_{ij} \in \mathcal{G}_{ij})},\ i = 1, \cdots, n;\ j = 1, \cdots, m, \tag{14}$$

where

$$\begin{aligned}
\tilde{\mu}_{ij} &= \boldsymbol{x}'_{ij}\boldsymbol{\beta}_j + \boldsymbol{\sigma}'_{(-j)}\boldsymbol{\Sigma}^{-1}_{(-j)} \left(\boldsymbol{z}_{i(-j)} - \boldsymbol{X}_{i(-j)}\boldsymbol{\beta}_{(-j)}\right) \\
\tilde{\sigma}_{jj} &= \sigma_{jj} - \boldsymbol{\sigma}'_{(-j)}\boldsymbol{\Sigma}^{-1}_{(-j)}\boldsymbol{\sigma}_{(-j)}.
\end{aligned}$$

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

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

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

$$p(\boldsymbol{\delta}_j|\boldsymbol{\beta}, \boldsymbol{\Sigma}, \boldsymbol{z}_{(-j)}, \boldsymbol{y}) \propto p(\boldsymbol{\gamma}_j|\boldsymbol{\beta}, \boldsymbol{\Sigma}, \boldsymbol{z}_{(-j)}, \boldsymbol{y}) \prod_{c=2}^{C_j-2} \frac{\big(1-\gamma_{j(c-1)}\big)\exp(\delta_{jc})}{\big(1+\exp(\delta_{jc})\big)^2}. \tag{15}$$

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 (15),

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

and ν is the degrees of freedom. The 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

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

References

Albert, J. H. and S. Chib, "Bayesian Analysis of Binary and Polychotomous Response Data," *Journal of the American Statistical Association*, 88(422), 669–679, 1993.

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

Chen, M.-H. and D. K. Dey, "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*, 133–157. New York: Marcel Dekker, 2000.

Chib, S. and E. Greenberg, "Analysis of Multivariate Probit Models," *Biometrika*, 85(2), 347–361, 1998.

Damien, P. and S. G. Walker, "Sampling Truncated Normal, Beta, and Gamma Densities," *Journal of Computational and Graphical Statistics*, 10(2), 206–215, 2001.

Di Tella, R., J. Haisken-De New, and R. MacCulloch, "Happiness Adaptation to Income and to Status in an Individual Panel," *Journal of Economic Behavior and Organization*, 76(3), 834–852, 2010.

Di Tella, R. and R. MacCulloch, "Some Uses of Happiness Data in Economics," *Journal of Economic Perspectives*, 20(1), 25–46, 2006.

Doornik, J. A., *An Object-oriented Matrix Programming Language Ox^{TM} 6.* London: Timberlake Consultants Ltd, 2009.

Easterlin, R. A., "Does Economic Growth Improve the Human Lot? Some Empirical Evidence," In P. A. David and M. W. Reder (Eds.), *Nations and Households in Economic Growth: Essays in Honor of Moses Abramowitz*, 89–125. New York: Academic Press, 1974.

Geweke, J., "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*, 169–193. Oxford: Oxford University Press, 1992.

Greene, W. H., *Econometric Analysis, 6th ed,* Upper Saddle River: Pearson Education, 2008.

Hasegawa, H., "Analyzing Tourists' Satisfaction: A Multivariate Ordered Probit Approach," *Tourism Management*, 31(1), 86–97, 2010.

Hasegawa, H. and K. Ueda, "Measuring Inequality of Subjective Well-being: A Bayesian Approach," *Journal of Socio-Economics*, 40(5), 700–708, 2011.

Hoff, P. D., *A First Course in Bayesian Statistical Methods*, Dordrecht: Springer-Verlag, 2009.

Ishida, H., "Industrialization, Class Structure, and Social Mobility in Postwar Japan," *British Journal of Sociology*, 52(4), 579–604, 2001.

Jeliazkov, I., J. Graves, and M. Kutzbach, “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, 115–156. Bingley: Emerald, 2009.

