Measuring Chronic and Transient Components of Poverty: A Bayesian Approach

Hikaru Hasegawa
Graduate School of Economics and Business Administration
Hokkaido University

Kazuhiro Ueda*
Faculty of Economics
Nihon Fukushi University

First version: November, 2004
Second version: November, 2005
Final version: May, 2006

Abstract

After the publication of Ravallion's (1988) seminal work on chronic and transient poverty, wide attention has been given to the components of poverty. We propose a Bayesian mixture model to measure poverty and divide it into chronic and the transient poverty using the Foster, Greer and Thorbecke (FGT) measure. These two types of poverty are illustrated using the Panel Study of Income Dynamics (PSID) data.

Keywords: Birth-death process, Foster, Greer and Thorbecke (FGT) measure, Gibbs sampling, Markov chain Monte Carlo (MCMC).

JEL Classification: **C11**, **C15**, **D63**

1 Introduction

Conquering poverty has been one of the most important problems for not only national but also international society. Although there is a huge amount of international and domestic aid for societies and people suffering from poverty, the fact that poverty does not vanish poses difficult questions about how to deal with it. It is likely that policies and measures aimed at eradicating poverty have not matched with types of poverty. Policy-makers should notice the characteristics of the type of poverty in question and adopt policies that correspond to it. It has been pointed out recently that there are two components of poverty observed. They are "transient poverty" and "chronic poverty." In Jalan and

*Address for correspondence: Faculty of Economics, Nihon Fukushi University, Okuda, Mihama-cho, Chita-gun, Aichi, 470-3295, Japan. (e-mail: ueda@n-fukushi.ac.jp)

Ravallion (1998), transient poverty is considered as poverty that arises due to variability in income or consumption over time and chronic poverty is considered as poverty that persists in mean consumption over time.

In the seminal study by Ravallion (1988) components of poverty are distinguished by examining the relationship between risk and poverty. Households in the agriculture sector, for example, face risk due to weather that causes variability in their income. When an index that measures poverty, the headcount ratio for example, rises above, or falls below, the poverty line as a result of social mobility in the population, we cannot say definitely whether the change is led by households which cross the poverty line temporarily because of the risk they face. Ravallion (1988) examines, "whether risk induced welfare variability increases or decreases aggregate poverty,"[1] and shows the importance of distinguishing transient poverty from chronic poverty. The article defines these two components of poverty and presents a method for measuring them empirically with panel data for households in dry regions of Central India by using the poverty measure proposed by Foster, Greer ant Thorbecke (FGT) (1984).

Thereafter, some empirical analyses to distinguish transient poverty from chronic poverty have been presented. Since considerable transient poverty is found in the consumption panel data of Hungarian households used in Ravallion *et al.* (1995) and that of the region of rural China in Jalan and Ravallion (1998), it has been confirmed that differences in the two types of poverty play an important role when choosing a policy for dealing with poverty. Gibson (2001) proposed a new method for decomposing cross-sectional poverty estimates into chronic and transient components using household survey data from Papua New Guinea.

Most of the earlier studies, however, only report values for three types of poverty measures descriptively, and do not use statistical inference. The purpose of this article is to introduce a statistical model and propose a Bayesian approach for estimating chronic poverty and transient poverty. We begin by introducing a multivariate distribution for panel data on income. Since the multivariate distribution may possess a complicated form, we use a finite mixture of normal distributions for income.[2] Ferguson (1983) suggests that an arbitrary probability density function on the real line can be approximated by a countable mixture of normal densities. We are, therefore, able to obtain a rich class of distributions for the income distribution. The finite mixture model in the Bayesian context was developed by Diebolt and Robert (1994). Diebolt and Robert (1994) assume that the number of components, say k, is known. Recently, Richardson and Green (1997) and Stephens (2000) consider the case where k is unknown. Richardson and Green (1997) use the reversible jump method (Green, 1995) that requires the complex calculation of a Jacobian matrix, while Stephens (2000) proposes an alternative to the reversible jump method by using the birth-death process. In this article, the method in Stephens (2000) is used for practical reasons.[3]

The Bayesian approach has several advantages when estimating poverty. In

[1] See Ravallion (1988, p.1172).

[2] For a mixture model for income distribution, see Pittau and Zelli (2004) and Pittau (2005).

[3] As Hurn *et al.* (2003) comment, both methods are equally valid on theoretical grounds. However, the birth-death procedure avoids the complex calculation of a Jacobian that is required in the reversible jump method. Furthermore, the implementation of the birth-death procedure is independent of the MCMC algorithm (Hurn *et al.*, 2003, p.70).

the context of survey data, observed income data usually suffers from survey errors, such as interviewer's error (Groves, 1989). This corresponds to the 'errors-in-variables (EIV) model' in econometrics. There are two ways of incorporating the necessary information: the non-Bayesian and Bayesian approaches. In the non-Bayesian approach, the EIV model can be estimated using instrumental variables (IV). In the Bayesian approach, no additional variation data (*i.e.*, IV) are required and a prior information on errors is incorporated instead.[4] This enables us to measure poverty based on not only observed data, but also the true income excluded by errors in the observed data. On the contrary, it is difficult to calculate poverty measure from unobserved true income using the non-Bayesian approach.

Our Bayesian approach also has the advantage that the model permits heterogeneity in individual incomes. This approach divides individual income into several income groups as part of the estimation procedure. In this sense, the possibility of dependency among individual incomes exists, although they are assumed to be independent.[5] Bayesian posterior analysis is often associated with a computational burden. However, using a recent development in the Markov chain Monte Carlo (MCMC) method, posterior analysis can be implemented easily.

We estimate the FGT poverty measures for total, chronic and transient poverty using samples from the posterior distributions of the observed and unobserved data. In order to verify the effectiveness of our method, we define the posterior intervals. These intervals include poverty measures by Ravallion's method. The chronic poverty measure based on the posterior results for unobserved true data can be regarded as a poverty measure for a stationary state. Comparing this chronic poverty and the total poverty derived from the observed data, we can measure the difference between total poverty in the data, which is predicted to be observed, and true poverty.

This article is organized as follows. In Section 2, we present a Bayesian mixture model for density estimation. In Section 3, the poverty measures for three types of poverty (total poverty, chronic poverty and transient poverty) are described. In Section 4, an application of our approach to the Panel Study of Income Dynamics (PSID) data is provided for illustration. Finally, in Section 5, a brief summary and some extensions of our approach are given.

2 Bayesian Mixture Model

2.1 The Basic Model and Ravallion's Decomposition

Let X_{it} denote the income of the ith individual (or household) at period t. We use x_{it} as a logarithm of X_{it}, that is $x_{it} = \log X_{it}$. It is assumed that x_{it} is generated by the following model:

$$x_{it} = \mu_i + u_{it},\ i = 1, \cdots, n;\ t = 1, \cdots, T. \tag{1}$$

In (1), μ_i is the ith individual's steady-state logarithm of income and is assumed to be constant over time. u_{it} is a transitory component that determines the deviation from the steady-state logarithm of income over time.

[4] See, for example, Florens *et al.* (1974), and Hasegawa and Kozumi (2001).

[5] See Hasegawa and Kozumi (2003, p.278) for more details on this point.

