Bayesian Estimation of Beta-type Distribution Parameters Based on Grouped Data

Kakamu, Kazuhiko

Nishino, Haruhisa

(Citation)
Computational Economics,54(2):625-645

(Issue Date)
2019-08

(Resource Type)
journal article

(Version)
Accepted Manuscript

Bayesian Estimation of Beta-type Distribution Parameters Based on Grouped Data [*]

Kazuhiko Kakamu [†] Haruhisa Nishino [‡]

August 7, 2018

[*]Previous versions of this paper were presented at the European Seminar on Bayesian Econometrics (ESOBE 2012) in Vienna, Computational and Financial Econometrics (CFE 2012) in Oviedo, Econometric Society of Australia Meeting (ESAM 2013) in Sydney, Japanese Joint Statistical Meeting (JJSM 2013) in Osaka, International Society for Bayesian Analysis World Meeting (ISBA 2014) in Cancun, First Eastern Asia Meeting on Bayesian Statistics (ISBA-EAC 2016) in Shanghai, and Tripartite Conference 2017 at Singapore Management University as well as the seminars at Kobe University, Tokyo Institute of Technology, Keio University, Hiroshima University, and Monash University. We would like to thank the seminar/conference participants, especially Sungbae An, Duangkamon Chotikapanich, William E. Griffiths, Gholamreza Hajargasht, Hideo Kozumi, Yasuhiro Omori, Scott Sisson, and Herman van Dijk. We would also like to thank the editor and an anonymous referee for their valuable comments and suggestions, which have improved the paper substantially. Part of this research was conducted while the author was visiting Institute for Economic Geography and GIScience, Vienna University of Economics and Business, whose hospitality is gratefully acknowledged. This work was partially supported by KAKENHI #26380266, #25245035, #16K03592, and #16KK0081.

[†]Corresponding author. Graduate School of Business Administration, Kobe University, 2-1, Rokkodai, Nada, Kobe 657-8501, Japan. Email: `kakamu@person.kobe-u.ac.jp`

[‡]Graduate School of Social Sciences, Hiroshima University, 1-2-1 Kagamiyama, Higashi-Hiroshima, Hiroshima 739-8525, Japan. Email: `hnishino@hiroshima-u.ac.jp`

Abstract

This study considers a method of estimating generalized beta (GB) distribution parameters based on grouped data from a Bayesian point of view and explores the possibility of the GB distribution focusing on the goodness-of-fit because the GB distribution is one of the most typical five-parameter distributions. It uses a tailored randomized block Metropolis–Hastings (TaRBMH) algorithm to estimate the GB distribution parameters and this method is then applied to one simulated and two real datasets. Moreover, the fit of the GB distribution is compared with those of the generalized beta distribution of the second kind (GB2 distribution) and Dagum (DA) distribution by using the marginal likelihood. The estimation results of simulated and real datasets show that the GB distributions are preferred to the DA distributions in general, while the GB2 distributions have similar performances to the GB distributions. In other words, the GB2 distribution could be adopted as well as the GB distribution in terms of the smallest possible number of parameters, although our TaRBMH algorithm can estimate the GB distribution parameters efficiently and accurately. The accuracy of the Gini coefficients also suggests the use of the GB2 distributions.

Key words: Dagum distribution; Generalized beta (GB) distribution; Generalized beta distribution of the second kind (GB2 distribution); Gini coefficient; Grouped data; Tailored randomized block Metropolis–Hastings (TaRBMH) algorithm.

1 Introduction

In seminal works estimating hypothetical income distributions, a number of alternative probability density functions (PDFs) have been proposed as models of the income distribution, as shown by Salem and Mount (1974). These PDFs provide a reasonably close approximation to the true distribution and their parameters should be sufficiently simple to interpret in an economically meaningful way. For example, the Pareto and lognormal distributions have parameters that can be directly related to inequality measures, but neither fits the full range of income data very well. Therefore, alternative flexible PDFs such as the double Pareto-lognormal distribution proposed by Reed and Jorgensen (2004) and κ-generalized distribution of Clementi et al. (2007, 2016) have been considered.

A family of beta-type size distributions can also be successfully fitted to income data. For example, the generalized beta (GB) distributions of the first and second kinds (GB1 and GB2 distributions, respectively) proposed by McDonald (1984), Dagum (DA) distribution proposed by Dagum (1977), Singh–Maddala distribution proposed by Singh and Maddala (1976), gamma distribution, chi-square distribution, and exponential distribution (see Figure 2 in McDonald and Xu, 1995, for details on the relationship of the distributions) have been considered to be hypothetical distributions in the family of beta-type size distributions. In particular, the successful empirical estimation of the GB2 distribution was complemented by Parker's (1999) theoretical models of income generation and Bordley et al. (1997) found that the GB2 distribution generally provides a significantly better fit than its nested distributions when fit to income data from the United States.

However, although such flexible distributions including the GB2 distribution provide us with an estimation approach because of computational progress over time, we must also consider the smallest possible number of parameters to ensure an adequate and meaningful representation (see Dagum, 1977). Therefore, several hypothetical distributions have been compared in studies such as Atoda et al. (1988); Bordley et al. (1997); Kloek and van Dijk (1978); McDonald and Mantrala (1995); McDonald and Ransom (1979b); Majumder and Chakravarty (1990); Salem and Mount (1974); Slottje (1984); Tachibanaki et al. (1997) and recent studies support the fit of the GB2 distribution. As a more flexible distribution, McDonald and Xu (1995) proposed the GB distribution, which should perform at least as well as other beta-type distributions because it includes the GB2 distribution as a special case. However, although McDonald and Xu (1995); Bordley et al. (1997); McDonald and Ransom (2008) empirically examined the GB distribution, the success of its estimation has not thus far been reported to our knowledge.

