ORIGINAL RESEARCH PAPER

Time use of married couples: Bayesian approach

Hikaru Hasegawa[a] and Kazuhiro Ueda[b]

[a]Department of Economics, Hokkaido University, Sapporo, Japan; [b]Faculty of Economics, Nihon Fukushi University, Aichi, Japan

ARTICLE HISTORY

Compiled March 26, 2019

ABSTRACT

The living hours data of individuals' time spent on daily activities are compositional and include many zeros because individuals do not pursue all activities every day. Thus, we should exercise caution in using such data for empirical analyses. The Bayesian method offers several advantages in analyzing compositional data. In this study, we analyze the time allocation of Japanese married couples using the Bayesian model. Based on the Bayes factors, we compare models that consider and do not consider the correlations between married couples' time use data. The model that considers the correlation shows superior performance. We show that the Bayesian method can adequately take into account the correlations of wives' and husbands' living hours, facilitating the calculation of partial effects that their activities' variables have on living hours. The partial effects of the model that considers the correlations between the couples' time use are easily calculated from the posterior results.

KEYWORDS

Bayesian statistics; compositional data; Markov chain Monte Carlo (MCMC); time allocation

JEL CLASSIFICATION

C11, J12, J22

CONTACT K, Ueda. Email: ueda@n-fukushi.ac.jp

Time use of married couples: Bayesian approach

ABSTRACT

The living hours data of individuals' time spent on daily activities are compositional and include many zeros because individuals do not pursue all activities every day. Thus, we should exercise caution in using such data for empirical analyses. The Bayesian method offers several advantages in analyzing compositional data. In this study, we analyze the time allocation of Japanese married couples using the Bayesian model. Based on the Bayes factors, we compare models that consider and do not consider the correlations between married couples' time use data. The model that considers the correlation shows superior performance. We show that the Bayesian method can adequately take into account the correlations of wives' and husbands' living hours, facilitating the calculation of partial effects that their activities' variables have on living hours. The partial effects of the model that considers the correlations between the couples' time use are easily calculated from the posterior results.

KEYWORDS

Bayesian statistics; compositional data; Markov chain Monte Carlo (MCMC); time allocation

JEL CLASSIFICATION

C11, J12, J22

1. Introduction

There is an increase in the availability of time use data and the studies using them. Time use data contain microlevel data of time that people use for various activities. The data can be used to analyze individual behavior, socio-economic background, and the relationship between time use and well-being (e.g., [10], [11]). In this study, we propose an empirical analysis method using time use data and analyze the daily time allocation of married couples. Especially, we investigate the factors with important effects on time use.

Time use data are not negative and sum up the time an individual spends on various activities in 24 hours ($1,440$ minutes). Thus, such living hours data are compositional in nature. In this study, we divide these data by $1,440$ minutes to obtain the ratio data. As the sum of the ratio data is one, we proceed with the empirical study while considering the data as compositional.

As is well known, we must use compositional data with caution. Although the Dirichlet distribution is used as a probability distribution of continuous variables whose sum is one, the covariances of the Dirichlet distribution are all negative, and applying them to compositional data is not appropriate [2, Section 3.4]. Multivariate lognormal distribution has been proposed as an alternative to the Dirichlet distribution [2, Chapter 6]. However, living hours data include many zeros, as people do not spend time in performing all the surveyed activities. For example, the work or commute time of people who do not work is recorded as zero. Therefore, it is not appropriate to use multivariate lognormal distribution for data that include many zeros. We must construct an appropriate model for such data. Models such as the Box-Cox transformation have been proposed to estimate models using compositional data that include zeros ([2, pp.308–309], [6, p.955]). Among the studies dealing with zeroes in the compositional data, [4] propose a general model that uses a latent Gaussian model. [12] provide a specification of their transformation function in the context of Bayesian statistics.

Although [12] estimate the model with spatial correlation, we estimate the model using compositional data with two correlations by extending their Bayesian method. The reason why we use the Bayesian method is that it can easily handle the following two types of correlations. Correlations between hours for some activities are supposed in individual living hours data. Thus, we have to estimate models with such correlations using compositional data, including zeros. It is easier to use the Bayesian method to estimate these models than frequentist approaches. Furthermore, there is another type of correlations. When the data are of couples' time use, correlations between hours for some activities in the time use of wives and husbands are also supposed. Thus, we need to use a method to estimate models with those correlations, using compositional data including many zeros. This can be easily done with the simulation-based Bayesian methods.

As an application of the estimation using compositional data including zeros, we analyze the living hours of married couples based on the Japanese Panel Survey of Consumers (JPSC) conducted by the Institute for Research on Household Economics. Although there are many previous studies on living hours, most consider only one type of activity hours, for example, the burden or time allocation of housework hours between husbands and wives [3, 15–17]. Few studies analyze multiple types of time data together, as in this study, except for [9].

Our study estimates the coefficients of explanatory variables using the Markov chain Monte Carlo (MCMC) algorithm. Furthermore, we demonstrate the partial effects of each explanatory variable on wives' and husbands' living hours in their time equations and the probability that such partial effects are positive. The advantage of the Bayesian method is that we can calculate the expected values of individual compositional data ([12, p.320]) and the partial effects that we define based on the MCMC [18, p.592].

This paper is organized as follows. In Section 2, we present the Bayesian model used in this study. In Section 3, we explain the time use data of the JPSC and derive the posterior results. In addition, we present the calculation and estimation of partial effects. Finally, in Section 4, we present a brief summary and future issues.

2. Bayesian Models

In this section, we use the Bayesian framework for compositional data as proposed in [12] and construct the Bayesian models, incorporating zero proportions for time use data.

2.1. Basic Compositional Data Model

Defining $Y_1, Y_2, \ldots, Y_{D-1}, Y_D$ as D proportional data (compositional data) on time use, we assume that

$$Y_1 + Y_2 + \cdots + Y_{D-1} + Y_D = 1, \ 0 \leq Y_k \leq 1, \ k = 1, \ldots, D.$$

In addition, we assume that the base time is positive, $Y_D > 0$. As relative ratios of $Y_k (k = 1, ..., D-1)$ to Y_D are not affected if we change the unit for measuring living hours, for example, by hour, minute, or second, scale invariance of the compositional data is ensured.[1] Incorporating the case of $Y_k = 0$, we use the following transformation from latent variables $\boldsymbol{z} = (z_1, \ldots, z_d)'$ to $\boldsymbol{Y} = (Y_1, \ldots, Y_d)'$, as proposed in [12, p.319]:

$$Y_k = \frac{(\max(0, z_k))^\gamma}{1 + \sum_{k'=1}^{d} (\max(0, z_{k'}))^\gamma}, \ k = 1, \ldots, d, \tag{1}$$

where $\gamma > 0$ is a known constant, and

$$d = D - 1, \ Y_D = \left[1 + \sum_{k'=1}^{d} (\max(0, z_{k'}))^\gamma \right]^{-1}.$$

The inverse transformation of (1) is

$$\begin{cases} z_k = (Y_k/Y_D)^{1/\gamma} & \text{if } Y_k > 0 \\ z_k \leq 0 & \text{if } Y_k = 0 \end{cases}, \ k = 1, \ldots, d. \tag{2}$$

Writing the mean of z_k as $\mathrm{E}(z_k) = \boldsymbol{x}'\boldsymbol{\beta}_k$ and the covariance matrix of $\boldsymbol{z}$ as $\boldsymbol{V}$, we set $\boldsymbol{z} \sim \mathrm{N}_d(\boldsymbol{B}'\boldsymbol{x}, \boldsymbol{V})$, where $\boldsymbol{B} = (\boldsymbol{\beta}_1, \ldots, \boldsymbol{\beta}_d)$ and $\boldsymbol{\beta}_k$ is a $p \times 1$ vector. Further, defining $y_k = (Y_k/Y_D)^{1/\gamma}$ and $\boldsymbol{y} = (y_1, \ldots, y_d)'$, (2) can be written as

$$\begin{cases} z_k = y_k & \text{if } y_k > 0 \\ z_k \leq 0 & \text{if } y_k = 0 \end{cases}, \ k = 1, \ldots, d. \tag{3}$$

Thus, (3) with $\boldsymbol{z} \sim \mathrm{N}_d(\boldsymbol{B}'\boldsymbol{x}, \boldsymbol{V})$ can be considered as a multivariate Tobit model. We partition the elements of $\boldsymbol{y}$ into two terms, $\boldsymbol{y}_+$ and $\boldsymbol{y}_\mathrm{o}$, which contain strictly positive elements and zero elements, respectively. For convenience of expression, we replace the elements of $\boldsymbol{y}_+$ as the first part of $\boldsymbol{y}$, that is, $\boldsymbol{y} = \begin{pmatrix} \boldsymbol{y}_+ \\ \boldsymbol{y}_\mathrm{o} \end{pmatrix}$. According to the order of