Ravallion (1988) proposed a method for decomposing total poverty into chronic and transient components. Following Ravallion (1988) and Jalan and Ravallion (1998), we define total, chronic and transient poverties. Let Υ denote a poverty line and $\upsilon = \log \Upsilon$. We also use the poverty measure proposed by Foster, Greer and Thorbecke (FGT) (1984). The discrete formulation of FGT at period t is defined as

$$\Pi(\Upsilon : t) = \frac{1}{n} \sum_{X_{it} < \Upsilon} \left(1 - \frac{X_{it}}{\Upsilon}\right)^{\zeta} = \frac{1}{n} \sum_{x_{it} < \upsilon} \left(1 - \mathrm{e}^{x_{it} - \upsilon}\right)^{\zeta}, t = 1, \cdots, T, \tag{2}$$

where $\zeta \geq 0$ is a sensitivity parameter of the poverty measure. Since the steady-state logarithm of income μ_i is unobservable, Ravallion's decomposition replaces μ_i with $\bar{X}_i = \sum_{t=1}^{T} X_{it}/T$. Thus, total, chronic and transient poverties are defined as follows:

$$\text{Total poverty: } \Pi_F(\Upsilon) = \frac{1}{T} \sum_{t=1}^{T} \Pi(\Upsilon : t) \tag{3}$$

$$\text{Chronic poverty: } \Pi_C(\Upsilon) = \frac{1}{n} \sum_{\bar{X}_i < \Upsilon} \left(1 - \frac{\bar{X}_i}{\Upsilon}\right)^{\zeta} \tag{4}$$

$$\text{Transient poverty: } \Pi_T(\Upsilon) = \Pi_F(\Upsilon) - \Pi_C(\Upsilon). \tag{5}$$

The poverty measures defined in (3), (4) and (5) are descriptive. Alternatively, in this article we begin by introducing a statistical model for x_{it}. Let define $x_i = [x_{i1}, \cdots, x_{iT}]'$. It is difficult to estimate a multivariate density of x_i using a single distribution. Therefore, we assume that the distribution of x_i is a mixture of normal distributions with k components. That is,

$$x_i | \{w_j, \mu_j, \Sigma_j\}_{j=1,\cdots,k} \sim \sum_{j=1}^{k} w_j \, \mathcal{N}(\mu_j, \Sigma_j), \sum_{j=1}^{k} w_j = 1, \; i = 1, \cdots, n. \tag{6}$$

Once we have estimated the income distribution of x_i, we can use the distribution for calculating Ravallion's poverty indices. In the next subsection, we extend the mixture model (6) to a model with errors-in-variables and present the estimation method by using simulation-based Bayesian statistics.

2.2 Bayesian Mixture Model with Errors-in-Variables

Observed income data are collected from surveys, so that the data usually suffers from survey errors, such as interviewer's errors and errors due to respondents.[6] We, therefore, introduce explicitly the errors in the observed income. Now, it is considered that x_{it} in (6) is logarithm of unobserved true income and observed income is denoted by Y_{it}. Let us define $y_{it} = \log Y_{it}$, and assume that y_{it} is generated by the following model:

$$y_{it} = x_{it} + \epsilon_{it}, \; i = 1, \cdots, n; \; t = 1, \cdots, T, \tag{7}$$

[6] See, for example, Groves (1989).

where ϵ_{it} is an error term.[7] Defining $y_i = [y_{i1}, \cdots, y_{iT}]'$ and $\epsilon_i = [\epsilon_{i1}, \cdots, \epsilon_{iT}]'$, (7) is summarized as

$$y_i = x_i + \epsilon_i, \ i = 1, \cdots, n. \tag{8}$$

Here x_i is not a vector of observations, but can be considered as a vector of unobserved parameters. So in our Bayesian context, (6) is regarded as a prior of x_i. We also assume that ϵ_i is normally distributed. That is,

$$x_i|\theta \sim \sum_{j=1}^{k} w_j \mathcal{N}(\mu_j, \Sigma_j), \ \sum_{j=1}^{k} w_j = 1, \ i = 1, \cdots, n \tag{9}$$

$$\epsilon_{it}|\theta \sim \mathcal{N}(0, \sigma_t^2), \ i = 1, \cdots, n; \ t = 1, \cdots, T, \tag{10}$$

where θ is a vector of other parameters except for x_i's. To complete the Bayesian model, we introduce the following prior distributions:[8]

$$\begin{cases} \mu_j \sim \mathcal{N}(\xi, \kappa^{-1}) \\ \Sigma_j^{-1}|\beta \sim \mathcal{W}(2\alpha, (2\beta)^{-1}) \\ \beta \sim \mathcal{W}(2g, (2h)^{-1}) \\ w \sim \mathcal{D}(\delta, \cdots, \delta) \\ \sigma_t^{-2} \sim \mathcal{G}a(a, b), \end{cases} \tag{11}$$

where $w = [w_1, \cdots, w_k]'$, $\mathcal{W}(a, A)$ denotes a Wishart distribution with degrees of freedom a and scale parameters matrix A, $\mathcal{D}(\delta_1, \cdots, \delta_k)$ denotes a Dirichlet distribution and $\mathcal{G}a(a, b)$ denotes a gamma distribution with shape parameter a and scale parameter b.

When considering a mixture distribution, it is convenient to introduce latent variables z_i with a probability mass function:

$$\mathsf{P}(z_i = j|\theta) = w_j, \ i = 1, \cdots, n; \ j = 1, \cdots, k.$$

Conditional on $z = [z_1, \cdots, z_n]'$, the distributions of $x_1, \cdots, x_n$ are assumed to be

$$p(x_i|z_i = j, \theta) = \mathcal{N}(x_i|\mu_j, \Sigma_j), \ i = 1, \cdots, n.$$

Integrating out z_i, we have the model (9):

$$p(x_i|\theta) = \sum_{j=1}^{k} \mathsf{P}(z_i = j|\theta) p(x_i|z_i = j, \theta) = \sum_{j=1}^{k} w_j \mathcal{N}(x_i|\mu_j, \Sigma_j).$$

Using the latent variables z, we can classify $x_1, \cdots, x_n$. Let us define a set $I_j = \{i|z_i = j\}$ and a number of observations allocated to class j, say $n_j = \#\{i|z_i = j\}$. We can derive the full conditional distributions (FCD) of θ, x_i and

[7] This specification is familiar in the literature on demand analysis. See, for example, Lewbel (1996) and Hasegawa and Kozumi (2001).

[8] These prior distributions are same as those in Stephens (2000).

z_i $(i = 1, \cdots, n)$ as follows: for $i = 1, \cdots, n;\ t = 1, \cdots, T;\ j = 1, \cdots, k$,

$$\mathsf{P}(z_i = j|\theta, x_i, y_i) = \frac{w_j \mathcal{N}(x_i|\mu_j, \Sigma_j)}{\sum_{l=1}^{k} w_l \mathcal{N}(x_i|\mu_l, \Sigma_l)} \tag{12}$$

$$\mu_j|\cdots \sim \mathcal{N}\left(\left(\kappa + n_j\Sigma_j^{-1}\right)^{-1}\left(\kappa\xi + n_j\Sigma_j^{-1}\bar{x}_j\right), \left(\kappa + n_j\Sigma_j^{-1}\right)^{-1}\right) \tag{13}$$

$$\Sigma_j^{-1}|\cdots \sim \mathcal{W}\left(2\alpha + n_j, \left[2\beta + \sum_{i\in I_j}(x_i - \mu_j)(x_i - \mu_j)'\right]^{-1}\right) \tag{14}$$

$$\beta|\cdots \mathcal{W}\left(2g + 2k\alpha, \left[2h + 2\sum_{j=1}^{k}\Sigma_j^{-1}\right]^{-1}\right) \tag{15}$$

$$w|\cdots \sim \mathcal{D}(\delta + n_1, \cdots, \delta + n_k) \tag{16}$$

$$\sigma_t^{-2}|\cdots \sim \mathcal{G}a\left(a + \frac{n}{2}, b + \frac{1}{2}\sum_{i=1}^{n}(y_{it} - x_{it})^2\right) \tag{17}$$

$$x_i|\cdots \sim \mathcal{N}\left(\left(\Sigma_j^{-1} + \Psi^{-1}\right)^{-1}\left(\Sigma_j^{-1}\mu_j + \Psi^{-1}y_i\right), \left(\Sigma_j^{-1} + \Psi^{-1}\right)^{-1}\right), \tag{18}$$

where $\Psi = \text{diag}\{\sigma_1^2, \cdots, \sigma_T^2\}$, $\bar{x}_j = \dfrac{1}{n_j}\sum_{i\in I_j} x_i$, and '$|\cdots$' denotes conditioning on the values of all other parameters.