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

Kalmijn, W. and R. Veenhoven, “Measuring Inequality of Happiness in Nations: In Search for Proper Statistics,” *Journal of Happiness Studies*, 6(4), 357–396, 2005.

Layard, R.. *Happiness: Lessons from a New Science*, London: Penguin Books, 2005.

Layard, R., G. Mayraz, and S. Nickell, “The Marginal Utility of Income,” *Journal of Public Economics*, 92(8–9), 1846–1857, 2008.

Luttmer, E. F. P., “Neighbors as Negatives: Relative Earnings and Well-being,” *Quarterly Journal of Economics*, 120(3), 963–1002, 2005.

Press, S. J., *Subjective and Objective Bayesian Statistics: Principles, Models, and Applications, 2nd ed.,* Hoboken: Wiley, 2003.

Ravallion, M. and M. Lokshin, “Identifying Welfare Effects from Subjective Questions,” *Economica*, 68(271), 335–357, 2001.

———, “Self-rated Economic Welfare in Russia,” *European Economic Review*, 46(8), 1453–1473, 2002.

Shirahase, S., “Japan as a Stratified Society: With a Focus on Class Identification,” *Social Science Japan Journal*, 13(1), 31–52, 2010.

Stevenson, B. and J. Wolfers. “Economic Growth and Subjective Well-being: Reassessing the Easterlin Paradox,” *Brookings Papers on Economic Activity*, Spring, 1–87, 2008.

Stutzer, A. and B. S. Frey, “Recent Advances in the Economics of Individual Subjective Well-being,” *Social Research*, 77(2), 679–714, 2010.

Tanaka, H., Review of Tax, Contribution and Social Welfare Benefit: Toward the Integration of Tax and Social Welfare System (in Japanese), *Discussion Paper Series No.481, the Center for Intergenerational Studies, the Institute of Economic Research, Hitotsubashi University*, 2010.

Wooldridge, J. M., *Introductory Econometrics, 4th ed.,* Mason: South-Western, 2008.

Table 1: Definitions of variables

Variables	Definition
SocPos10	Self-assessed position in the society in 10 level that takes ten ordinal choices from 1 (Bottom) to 10 (Top). In the empirical analysis, rungs 1–2 and 8–10 were aggregated as two choices because of small number of responses.
SatFam	Satisfaction with family life that takes five ordinal choices from 1 (Dissatisfied) to 5 (Satisfied).
SatBudget	Satisfaction with household budget situation that takes five ordinal choices from 1 (Dissatisfied) to 5 (Satisfied).
SatHealth	Satisfaction with health condition that takes five ordinal choices from 1 (Dissatisfied) to 5 (Satisfied).
PovLine	Poverty line for equivalent income. Equivalent income is defined as annual household income in ten thousand yen divided by the square root of the number of family members.
DPov	A dummy variable that takes 1 if equivalent income is less than poverty line, and 0, otherwise.
Male	A dummy variable that takes 1 if the respondent is male, and 0 if the respondent is female.
log(Income)	Logarithm of equivalent income. In the case of zero income, we set log(Income) equal to zero.
$\log\left(\frac{\texttt{Income}}{\texttt{PovLine}}\right)$	Logarithm of equivalent income deflated by poverty line.
Age	Age of the respondent.
Educ	Years of education of the respondent.

Table 2 (a): Summary statistics (Position in the society in 10 level, `SocPos10`)[a]

Bottom ≤ 2	3	4	5	6	7	Top ≥ 8
106	187	284	367	855	251	170
(4.77)	(8.42)	(12.79)	(16.53)	(38.51)	(11.31)	(7.66)

a: Values in parentheses denote percentage.

Table 2 (b): Summary statistics (Satisfaction)[a]

Satisfaction with	Dissatisfied 1	2	3	4	Satisfied 5
Family Life	37	144	665	730	644
(`SatFam`)	(1.67)	(6.49)	(29.95)	(32.88)	(29.01)
Household Budget Situation	171	423	795	516	315
(`SatBudget`)	(7.7)	(19.05)	(35.81)	(23.24)	(14.19)
Health Condition	61	309	800	656	394
(`SatHealth`)	(2.75)	(13.92)	(36.04)	(29.55)	(17.75)

a: Values in parentheses denote percentage.