Grouped data has been widely used to estimate income distributions because they are available in many countries (see Kleiber and Kotz, 2003, Appendix B). [1] In grouped data, suppose that the income units are grouped into $(k+1)$ income classes, viz., (x_0 to x_1), (x_1 to x_2), ..., (x_k to x_{k+1}) with $x_0 = 0$ and $x_{k+1} \leq \infty$: let n be the total number of units and $n_i^* = n_i - n_{i-1}$ with $n_0 = 0$ and $n_{k+1} = n$ be the amount of earning income in the intervals x_{i-1} and x_i, which varies from 1 to $k+1$. In this situation, the estimation of the hypothetical distribution parameters explores the distribution (i.e., the dotted line), which is fitted to the data, from the histogram shown in Figure 1.

Although maximum likelihood estimation (MLE) is one of the most common methods of estimating the parameters of hypothetical distributions from grouped data (see McDonald and Ransom, 2008), several other estimation methods have been used, including the minimum chi-square, scoring, least squares, method of moments, least lines, generalized method of moments, generalized least squares, and Markov chain Monte Carlo (MCMC) methods (see e.g. Chotikapanich and Griffiths, 2000; Chotikapanich et al., 2007; McDonald and Ransom, 1979b,a; Nishino and Kakamu, 2011; van Dijk and Kloek, 1980). Moreover, the GB distribution has already been estimated using MLE by McDonald and Xu (1995), and Bordley et al. (1997) applied the method to the U.S. income data. However, the MCMC method is a powerful tool not only for the estimation but also for the inference as an alternative to MLE. As shown by Gelfand et al. (1990), the MCMC method can yield valid inferences on the nonlinear functions of the parameters, such as Gini coefficients. For example, the credible intervals are obtained not only for the parameters but also for the Gini coefficients. Moreover, as stated by McDonald and Ransom (1979b) and Kakamu (2016), it is sometimes difficult to find the mode of the parameters

[1] Studies of the estimation of the income distribution from individual data exist, such as Chotikapanich and Griffiths (2008); Hasegawa and Kozumi (2003); Tachibanaki et al. (1997). Another type of grouped dataset includes population shares and group mean incomes, and Hajargasht et al. (2012) considered this situation.

as their number increases by using MLE. Therefore, it would be worthwhile taking a Bayesian approach to reexamine the parameter estimation of the GB distribution.

This study considers an efficient Bayesian estimation procedure of the GB distribution by using the tailored randomized block Metropolis–Hastings (TaRBMH) algorithm, [2] which was proposed by Chib and Ramamurthy (2010) over the existing procedure, namely the random walk Metropolis–Hastings (RWMH) algorithm, to focus on the efficiency of the algorithm and accuracy of the Bayesian estimates. Moreover, by using the proposed algorithm, the goodness-of-fit of the GB distribution is explored, using a simulated dataset and two real datasets from the United States and Japan.

First, from the simulated dataset, we confirm that this procedure can estimate the parameters more efficiently than the existing method in terms of mixing and that the parameters are estimated precisely from the grouped data. When we compare the fit of the distributions, the GB distribution is preferred to the DA distribution. However, the fit of the GB2 distribution is almost equivalent to that of the GB distribution. In other words, the use of the GB2 distribution is preferred to the GB distribution in terms of the smallest possible number of parameters, even if the parameters of the GB distribution are estimated accurately. Second, by using the two real datasets, the empirical results of the U.S. data show that the GB2 distributions are preferred to the GB distribution, although the maximum likelihood estimates are slightly different from the Bayesian estimates. We should notice that the Baysian estimates of c of the GB distribution in Table 3 are sometimes substantially different from 1, and it suggests that the GB distribution could be more suitable for the data than the GB2 because in the case of $c = 1$ the GB distribution is reduced to the GB2 distribution. The empirical results of the Japanese data also confirm that the GB2 distributions are preferred to the GB distribution, although the GB distributions are preferred to the DA distribution. Moreover, the accuracy of the Gini coefficients also suggests the use of the GB2 distribution.

The rest of this paper is organized as follows. In the next section, we introduce the features of the GB

[2] Although this method has been proposed for dynamic stochastic general equilibrium models, whose identification is weak, the TaRBMH algorithm performs well in Tobit models with an asymmetric Laplace distribution according to Wichitaksorn and Tsurumi (2013). As shown later, the GB distribution encounters the same situation.

distribution, including the PDF, cumulative distribution function (CDF), and likelihood function, and explain the MCMC estimation procedure for this distribution. In Section 3, we examine the numerical example of the simulated dataset focusing on the efficiency of the MCMC draws and accuracy of the Bayesian estimates. Moreover, the goodness-of-fit of the GB distribution is examined. In Section 4, three real datasets that include income data from the United States and Japan are examined to confirm the results in Section 3 from an empirical point of view. In the Japanese dataset, the performance of the Gini coefficients is also discussed. In Section 5, we conclude the discussion and state the remaining issues.