Since the data y and the predictive data $\tilde{y}$ are conditionally independent given $\theta^* = \{\theta, x, z\}$, the predictive distribution can be written as follows:

$$p(\tilde{y}|y) = \int p(\tilde{y}|\theta^*)p(\theta^*|y)d\theta^* \tag{19}$$

Therefore, $p(\tilde{y}|\theta^*)$ can be used to estimate poverty indices for y.

The degrees of heterogeneity of income are usually unknown. The number of components in the mixture model k is, therefore, also unknown. When k is unknown, we can use the reversible jump method (Richardson and Green, 1997) and its alternative proposed by Stephens (2000), which is based on the birth-death process. Following Hurn *et al.* (2003), we use the birth-death process approach for practical considerations (see the footnote 4). Stephens' (2000) algorithm will be discussed briefly in the appendix.

3 Poverty Measures

We use the poverty measure proposed by Foster, Greer and Thorbecke (FGT) (1984), as mentioned in Section 2.1. The continuous formulation of FGT is defined as

$$\Pi(\upsilon) = \int_{-\infty}^{\upsilon}(1 - \mathrm{e}^{y-\upsilon})^{\zeta}dF(y), \tag{20}$$

where Υ denotes a poverty line and $\upsilon = \log \Upsilon$. The discrete formulation of FGT with sample size N corresponding to (20) is defined as

$$\Pi(\upsilon) = \frac{1}{N} \sum_{y_i < \upsilon} \left(1 - \mathrm{e}^{y_i - \upsilon}\right)^{\zeta} = \frac{1}{N} \sum_{Y_i < \Upsilon} \left(1 - \frac{Y_i}{\Upsilon}\right)^{\zeta}. \tag{21}$$

Applying the discrete formulation of FGT in (21) to Y_{it} $(i = 1, \cdots, n; t = 1, \cdots, T)$, we define the sample FGT on Y at period t by

$$\Pi_y(\Upsilon : t) = \frac{1}{n} \sum_{Y_{it} < \Upsilon} \left(1 - \frac{Y_{it}}{\Upsilon}\right)^{\zeta}, \ t = 1, \cdots, T. \tag{22}$$

Ravallion's total, chronic and transient poverties on Y are written as follows:

$$\text{Total poverty: } \Pi_F^{em}(\Upsilon) = \frac{1}{T} \sum_{t=1}^{T} \Pi_y(\Upsilon : t) \tag{23}$$

$$\text{Chronic poverty: } \Pi_C^{em}(\Upsilon) = \frac{1}{n} \sum_{\bar{Y}_i < \Upsilon} \left(1 - \frac{\bar{Y}_i}{\Upsilon}\right)^{\zeta} \tag{24}$$

$$\Pi_T^{em}(\Upsilon) = \Pi_F^{em}(\Upsilon) - \Pi_C^{em}(\Upsilon), \tag{25}$$

where $\bar{Y}_i = \sum_{t=1}^{T} Y_{it}/T$.

Now, we define the total, chronic and transient poverties using the posterior results. First, we simulate $\tilde{x} = [\tilde{x}_1, \cdots, \tilde{x}_T]'$ from its FCD and define the poverty measure based on $\tilde{x}_t$ at period t as follows:

$$\Pi_x(\upsilon : t) = \int_{-\infty}^{\upsilon} (1 - \mathrm{e}^{\tilde{x}_t - \upsilon})^{\zeta} dF(\tilde{x}_t), \ t = 1, \cdots, T.$$

Using $\Pi_x(\upsilon : t)$, we define total poverty as follows:

$$\text{Total poverty based on } x\text{: } \Pi_F^{post}(\upsilon) = \frac{1}{T} \sum_{t=1}^{T} \Pi_x(\upsilon : t). \tag{26}$$

Furthermore, we calculate $\bar{\tilde{x}} = \sum_{t=1}^{T} \tilde{x}_t/T$ and define chronic poverty as the FGT of $\bar{\tilde{x}}$:

$$\text{Chronic poverty: } \Pi_C^{post}(\upsilon) = \int_{-\infty}^{\upsilon} (1 - \mathrm{e}^{\bar{\tilde{x}} - \upsilon})^{\zeta} dF(\bar{\tilde{x}}) \tag{27}$$

Finally, we define transient poverty in a manner similar to (25):

$$\text{Transient poverty based on } x\text{: } \Pi_T^{post}(\upsilon) = \Pi_F^{post}(\upsilon) - \Pi_C^{post}(\upsilon). \tag{28}$$

The poverty measures based on y are also available from the predictive distribution. That is, we simulate $\tilde{y} = [\tilde{y}_1, \cdots, \tilde{y}_T]'$ from $p(\tilde{y}|\theta^*)$ and define a poverty measure at period t as follows:

$$\Pi_y(\upsilon : t) = \int_{-\infty}^{\upsilon} (1 - \mathrm{e}^{\tilde{y}_t - \upsilon})^{\zeta} dF(\tilde{y}_t), \ t = 1, \cdots, T.$$

Using $\Pi_y(\upsilon : t)$, we define total poverty as follows:

$$\text{Total poverty based on } y\text{: } \Pi_{Fy}^{post}(\upsilon) = \frac{1}{T}\sum_{t=1}^{T}\Pi_y(\upsilon : t). \tag{29}$$

We define the transient poverty based on y:

$$\text{Transient poverty based on } y\text{: } \Pi_{Ty}^{post}(\upsilon) = \Pi_{Fy}^{post}(\upsilon) - \Pi_{C}^{post}(\upsilon). \tag{30}$$

The discrete formulation of FGT with sample size N defined in (21) converges to the continuous formulation defined in (20) as $N \to \infty$. Therefore, we calculate the discrete formulations of (26) to (30) for a large number of N. Furthermore, since the values of (26) to (30) are obtained for each iteration in the MCMC simulation, statistical inference can be carried out on the three types of poverties.

4 Empirical Example

4.1 Model Specification and Data

For an illustration of our Bayesian approach, we use the Panel Study of Income Dynamics (PSID) for waves 21–26 (1988–1993) of family data.[9] There are $5,371$ records for the heads of family interviewed successively during the period, and we select $1,532$ records satisfying the following conditions from them : a) the record is for a head of family, b) the data for annual family income, family composition, and the number of children by sex and age are available, c) the family consists of a husband, a wife and a child/children, d) the family income in a record is greater than 1 and an exact value.

We use "total family money income" as family income, which is sum of the taxable incomes of the head and wife, total transfers to the head and wife, taxable prorated income of others, and total prorated transfers of others. Although it is common that heads in the records we pick up have a family that consists of a married-couple and a single child or children, family size can vary with the number of children. This raises a problem in how to compare the poverty of people whose family sizes are different, and thus we have to count how many adults a child or two children correspond to. Therefore, it is required that the per capita income be adjusted for the demographic differences to overcome such demographic difficulties. In order to calculate per capita income deflated by equivalence scale from family income, we use the following formulation for equivalence scale proposed by Citro and Michael (1995):

$$\text{scale value} = (A + PK)^F, \tag{31}$$

where A is the number of adults in a family, K is the number of children, each of whom is treated as a proportion P of an adult, and F is the scale economy factor.[10] The values we use for both P and F are 0.7, as recommended in Citro

[9] We use the data for 1988–93 to estimate poverty measures for 1988–92, since we need the income data for 1988–1992, but the PSID income data for a given year give data for the previous year.