Table 2 (c): Summary statistics (Explanatory variables)

Gender (`Male`)	
Female	Male
1085	1135

Poverty Dummy (`DPov`)	
0	1
1992	228

Table 2 (d): Summary statistics (Explanatory variables)[a]

Household Income (thousand yen)						
Mean	SD	Min	25%	50%	75%	Max
683.91	488.86	0	400	600	800	3450

Number of Family Members						
Mean	SD	Min	25%	50%	75%	Max
3.59	1.39	2	2	4	4	9

Age (`Age`)						
Mean	SD	Min	25%	50%	75%	Max
53.54	14.20	21	42	55	65	89

Education (`Educ`)						
Mean	SD	Min	25%	50%	75%	Max
12.60	2.41	6	12	12	14	18

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 3: Posterior results of multivariate ordered probit model (coefficient parameters)[a]

Position in the Society in 10 Level (`SocPos10`)					
		Mean	SD	Median[b]	P-value
`Intercept`	β_{11}	0.1272	0.0552	0.1276^{**}	0.5057
`DPov`	β_{12}	-0.2319	0.0352	-0.2315^{**}	0.8831
`Male`	β_{13}	0.0146	0.0135	0.0147	0.1645
$\log\left(\frac{\texttt{Income}}{\texttt{PovLine}}\right)$	β_{14}	0.1923	0.0159	0.1922^{**}	0.8782
$\texttt{DPov}\times\log\left(\frac{\texttt{Income}}{\texttt{PovLine}}\right)^{c}$	β_{15}	-0.2301	0.0302	-0.2300^{**}	0.5872
`Age`	β_{16}	0.0021	0.0005	0.0021^{**}	0.7418
`Educ`	β_{17}	0.0125	0.0032	0.0125^{**}	0.1975
	γ_{12}	0.1802	0.0109	0.1802^{**}	0.7778
	γ_{13}	0.3342	0.0109	0.3342^{**}	0.4707
	γ_{14}	0.4805	0.0105	0.4806^{**}	0.3388
	γ_{15}	0.8235	0.0094	0.8237^{**}	0.7643

Satisfaction with Family Life (`SatFam`)					
		Mean	SD	Median	P-value
`Intercept`	β_{21}	0.6648	0.0703	0.6648^{**}	0.4476
`DPov`	β_{22}	-0.0216	0.0460	-0.0216	0.3594
`Male`	β_{23}	0.1075	0.0177	0.1075^{**}	0.7638
$\log\left(\frac{\texttt{Income}}{\texttt{PovLine}}\right)$	β_{24}	0.0642	0.0203	0.0641^{**}	0.3672
$\texttt{DPov}\times\log\left(\frac{\texttt{Income}}{\texttt{PovLine}}\right)$	β_{25}	-0.0292	0.0395	-0.0290	0.9054
`Age`	β_{26}	0.0010	0.0007	0.0010	0.8471
`Educ`	β_{27}	-0.0018	0.0041	-0.0018	0.1828
	γ_{22}	0.2709	0.0160	0.2708^{**}	0.8102
	γ_{23}	0.6760	0.0104	0.6761^{**}	0.5682

a: "Mean," "SD," and "Median" denote the posterior mean, posterior standard deviation, and posterior median, respectively. "P-value" denotes the p-values for the convergence diagnostic statistic (CD) proposed by Geweke (1992).

b: "$**$" and "$*$" denote that zero is not included in the 95% and 90% credible intervals, respectively.

c: $\log\left(\frac{\texttt{Income}}{\texttt{PovLine}}\right)$ denotes $\left[\log\left(\frac{\texttt{Income}}{\texttt{PovLine}}\right) - \overline{\log\left(\frac{\texttt{Income}}{\texttt{PovLine}}\right)}\right]$.