[1] Scale invariance and subcompositional coherence are required in compositional data analysis ([2, p.418], [13, p.43]). However, subcompositional coherence is not ensured in our study, because zeros are included in $Y_1, \ldots, Y_{D-1}$.

the elements of $\boldsymbol{y}$, we partition $\boldsymbol{B}$ and $\boldsymbol{V}$ as follows:

$$\boldsymbol{B} = (\boldsymbol{B}_+, \boldsymbol{B}_\circ), \ \boldsymbol{V} = \begin{pmatrix} \boldsymbol{V}_{++} & \boldsymbol{V}_{+\circ} \\ \boldsymbol{V}_{\circ +} & \boldsymbol{V}_{\circ\circ} \end{pmatrix}.$$

Then, the density function of $\boldsymbol{y}$ will be

$$\begin{aligned} p(\boldsymbol{y}|\boldsymbol{B}, \boldsymbol{V}) &= p(\boldsymbol{y}_+, \boldsymbol{y}_\circ|\boldsymbol{B}, \boldsymbol{V}) = p(\boldsymbol{y}_+|\boldsymbol{B}, \boldsymbol{V}) \Pr(\boldsymbol{y}_\circ|\boldsymbol{y}_+, \boldsymbol{B}, \boldsymbol{V}) \\ &= \phi(\boldsymbol{y}_+|\boldsymbol{B}'_+\boldsymbol{x}, \boldsymbol{V}_{++})\Phi\left(\mathrm{E}(\boldsymbol{y}_\circ|\boldsymbol{y}_+, \boldsymbol{B}, \boldsymbol{V}), \mathrm{var}(\boldsymbol{y}_\circ|\boldsymbol{y}_+, \boldsymbol{B}, \boldsymbol{V})\right), \end{aligned} \tag{4}$$

where

$$\begin{aligned} \mathrm{E}(\boldsymbol{y}_\circ|\boldsymbol{y}_+, \boldsymbol{B}, \boldsymbol{V}) &= \boldsymbol{B}'_\circ\boldsymbol{x} + \boldsymbol{V}_{\circ +}\boldsymbol{V}_{++}^{-1}(\boldsymbol{y}_+ - \boldsymbol{B}'_+\boldsymbol{x}) \\ \mathrm{var}(\boldsymbol{y}_\circ|\boldsymbol{y}_+, \boldsymbol{B}, \boldsymbol{V}) &= \boldsymbol{V}_{\circ\circ} - \boldsymbol{V}_{\circ +}\boldsymbol{V}_{++}^{-1}\boldsymbol{V}_{+\circ}, \end{aligned}$$

$\phi(\boldsymbol{w}|\boldsymbol{c}, \boldsymbol{C})$ is a density function of multivariate normal distribution $\mathrm{N}(\boldsymbol{c}, \boldsymbol{C})$, and

$$\Phi(\boldsymbol{c}, \boldsymbol{C}) = \int_{-\infty}^{0} \cdots \int_{-\infty}^{0} \phi(\boldsymbol{w}|\boldsymbol{c}, \boldsymbol{C}) d\boldsymbol{w}.$$

Now, let $\boldsymbol{y}_1, \ldots, \boldsymbol{y}_n$ denote an i.i.d. sample of $\boldsymbol{y}$ and $\boldsymbol{z}_1, \ldots, \boldsymbol{z}_n$ be the vectors of $\boldsymbol{z}$ corresponding to $\boldsymbol{y}_1, \ldots, \boldsymbol{y}_n$. Further, we specify the prior distributions as follows:

$$\boldsymbol{z}_i \sim \mathrm{N}_d(\boldsymbol{B}'\boldsymbol{x}_i, \boldsymbol{V}), \ i = 1, \ldots, n \tag{5}$$

$$\mathrm{vec}\, \boldsymbol{B} \sim \mathrm{N}_{dp}\left(\mathrm{vec}\, \boldsymbol{B}_0, \lambda \boldsymbol{I}_{dp}\right) \tag{6}$$

$$\boldsymbol{V} \sim \mathrm{IW}_d(m, \boldsymbol{M}^{-1}), \tag{7}$$

where $\mathrm{IW}_d(m, \boldsymbol{M}^{-1})$ denotes an inverse Wishart distribution with degrees of freedom m and scale matrix $\boldsymbol{M}$ [8, pp.576–577], and $\mathrm{vec}\, \boldsymbol{B} \sim \mathrm{N}_{d(p+1)}\left(\mathrm{vec}\,(\mu \mathbf{1}_{d\times 1}, \mathbf{0}_{d\times p})', \lambda \boldsymbol{I}_{d(p+1)}\right)$ in [12].

After some manipulation, we derive the following algorithm. Appendix A.1 in Supplemental material presents details of the derivation.

Algorithm 1. Define $\boldsymbol{X} = (\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n)'$ and $\boldsymbol{Z} = (\boldsymbol{z}_1, \ldots, \boldsymbol{z}_n)'$. The full conditional distribution (FCD) of $\operatorname{vec}\boldsymbol{B}$ is

$$\operatorname{vec}\boldsymbol{B}| \cdots \sim \mathrm{N}\left(\boldsymbol{\beta}^*, \boldsymbol{\Omega}^*\right),$$

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

$$\boldsymbol{\Omega}^* = \left[\lambda^{-1}\boldsymbol{I} + (\boldsymbol{V}^{-1} \otimes \boldsymbol{X}'\boldsymbol{X}_j)\right]^{-1}$$
$$\boldsymbol{\beta}^* = \boldsymbol{\Omega}^* \left[\lambda^{-1} \operatorname{vec}\boldsymbol{B}_0 + (\boldsymbol{V}^{-1} \otimes \boldsymbol{I}) \operatorname{vec}(\boldsymbol{X}'\boldsymbol{Z})\right].$$

The FCD of $\boldsymbol{V}$ is

$$\boldsymbol{V}| \cdots \sim \mathrm{IW}\left(m^*, \boldsymbol{M}^{*-1}\right),$$

where

$$m^* = m + n,\ \boldsymbol{M}^* = \boldsymbol{M}_j + \sum(\boldsymbol{z}_i - \boldsymbol{B}'\boldsymbol{x}_i)(\boldsymbol{z}_i - \boldsymbol{B}'\boldsymbol{x}_i)'.$$

We partition the elements of $\boldsymbol{z}_i$ into two terms, $\boldsymbol{z}_{i+}$ and $\boldsymbol{z}_{i\,\mathrm{o}}$, which contain strictly positive elements and non-positive elements, respectively. For convenience of expression, we replace the elements of $\boldsymbol{z}_{i+}$ as the first part of $\boldsymbol{z}_i$, that is, $\boldsymbol{z}_i = \begin{pmatrix} \boldsymbol{z}_{i+} \\ \boldsymbol{z}_{i\,\mathrm{o}} \end{pmatrix}$. The FCD of $\boldsymbol{z}_{i\,\mathrm{o}}$ is

$$p(\boldsymbol{z}_{i\,\mathrm{o}}| \cdots) \propto \begin{cases} \mathrm{N}\Big(\boldsymbol{z}_{i\,\mathrm{o}}|\, \mathrm{E}(\boldsymbol{z}_{i\,\mathrm{o}}|\boldsymbol{z}_{i+}, \boldsymbol{B}, \boldsymbol{V}), \operatorname{var}(\boldsymbol{z}_{i\,\mathrm{o}}|\boldsymbol{z}_{i+}, \boldsymbol{B}, \boldsymbol{V})\Big) \\ \qquad\qquad\qquad\qquad \text{if some elements of } \boldsymbol{y}_i \text{ are zero} \\ \mathrm{N}(\boldsymbol{B}'\boldsymbol{x}_i, \boldsymbol{V})1(\boldsymbol{z}_{i\,\mathrm{o}} \leq \boldsymbol{0}) \quad \text{if } \boldsymbol{y}_i = \boldsymbol{0}, \end{cases}$$

where $1(\cdot)$ is an indicator function, and

$$\mathrm{E}(\boldsymbol{z}_{i\,\mathrm{o}}|\boldsymbol{z}_{i+},\boldsymbol{B},\boldsymbol{V}) = \boldsymbol{B}'_{i\,\mathrm{o}}\boldsymbol{x}_i + \boldsymbol{V}_{i\,\mathrm{o}+}\boldsymbol{V}_{i++}^{-1}(\boldsymbol{z}_{i+} - \boldsymbol{B}'_{i+}\boldsymbol{x}_i)$$
$$\mathrm{var}(\boldsymbol{z}_{i\,\mathrm{o}}|\boldsymbol{z}_{ij+},\boldsymbol{B},\boldsymbol{V}) = \boldsymbol{V}_{i\,\mathrm{oo}} - \boldsymbol{V}_{i\,\mathrm{o}+}\boldsymbol{V}_{i++}^{-1}\boldsymbol{V}_{i+\mathrm{o}}. \;\; \|$$