[10] See Citro and Michael (1995, p.161).

and Michael (1995). The income for estimation is converted into real income using the Consumer Price Index (CPI).[11]

Tables 1 and 2 show summary statistics and the correlation matrix of the data, respectively. From Table 2, we can observe that the logarithms of the real income deflated by equivalence scale for the respective years are highly correlated. This fact justifies the formulation of a multivariate distribution of x in (6) or (9). Figure 1 shows the histograms of $y = \log Y$ for each year.[12] Table 3 shows the total, the chronic and the transient poverty in the data using the method in Ravallion (1988). In this table, we use $\zeta = 0, 1$ and 2, which correspond to the headcount ratio, the poverty gap and the squared poverty gap, respectively.

We set the hyperparameter values as follows:[13]

$$\begin{cases} \xi = \left[\min_{y_1} + \dfrac{R_1}{2}, \cdots, \min_{y_T} + \dfrac{R_T}{2}\right]' \quad \kappa = \operatorname{diag}\left\{\dfrac{0.1}{R_1^2}, \cdots, \dfrac{0.1}{R_T^2}\right\} \\ \alpha = 6.0,\ g = 0.1\alpha,\ h = \operatorname{diag}\left\{\dfrac{100g}{\alpha R_1^2}, \cdots, \dfrac{100g}{\alpha R_T^2}\right\} \\ \delta = 1.0,\ a = 2.0,\ b = 0.002,\ \lambda = 1 \text{ (birth rate)}, \end{cases} \tag{32}$$

where

$$\max_{y_t} = \max\{y_{1t}, \cdots, y_{nt}\},\ \min_{y_t} = \min\{y_{1t}, \cdots, y_{nt}\},\ R_t = \max_{y_t} - \min_{y_t}.$$

The initial value of k is set as $k = 1$. For a large sample size N, the discrete formulation of FGT converges to the continuous formulation. We can simulate the samples with a large N from the posterior predictive distributions. We set $N = 20,000$. The MCMC simulation was run for $40,000$ iterations and the first $10,000$ samples were discarded as a burn-in period.

4.2 Posterior Results

Table 4 shows the posterior results of β, σ_t^2 and k. The CD denotes the convergence diagnostic statistic proposed by Geweke (1992).[14] The sixth column of Table 4 shows the p values for CD and the convergence of the posterior distribution would be confirmed from them. Further, the posterior mean and the

[11] We obtained the data for the CPI from the web-site of the Bureau of Labor Statistics, U.S. Department of Labor (`http://stats.bls.gov/`).

[12] The posterior results obtained afterward are generated by using DIGITAL Visual Fortran version 6.6 and all figures are drawn using Ox version 3.40 (Doornik 2001).

[13] These values are similar to those of Stephens (2000).

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

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

where

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

and $\hat{S}_A(0)$ and $\hat{S}_B(0)$ denote consistent spectral density estimates. In this article, we set $n_A = 3,000$ and $n_B = 15,000$, and calculate $\hat{S}_A(0)$ and $\hat{S}_B(0)$ using Parzen windows with bandwidths of 300 and $1,500$, respectively.

standard deviation of the number of components in the mixture k is about 23.7 and 3.8, respectively. Using the following posterior interval

$$\begin{aligned}&(\text{posterior mean} - 2 \times \text{posterior standard deviation},\\ &\quad \text{posterior mean} + 2 \times \text{posterior standard deviation}),\end{aligned} \tag{33}$$

roughly speaking, the number of components k varies from 16 to 31. Figure 2 shows the estimated densities of x. It seems that the densities fit the histograms of the data appropriately.

Table 5 shows the result of sensitivity analysis to choice of hyperparameter values and to the number of distributions k in the mixture. In the table, the case (a) has the hyperparameter values given in (32). The other cases change the hyperparameter values as follows:

$$\begin{aligned}&\text{(b) } \kappa = \text{diag}\{1/R_t^2\},\ \text{(c) } \kappa = \text{diag}\{0.01/R_t^2\},\ \text{(d) } h = \text{diag}\{200g/\alpha R_t^2\}\\ &\text{(e) } h = \text{diag}\{50g/\alpha R_t^2\},\ \text{(f) } b = 0.2,\ \text{(g) } \lambda = 3.\end{aligned}$$

The cases (b) to (e) give the similar value of k as in the case (a). Therefore, whether κ and h are more informative or less informative does not affect the posterior results. The case (f) of more informative value of $b = 0.2$ gives smaller number of k than that in the case (a) ($b = 0.002$). The case (g) of $\lambda = 3$ gives larger number of k than that in the case (a) ($\lambda = 1$). This result is the same as a result in Stephens (2000, p.57). Figure 3 shows the estimated densities of x in 1988 for the cases (b) to (g).[15] From this figure, we can observe that the estimated densities are very similar, although the estimated density of case (f) is slightly steeper than that of the case (a).

Table 6 shows the posterior results for total, chronic and transient poverty. The higher the poverty line is, the higher the ratio of mean to standard deviation for the poverty measures (mean/sd) is. When the poverty line is higher than $3,000$, which is the minimal poverty line we set, all the ratios are higher than 2. The range of poverty line in the table is not unrealistic, because the actual poverty line in USA lies within the range we set, according to the web-site of US Census Bureau.[16] This implies that the estimated poverty measures are significant and our method is appropriate for measuring poverty.

Figure 4 reports total, chronic and transient poverty for the data and posterior results. The horizontal and the vertical axes denote the poverty line and the poverty measure. The solid lines, with the symbols '$\bullet$' and '$+$', denote the poverty measure computed by Ravallion's method using the observed data and the posterior mean of poverty measure, respectively. The dotted lines denote the upper and the lower bounds of the posterior intervals (33). There are 5 graphs for each ζ, which show the total and transient poverty derived from y and x, and the chronic poverty from x.

We compare the three types of poverty measure calculated by our method with those calculated by Ravallion's method. It is observed, according to Figure 4, that the transient poverty calculated by Ravallion's method fluctuates and

[15] We have the same results of estimated densities in the other years as those in 1988.

[16] The poverty threshold of a family that consists of two adults and one child, for example, in 1988 is $9,522$ dollars according to the web-site. Using (31) and CPI, we obtain its real per capita value, deflated by equivalence scale, as approximately $4,016$ dollars. See `http://www.census.gov/`.

at some points jumps out of the posterior interval based on both y and x, when $\zeta = 0$.

FGT poverty measures with $\zeta = 0$ or 1 are well-known poverty measures: the headcount ratio and the poverty gap. However, it is pointed out by Kurosaki (2005) that these measures have faults in that they do not reflect the degree of poverty well, and that the adoption of the FGT poverty measure with $\zeta > 1$ implies that the social planner is assumed to be risk-averse and inequality averse. Therefore, the FGT is often used as $\zeta = 2$ (the squared poverty gap) in the literature of development economics.[17]

Figure 5 shows the increment of the total and the chronic poverty measures by Ravallion's method defined as

$$\Delta\Pi_F^{em}(\Upsilon_l) = \Pi_F^{em}(\Upsilon_l) - \Pi_F^{em}(\Upsilon_{l-1}),\ l = 2, \cdots, L \tag{34}$$

$$\Delta\Pi_C^{em}(\Upsilon_l) = \Pi_C^{em}(\Upsilon_l) - \Pi_C^{em}(\Upsilon_{l-1}),\ l = 2, \cdots, L, \tag{35}$$

where $\Pi_F^{em}(\Upsilon_l)$ and $\Pi_C^{em}(\Upsilon_l)$ are defined in (23) and (24), and L is the number of poverty lines. The total poverty and chronic poverty measures for $\zeta = 0$ fluctuate more than for $\zeta = 1$ or 2 in the figure. It is interesting that the chronic poverty measure, which ought to be constant or move stably, fluctuates according to the poverty line we assume when $\zeta = 0$. These fluctuations cause the unstable movement of the transient poverty measures by Ravallion's method in Figure 4. Figures 4 and 5 suggest that the headcount ratio (the case of $\zeta = 0$) may not capture poverty sufficiently, especially transient poverty.