2 The GB Distribution

2.1 PDF, CDF, and Likelihood Function

While various probability distributions are used to estimate the parameters of a hypothetical income distribution, the GB distribution includes various probability distributions as special or limiting cases (see McDonald and Xu, 1995), most of which were presented in the previous section. Therefore, it is reasonable to estimate the parameters of the GB distribution because this distribution performs at least as well as the special or limiting distributions.

The GB distribution has five parameters $(a,\ b,\ c,\ p,\ q)$, and its PDF is written as

$$f(x) \quad = \quad \frac{|a|x^{ap-1}\left[1-(1-c)\left(\frac{x}{b}\right)^{a}\right]^{q-1}}{b^{ap}B(p,q)\left[1+c\left(\frac{x}{b}\right)^{a}\right]^{p+q}}, \quad 0 < x^a < \frac{b^a}{1-c}, \tag{1}$$

where $B(p,q)$ is a beta function. For example, if we set $c=0$ or $c=1$, the distributions are reduced to GB1 and GB2 distributions, respectively. Moreover, if we set $c=1$ and $q=1$, it is reduced to a DA distribution (see Dagum, 1977), which is confirmed as a desirable distribution in many empirical applications (see Kleiber, 2008). The detailed relationships among the class of distributions are summarized in McDonald and Xu (1995).

To introduce the CDF, we provide the following function:

$$I_x(p,q) = \frac{B_x(p,q)}{B(p,q)},$$

where $B_x(p,q)$ is an incomplete beta function. Then, the CDF is written as

$$F(x) = I_z(p,q), \text{ where } z = \frac{\left(\frac{x}{b}\right)^a}{1 + c\left(\frac{x}{b}\right)^a}. \tag{2}$$

Given the PDF and CDF, we define the likelihood function following Nishino and Kakamu (2011), which is based on the concept of selected order statistics. [3] To explain the likelihood function, let $\boldsymbol{\theta} = (a,\ b,\ c,\ p,\ q)'$ be the vector of parameters and let $\mathbf{x} = (x_1,\ x_2, \ldots,\ x_k)'$ be the vector of observations. Then, the likelihood function is defined as follows:

$$L(\mathbf{x}|\boldsymbol{\theta}) = n! \frac{F(x_1)^{n_1-1}}{(n_1-1)!} f(x_1) \left\{ \prod_{i=2}^{k} \frac{(F(x_i) - F(x_{i-1}))^{n_i - n_{i-1} - 1}}{(n_i - n_{i-1} - 1)!} f(x_i) \right\} \frac{(1 - F(x_k))^{n - n_k}}{(n - n_k)!}. \tag{3}$$

If we substitute (1) and (2) for (3), it becomes the likelihood function for the GB distribution. [4]

2.2 Posterior Analysis

Because we adopt a Bayesian approach, we complete the model by specifying the prior distribution over the parameters. [5] We apply the following prior:

$$\pi(\boldsymbol{\theta}) = \pi(a)\pi(b)\pi(c)\pi(p)\pi(q).$$

Given a prior density $\pi(\boldsymbol{\theta})$ and the likelihood function in (3), the joint posterior distribution can be expressed as

$$\pi(\boldsymbol{\theta}|\mathbf{x}) \propto \pi(\boldsymbol{\theta}) L(\mathbf{x}|\boldsymbol{\theta}). \tag{4}$$

[3] In MLE, the likelihood (which is based on the multinomial distribution) is widely used. Nishino and Kakamu (2011) applied a likelihood based on selected order statistics to the lognormal distribution, which is more exact than that based on the multinomial distribution. Therefore, we use this likelihood and the concept of selected order statistics as summarized in, for example, David and Nagaraja (2003). However, which likelihood we use is not critical, because the difference is small.

[4] If the PDF and CDF for the distribution in question are available, we can apply this likelihood function to the distribution. The PDFs and CDFs for the beta-type distributions are summarized in, for example, Hajargasht et al. (2012); Kleiber and Kotz (2003).

[5] A Bayesian approach was first proposed by Chotikapanich and Griffiths (2000) using an RWMH algorithm and modified by Kakamu (2016) for more general cases. In this paper, we compare the performance of their algorithm with our proposed algorithm in the numerical examples.

Finally, we assume the following prior distributions:

$$a \sim \mathcal{N}(\mu_0, \tau_0^2),\ b \sim \mathcal{G}(\alpha_0, \beta_0),\ c \sim \mathcal{B}(\gamma_0, \delta_0),\ p \sim \mathcal{G}(\epsilon_0, \zeta_0),\ q \sim \mathcal{G}(\eta_0, \nu_0),$$

where $\mathcal{G}(a, b)$ and $\mathcal{B}(a, b)$ denote the gamma and beta distributions, respectively, because these prior distributions satisfy the parameter constraints.

To obtain the posterior estimates, we implement the TaRBMH algorithm proposed by Chib and Ramamurthy (2010) as follows.

1. Separate $\boldsymbol{\theta}$ into a 3×1 vector $\boldsymbol{\theta}_1^{(m-1)}$ and a 2×1 vector $\boldsymbol{\theta}_2^{(m-1)}$ randomly.