2.2. Two Independent Compositional Data

We now consider having two independent compositional data (e.g., time use data for females and males), that is, $\{Y_{jk}\}$, $j = 1, 2$. We assume that

$$Y_{j1} + Y_{j2} + \cdots + Y_{j,D-1} + Y_{jD} = 1,\;\; 0 \le Y_{jk} \le 1,\;\; k = 1, \ldots, D,\;\; j = 1, 2.$$

By introducing latent variables $\boldsymbol{z}_j = (z_{j1}, \ldots, z_{jd})'$ into $\boldsymbol{Y}_j = (Y_{j1}, \ldots, Y_{jd})'$, we have

$$Y_{jk} = \frac{(\max(0, z_{jk}))^{\gamma}}{1 + \sum_{k'=1}^{d}(\max(0, z_{jk'}))^{\gamma}},\;\; k = 1, \ldots, d. \tag{8}$$

Writing the mean of z_{jk} as $\mathrm{E}(z_{jk}) = \boldsymbol{x}'_j\boldsymbol{\beta}_{jk}$ and the covariance matrix of $\boldsymbol{z}_j$ as $\boldsymbol{V}_j$, we set $\boldsymbol{z}_j \sim \mathrm{N}_d(\boldsymbol{B}'_j\boldsymbol{x}_j, \boldsymbol{V}_j)$, where $\boldsymbol{B}_j = (\boldsymbol{\beta}_{j1}, \ldots, \boldsymbol{\beta}_{jd})$ and $\boldsymbol{\beta}_{jk}$ is a $p \times 1$ vector. Further, we define $y_{jk} = (Y_{jk}/Y_{jD})^{1/\gamma}$ and $\boldsymbol{y}_j = (y_{j1}, \ldots, y_{jd})'$.

Now, let $\boldsymbol{y}_{1j}, \ldots, \boldsymbol{y}_{nj}$ denote an i.i.d. sample of $\boldsymbol{y}_j$ and $\boldsymbol{z}_{1j}, \ldots, \boldsymbol{z}_{nj}$ be the vectors of $\boldsymbol{z}_j$ corresponding to $\boldsymbol{y}_{1j}, \ldots, \boldsymbol{y}_{nj}$. By specifying the prior distributions of the models for $j = 1, 2$, as in (5) to (7),

$$\boldsymbol{z}_{ij} \sim \mathrm{N}_d(\boldsymbol{B}'_j\boldsymbol{x}_{ij}, \boldsymbol{V}_j),\;\; i = 1, \ldots, n \tag{9}$$
$$\mathrm{vec}\,\boldsymbol{B}_j \sim \mathrm{N}_{dp}\,(\mathrm{vec}\,\boldsymbol{B}_{j0}, \lambda_j\boldsymbol{I}_{dp}) \tag{10}$$
$$\boldsymbol{V}_j \sim \mathrm{IW}_d(m_j, \boldsymbol{M}_j^{-1}), \tag{11}$$

we can apply Algorithm 1 to the models for $j = 1, 2$.

Further, by defining $\boldsymbol{y}_i = \begin{pmatrix} \boldsymbol{y}_{i1} \\ \boldsymbol{y}_{i2} \end{pmatrix}$, the likelihood function can be written as

$$
\begin{aligned}
l(\boldsymbol{B}_1, \boldsymbol{B}_2, \boldsymbol{V}_1, \boldsymbol{V}_2 | \boldsymbol{y}_1, \ldots, \boldsymbol{y}_n) &= \prod_{i=1}^{n} \prod_{j=1}^{2} p(\boldsymbol{y}_{ij} | \boldsymbol{B}_j, \boldsymbol{V}_j) \\
&= \prod_{i=1}^{n} \prod_{j=1}^{2} \phi(\boldsymbol{y}_{ij+} | \boldsymbol{B}'_{ij+} \boldsymbol{x}_{ij}, \boldsymbol{V}_{ij++}) \\
&\times \Phi \left(\mathrm{E}(\boldsymbol{y}_{ij\,\mathrm{o}} | \boldsymbol{y}_{ji+}, \boldsymbol{B}_{ij}, \boldsymbol{V}_{ij}), \mathrm{var}(\boldsymbol{y}_{ij\,\mathrm{o}} | \boldsymbol{y}_{ij+}, \boldsymbol{B}_{ij}, \boldsymbol{V}_{ij}) \right),
\end{aligned} \tag{12}
$$

where

$$
\begin{aligned}
\boldsymbol{y}_{ij} &= \begin{pmatrix} \boldsymbol{y}_{ij+} \\ \boldsymbol{y}_{ij\,\mathrm{o}} \end{pmatrix}, \ \boldsymbol{B}_{ij} = (\boldsymbol{B}_{ij+}, \boldsymbol{B}_{ij\,\mathrm{o}}), \ \boldsymbol{V}_{ij} = \begin{pmatrix} \boldsymbol{V}_{ij++} & \boldsymbol{V}_{ij+\mathrm{o}} \\ \boldsymbol{V}_{ij\,\mathrm{o}+} & \boldsymbol{V}_{ij\,\mathrm{oo}} \end{pmatrix} \\
\mathrm{E}(\boldsymbol{y}_{ij\,\mathrm{o}} | \boldsymbol{y}_{ij+}, \boldsymbol{B}_{ij}, \boldsymbol{V}_{ij}) &= \boldsymbol{B}'_{ij\,\mathrm{o}} \boldsymbol{x}_{ij} + \boldsymbol{V}_{ij\,\mathrm{o}+} \boldsymbol{V}^{-1}_{ij++} (\boldsymbol{y}_{ij+} - \boldsymbol{B}'_{ij+} \boldsymbol{x}_{ij}) \\
\mathrm{var}(\boldsymbol{y}_{ij\,\mathrm{o}} | \boldsymbol{y}_{ij+}, \boldsymbol{B}_{ij}, \boldsymbol{V}_{ij}) &= \boldsymbol{V}_{ij\,\mathrm{oo}} - \boldsymbol{V}_{ij\,\mathrm{o}+} \boldsymbol{V}^{-1}_{ij++} \boldsymbol{V}_{ij+\mathrm{o}}.
\end{aligned}
$$

Note that the partitions of $\boldsymbol{B}_{ij}$ and $\boldsymbol{V}_{ij}$ depend on the partition of $\boldsymbol{y}_{ij}$.

2.3. Two Dependent Compositional Data

In this subsection, we consider a case in which $\boldsymbol{z}_1$ and $\boldsymbol{z}_2$ are correlated. We assume that $\boldsymbol{z}_i = \begin{pmatrix} \boldsymbol{z}_{i1} \\ \boldsymbol{z}_{i2} \end{pmatrix}$ is normally distributed as follows:

$$
\boldsymbol{z}_i = \begin{pmatrix} \boldsymbol{z}_{i1} \\ \boldsymbol{z}_{i2} \end{pmatrix} \sim \mathrm{N}_{2d} \left(\begin{pmatrix} \boldsymbol{B}'_1 \boldsymbol{x}_{i1} \\ \boldsymbol{B}'_2 \boldsymbol{x}_{i2} \end{pmatrix}, \boldsymbol{V} \right), \tag{13}
$$

where $\boldsymbol{V}$ is a $(2d \times 2d)$ covariance matrix. For the prior distribution specification of $\boldsymbol{B}_j$, we use prior distribution (10), and for that of $\boldsymbol{V}$, we use

$$
\boldsymbol{V} \sim \mathrm{IW}_{2d}(m, \boldsymbol{M}^{-1}) \tag{14}
$$

instead of (11).