5 Concluding Remarks

As shown in Ravallion's seminal work, poverty can be decomposed into two components, chronic poverty and transient poverty. This article provided an alternative empirical method to measure them. Specifically, we proposed the Bayesian mixture model with an unknown number of mixtures developed by Stephens (2000). Our Bayesian approach in this article has the following merits:

1. No additional data (*i.e.*, IV) are required for the estimation of EIV.
2. The Bayesian model in this article permits heterogeneity in individual income by using the mixture model.
3. Unobserved incomes can be derived from the posterior results, and using them, we can use statistical inference on the total, chronic and transient poverties.

In our real example using the PSID data, the ratios of posterior mean to posterior standard deviation are fairly high for the actual level of the poverty line. Furthermore, in most cases the posterior intervals include poverty measures calculated using Ravallion's descriptive statistics. Regarding that Ravallion's method has been supported, it shows that the plausible poverty measures can be obtained by our method.

Acknowledgements The authors are grateful to Professor Khondaker Mizanur Rahman of Nanzan University, Professor Fitzenberger, the editor, and three

[17] See, for example, Jalan and Ravallion (1998).

anonymous referees for their helpful comments, which improved an earlier version of this article. They also acknowledge the financial support of the Japanese Ministry of Education, Culture, Sports, Science and Technology under Grant-in-Aid for Scientific Research No.13630024 and No.14653004, respectively.

Appendix

Stephens' birth-death algorithm is as follows:[18]

Stephens' birth-death algorithm Let define $\phi_j = (\mu_j, \Sigma_j)$ and $\psi = \{(w_1, \phi_1), \cdots, (w_k, \phi_k)\}$ and the birth rate as λ.

1. Calculate the death rate for each component. The death rate for the jth component is given by

$$\nu_j = \lambda \frac{L(\psi \setminus (w_j, \phi_j))}{L(\psi)} \frac{p(k-1)}{kp(k)}, \ j = 1, \cdots, k,$$

   where $L(\cdot)$ is the likelihood function.
2. Calculate the total death rate as $\nu = \sum_{j=1}^{k} \nu_j$.
3. Simulate the time to the next jump from an exponential distribution with mean $\dfrac{1}{\lambda + \nu}$.
4. Simulate the type of jump: birth or death with respective probabilities

$$\mathsf{P}(\text{birth}) = \frac{\lambda}{\lambda + \nu}, \ \mathsf{P}(\text{death}) = \frac{\nu}{\lambda + \nu}.$$

5. Adjust ψ to reflect the birth or death:

$$\begin{aligned}
&\text{birth}: \ \psi \cup (w_b, \phi_b) \\
&\quad = \{(w_1(1-w_b), \phi_1), \cdots, (w_k(1-w_b), \phi_k), (w_b, \phi_b)\} \\
&\text{death}: \ \psi \setminus (w_j, \phi_j) = \left\{ \left(\frac{w_1}{1-w_j}, \phi_1 \right), \cdots, \left(\frac{w_{j-1}}{1-w_j}, \phi_{j-1} \right), \right. \\
&\quad \left. \left(\frac{w_{j+1}}{1-w_j}, \phi_{j+1} \right), \cdots, \left(\frac{w_k}{1-w_j}, \phi_k \right) \right\}.
\end{aligned}$$

   Birth: Simulate w_b and ϕ_b independently from densities $k(1-w_b)^{k-1}$ and $p(\phi_b|\theta)$, respectively.
   Death: Select a component to die: $(w_j, \phi_j) \in \psi$ being selected with probability $\dfrac{\nu_j}{\nu}$ for $j = 1, \cdots, k$
6. Return to step 1.

In the birth-death process, we use a truncated Poisson prior with λ on the number of components k (Stephens, p.50, 2000):

$$p(k) \propto \frac{\lambda^k}{k!}, \ k = 1, \cdots, k_{\max} = 100.$$

[18] Algorithm 3.1 in Stephens (2000).

References

Citro CF, and Michael RT (1995) Adjusting poverty thresholds. In: Citro CF, and Michael RT (eds.) Measuring Poverty: A New Approach. National Academic Press, Washington; 159–201

Diebolt J, Robert CP (1994) Estimation of finite mixture distributions through Bayesian sampling. Journal of the Royal Statistical Society B56:363–375

Doornik JA (2001) Ox 3.0: An object-oriented matrix programming language. Timberlake Consultants Ltd., London

Ferguson TS (1983) Bayesian density estimation by mixtures of normal distributions. In: Rizvi MH, Rustagi JS (eds.) Recent Advantages in Statistics: Papers in Honor of Herman Chernoff on His Sixtieth Birthday. Academic Press, New York; 287–302

Florens JP, Mouchart M, Richard JF (1974) Bayesian inference in error-in-variables models. Journal of Multivariate Analysis 4:419–452

Foster J, Greer J, Thorbecke E (1984) A class of decomposable poverty measures. Econometrica 52:761–766

Geweke J (1992) Evaluating the accuracy of sampling–based approaches to the calculation of posterior moments. In: Bernardo JM, Berger JO, Dawid AP, Smith AFM (ed.) Bayesian statistics 4. Oxford University Press, Oxford, 169–193

Gibson J (2001) Measuring chronic poverty without a panel. Journal of Development Economics 65:243–266

Green PJ (1995) Reversible jump Markov chain Monte Carlo computation and Bayesian model determination. Biometrika 82:711–732

Groves RM (1989) Survey errors and survey costs. Wiley, New York

Hasegawa H, Kozumi H (2001) Bayesian analysis on Engel curves estimation with measurement errors and an instrumental variable. Journal of Business & Economic Statistics 19:292–298

Hasegawa H, Kozumi H (2003) Estimation of Lorenz curves: A Bayesian nonparametric approach. Journal of Econometrics 115:277–291

Hurn M, Justel A, Robert CP (2003) Estimating mixtures of regressions. Journal of Computational and Graphical Statistics 12:55–79

Jalan J, Ravallion M (1998) Transient poverty in postreform rural China. Journal of Comparative Economics 26:338–357

Kurosaki T, (2005) The Measurement of transient poverty: Theory and application to Pakistan. Institute of Economic Research, Hitotsubashi University

Lewbel A (1996) Demand estimation with expenditure measurement errors on the left and right hand side. The Review of Economics and Statistics 78:718–725

Pittau MG (2005) Fitting regional income distributions in the European Union. Oxford Bulletin of Economics and Statistics 67:135–161

Pittau MG, Zelli R (2004) Testing for changing shapes of income distribution: Italian evidence in the 1990s from kernel density estimates. Empirical Economics 29:415–430

Ravallion M (1988) Expected poverty under risk-induced welfare variability. The Economic Journal 98:1171–1182

Ravallion M, van de Walle D, Gautam M (1995) Testing a social safety net. Journal of Public Economics 57:175–199

Richardson S, Green PJ (1997) On Bayesian analysis of mixtures with an unknown number of components (with discussion). Journal of the Royal Statistical Society B59:731–792

Stephens M (2000) Bayesian analysis of mixture models with an unknown number of components — An alternative to reversible jump methods. The Annals of Statistics 28:40–74