2. For $j = 1,\ 2$, implement the following Metropolis–Hastings steps.

   (a) Generate $\boldsymbol{\theta}_j^{new}$ from a multivariate t distribution, $t(\hat{\boldsymbol{\theta}}_j, \boldsymbol{\Sigma}_j, \nu)$, with mean $\hat{\boldsymbol{\theta}}_j$, covariance $\boldsymbol{\Sigma}_j$, and ν degrees of freedom. [6] Here,

$$\hat{\boldsymbol{\theta}}_j = \underset{\boldsymbol{\theta}_j}{\arg\max}\ \log\left\{\pi(\boldsymbol{\theta})L(\mathbf{x}|\boldsymbol{\theta})\right\},$$

   $\boldsymbol{\theta} = (\boldsymbol{\theta}_1', \boldsymbol{\theta}_2^{(m-1)\prime})'$ for $j = 1$, and $\boldsymbol{\theta} = (\boldsymbol{\theta}_1^{(m)\prime}, \boldsymbol{\theta}_2')'$ for $j = 2$ using simulated annealing by Goffe et al. (1994). Here,

$$\boldsymbol{\Sigma}_j = \left(- \frac{\partial^2 \log\left\{\pi(\boldsymbol{\theta})L(\mathbf{x}|\boldsymbol{\theta})\right\}}{\partial \boldsymbol{\theta}_j \partial \boldsymbol{\theta}_j'} \right)^{-1} \Bigg|_{\boldsymbol{\theta}_j = \hat{\boldsymbol{\theta}}_j}.$$

   (b) If $j = 1$, compute

$$\alpha_1(\boldsymbol{\theta}_1^{m-1}, \boldsymbol{\theta}_1^{new}) = \min\left\{ \frac{\pi(\boldsymbol{\theta}_1^{new}|\boldsymbol{\theta}_2^{(m-1)}, \mathbf{x})q(\boldsymbol{\theta}_1^{(m-1)}|\hat{\boldsymbol{\theta}}_1, \boldsymbol{\Sigma}_1)}{\pi(\boldsymbol{\theta}_1^{(m-1)}|\boldsymbol{\theta}_2^{(m-1)}, \mathbf{x})q(\boldsymbol{\theta}_1^{new}|\hat{\boldsymbol{\theta}}_1, \boldsymbol{\Sigma}_1)}, 1 \right\},$$

   and if $j = 2$, compute

$$\alpha_2(\boldsymbol{\theta}_2^{m-1}, \boldsymbol{\theta}_2^{new}) = \min\left\{ \frac{\pi(\boldsymbol{\theta}_2^{new}|\boldsymbol{\theta}_1^{(m)}, \mathbf{x})q(\boldsymbol{\theta}_2^{(m-1)}|\hat{\boldsymbol{\theta}}_2, \boldsymbol{\Sigma}_2)}{\pi(\boldsymbol{\theta}_2^{(m-1)}|\boldsymbol{\theta}_1^{(m)}, \mathbf{x})q(\boldsymbol{\theta}_2^{new}|\hat{\boldsymbol{\theta}}_2, \boldsymbol{\Sigma}_2)}, 1 \right\},$$

   where $q(\boldsymbol{\theta}_j^{new}|\hat{\boldsymbol{\theta}}_j, \boldsymbol{\Sigma}_j)$ is a multivariate t distribution given in (a).

[6] In the numerical examples discussed below, we set $\nu = 15$ as recommended by Chib and Ramamurthy (2010). In addition, as in Chib and Ramamurthy (2010), the inverse of the negative Hessian, which is a covariance matrix $\boldsymbol{\Sigma}_j$, may not be positive definite. Thus, we also compute the modified Cholesky algorithm proposed by Nocedal and Wright (2000).

(c) Generate a value u_j from $\mathcal{U}(0,1)$, where $\mathcal{U}(a,b)$ is a uniform distribution on the interval (a,b).

(d) If $u_j \leq \alpha_j\left(\boldsymbol{\theta}_j^{(m-1)}, \boldsymbol{\theta}_j^{new}\right)$, set $\boldsymbol{\theta}_j^{(m)} = \boldsymbol{\theta}_j^{new}$, otherwise $\boldsymbol{\theta}_j^{(m)} = \boldsymbol{\theta}_j^{(m-1)}$.

3. Return to step 1, and set m to $m+1$.

In all the numerical examples discussed in the next section, we set the hyper-parameters to $\mu_0 = 0$, $\tau_0^2 = 100$, and $\gamma_0 = \delta_0 = \epsilon_0 = \zeta_0 = \eta_0 = \nu_0 = 1.0$. The results reported in the next section are generated by using Ox version 7.10 (OS_X_64/U) (see Doornik, 2013) and the figures are generated by using R version 3.4.3 (see R Core Team, 2017).

3 Numerical Example with the Simulated Data

To illustrate the Bayesian approach discussed in the previous section, we first compare the algorithm with that proposed by Chotikapanich and Griffiths (2000) and Kakamu (2016) by using the simulated dataset. We set the number of observations to $n = 5,000$ and assume a decile number of groups ($k = 9$). Given n and k, we assume that the true data-generating process (DGP) is a GB distribution, [7] where the parameters are $a = 5$, $b = 30$, $c = 0.95$, $p = 0.5$, and $q = 0.8$. Furthermore, generate x_j for $j = 1, 2, \ldots, n$. The generated random numbers are sorted in ascending order, and x_i corresponds to the n_ith observation, where $n_i = n \times \dfrac{i}{k+1}$. Then, $i = 1, 2, \ldots, k$ is picked up and $\mathbf{x} = (x_1, x_2, \ldots, x_k)'$ are collected. Given the dataset, we run an RWMH algorithm using $800,000$ iterations and discarding the first $300,000$, while we run a TaRBMH algorithm using $11,000$ iterations and discarding the first $1,000$.