Table 3: Continued

Satisfaction with Household Budget Situation (`SatBudget`)					
		Mean	SD	Median	P-value
`Intercept`	β_{31}	-0.0110	0.0693	-0.0105	0.4304
`DPov`	β_{32}	-0.1477	0.0452	-0.1471^{**}	0.6507
`Male`	β_{33}	-0.0086	0.0173	-0.0087	0.7484
$\log\left(\frac{\texttt{Income}}{\texttt{PovLine}}\right)$	β_{34}	0.2239	0.0203	0.2238^{**}	0.1585
$\texttt{DPov} \times \log\left(\frac{\texttt{Income}}{\texttt{PovLine}}\right)$	β_{35}	-0.1849	0.0391	-0.1843^{**}	0.7276
`Age`	β_{36}	0.0070	0.0006	0.0070^{**}	0.7318
`Educ`	β_{37}	0.0032	0.0041	0.0032	0.2363
	γ_{32}	0.3089	0.0110	0.3087	0.0966
	γ_{33}	0.6858	0.0103	0.6859	0.6364

Satisfaction with Health Condition (`SatHealth`)					
		Mean	SD	Median	P-value
`Intercept`	β_{41}	0.7226	0.0656	0.7230^{**}	0.4245
`DPov`	β_{42}	-0.0477	0.0426	-0.0478	0.2848
`Male`	β_{43}	0.0009	0.0163	0.0008	0.9725
$\log\left(\frac{\texttt{Income}}{\texttt{PovLine}}\right)$	β_{44}	0.0612	0.0186	0.0612^{**}	0.6009
$\texttt{DPov} \times \log\left(\frac{\texttt{Income}}{\texttt{PovLine}}\right)$	β_{45}	-0.0716	0.0365	-0.0712^{**}	0.3669
`Age`	β_{46}	-0.0011	0.0006	-0.0011^{*}	0.8590
`Educ`	β_{47}	-0.0034	0.0038	-0.0034	0.2365
	γ_{42}	0.3245	0.0128	0.3245^{**}	0.5604
	γ_{43}	0.6889	0.0098	0.6890^{**}	0.8262

Table 4: **Posterior results of multivariate ordered probit model ($\boldsymbol{\Sigma}$)**[a]

	Mean	SD	Median[b]	P-value
σ_{11}	0.0891	0.0033	0.0890**	0.6286
σ_{21}	0.0169	0.0027	0.0169**	0.5855
σ_{31}	0.0314	0.0027	0.0314**	0.6192
σ_{41}	0.0124	0.0025	0.0123**	0.8946
σ_{22}	0.1406	0.0067	0.1405**	0.4510
σ_{32}	0.0873	0.0044	0.0873**	0.9732
σ_{42}	0.0644	0.0039	0.0643**	0.3351
σ_{33}	0.1415	0.0055	0.1414**	0.2567
σ_{43}	0.0532	0.0035	0.0531**	0.2311
σ_{44}	0.1260	0.0051	0.1259**	0.4644

a: "Mean," "SD," and "Median" denote the posterior mean, posterior standard deviation, and posterior median, respectively. "P-value" denotes the p-values for the convergence diagnostic statistic (CD) proposed by Geweke (1992).

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

Table 5: Posterior results of model with ordinal explanatory variables (`SocPos10`)[a]

		Mean	SD	Median[b]	P-value
`SatFam`	α_2	-0.0336	0.0264	-0.0336	0.5958
`SatBudget`	α_3	0.2369	0.0243	0.2369^{**}	0.9196
`SatHealth`	α_4	0.0153	0.0226	0.0152	0.8448
`Intercept`	β_1	0.1411	0.0567	0.1415^{**}	0.5518
`DPov`	β_2	-0.1969	0.0341	-0.1968^{**}	0.9712
`Male`	β_3	0.0203	0.0134	0.0203	0.1180
$\log\left(\frac{\texttt{Income}}{\texttt{PovLine}}\right)$	β_4	0.1405	0.0159	0.1405^{**}	0.4515
$\texttt{DPov} \times \log\left(\frac{\texttt{Income}}{\texttt{PovLine}}\right)^c$	β_5	-0.1862	0.0294	-0.1860^{**}	0.4534
`Age`	β_6	0.0004	0.0005	0.0004	0.8395
`Educ`	β_7	0.0117	0.0031	0.0117^{**}	0.3101