We partition the elements of $\boldsymbol{y}_i$ into two terms, $\boldsymbol{y}_{i+}$ and $\boldsymbol{y}_{i\,\mathrm{o}}$, which contain strictly positive elements and zero elements, respectively. For convenience of expression, we replace the elements of $\boldsymbol{y}_{i+}$ as the first part of $\boldsymbol{y}_i$, that is, $\boldsymbol{y}_i = \begin{pmatrix} \boldsymbol{y}_{i+} \\ \boldsymbol{y}_{i\,\mathrm{o}} \end{pmatrix}$. According to the order of elements of $\boldsymbol{y}_i$, we partition $\boldsymbol{B} = (\boldsymbol{B}_1, \boldsymbol{B}_2)$ and $\boldsymbol{V}$ as follows:

$$\boldsymbol{B}_i = (\boldsymbol{B}_{i1+}, \boldsymbol{B}_{i2+}, \boldsymbol{B}_{i1\,\mathrm{o}}, \boldsymbol{B}_{i2\,\mathrm{o}}),\ \ \boldsymbol{V}_i = \begin{pmatrix} \boldsymbol{V}_{i++} & \boldsymbol{V}_{i+\mathrm{o}} \\ \boldsymbol{V}_{i\,\mathrm{o}+} & \boldsymbol{V}_{i\,\mathrm{o}\,\mathrm{o}} \end{pmatrix}.$$

Note that the partitions of $\boldsymbol{B}$ and $\boldsymbol{V}$ depend on the partition of $\boldsymbol{y}_i$. Then, the density function of $\boldsymbol{y}_i$ will be

$$\begin{aligned} p(\boldsymbol{y}_i|\boldsymbol{B}_i, \boldsymbol{V}_i) = p(\boldsymbol{y}_{i+}, \boldsymbol{y}_{i\,\mathrm{o}}|\boldsymbol{B}_i, \boldsymbol{V}_i) &= p(\boldsymbol{y}_{i+}|\boldsymbol{B}_i, \boldsymbol{V}_i)\,\mathrm{Pr}(\boldsymbol{y}_{i\,\mathrm{o}}|\boldsymbol{y}_{i+}, \boldsymbol{B}_i, \boldsymbol{V}_i) \\ &= \phi\left(\boldsymbol{y}_{i+} \mid \begin{pmatrix} \boldsymbol{B}'_{i1+}\boldsymbol{x}_{i1} \\ \boldsymbol{B}'_{i2+}\boldsymbol{x}_{i2} \end{pmatrix}, \boldsymbol{V}_{i++}\right) \\ &\times \Phi\left(\mathrm{E}(\boldsymbol{y}_{i\,\mathrm{o}}|\boldsymbol{y}_{j+}, \boldsymbol{B}_i, \boldsymbol{V}_i), \mathrm{var}(\boldsymbol{y}_{i\,\mathrm{o}}|\boldsymbol{y}_{j+}, \boldsymbol{B}_i, \boldsymbol{V}_i)\right), \end{aligned} \tag{15}$$

where

$$\begin{aligned} \mathrm{E}(\boldsymbol{y}_{i\,\mathrm{o}}|\boldsymbol{y}_{i+}, \boldsymbol{B}_i, \boldsymbol{V}_i) &= \begin{pmatrix} \boldsymbol{B}'_{i1\,\mathrm{o}}\boldsymbol{x}_{i1} \\ \boldsymbol{B}'_{i2\,\mathrm{o}}\boldsymbol{x}_{i2} \end{pmatrix} + \boldsymbol{V}_{i\,\mathrm{o}+}\boldsymbol{V}_{i++}^{-1}\left(\boldsymbol{y}_{i+} - \begin{pmatrix} \boldsymbol{B}'_{i1+}\boldsymbol{x}_{i1} \\ \boldsymbol{B}'_{i2+}\boldsymbol{x}_{i2} \end{pmatrix}\right) \\ \mathrm{var}(\boldsymbol{y}_{i\,\mathrm{o}}|\boldsymbol{y}_{i+}, \boldsymbol{B}_i, \boldsymbol{V}_i) &= \boldsymbol{V}_{i\,\mathrm{o}\,\mathrm{o}} - \boldsymbol{V}_{i\,\mathrm{o}+}\boldsymbol{V}_{i++}^{-1}\boldsymbol{V}_{i+\mathrm{o}}. \end{aligned}$$

Further, the likelihood function can be written as

$$\begin{aligned} l(\boldsymbol{B}_1, \boldsymbol{B}_2, \boldsymbol{V}|\boldsymbol{y}_1, \ldots, \boldsymbol{y}_n) &= \prod_{i=1}^{n} p(\boldsymbol{y}_i|\boldsymbol{B}_1, \boldsymbol{B}_2, \boldsymbol{V}) \\ &= \prod_{i=1}^{n} \phi\left(\boldsymbol{y}_{i+} \mid \begin{pmatrix} \boldsymbol{B}'_{i1+}\boldsymbol{x}_{i1} \\ \boldsymbol{B}'_{i2+}\boldsymbol{x}_{i2} \end{pmatrix}, \boldsymbol{V}_{i++}\right) \\ &\times \Phi\left(\mathrm{E}(\boldsymbol{y}_{i\,\mathrm{o}}|\boldsymbol{y}_{j+}, \boldsymbol{B}_i, \boldsymbol{V}_i), \mathrm{var}(\boldsymbol{y}_{i\,\mathrm{o}}|\boldsymbol{y}_{j+}, \boldsymbol{B}_i, \boldsymbol{V}_i)\right). \end{aligned} \tag{16}$$

After some manipulation, we derive the following algorithm. Appendix A.2 in Supplemental material presents details of the derivation.

Algorithm 2. Define $\boldsymbol{X}_j = (\boldsymbol{x}_{1j}, \ldots, \boldsymbol{x}_{nj})'$, $\boldsymbol{Z}_j = (\boldsymbol{z}_{1j}, \ldots, \boldsymbol{z}_{nj})'$ for $j = 1, 2$, and $\boldsymbol{V}^{-1} = \begin{pmatrix} \boldsymbol{V}^{11} & \boldsymbol{V}^{12} \\ \boldsymbol{V}^{12\prime} & \boldsymbol{V}^{22} \end{pmatrix}$. The FCD of $\operatorname{vec} \boldsymbol{B}_1$ is

$$\operatorname{vec} \boldsymbol{B}_1 | \cdots \sim \mathrm{N}\left(\boldsymbol{\beta}_1^*, \boldsymbol{\Omega}_1^*\right),$$

where

$$\begin{aligned} \boldsymbol{\Omega}_1^* &= \left[\lambda_1^{-1} \boldsymbol{I} + (\boldsymbol{V}^{11} \otimes \boldsymbol{X}_1' \boldsymbol{X}_1)\right]^{-1} \\ \boldsymbol{\beta}_1^* &= \boldsymbol{\Omega}_1^* \left[\lambda_1^{-1} \operatorname{vec} \boldsymbol{B}_{10} + (\boldsymbol{V}^{11} \otimes \boldsymbol{I}) \operatorname{vec}(\boldsymbol{X}_1' \boldsymbol{Z}_1) \right. \\ &\quad \left. + (\boldsymbol{V}^{12} \otimes \boldsymbol{I}) \operatorname{vec}[\boldsymbol{X}_1' (\boldsymbol{Z}_2 - \boldsymbol{X}_2 \boldsymbol{B}_2)]\right]. \end{aligned}$$

The FCD of $\operatorname{vec} \boldsymbol{B}_2$ is

$$\operatorname{vec} \boldsymbol{B}_2 | \cdots \sim \mathrm{N}\left(\boldsymbol{\beta}_2^*, \boldsymbol{\Omega}_2^*\right),$$

where

$$\begin{aligned} \boldsymbol{\Omega}_2^* &= \left[\lambda_2^{-1} \boldsymbol{I} + (\boldsymbol{V}^{22} \otimes \boldsymbol{X}_2' \boldsymbol{X}_2)\right]^{-1} \\ \boldsymbol{\beta}_2^* &= \boldsymbol{\Omega}_2^* \left[\lambda_2^{-1} \operatorname{vec} \boldsymbol{B}_{20} + (\boldsymbol{V}^{22} \otimes \boldsymbol{I}) \operatorname{vec}(\boldsymbol{X}_2' \boldsymbol{Z}_2) \right. \\ &\quad \left. + (\boldsymbol{V}^{12\prime} \otimes \boldsymbol{I}) \operatorname{vec}[\boldsymbol{X}_2' (\boldsymbol{Z}_1 - \boldsymbol{X}_1 \boldsymbol{B}_1)]\right]. \end{aligned}$$