Table 1 shows the posterior estimates of the parameters using the RWMH and TaRBMH algorithms. From Table 1, we first observe that the posterior means are close to each other and share similar 95% credible intervals. In addition, all the parameters include true values in the 95% credible intervals. Therefore, we conclude that both Bayesian approaches work well in the parameter estimation of the GB distribution. However, a greater

[7] To generate a random number from the GB distribution with the parameters a, b, c, p, and q, first generate Z from a beta distribution with the parameters p and q. Then, $X = b\left(\dfrac{Z}{1-cZ}\right)^{\frac{1}{a}}$ becomes the random number from the GB distribution.

than 80-times difference appears in the inefficiency factors, which approximate the ratio of the numerical variance of the estimate from the MCMC chain relative to that from the hypothetical i.i.d. draws.

To illustrate the differences in the mixing of the two algorithms over the parameter space, Figure 2 plots the sample draws from the posterior distribution for the parameters together with the auto-correlation functions (ACFs). The top panel corresponds to the draws from the TaRBMH algorithm and the bottom panel corresponds to those from the RWMH algorithm. As these plots show, the RWMH chain is highly persistent with the ACFs retaining significant mass even at a lag length of 1000. On the contrary, the TaRBMH ACFs decay quickly (within a lag length of 200 for all the parameters).

To explore why the convergence of the MCMC chain is slow, Figure 3 shows the contour plots of the marginal joint posterior distributions of each parameter. This figure implies that the correlations among a, c, p, and q are very high. In the GB distribution, the shape parameters are a, p, and q, and several combinations of these three parameters might lead to similar distribution shapes. Therefore, such high correlations lead to an identification problem and slow the convergence of the MCMC chain. However, because the TaRBMH algorithm speeds up convergence, we can conclude that our algorithm is superior to the RWMH algorithm in terms of mixing.

Having confirmed the advantages of our TaRBMH algorithm and the accuracy of the Bayesian estimates from the GB distribution, we next consider the GB distribution. To examine the goodness-of-fit of the GB distribution, we compare it with the GB2 distribution, where $c = 1$ is assumed in the GB distribution and DA distribution, where $q = 1$ is also assumed in the GB2 distribution. As mentioned in the Introduction, the successful empirical estimation of the GB2 distribution was complemented by Parker's (1999) theoretical models of income generation and Bordley et al. (1997) found that the GB2 distribution generally provided a significantly better fit than its nested distributions when fit to income data from the United States. Moreover, among the three parameter distributions, the DA distribution fits the real data well (see Kleiber, 2008). Therefore, we compare these two distributions with the GB distribution. [8]

[8] To estimate these two distributions, the RWMH algorithm using $40,000$ iterations and discarding the first $20,000$ is adopted for the DA distribution, while the RWMH algorithm using $300,000$ iterations and discarding the first $100,000$ is adopted for the GB2

Table 2 shows the log of marginal likelihood using the harmonic mean estimator proposed by Newton and Raftery (1994) with the standard errors. It confirms that the highest log of marginal likelihood can be found in the GB distribution using the TaRBMH algorithm. However, if the standard errors are taken into account, we can conclude that the fit of the GB distribution is almost equivalent to the fit of the GB2 distribution, while the GB and GB2 distributions fit better than the DA distribution. Therefore, the use of the GB2 distribution is recommended, even though the true DGP is the GB distribution and its parameters are estimated accurately. On the contrary, the estimation of the GB2 distribution has the same problem as that of the GB distribution, namely the slow convergence of the MCMC chain (not reported in the paper to save the space). Thus, the use of the TaRBMH algorithm, as proposed in this paper, is also desirable in terms of mixing. [9]

4 Numerical Examples with the Real Data

Having confirmed that the proposed algorithm can estimate the GB distribution parameters more efficiently than the RWMH algorithm but that the use of the GB2 distribution is preferred to the GB distribution, we next confirm the results from an empirical point of view by using two real datasets.

4.1 U.S. Family Income Data

In this subsection, we examine the empirical performance of the GB distribution focusing on the performance of the TaRBMH algorithm over the RWMH algorithm and MLE and the accuracy of the Bayesian estimator, using U.S. data. Moreover, the fit of the GB distribution in the U.S. data is examined.

The literature McDonald and Xu (1995) and Bordley et al. (1997) has greatly contributed to the estimation of the GB distribution for grouped income data. While McDonald and Xu (1995) did not seem to estimate parameter c and estimated the GB2 distribution instead, Bordley et al. (1997) estimated all parameters including c every five years from 1970 to 1990. Therefore, it is reasonable to compare the results from Bordley et al.

distribution. The prior distributions and hyper-parameters are the same as in the GB distribution.

[9] If the standard errors are taken into account, it is not important which algorithm we use in the GB distribution. Therefore, we also use the log of marginal likelihood from the TaRBMH algorithm in the next section.

(1997) with the Bayesian approaches; the U.S. income data were estimated in Bordley et al. (1997). To estimate the parameters of the GB distribution from these data, we run an RWMH algorithm using $200,000$ iterations and discarding the first $100,000$, while we run a TaRBMH algorithm using $11,000$ iterations and discarding the first $1,000$ for every year.