a: "Mean," "SD," and "Median" denote the posterior mean, posterior standard deviation, and posterior median, respectively. "P-value" denotes the p-values for the convergence diagnostic statistic (CD) proposed by Geweke (1992).

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

c: $\log\left(\frac{\texttt{Income}}{\texttt{PovLine}}\right)$ denotes $\left[\log\left(\frac{\texttt{Income}}{\texttt{PovLine}}\right) - \overline{\log\left(\frac{\texttt{Income}}{\texttt{PovLine}}\right)}\right]$.

Table 6: Number of respondents choosing $y_1 = 1, \cdots, 7$ in the posterior result

above the poverty line		below the poverty line	
y_1	number	y_1	number
1	58	1	48
2	145	2	42
3	246	3	38
4	325	4	42
5	777	5	78
6	237	6	14
7	165	7	5
total	1953	total	267

Table 7: Posterior results of regretted position[a]

		Mean	SD	Median[b]
`regret1`	$y_1 = 1$	0.0558	0.0389	0.0507
`regret2`	$y_1 = 1$	0.2026	0.0963	0.2065
`regret2`	$y_1 = 2$	0.0630	0.0280	0.0644
`regret2`	$y_1 \leq 2$	0.1135	0.0914	0.0804
`regret3`	$y_1 = 1$	0.3791	0.1292	0.4004
`regret3`	$y_1 = 2$	0.2276	0.0695	0.2430
`regret3`	$y_1 = 3$	0.0736	0.0220	0.0759
`regret3`	$y_1 \leq 3$	0.1797	0.1359	0.1050
`regret4`	$y_1 = 1$	0.5604	0.1282	0.5972
`regret4`	$y_1 = 2$	0.4094	0.0852	0.4433
`regret4`	$y_1 = 3$	0.2515	0.0458	0.2672
`regret4`	$y_1 = 4$	0.0869	0.0160	0.0932
`regret4`	$y_1 \leq 4$	0.2535	0.1738	0.2419
`regret5`	$y_1 = 1$	0.8592	0.0425	0.8744
`regret5`	$y_1 = 2$	0.7417	0.0542	0.7662
`regret5`	$y_1 = 3$	0.6201	0.0597	0.6444
`regret5`	$y_1 = 4$	0.4667	0.0597	0.4868
`regret5`	$y_1 = 5$	0.1995	0.0298	0.2121
`regret5`	$y_1 \leq 5$	0.4156	0.2326	0.3863
`regret6`	$y_1 = 1$	0.9174	0.0250	0.9284
`regret6`	$y_1 = 2$	0.8150	0.0668	0.8393
`regret6`	$y_1 = 3$	0.7125	0.0887	0.7427
`regret6`	$y_1 = 4$	0.5686	0.0986	0.5880
`regret6`	$y_1 = 5$	0.3199	0.0773	0.3359
`regret6`	$y_1 = 6$	0.0730	0.0286	0.0747
`regret6`	$y_1 \leq 6$	0.4646	0.2562	0.4061
`regret7`	$y_1 = 1$	0.9442	0.0388	0.9488
`regret7`	$y_1 = 2$	0.8533	0.0910	0.8665
`regret7`	$y_1 = 3$	0.7653	0.1212	0.7841
`regret7`	$y_1 = 4$	0.6317	0.1450	0.6322
`regret7`	$y_1 = 5$	0.4006	0.1337	0.4037
`regret7`	$y_1 = 6$	0.1914	0.1153	0.1717
`regret7`	$y_1 = 7$	0.0752	0.0596	0.0624
`regret7`	$y_1 \leq 7$	0.5010	0.2738	0.4808