The FCD of $\boldsymbol{V}$ is

$$\boldsymbol{V} | \cdots \sim \mathrm{IW}\left(m^*, \boldsymbol{M}^{*-1}\right),$$

where

$$m^* = m + n,\ \ \boldsymbol{M}^* = \boldsymbol{M} + \sum \begin{pmatrix} \boldsymbol{z}_{i1} - \boldsymbol{B}_1'\boldsymbol{x}_{i1} \\ \boldsymbol{z}_{i2} - \boldsymbol{B}_2'\boldsymbol{x}_{i2} \end{pmatrix} \begin{pmatrix} \boldsymbol{z}_{i1} - \boldsymbol{B}_1'\boldsymbol{x}_{i1} \\ \boldsymbol{z}_{i2} - \boldsymbol{B}_2'\boldsymbol{x}_{i2} \end{pmatrix}'.$$

We partition the elements of $\boldsymbol{z}_i$ into two terms, $\boldsymbol{z}_{i+}$ and $\boldsymbol{z}_{i\,\mathrm{o}}$, which contain strictly positive elements and non-positive elements, respectively. For convenience of expression, we replace the elements of $\boldsymbol{z}_{i+}$ as the first part of $\boldsymbol{z}_i$, that is, $\boldsymbol{z}_i = \begin{pmatrix} \boldsymbol{z}_{i+} \\ \boldsymbol{z}_{i\,\mathrm{o}} \end{pmatrix}$. The FCD of $\boldsymbol{z}_{i\,\mathrm{o}}$ is

$$p(\boldsymbol{z}_{i\,\mathrm{o}}|\cdots) \propto \begin{cases} \mathrm{N}\Big(\boldsymbol{z}_{i\,\mathrm{o}}|\,\mathrm{E}(\boldsymbol{z}_{i\,\mathrm{o}}|\boldsymbol{z}_{i+}, \boldsymbol{B}_i, \boldsymbol{V}_i), \mathrm{var}(\boldsymbol{z}_{i\,\mathrm{o}}|\boldsymbol{z}_{i+}, \boldsymbol{B}_i, \boldsymbol{V}_i)\Big) & \\ & \text{if some elements of } \boldsymbol{y}_i \text{ are zero} \\ \mathrm{N}(\boldsymbol{B}_i'\boldsymbol{x}_i, \boldsymbol{V}_i)1(\boldsymbol{z}_{i\,\mathrm{o}} \leq \boldsymbol{0}) & \text{if } \boldsymbol{y}_i = \boldsymbol{0}, \end{cases}$$

where

$$\mathrm{E}(\boldsymbol{z}_{i\,\mathrm{o}}|\boldsymbol{z}_{i+}, \boldsymbol{B}_i, \boldsymbol{V}_i) = \begin{pmatrix} \boldsymbol{B}_{i1\,\mathrm{o}}'\boldsymbol{x}_{i1} \\ \boldsymbol{B}_{i2\,\mathrm{o}}'\boldsymbol{x}_{i2} \end{pmatrix} + \boldsymbol{V}_{ij\,\mathrm{o}+}\boldsymbol{V}_{ij++}^{-1}\left(\boldsymbol{z}_{i+} - \begin{pmatrix} \boldsymbol{B}_{i1+}'\boldsymbol{x}_{i1} \\ \boldsymbol{B}_{i2+}'\boldsymbol{x}_{i2} \end{pmatrix}\right)$$

$$\mathrm{var}(\boldsymbol{z}_{i\,\mathrm{o}}|\boldsymbol{z}_{i+}, \boldsymbol{B}_i, \boldsymbol{V}_i) = \boldsymbol{V}_{i\,\mathrm{oo}} - \boldsymbol{V}_{i\,\mathrm{o}+}\boldsymbol{V}_{i++}^{-1}\boldsymbol{V}_{i+\mathrm{o}}. \ \ \|$$

2.4. Marginal Likelihood Functions and the Bayes Factor

The Bayes factor is a model selection tool in Bayesian analysis. In this subsection, we derive the estimated Bayes factor for two models: the model of two independent compositional data, which we denote model 1, $\mathcal{M}_1$, and the model of two dependent compositional data, which we denote model 2, $\mathcal{M}_2$. Using Chib's method [5], we can evaluate the marginal likelihood in both models.

In general, if $\boldsymbol{\theta}$ denotes the vector of parameters included in model $\mathcal{M}$, the marginal likelihood will be defined as

$$m(\boldsymbol{y}|\mathcal{M}) = \frac{p(\boldsymbol{y}|\boldsymbol{\theta}, \mathcal{M})p(\boldsymbol{\theta}|\mathcal{M})}{p(\boldsymbol{\theta}|\boldsymbol{y}, \mathcal{M})}, \tag{17}$$

where $p(\boldsymbol{y}|\boldsymbol{\theta}, \mathcal{M})$, $p(\boldsymbol{\theta}|\mathcal{M})$, and $p(\boldsymbol{\theta}|\boldsymbol{y}, \mathcal{M})$ denote the likelihood function and the prior and posterior densities under model $\mathcal{M}$, respectively. Taking the logarithm of (17) and evaluating it at a given value of $\boldsymbol{\theta}$, $\hat{\boldsymbol{\theta}}$, we have

$$\log m(\boldsymbol{y}|\mathcal{M}) = \log p(\boldsymbol{y}|\hat{\boldsymbol{\theta}}, \mathcal{M}) + \log p(\hat{\boldsymbol{\theta}}|\mathcal{M}) - \log p(\hat{\boldsymbol{\theta}}|\boldsymbol{y}, \mathcal{M}). \tag{18}$$

We derive the estimates of the marginal likelihood $m(\boldsymbol{y}|\mathcal{M}_1)$ and $m(\boldsymbol{y}|\mathcal{M}_2)$, that is, $\hat{m}(\boldsymbol{y}|\mathcal{M}_1)$ and $\hat{m}(\boldsymbol{y}|\mathcal{M}_2)$ (Appendix B in Supplemental material).

The Bayes factor of $\mathcal{M}_2$ against $\mathcal{M}_1$ is defined as

$$B_{21}(\boldsymbol{y}) = \frac{m(\boldsymbol{y}|\mathcal{M}_2)}{m(\boldsymbol{y}|\mathcal{M}_1)}. \tag{19}$$

By substituting the estimates of $m(\boldsymbol{y}|\mathcal{M}_1)$ and $m(\boldsymbol{y}|\mathcal{M}_2)$ into (19), we have the following estimated Bayes factor:

$$\hat{B}_{21}(\boldsymbol{y}) = \frac{\hat{m}(\boldsymbol{y}|\mathcal{M}_2)}{\hat{m}(\boldsymbol{y}|\mathcal{M}_1)}. \tag{20}$$

In the next section, we calculate the estimated Bayes factor (20) using time use data.

3. Empirical Studies on Time Use Data

3.1. Data

We use microlevel data from the JPSC, wave 18 (2010), conducted by the Institute for Research on Household Economics. The JPSC is a panel survey of many aspects of women's lifestyles, such as marriage, family relationships, income, expenditures, employment status, education, and life consciousness. If the respondents are married, their husbands' age, income, employment status, and so on are also surveyed. In the

JPSC, the living hours of respondents, and of their husbands if they are married, are surveyed per workday and non-workday. The daily living hours are classified into "commuting," "work," "schoolwork (studies)," "housekeeping and childcare," "hobby, leisure, social interaction, etc.," and "other activities such as sleeping, meals, taking a bath, etc.," which hereafter we call "basic living hours."

In this study, we use the workday living hours data of married respondents and their husbands. These data are represented in hours and minutes in the JPSC. We convert these data into minutes and choose data representing the total activities in $1,440$ minutes, that is, 24 hours. For the categories of "commuting," "work," "schoolwork (studies)," "housekeeping and childcare," and "hobby, leisure, social interaction, etc.," if the respondents' data values minus their means are greater than three times their standard deviations, the data of such respondents are treated as outliers. For the "basic living hours" data, if the absolute deviation from the mean is more than three times the standard deviation, such data are also treated as outliers. Such data processing is often used for JPSC data (e.g. [7]). We obtain the living hours ratio data represented in minutes by dividing each activity's hours by $1,440$ minutes.

Furthermore, we use data of the number of children aged three years or less, respondents' and their husbands' incomes, whether the respondents live with their own parent/parents and/or their husbands' parent/parents (hereinafter, "living with parents"), and the education level of the couples (hereinafter, "education level"). The JPSC surveys the annual income of the respondents and their husbands in the previous year. We use this as the income data. For the data of "living with parents," the variable takes the value of 0 when the respondent and her husband do not live with their parents and 1 when they do. The respondents' and their husbands' education level data are calculated using their highest academic achievements as per the JPSC. The education level is classified as junior high school, high school, vocational school, junior college, specialized high school, college (includes medical and dental school), and graduate school, with values of 9, 12, 15, 14, 14, 16, and 18, respectively, as data for the education level. Table 1 shows the summary statistics.