Table 3 shows the estimation results from Bordley et al. (1997) and the posterior estimates of the parameters using the RWMH and TaRBMH algorithms. First, if we compare the posterior estimates of both algorithms, we see that the 95% credible intervals from the TaRBMH algorithm are slightly wider than those from the RWMH algorithm. However, the posterior means are generally close to each other and the inefficiency factor from the TaRBMH algorithm is much smaller than that from the RWMH algorithm. Therefore, the TaRBMH algorithm also searches through wider parameter spaces efficiently in a real dataset. Hence, we focus on the results from the TaRBMH algorithm hereafter.

Next, we compare the estimation results from Bordley et al. (1997) with the posterior estimates. Focusing on the results for 1970 shows that all maximum likelihood estimates are close to the posterior means and are included in the 95% credible intervals. The same situation occurs for the results for 1975. However, for the results for 1980, the maximum likelihood estimates of the parameters b, c, and q are not within the 95% credible intervals of the Bayesian estimates. In addition, none of the parameters in 1985 and 1990 is within the 95% credible intervals. Therefore, these results imply that the maximum likelihood estimates may not achieve the optima.

To confirm the effects of parameter c, Figure 4 displays its marginal posterior distributions. This figure shows a truncation of the posterior distribution at 1.0 in 1975. However, the truncation is not clear in the other years, especially in 1970 and 1990. Therefore, the estimates of c in 1970, 1980, 1985, and 1990 do not seem to be 1.0. The differences in the estimates may be caused by the failure of the global maximization in our MLE. Moreover, parameter c plays an important role in the estimation of the income distribution in the United States because it is estimated to be different from 1.0 if we assume the GB distribution.

Although Bordley et al. (1997) confirmed the fit of the GB2 distribution using MLE, the estimates are different from the Bayesian estimates. Therefore, it is also reasonable to reexamine the fit of the distributions by using the Bayesian estimates; we also examine the fit of the distributions by using the harmonic mean estimator as with the simulated data in the previous section. Table 4 presents the log of marginal likelihood with the standard errors, showing that the highest log can be found in the GB distribution in 1975 as well as in the GB2 distribution in 1980, 1985, 1990, and 1995. Even with the standard errors taken into account, the fit of the GB distribution is superior to that of the GB2 distribution in 1975 and the fit of the GB2 distribution is superior to that of the GB distribution in 1985 and 1990. In 1985 and 1995, the fit of the GB distribution is almost equivalent to that of the GB2 distribution. In addition, the fit of the GB and GB2 distributions is superior to that of the DA distribution from 1975 to 1990, whereas the fit of the GB distribution is not in 1995.

To summarize the results above, we obtain the same results here we did from the simulated data, reconfirming that our algorithm is superior to the RWMH algorithm in terms of mixing. Moreover, the maximum likelihood estimates differ from the Bayesian estimates and both Bayesian estimates are close to each other. Therefore, the Bayesian estimates might be able to estimate the GB distribution parameters more accurately than MLE. In addition, the GB distribution is superior to the GB2 and DA distributions in terms of goodness-of-fit when using the log of marginal likelihood only in 1975. Although the estimates are accurate, the GB2 distribution fits as well as the GB distribution in general. Therefore, if we take the smallest possible number of parameters into account, the use of the GB2 distribution is also recommended empirically in general and these results are consistent with those of Bordley et al. (1997). Finally, by using the Bayesian estimates from the TaRBMH algorithm, we examine the empirical findings briefly.

Once the parameters of the GB distribution are estimated, we can draw the income distribution and calculate the Gini coefficients. [10] Figure 5 shows the changes in income distributions $f(x)$ evaluated at $\bar{\boldsymbol{\theta}}$, which is the

[10] No analytical expression of the Gini coefficients for the GB distribution differs from that for the other beta-type distributions. Therefore, we calculate the Gini coefficients from the GB distribution by using numerical integration as follows:

$$G = 1 - \frac{\int_0^\infty (1 - F(x))^2 dx}{\int_0^\infty (1 - F(x)) dx}.$$

See Dorfman (1979).

posterior means. This figure illustrates that the modes slightly lower and the shapes of the lower and upper tails change, showing the flexibility of the GB distribution. As these changes lead to changes in the Gini coefficients, Figure 5 also shows the Gini coefficients of the U.S. family income data. The results show that the inequalities have certain increasing trends because the 95% credible intervals do not overlap. Therefore, we can conclude that U.S. income inequality is increasing.

4.2 Family Income and Expenditure Survey in Japan

In this subsection, we discuss the use of the GB distribution compared with its special distributions by using the Family Income and Expenditure Survey (FIES) in 2009, a Japanese household survey compiled by the Statistics Bureau of the Ministry of Internal Affairs and Communications, to consider the income distribution in Japan. In addition to this information, which is required to estimate the parameters of the income distribution, the class income means are also available. In this situation, Gastwirth (1972) proposed a calculation method for the nonparametric lower and upper bounds of the Gini coefficients. Therefore, it is possible to compare not only the fit of the distribution by using the marginal likelihoods but also the accuracy of the Gini coefficients. Moreover, by using data in the form of quintiles and deciles, it is also possible to examine the effect of the number of groups.