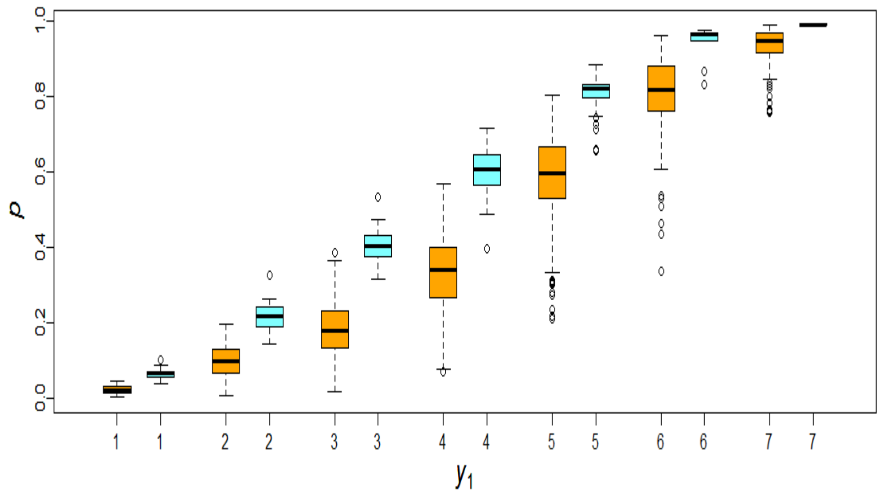
a: "Mean," "SD," and "Median" denote the sample mean, sample standard deviation, and sample median, respectively.

Table 8: Gini coefficients[a]

	Mean	SD	Median[b]
regret1	0.9771	0.0012	0.9771**
regret2	0.9300	0.0023	0.9300**
regret3	0.8532	0.0030	0.8532**
regret4	0.7442	0.0036	0.7442**
regret5	0.4692	0.0052	0.4692**
regret6	0.3865	0.0063	0.3865**
regret7	0.3313	0.0066	0.3313**

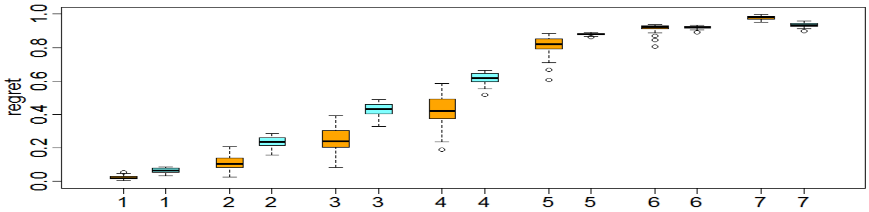
a: "Mean," "SD," and "Median" denote the posterior mean, posterior standard deviation, and posterior median, respectively.

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

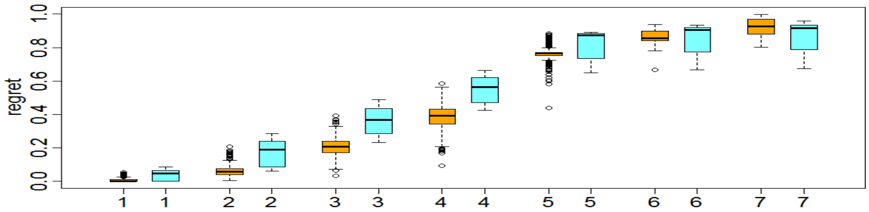


Orange boxplots show the probability of respondents whose income is above the poverty line and the blue boxplots show the probability of respondents whose income is below the poverty line.