Table 1: Summary statistics

equivalence scale					
	1988	1989	1990	1991	1992
mean	2.37395	2.38834	2.39820	2.39844	2.38234
sd[a]	0.31713	0.31249	0.30965	0.30923	0.30923
m/sd	7.48563	7.64290	7.74484	7.75616	7.70405
min	2.00426	2.00426	2.00426	2.00426	2.00426
max	4.39901	4.39901	4.39901	4.13590	4.13590
2.50%[b]	2.00426	2.00426	2.00426	2.00426	2.00426
5%	2.00426	2.00426	2.00426	2.00426	2.00426
25%	2.00426	2.00426	2.35524	2.35524	2.00426
50%	2.35524	2.35524	2.35524	2.35524	2.35524
75%	2.68503	2.68503	2.68503	2.68503	2.68503
95%	2.99826	2.99826	2.99826	2.99826	2.99826
97.50%	2.99826	2.99826	2.99826	2.99826	2.99826
Y					
	1988	1989	1990	1991	1992
mean	16218.5	16383.5	16719.9	16802.8	17623.8
sd	11531.0	12112.2	13057.1	13809.0	12673.8
m/sd	1.40651	1.35264	1.28052	1.2168.0	1.39058
min	704.8	700.4	190.9	280.6	613.3
max	167643.3	190723.1	164896.3	266611.5	143240.4
2.50%	3286.4	2968.1	2777.7	2732.6	2920.0
5%	4307.1	4133.4	4008.3	4265.8	4391.9
25%	9152.1	9655.9	9518.2	9662.9	9981.2
50%	14060.9	14038.7	14287.8	13957.0	15010.3
75%	19733.5	19839.4	20569.8	20574.6	22330.8
95%	33961.1	35062.5	34970.6	34874.0	37624.7
97.5%	42243.2	43848.7	43327.7	43959.2	48542.3
$y = \log Y$					
	1988	1989	1990	1991	1992
sd	0.63807	0.64952	0.66858	0.67004	0.67972
m/sd	14.88860	14.63348	14.23047	14.20222	14.07497
min	6.55796	6.55166	5.25160	5.63680	6.41890
max	12.02959	12.15858	12.01307	12.49355	11.87228
2.50%	8.09756	7.99568	7.92939	7.91300	7.97935
5%	8.36802	8.32686	8.29612	8.35838	8.38752
25%	9.12174	9.17532	9.16097	9.17605	9.20846
50%	9.55116	9.54957	9.56716	9.54374	9.61649
75%	9.89007	9.89543	9.93158	9.93181	10.01372
95%	10.43297	10.46489	10.46226	10.45950	10.53542
97.50%	10.65120	10.68850	10.67655	10.69102	10.79019

a: 'sd' denotes the standard deviation.

b: 'c%' denotes the c percent point of data.

Table 2: Correlation matrix of y

	1988	1989	1990	1991	1992
1988	1	0.87963	0.82729	0.7886	0.72356
1989	0.87963	1	0.86713	0.82506	0.75361
1990	0.82729	0.86713	1	0.86283	0.77849
1991	0.7886	0.82506	0.86283	1	0.82324
1992	0.72356	0.75361	0.77849	0.82324	1

Table 3: Total, chronic and transient poverty measures from data

	$\zeta = 0$		
$\Upsilon = \mathrm{e}^{v}$	Total	Chronic	Transient
3000	0.02532637	0.01436031	0.01096606
4000	0.04464752	0.03067885	0.01396867
5000	0.06657963	0.04699739	0.01958225
6000	0.09778068	0.07832898	0.0194517
7000	0.13250653	0.11031332	0.02219321
8000	0.17062663	0.15274151	0.01788512
9000	0.21710183	0.19843342	0.01866841
10000	0.2689295	0.2421671	0.02676240
	$\zeta = 1$		
$\Upsilon = \mathrm{e}^{v}$	Total	Chronic	Transient
3000	0.00811119	0.0035336	0.0045776
4000	0.01463314	0.00796997	0.00666318
5000	0.02280064	0.01412318	0.00867746
6000	0.03268208	0.02216516	0.01051692
7000	0.04435189	0.03237172	0.01198016
8000	0.0577535	0.04464175	0.01311175
9000	0.07286232	0.05902915	0.01383316
10000	0.08994078	0.07520106	0.01473972
	$\zeta = 2$		
$\Upsilon = \mathrm{e}^{v}$	Total	Chronic	Transient
3000	0.00392037	0.00132993	0.00259043
4000	0.0071759	0.00323238	0.00394351
5000	0.01135056	0.00606911	0.00528144
6000	0.01634945	0.0097172	0.00663225
7000	0.0222336	0.01436342	0.00787017
8000	0.02899371	0.02000952	0.00898419
9000	0.03661186	0.0266792	0.00993266
10000	0.04512168	0.03436908	0.0107526

Table 4: Posterior results

	mean[a]	sd[a]	mean/sd	CD[b]	CD p-value
β_{11}	1.72674	0.28653	6.02646	0.14767	0.88260
β_{12}	1.58823	0.2767	5.73981	-0.078920	0.93710
β_{13}	1.51405	0.26663	5.67847	-0.032986	0.97369
β_{14}	1.45526	0.26535	5.48425	0.14685	0.88325
β_{15}	1.36285	0.26095	5.22267	-0.012566	0.98997
β_{22}	1.71669	0.29081	5.90321	-0.48495	0.62771
β_{23}	1.58672	0.2783	5.70149	-0.19332	0.84670
β_{24}	1.52712	0.27597	5.53357	0.016996	0.98644
β_{25}	1.43597	0.2722	5.27534	-0.34426	0.73065
β_{33}	1.73896	0.29052	5.98561	-0.38570	0.69972
β_{34}	1.63252	0.28246	5.77963	-0.23942	0.81078
β_{35}	1.52599	0.27601	5.52869	-0.33643	0.73655
β_{44}	1.79868	0.29676	6.06108	-0.34411	0.73077
β_{45}	1.58585	0.28139	5.63577	-0.27149	0.78602
β_{55}	1.74932	0.2951	5.92786	-0.40052	0.68878
σ_1^2	0.0011	0.00067	1.63995	0.25149	0.80143
σ_2^2	0.00084	0.00044	1.88625	-1.8229	0.068325
σ_3^2	0.00089	0.00053	1.6649	-0.83710	0.40254
σ_4^2	0.00083	0.00043	1.94973	1.7216	0.085139
σ_5^2	0.00108	0.00072	1.4979	1.3147	0.18861
k	23.7454	3.79746	6.25297	1.0764	0.28173

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

Table 5: Comparison of posterior results for k

	mean[a]	sd[a]	mean/sd
(a)[b]	23.7454	3.79746	6.25297
(b)	23.1065	3.40147	6.79309
(c)	23.9645	3.33524	7.18525
(d)	22.9075	3.50345	6.53855
(e)	23.2462	3.44160	6.75449
(f)	19.4390	3.74489	5.19080
(g)	29.1526	3.69131	7.89763

a: 'mean' and 'sd' denote the posterior mean and the posterior standard deviation, respectively.
b: The case (a) has the hyperparameter values given in (32). The other cases change the hyperparameter values as follows:

(b) $\kappa = \text{diag}\{1/R_t^2\}$, (c) $\kappa = \text{diag}\{0.01/R_t^2\}$, (d) $h = \text{diag}\{200g/\alpha R_t^2\}$
(e) $h = \text{diag}\{50g/\alpha R_t^2\}$, (f) $b = 0.2$, (g) $\lambda = 3$.