The datasets consists of the yearly pretax income, which is surveyed for the households of workers employed as clerks or wage earners in public or private enterprises, such as government offices, private companies, factories, schools, hospitals, and shops. Therefore, individual proprietors and households whose heads are merchants, artisans, or administrators of unincorporated enterprises are excluded from the datasets. The sample size for each dataset is $n = 10,000$; therefore, $n_i^* = 2,000$ for the quintile data and $n_i^* = 1,000$ for the decile data. [11] To estimate the parameters of the GB distribution from the FIES data, we run the TaRBMH algorithm using $21,000$ iterations and discarding the first $1,000$.

Table 5 shows the log of marginal likelihood with the standard errors. [12] The table confirms that the highest

[11] For more details, see `http://www.stat.go.jp/english/`.

[12] To estimate the DA and GB2 distributions, the RWMH algorithm using $40,000$ iterations and discarding the first $20,000$ is

marginal likelihoods can be found in the GB2 distribution in the quintile data as well as in the GB distribution in the decile data. However, if we take account of the standard errors, the fit of the GB distribution is almost equivalent to that of the GB2 distribution in both datasets. Moreover, the fit of the GB and GB2 distributions in the decile data is much higher than that in the decile data, while that of the DA distribution is no different in both datasets. To visualize the differences and similarities, Figure 6 displays the distributions fitted to the representative sample, which shows the differences in the GB and GB2 distributions in the quintile data and the similarities in the GB and GB2 distributions. In addition, the DA distributions are similar in both datasets. Therefore, we can also confirm the importance of the number of groups from these results.

To examine the fit of the distributions from another point of view, we also examine the Gini coefficients from the assumed hypothetical distribution as in the case of the U.S. family income data to confirm the accuracy of the Gini coefficients. As stated above, we can calculate Gastwirth's (1972) nonparametric bounds in the FIES data because the class income means are also available. We can then also confirm the accuracy of the Gini coefficients in this dataset. Table 6 shows the posterior estimates of the Gini coefficients along with Gastwirth's (1972) nonparametric bounds, and Figure 7 shows the posterior distributions with the posterior means and nonparametric bounds, where the shaded area represents the region inside the nonparametric bounds. From the table and figure, we first confirm that the 95% credible interval from the quintile data is wider than that from the decile data. As shown by Kakamu (2016), [13] the effect of the number of groups may also appear in the distribution variance of the Gini coefficients. Second, the posterior means from the GB and GB2 distributions are included in the nonparametric bounds in the quintile data, while only that from the GB2 distribution is included in the decile data. In addition, the posterior distribution of the Gini coefficients from the GB distribution in the quintile data is wider than the nonparametric bounds. Therefore, we can conclude

adopted for the DA distribution, while the RWMH algorithm using $300,000$ iterations and discarding the first $100,000$ is adopted for the GB2 distribution. The prior distributions and hyper-parameters are the same as for the GB distribution.

[13] Kakamu (2016) examined the performance of the Gini coefficients assuming the Singh–Maddala and DA distributions by using Monte Carlo experiments and showed that the effects of not only the number of observations but also the number of groups appear in the Gini coefficients in terms of the root mean square errors.

that the use of the GB2 distribution is also recommended in terms of the accuracy of the Gini coefficients. Moreover, the GB and GB2 distributions outperform the DA distribution even when the datasets are in the form of quintile and decile data.

Finally, we examine the estimates from the FIES data. Table 7 shows the posterior estimates and implies that the estimates of c differ from 1.0. Therefore, parameter c also plays an important role in the income distribution in Japan if we assume the GB distribution. We find that the posterior means and 95% credible intervals differ slightly. This is especially the case for parameter p, where only the number of groups differs. As a result, we can conclude that the use of the GB2 distribution is recommended even if the parameters of the GB distribution are estimated accurately.

5 Conclusions

This study considered the estimation of the GB distribution parameters for grouped data from a Bayesian point of view and the importance of the GB distribution, using a simulated dataset and two real datasets from the United States and Japan. First, to estimate the parameters of the distribution, we used the TaRBMH algorithm proposed by Chib and Ramamurthy (2010). From the simulated data, we confirmed that the TaRBMH algorithm is more efficient than the RWMH algorithm proposed by Chotikapanich and Griffiths (2000) and Kakamu (2016) in terms of mixing. Moreover, from the U.S. dataset, we found that the estimated parameters of our Bayesian approach may differ from those of MLE. Therefore, we can conclude that our Bayesian procedure works well in terms of mixing and that the Bayesian estimates are accurate, although the TaRBMH algorithm is much more time-consuming than other methods. However, it should be emphasized that the shorter iterations of the RWMH algorithm might lead to an inaccurate estimation of the true parameters, while the TaRBMH algorithm can avoid this pitfall, even for shorter iterations.

Second, the goodness-of-fit of the GB distribution was compared with the GB2 and DA distributions. From the simulated and real datasets, we found that the fit of the GB distribution is equivalent to that of the GB2 distribution in many cases. On the contrary, the fit of the GB and GB2 distributions is better than that of the DA distribution in general. Moreover, the effect of the number of groups was examined by using quintile and decile data. Although the estimates are slightly different, similar distributions can be obtained from both types of data. The fit of the GB2 distribution is the best and this improves as the number of groups increases. In addition, we examined the accuracy of the Gini coefficients by using the Japanese dataset, finding that the GB2 distribution can estimate most accurately even if the number of groups has changed. Therefore, the GB2 distribution is recommended for examining beta-type distributions. These results are consistent with those from Bordley et al. (1997), even though the estimates are different.