3.2. Empirical Analysis

The living hours data consist of six categories of daily activities in this empirical analysis. Explained variables Y_1, Y_2, Y_3, Y_4, Y_5, and Y_6 denote "commuting" (`commute`), "work" (`work`), "schoolwork (studies)" (`study`), "housekeeping and childcare" (`housework`), "hobby, leisure, social interaction, etc." (`leisure`), and "basic living hours" (`basis`), respectively. The data for these variables are proportional data.

For estimation, we use the following time equations pertaining to latent variables z_k of wives' and husbands' living hours:

$$\mathrm{E}(z_k) = \beta_{k1} + \beta_{k2}\texttt{age} + \beta_{k3}\texttt{nc3} + \beta_{k4}\log(\texttt{w\_inc}) + \beta_{k5}\log(\texttt{h\_inc}) \\ +\beta_{k6}\texttt{educ} + \beta_{k7}\texttt{lwp},\ k = 1, \dots, 5, \quad (21)$$

where `age` and `educ` denote the age and education level, respectively, of wives and husbands.

The latent variables z_k for $\gamma = 2$ are defined as follows:

$$z_1 = \sqrt{\frac{\texttt{commute}}{\texttt{basis}}},\ z_2 = \sqrt{\frac{\texttt{work}}{\texttt{basis}}},\ z_3 = \sqrt{\frac{\texttt{study}}{\texttt{basis}}},\ z_4 = \sqrt{\frac{\texttt{housework}}{\texttt{basis}}} \\ z_5 = \sqrt{\frac{\texttt{leisure}}{\texttt{basis}}}.$$

The explanatory variables are the age of wives (`w_age`), age of husbands (`h_age`), number of children up to three years of age (`nc3`), income of wives (`w_inc`), income of husbands (`h_inc`), wives' education level (`w_educ`), husbands' education level (`h_educ`), and dummy variables for living with parents (`lwp`).

For the MCMC simulation, we set the hyperparameters of prior distribution in the two independent compositional data model ($\mathcal{M}_1$) as follows:

$$\lambda_1 = 10,\ \lambda_2 = 10,\ m = D + 1,\ M = 2\boldsymbol{I}_d.$$

Similarly, we set those in the two dependent compositional data model ($\mathcal{M}_2$) as follows:

$$\lambda_1 = 10,\ \lambda_2 = 10,\ m = 2(D+1),\ M = 2\boldsymbol{I}_{2d}.$$

We run the MCMC simulation for $30,000$ iterations; the first $10,000$ samples are discarded as the burn-in period, and the remaining $20,000$ samples are used to analyze the posterior results.

3.2.1. Estimation results

We estimate the time equations (21) for models $\mathcal{M}_1$ and $\mathcal{M}_2$ and examine which model should be selected. Table 2 shows the marginal likelihood and Bayes factor from their posterior results. According to this table, $\log_{10} \hat{B}_{21} = 49.994$. This implies that the evidence against $\mathcal{M}_1$ is *decisive* [14, p.228]. Therefore, we select the two dependent compositional data model ($\mathcal{M}_2$).

Table 3 shows the estimated results. From a comparison of the number of explanatory variables' coefficients whose 95% or 90% credible intervals (CIs) do not include zeros, wives show more such coefficients than husbands. This could be because wives responded to the questions on behalf of their husbands, as the JPSC is a survey of women. This may also suggest that the factors determining the wives' hours for each activity are different from those of husbands. Especially, the income of wives affects two activity hours ratios of husbands, in addition to all the five activity hours ratios of wives.

Furthermore, there are many coefficients whose 95% or 90% CIs do not include zeros in the time equations of the study hours (`study`) and housekeeping hours (`housework`) of wives. These activity hours are affected by their own age, households' economic situation such as wives' and husbands' income, educational level, and whether they live with young children or their parents.

Table 4 shows the posterior results of the partial correlation coefficients between the latent variables. In this table, the corresponding latent variables of wives and husbands, such as wives' `commute` and husbands' `commute`, except `housework`, are correlated. Wives' `housework` is correlated with their husbands' `work` and `leisure`.

3.2.2. Partial effects

We showed the effects of various attributes of wives and husbands on the time use ratio of their activities to basic living hours. The model selection in Table 2 shows that the time spent on activities by wives is correlated with that by their husbands. Therefore, the variables for wives, such as age and income, are likely to have some effect on not only their activity hours but also their husbands' activity hours, as is previously indicated. Thus, we estimate the partial effect of change in the explanatory variables in time equations (21) on their activity hours and the probability that such partial effect is positive. This probability is calculated from the average partial effect (APE), as shown in Appendix C in Supplemental material. The result is shown in Table 5.

When the correlations between wives' and husbands' time use data are considered, the probabilities are shown outside of the parentheses. We also show the probabilities in parentheses when those correlations are not considered. Comparing the two cases, the probabilities of some variables such as wives' education level differ by a few percentage points between the two cases. Furthermore, we can find whether the explanatory variables not included in the wives' equation, such as husbands' age and education level, influence the wives' activity hours.

According to the results of the case that considers the correlations between wives' and husbands' time use data, the probability of the number of children up to three years of age (`nc3`) having positive partial effects on the hours of wives' activities and the probability of husbands' activities are either high or low in many cases. On the other hand, we find a bias that the probability of living with parents (`lwp`) having positive partial effects on the hours of wives' activities is high but that on the hours of husbands' activities is low, or vice versa in some cases. Further, we cannot find high or low probabilities of wives' and husbands' education level (`w_educ` and `h_educ`) having positive partial effects on the hours of their spouses' activities, respectively, as the probabilities of `w_educ` and `h_educ` having positive partial effects are about 0.5 in many cases.

Table 6 summarizes the information in Tables 3 and 5. The first column shows the explanatory variables in the time equation (21). The second and third columns present a summary of Table 3. Wives' and husbands' activities, for which 95% or 90% CIs of the explanatory variables' coefficients do not include zero, are listed. "+" in a parenthesis denotes that the sign of the explanatory variable's coefficient is positive, and "-" in a parenthesis denotes the opposite. The fourth and fifth columns present the summary of Table 5. They show activities for which the probability that the APE due to the change in explanatory variables is positive is high or low. "H" denotes that the probability of a positive APE is higher than 0.7, and "L" denotes that the probability is lower than 0.3.

According to Table 6, we obtained some noteworthy findings. Although the coefficients of the wives' and husbands' ages and incomes (`w_age`, `w_inc`, `h_age`, and `h_inc`) are positive or negative in time equations (21), those explanatory variables do not have obvious effects on the probability that the APEs are positive.

The coefficients of the number of children up to three years of age (`nc3`) are positive in the time equations of the wives' and husbands' housekeeping hours (`housework`) and negative in the time equations of their leisure hours (`leisure`), based on both time equations and the probability of positive APEs. In addition, the probability of a positive APE on the basic living hours is low and that on the wives' work and study hours (`work` and `study`) is high. Furthermore, the coefficient of living with parents (`lwp`) is positive in the time equation of wives' housekeeping hours (`housework`), and the probability of a positive APE on their `housework` is high.

These might be results that most people acknowledge, but the following points give us further tasks to examine. An increase in the number of younger children seems to reduce the leisure hours of wives and husbands to compensate for their housekeeping and childcare hours. However, the probability of an increase in wives' work and study hours is high, but that for husbands' study hours is low. The cause of such different behavior is a subject worth investigating.

With regard to wives' and husbands' education level, wives seem to increase their study and leisure hours instead of housekeeping hours. On the other hand, husbands decrease their leisure hours, but their work and study hours seem to have a slightly higher probability of increasing, as the probabilities of husbands' work and study hours increasing are 0.606 and 0.572, respectively, as shown in Table 5. We observe differences in responses to increasing the education level between wives and husbands.

We can confirm that living with parents still takes a toll on the housekeeping hours of wives but does not affect husbands. [16, p.217] state that "Japanese husbands are more likely to benefit from co-residence with parents than their wives and other household members," because living with parents decreases husbands' share in housework in Japan, but not in the United States or Korea. In our study, we also observe this asymmetry in the burden that wives and husbands shoulder.

4. Concluding remarks

This study provided a statistical method to use compositional data including zeros by extending the Bayesian method suggested in [12] and analyzed the living hours of married couples based on their time use data.

The distinguishing feature of our empirical analysis is that we estimated the compositional data including many zeros with two correlations using the Bayesian method. We suppose that there are correlations among not only the time use data of individual activities but also activities of wives and husbands. We used the Bayesian method based on [12] to estimate the model with such correlations, as this estimation is difficult using the frequentist methods.