Table 6: Total, chronic and transient poverty measures from posterior results

$\zeta = 0$						
$\Upsilon = \mathrm{e}^{\upsilon}$		Total (y)	Total (x)	Chronic	Transient (y)	Transient (x)
3000	mean[a]	0.02731893	0.02727368	0.01590985	0.01140908	0.01136383
	sd[a]	0.00317278	0.00317286	0.00291239	0.00122075	0.00121899
4000	mean	0.04681263	0.04672895	0.03166074	0.01515189	0.01506822
	sd	0.00421828	0.00421858	0.00407234	0.00141076	0.0014077
5000	mean	0.07173526	0.07159818	0.05333663	0.01839863	0.01826155
	sd	0.00523897	0.00524136	0.00519019	0.00156528	0.00156304
6000	mean	0.10265384	0.10245039	0.08156317	0.02109067	0.02088722
	sd	0.00624255	0.00624686	0.00629741	0.00167403	0.00166991
7000	mean	0.13973727	0.13947279	0.11671308	0.02302419	0.02275971
	sd	0.00723133	0.00724068	0.00741385	0.00174528	0.00173844
8000	mean	0.18254901	0.1822352	0.15853889	0.02401011	0.0236963
	sd	0.00815178	0.00816169	0.00849056	0.00183797	0.00183053
9000	mean	0.23008393	0.22974655	0.20614708	0.02393684	0.02359947
	sd	0.00895475	0.00896629	0.0094533	0.00191798	0.00191156
10000	mean	0.28094416	0.28061214	0.2580891	0.02285506	0.02252304
	sd	0.00959933	0.00961201	0.01022574	0.00199301	0.00198606
$\zeta = 1$						
$\Upsilon = \mathrm{e}^{\upsilon}$		Total (y)	Total (x)	Chronic	Transient (y)	Transient (x)
3000	mean	0.00955136	0.0095358	0.00451094	0.00504042	0.00502486
	sd	0.00148649	0.00148617	0.00117329	0.00063994	0.0006394
4000	mean	0.0163228	0.01629547	0.00921085	0.00711195	0.00708462
	sd	0.0019771	0.00197681	0.00168675	0.00072034	0.00071965
5000	mean	0.02481727	0.02477371	0.01576494	0.00905233	0.00900877
	sd	0.00245645	0.00245632	0.00219325	0.00077997	0.00077913
6000	mean	0.03512734	0.03506278	0.02428456	0.01084278	0.01077822
	sd	0.00292255	0.0029226	0.00268945	0.00082228	0.00082146
7000	mean	0.04735069	0.04726171	0.03489508	0.01245561	0.01236662
	sd	0.00337605	0.00337656	0.00317369	0.00084773	0.00084691
8000	mean	0.06151954	0.06140511	0.04767046	0.01384907	0.01373465
	sd	0.00381591	0.00381692	0.0036488	0.00085892	0.00085838
9000	mean	0.07757048	0.07743194	0.06258645	0.01498403	0.01484549
	sd	0.00423826	0.00424002	0.00411166	0.00086173	0.00086155
10000	mean	0.0953435	0.09518496	0.0795098	0.0158337	0.01567516
	sd	0.00463689	0.00463943	0.0045558	0.0008606	0.00086084

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

Table 6: Continued

$\zeta = 2$						
$\Upsilon = \mathrm{e}^{\upsilon}$		Total (y)	Total (x)	Chronic	Transient (y)	Transient (x)
3000	mean	0.00496513	0.00495709	0.00197067	0.00299447	0.00298642
	sd	0.00096918	0.0009689	0.00067769	0.00049341	0.00049311
4000	mean	0.0084615	0.00844762	0.00409298	0.00436852	0.00435464
	sd	0.00128968	0.00128938	0.00099506	0.00055774	0.00055736
5000	mean	0.01282438	0.01280277	0.0071018	0.00572257	0.00570097
	sd	0.00160591	0.00160562	0.00131773	0.00060761	0.00060715
6000	mean	0.01806409	0.01803258	0.01103818	0.00702591	0.0069944
	sd	0.00191617	0.00191592	0.00163886	0.00064684	0.00064636
7000	mean	0.02421111	0.02416756	0.01594876	0.00826235	0.0082188
	sd	0.00221942	0.00221927	0.00195558	0.00067708	0.0006766
8000	mean	0.03129402	0.03123677	0.02187725	0.00941676	0.00935952
	sd	0.00251543	0.00251544	0.00226671	0.00069907	0.00069864
9000	mean	0.03932323	0.03925135	0.02884997	0.01047326	0.01040139
	sd	0.00280377	0.00280401	0.00257209	0.00071401	0.00071372
10000	mean	0.04828269	0.04819615	0.03686562	0.01141708	0.01133054
	sd	0.00308371	0.00308423	0.00287133	0.00072317	0.00072309

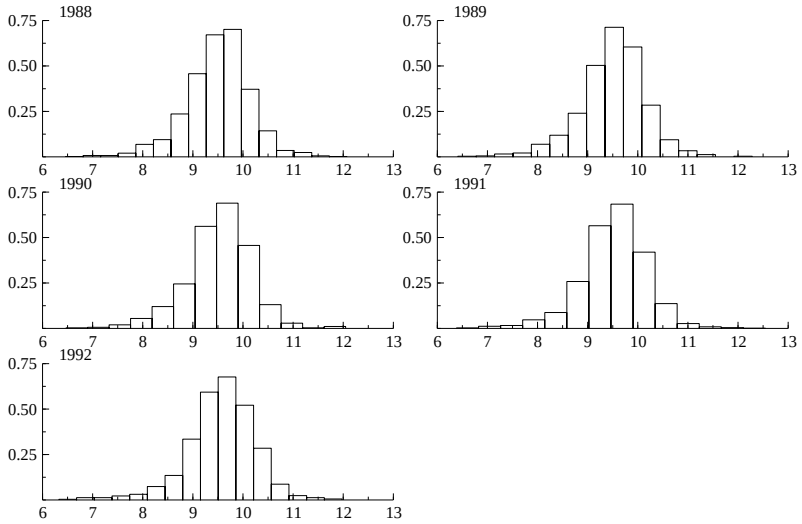

Figure 1: Histograms of $y = \log Y$

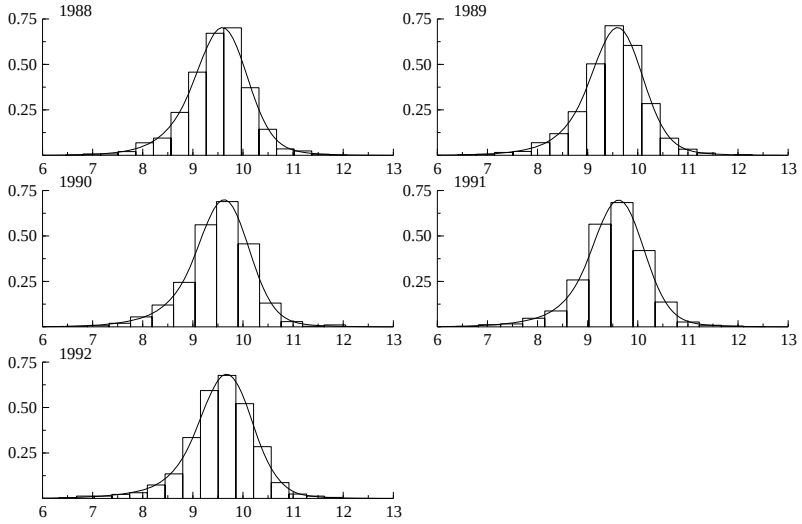

Figure 2: Estimated densities of x

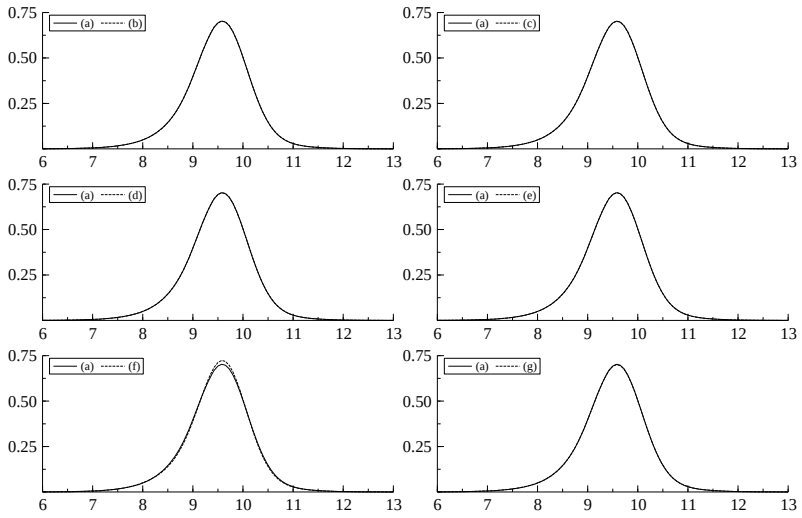

Figure 3: Comparison of estimated densities of $\boldsymbol{x}$ [a]

a: The case (a) has the hyperparameter values given in (32). The other cases change the hyperparameter values as follows:

(b) $\kappa = \mathrm{diag}\{1/R_t^2\}$, (c) $\kappa = \mathrm{diag}\{0.01/R_t^2\}$, (d) $h = \mathrm{diag}\{200g/\alpha R_t^2\}$
(e) $h = \mathrm{diag}\{50g/\alpha R_t^2\}$, (f) $b = 0.2$, (g) $\lambda = 3$.

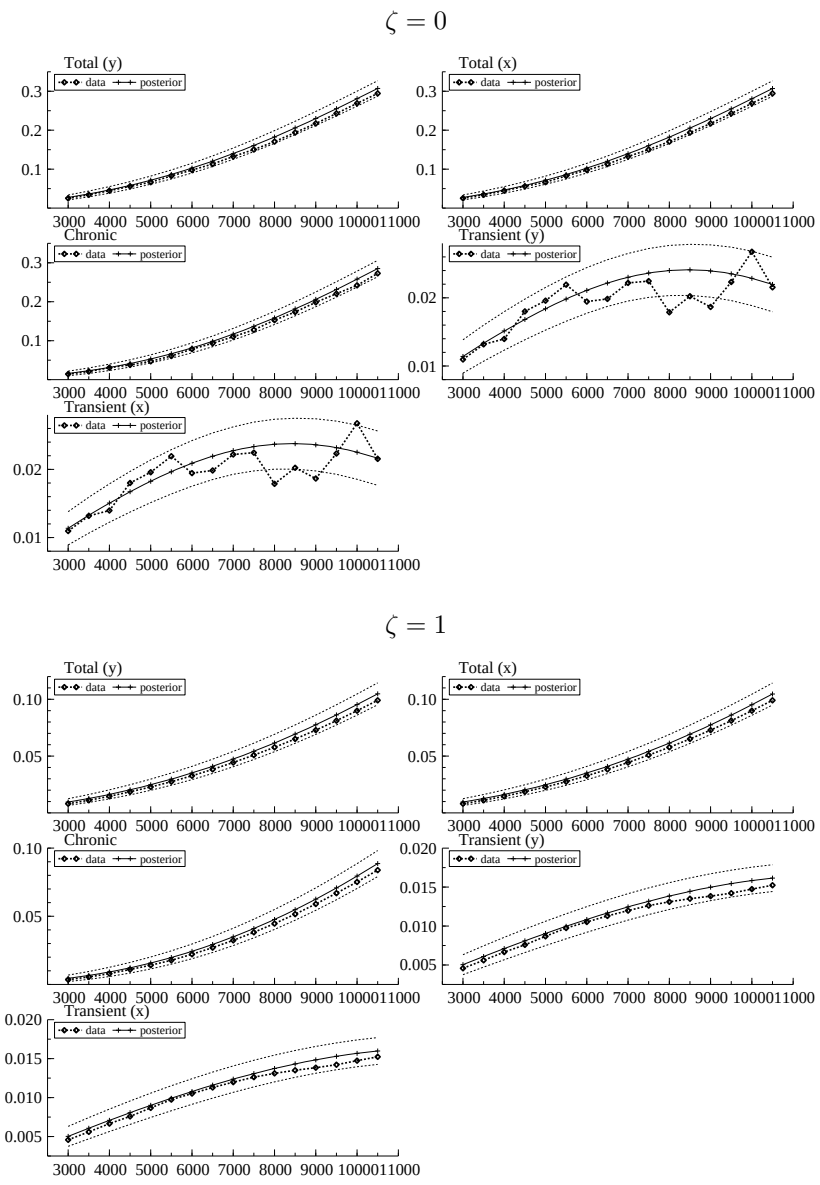

Figure 4: Total, chronic and transient poverty[a]

a: The dotted lines denote the following posterior interval:

$$(\text{posterior mean} - 2 \times \text{posterior sd},\ \text{posterior mean} + 2 \times \text{posterior sd}).$$

$\zeta = 2$

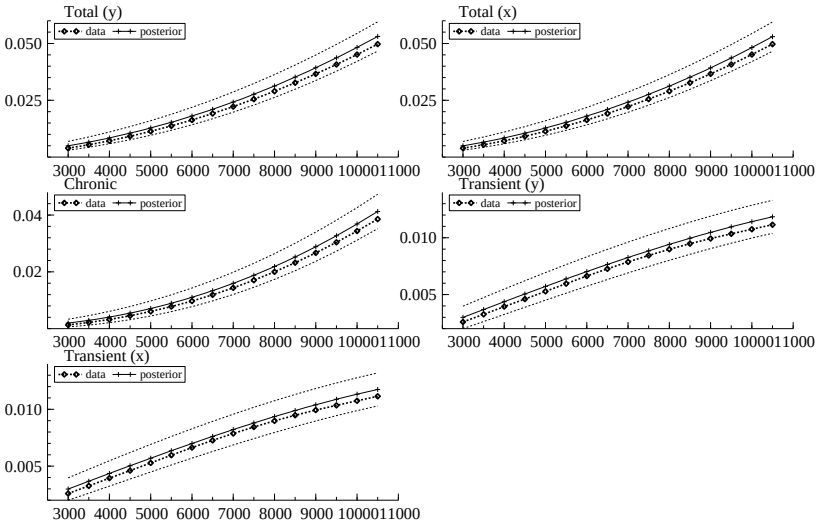

Figure 4: Continued

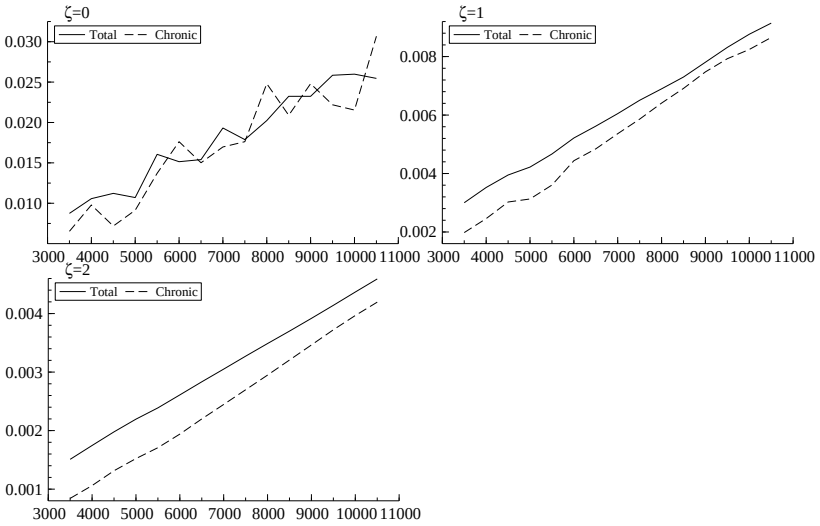

Figure 5: Fluctuation of total and chronic poverty on data by Ravallion's method