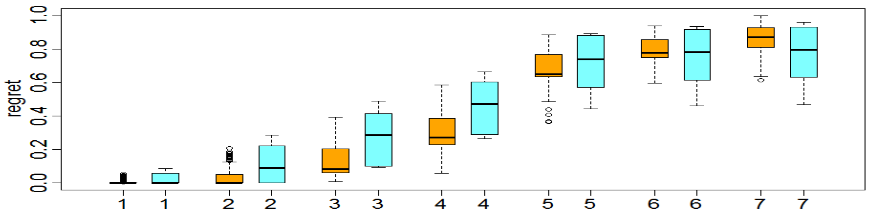
Figure 1: Boxplots of p



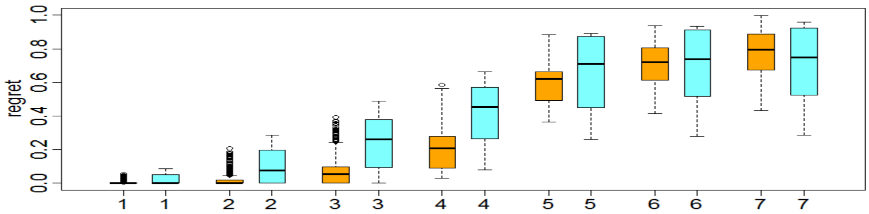
$y_1 = 1$ $\quad$ $\mathtt{A}^a$ $(58)^b$, $\quad$ $\mathtt{B}$ (48)



$y_1 \leq 2$ $\quad$ $\mathtt{A}$ (203) $\quad$ $\mathtt{B}$ (90)



$y_1 \leq 3$ $\quad$ $\mathtt{A}$ (449) $\quad$ $\mathtt{B}$ (128)

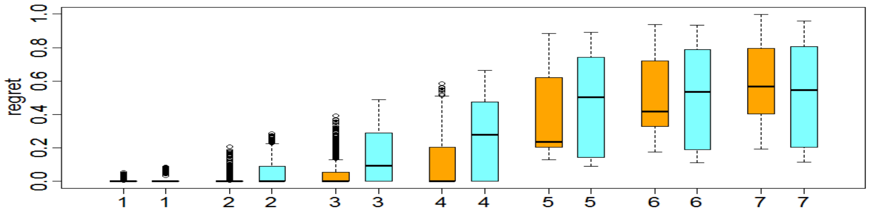


$y_1 \leq 4$ $\quad$ $\mathtt{A}$ (774) $\quad$ $\mathtt{B}$ (170)

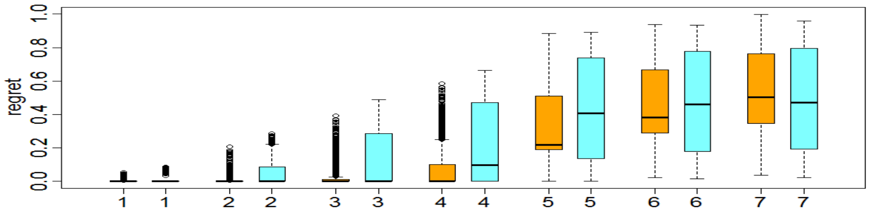
a: Left boxplots show the distributions of regret for respondents "above the poverty line (case A)" and right boxplots show those for respondents "below the poverty line (case B)."

b: Parenthetic numbers denote the number of respondents in the case `A` and `B`.

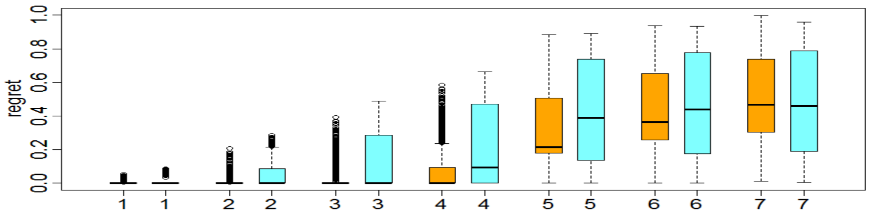
Figure 2: Boxplots of regretted self-assessed social position



$y_1 \leq 5$ `A` (1551) `B` (248)



$y_1 \leq 6$ `A` (1788) `B` (262)



$y_1 \leq 7$ `A` (1953) `B` (267)

Figure 2: Continued