In the posterior results on living hours, the signs of the explanatory variables' coefficients show that an increase in the number of younger children increases the housekeeping hours and decreases the leisure hours of both wives and husbands. Living with parents increases the housekeeping hours of wives. These are thought to be generally anticipated results. In addition, increases in the income of wives and husbands have positive effects on their commuting hours. This implies that an increase in commuting hours is accepted as leading to higher income. The negative effect of an increase in wives' income on their housekeeping hours suggests that it enables them to purchase more housekeeping services. Regarding effects of living with parents, we do not consider whose parents the couples live with. Whether living with a spouse's parents would make a difference is a topic for future study.

As the APE, we calculated the probability that the activity hours of wives and husbands increase by one unit of explanatory variables in the time equations. We obtained the following results, which are similar to those in the estimation of time equations. For example, the probability that an increase in the number of younger children extends the housekeeping hours of wives and husbands is high, whereas the probability that it extends their leisure hours is low. The probability that living with parents increases the housekeeping hours of wives is high.

Acknowledgement

The authors appreciate the comments of two anonymous referees, which improve the article greatly. They are grateful to the Institute for Research on Household Economics for providing the micro-data of the Japanese Panel Survey of Consumers (JPSC). Further, the work of the first author was supported in part by a Grant-in-Aid for Scientific Research (No.16K03589) 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.17K03668) from the JSPS.

Disclosure Statement

Funding: The work of the first author was supported in part by a Grant-in-Aid for Scientific Research (No.16K03589) 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.17K03668) from the JSPS.

Conflict of Interest: The authors declare that they have no conflict of interest.

References

[1] K.M. Abadir and J.R. Magnus, *Matrix Algebra*, Cambridge University Press, Cambridge, 2005.

[2] J. Aitchison, *The Statistical Analysis of Compositional Data*, Chapman and Hall, London. 1986.

[3] J. Baxter, B. Hewitt, and M. Haynes, *Life course transitions and housework: Marriage, parenthood, and time on housework*, J. Marriage. and Fam. 70 (2008), pp. 259–272.

[4] A. Butler, and C. Glasbey. *A latent Gaussian model for compositional data with zeros*, J. R. Stat. Soc. Ser. C. Appl. Stat. 57 (2008), pp. 505–520.

[5] S. Chib, *Marginal likelihood from the Gibbs output*, J. Amer. Statist. Assoc. 90 (1995), pp. 1313–1321.

[6] J. M. Fry, T. R. L. Fry, and K. R. McLaren, *Compositional data analysis and zeros in micro data*, Appl Econ 32 (2000), pp. 953–959.

[7] S. Fukuda, *Determinants of the time spent on childcare and household tasks in the life courses of Japanese men and women: An evidence from pooled time-series analyses using a 14-wave panel data (in Japanese)*, Japanese Journal of Research on Household Economics 76 (2007), pp. 26–36.

[8] A. Gelman, J. B. Carlin, H. S. Stern, D. B. Dunson, A. Vehtari, and D. B. Rubin, *Bayesian Data Analysis, 3rd ed.*, CRC Press, Boca Raton, 2014.

[9] J. González Chapela, *Things that make us different: Analysis of deviance with time-use data*, J. Appl. Stat. 40 (2013), pp. 1572–1585.

[10] D. S. Hamermesh, and G. A. Pfann, *Time-use data in economics*, Eur. Econ. Rev. 49 (2005), pp. 1–7.

[11] A. B. Krueger, D. Kahneman, C. Fischler, D. Schkade, N. Schwarz, and A. A. Stone, *Time use and subjective well-being in France and the U.S.*, Soc. Indic. Res. 93 (2009), pp. 7–18.

[12] T. J. Leininger, A. E. Gelfand, J. M. Allen, and J. A. Silander Jr, *Spatial regression modeling for compositional data with many zeros*, J. Agric. Biol. Environ. Stat. 18 (2013), pp. 314–334.

[13] J. A. Martín-Fernández, J. A., J. Palarea-Albaladejo, and R. A. Olea, *Dealing with zeros*, in *Compositional Data Analysis: Theory and Applications*, V. Pawlowsky-Glahn and A. Buccianti eds., Wiley, Chichester, 2011, pp. 43–58.

[14] C. P. Robert, *The Bayesian Choice: From Decision-Theoretic Foundations to Computa-*

tional Implementation, 2nd ed., Springer, New York. 2001

[15] L. S. Stratton, *The role of preferences and opportunity costs in determining the time allocated to housework*. Am. Econ. Rev.: Papers and Proceedings 102 (2012), pp. 606–611.

[16] N. O. Tsuya, L. L. Bumpass, and M. K. Choe, *Gender, employment, and housework in Japan, South Korea, and the United States*, Review of Population and Social Policy 9 (2000), pp. 195–220.

[17] A. Ueda, *Intrafamily time allocation of housework: Evidence from Japan*, J. Jpn. Int. Econ. 19 (2005), pp. 1–23.

[18] J. M. Wooldridge, *Introductory Econometrics: A Modern Approach, 6th ed.*, Cengage Learning, Boston, 2016.

Table 1. Summary statistics

Living hours of activities in a workday (minutes): respondents	Mean[a]	SD	Min	Q1	Median	Q3	Max
commuting	39.96	30.97	0	20	30	60	130
work	416.77	131.09	0	300	420	480	900
school work	7.32	19.51	0	0	0	0	120
housekeeping and childcare	252.31	138.29	0	160	240	320	770
hobby, leisure, social interaction, etc.	112.23	100.00	0	40	110	180	530
basic living hours	611.41	117.36	300	540	600	690	1,000

Living hours of activities in a workday (minutes): husbands	Mean	SD	Min	Q1	Median	Q3	Max
commuting	63.06	44.58	0	30	60	90	180
work	603.64	109.64	240	540	600	660	1,020
school work	6.92	18.17	0	0	0	0	90
housekeeping and childcare	29.89	43.70	0	0	0	60	200
hobby, leisure, social interaction, etc.	131.26	96.35	0	60	120	180	420
basic living hours	605.23	110.62	300	540	600	680	960

Other variables: respondents	Mean	SD	Min	Q1	Median	Q3	Max
age	40.37	7.18	26	35	41	47	51
income (ten thousand yen)	183.71	158.24	3	84	120	240	960
education level (year)	14.55	1.28	9	14	15	15	18

Other variables: husbands	Mean	SD	Min	Q1	Median	Q3	Max
age	42.41	8.08	24	36	43	49	61
income (ten thousand yen)	514.91	290.89	70	350	460	622	4,300
education level (year)	14.71	1.90	9	14	15	16	18

Numbers of children equal to or less than three years of age	Mean	SD	0	1	2
	0.12	0.34	474	56	3

Living with parents: Yes	Living with parents: No
149	384

[a] "Mean," "SD," "Min," "Q1," "Median," "Q3," and "Max" denote sample mean, sample standard deviation, minimum value, the first quartile, median, the third quartile, and maximum value, respectively.

Table 2. Marginal likelihood and Bayes factor

Model	$\log \hat{m}(\boldsymbol{y}\|\mathcal{M})$	$\log_{10} \hat{B}_{21}$
$\mathcal{M}_1$	-211.428	
$\mathcal{M}_2$	-96.313	49.994

Table 3. Posterior results of $\boldsymbol{B}_1$ and $\boldsymbol{B}_2$: Two dependent compositional data

		Wives			Husbands		
		Mean[a]	SD	Median	Mean	SD	Median
				commute			
intercept	β_{11}	−0.0141	0.1019	−0.0140	0.0069	0.0989	0.0065
age[b]	β_{12}	−0.0023	0.0009	−0.0023**[c]	−0.0004	0.0009	−0.0004
nc3	β_{13}	0.0067	0.0185	0.0067	0.0101	0.0207	0.0102
log(w_inc)	β_{14}	0.0338	0.0069	0.0339**	0.0018	0.0077	0.0018
log(h_inc)	β_{15}	0.0299	0.0130	0.0299**	0.0556	0.0151	0.0557**
educ[c]	β_{16}	−0.0005	0.0044	−0.0005	−0.0018	0.0034	−0.0018
lwp	β_{17}	−0.0134	0.0134	−0.0135	−0.0450	0.0153	−0.0451**
				work			
intercept	β_{21}	0.5118	0.1273	0.5113**	0.8197	0.1167	0.8196**
age	β_{22}	−0.0022	0.0011	−0.0022*	−0.0035	0.0011	−0.0035**
nc3	β_{23}	0.0213	0.0230	0.0213	0.0233	0.0241	0.0232
log(w_inc)	β_{24}	0.1029	0.0087	0.1029**	0.0063	0.0090	0.0062
log(h_inc)	β_{25}	−0.0121	0.0162	−0.0120	0.0504	0.0178	0.0502**
educ	β_{26}	−0.0019	0.0056	−0.0019	0.0001	0.0040	0.0001
lwp	β_{27}	0.0084	0.0167	0.0085	−0.0088	0.0177	−0.0087
				study			
intercept	β_{31}	−2.7968	0.6185	−2.7728**	−1.9811	0.5597	−1.9623**
age	β_{32}	0.0099	0.0051	0.0099**	0.0037	0.0049	0.0036
nc3	β_{33}	0.0716	0.0972	0.0721	−0.1235	0.1282	−0.1181
log(w_inc)	β_{34}	0.0874	0.0365	0.0866**	0.0864	0.0417	0.0852**
log(h_inc)	β_{35}	0.1125	0.0700	0.1113*	0.1141	0.0834	0.1119
educ	β_{36}	0.0558	0.0278	0.0550**	0.0112	0.0172	0.0109
lwp	β_{37}	−0.1419	0.0753	−0.1409*	−0.1047	0.0860	−0.1025
				housework			
intercept	β_{41}	1.2951	0.1542	1.2959**	0.0117	0.2457	0.0116
age	β_{42}	−0.0027	0.0014	−0.0027*	−0.0013	0.0023	−0.0013
nc3	β_{43}	0.0832	0.0280	0.0832**	0.2418	0.0477	0.2413**
log(w_inc)	β_{44}	−0.0992	0.0102	−0.0992**	0.0675	0.0187	0.0676**
log(h_inc)	β_{45}	0.0139	0.0196	0.0140	−0.0469	0.0375	−0.0470
educ	β_{46}	−0.0121	0.0068	−0.0120*	−0.0015	0.0088	−0.0015
lwp	β_{47}	0.0414	0.0203	0.0414**	−0.0283	0.0380	−0.0283
				leisure			
intercept	β_{51}	−0.3163	0.2253	−0.3146	0.7355	0.1858	0.7351**
age	β_{52}	0.0024	0.0020	0.0024	0.0004	0.0016	0.0004
nc3	β_{53}	−0.1855	0.0427	−0.1856**	−0.0680	0.0389	−0.0683*
log(w_inc)	β_{54}	−0.0389	0.0154	−0.0388**	−0.0070	0.0144	−0.0069
log(h_inc)	β_{55}	0.0909	0.0294	0.0905**	−0.0225	0.0280	−0.0228
educ	β_{56}	0.0149	0.0091	0.0148	−0.0115	0.0058	−0.0115**
lwp	β_{57}	−0.0142	0.0304	−0.0141	0.0128	0.0286	0.0130

[a] "Mean," "SD," and "Median" denote posterior mean, posterior standard deviation, and posterior median, respectively.
[b] age and educ denote the age and education level of wives in case of wives, and those of husbands in case of husbands.
[c] "**" and "*" denote that zero is not included in the 95% and 90% credible intervals, respectively.

Table 4. Partial correlation matrix of latent variables (posterior median)

		Wives				
		commute	work	study	housework	leisure
Wives	commute					
	work	0.2315**[a]				
	study	0.0159	−0.0414			
	housework	−0.0243	0.1407**	−0.0691		
	leisure	0.0480	−0.0026	0.0454	−0.1554**	
Husbands	commute	0.2468**	−0.0736*	0.0563	0.0755*	0.0605
	work	−0.0104	0.1861**	0.0097	0.1415**	0.1759**
	study	0.0098	0.0030	0.5572**	0.0946	−0.0602
	housework	0.0076	0.0544	−0.0329	0.0617	0.0239
	leisure	−0.0366	0.0333	−0.0396	0.1744**	0.5117**
		Husbands				
		commute	work	study	housework	leisure
Wives	commute					
	work					
	study					
	housework					
	leisure					
Husbands	commute					
	work	0.1388**				
	study	0.0273	0.0526			
	housework	−0.0405	−0.0650	0.1959**		
	leisure	0.0161	−0.0280	0.0982	−0.0559	

[a] "**" and "*" denote that zero is not included in the 95% and 90% credible intervals, respectively.

Table 5. Probability of the positive partial effect of the two models

	w_age		nc3	
	Wives	Husbands	Wives	Husbands
commute	0.414 (0.416)[a]	0.493	0.571 (0.582)	0.569 (0.552)
work	0.434 (0.437)	0.502	0.733 (0.732)	0.600 (0.596)
study	0.571 (0.579)	0.504	0.735 (0.724)	0.175 (0.114)
housework	0.424 (0.431)	0.506	0.999 (0.998)	1.000 (1.000)
leisure	0.571 (0.559)	0.503	0.000 (0.000)	0.026 (0.025)
basis	0.568 (0.574)	0.498	0.209 (0.201)	0.095 (0.106)
	w_inc		**h_inc**	
	Wives	Husbands	Wives	Husbands
commute	0.517 (0.518)	0.496 (0.499)	0.495 (0.504)	0.504 (0.504)
work	0.558 (0.554)	0.496 (0.500)	0.499 (0.496)	0.500 (0.507)
study	0.506 (0.507)	0.510 (0.507)	0.504 (0.504)	0.504 (0.501)
housework	0.441 (0.452)	0.511 (0.506)	0.499 (0.500)	0.504 (0.500)
leisure	0.492 (0.484)	0.496 (0.495)	0.509 (0.506)	0.497 (0.496)
basis	0.504 (0.502)	0.501 (0.498)	0.494 (0.498)	0.493 (0.493)
	w_educ		**lwp**	
	Wives	Husbands	Wives	Husbands
commute	0.489 (0.440)	0.500	0.151 (0.149)	0.007 (0.007)
work	0.434 (0.448)	0.500	0.545 (0.524)	0.523 (0.521)
study	0.811 (0.871)	0.502	0.068 (0.105)	0.180 (0.223)
housework	0.163 (0.201)	0.504	0.963 (0.961)	0.302 (0.307)
leisure	0.769 (0.798)	0.500	0.298 (0.302)	0.720 (0.723)
basis	0.536 (0.401)	0.493	0.232 (0.223)	0.791 (0.768)
	h_age		**h_educ**	
	Wives	Husbands	Wives	Husbands
commute	0.505	0.503 (0.501)	0.495	0.453 (0.437)
work	0.500	0.381 (0.378)	0.503	0.606 (0.606)
study	0.506	0.533 (0.533)	0.506	0.572 (0.552)
housework	0.502	0.500 (0.485)	0.497	0.499 (0.476)
leisure	0.503	0.532 (0.526)	0.508	0.274 (0.279)
basis	0.496	0.585 (0.604)	0.495	0.611 (0.627)

[a] The values denote the probability of the positive partial effect in case of two dependent compositional data models, and those in the parentheses are the probability of the positive partial effect in case of two independent compositional data models.

Table 6. Summary of posterior results of estimation

	Coefficient of time equation[a]		Probability of the positive partial effect[b]	
	Wives' activities	Husbands' activities	Wives' activities	Husbands' activities
w_age	study(+)[c] commute(-) [work(-), housework(-)]			
h_age		work(-)		
nc3	housework(+) leisure(-)	housework(+) [leisure(-)]	work(H)[d], study(H), housework(H) leisure(L), basis(L)	housework(H) study(L), leisure(L) basis(L)
w_inc	commute(+), work(+), study(+) housework(-), leisure(-)	study(+), housework(+)		
h_inc	commute(+), leisure(+) [study(+)]	commute(+), work(+)		
w_educ	study(+) [housework(-)]		study(H), leisure(H) housework(L)	
h_educ		leisure(-)		leisure(L)
lwp	housework(+) [study(-)]	 commute(-)	housework(H) commute(L), study(L), leisure(L), basis(L)	leisure(H), basis(H) commute(L), study(L)

[a] In "Coefficient of time equation," the activities of wives and husbands whose 95% credible intervals do not include zero are listed for each explanatory variable in the time equation (21). The activities whose 90% credible intervals do not include zero are in square brackets.
[b] In "Probability of positive average partial effect," activities whose probability of positive average partial effect is high or low are listed when the variables in the first column changes.
[c] "+" denotes that the sign of the coefficient of explanatory variable in the first column is positive, and "-" denotes the opposite.
[d] "H" denotes that the probability of positive average partial effect is greater than 0.7, and "L" denotes that the probability is less than 0.3.