Finally, we discuss remaining issues. As stated earlier, improving the efficiency of the MCMC chain and overcoming time-consuming problems are left to future work. The availability of individual data in the GB distribution is another problem. In our experience, it is difficult to estimate the parameters of the GB distribution from individual data. Therefore, creating an estimation procedure for the GB distribution and the possibility of individual data is also left to future work.

References

Atoda N, Suruga T, Tachibanaki T. 1988. Statistical inference of functional forms for income distribution. *The Economic Studies Quarterly* **39**: 14–40.

Bordley RF, McDonald JB, Mantrala A. 1997. Something new, something old: Parametric models for the size of distribution of income. *Journal of Income Distribution* **6**: 91–103.

Chib S, Ramamurthy S. 2010. Tailored randomized block MCMC methods with application to DSGE models. *Journal of Econometrics* **155**: 19–38.

Chotikapanich D, Griffiths WE. 2000. Posterior distributions for the Gini coefficient using grouped data. *Australian & New Zealand Journal of Statistics* **42**: 383–392.

Chotikapanich D, Griffiths WE. 2008. Estimating income distributions using a mixture of gamma densities. In Chotikapanich D (ed.) *Modeling Income Distributions and Lorenz Curves*. New York: Springer, 285–302.

Chotikapanich D, Rao DSP, Tang KK. 2007. Estimating income inequality in China using grouped data and the generalized beta distribution. *Review of Income and Wealth* **53**: 127–147.

Clementi F, Gallegati M, Kaniadakis G. 2007. κ-generalized statistics in personal income distribution. *The European Physical Journal B* **57**: 187–193.

Clementi F, Gallegati M, Kaniadakis G, Landini S. 2016. κ-generalized models of income and wealth distributions: A survey. *The European Physical Journal Special Topics* **225**: 1959–1984.

Dagum C. 1977. A new model of personal income distribution: Specification and estimation. *Economie Appliquée* **30**: 413–437.

David H, Nagaraja H. 2003. *Order Statistics, 3rd Edition*. New York: Wiley.

Doornik JA. 2013. *Ox™ 7: An Object-Oriented Matrix Programming Language*. London: Timberlake Consultants Press.

Dorfman R. 1979. A formula for the Gini coefficient. *The Review of Economics and Statistics* **61**: 146–149.

Gastwirth JL. 1972. The estimation of the Lorenz curve and Gini index. *The Review of Economics and Statistics* **54**: 306–316.

Gelfand AE, Hills SE, Racine-Poon A, Smith AFM. 1990. Illustration of Bayesian inference in normal data models using Gibbs sampling. *Journal of the American Statistical Association* **85**: 972–985.

Goffe WL, Ferrier GD, Rogers J. 1994. Global optimization of statistical functions with simulated annealing. *Journal of Econometrics* **60**: 65–99.

Hajargasht G, Griffiths WE, Brice J, Rao DP, Chotikapanich D. 2012. Inference for income distributions using grouped data. *Journal of Business & Economic Statistics* **30**: 563–575.

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

Kakamu K. 2016. Simulation studies comparing Dagum and Singh–Maddala income distributions. *Computational Economics* **48**: 593–605.

Kleiber C. 2008. A guide to the Dagum distributions. In Chotikapanich D (ed.) *Modeling Income Distributions and Lorenz Curves*. New York: Springer, 97–117.

Kleiber C, Kotz S. 2003. *Statistical Size Distributions in Economics and Actuarial Sciences*. New York: Wiley.

Kloek T, van Dijk HK. 1978. Efficient estimation of income distribution parameters. *Journal of Econometrics* **8**: 61–74.

Majumder A, Chakravarty SR. 1990. Distribution of personal income: Development of a new model and its application to U.S. income data. *Journal of Applied Econometrics* **5**: 189–196.

McDonald JB. 1984. Some generalized functions for the size distribution of income. *Econometrica* **52**: 647–663.

McDonald JB, Mantrala A. 1995. The distribution of personal income: Revisited. *Journal of Applied Econometrics* **10**: 201–204.

McDonald JB, Ransom M. 2008. The generalized beta distribution as a model for the distribution of income: Estimation of related measures of inequality. In Chotikapanich D (ed.) *Modeling Income Distributions and Lorenz Curves*. New York: Springer, 147–166.

McDonald JB, Ransom MR. 1979a. Alternative parameter estimators based upon grouped data. *Communications in Statistics - Theory and Methods* **8**: 899–917.

McDonald JB, Ransom MR. 1979b. Functional forms, estimation techniques and the distribution of income. *Econometrica* **47**: 1513–1525.

McDonald JB, Xu YJ. 1995. A generalization of the beta distribution with applications. *Journal of Econometrics* **66**: 133–152.

Newton MA, Raftery AE. 1994. Approximate Bayesian inference with the weighted likelihood bootstrap (with discussion). *Journal of the Royal Statistical Society. Series B (Methodological)* **56**: 3–48.

Nishino H, Kakamu K. 2011. Grouped data estimation and testing of Gini coefficients using lognormal distributions. *Sankhya B* **73**: 193–210.

Nocedal J, Wright S. 2000. *Numerical Optimization, 2nd Edition*. New York: Springer.

Parker SC. 1999. The generalised beta as a model for the distribution of earnings. *Economics Letters* **62**: 197–200.

R Core Team. 2017. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.
URL `https://www.R-project.org/`

Reed WJ, Jorgensen M. 2004. The double Pareto-lognormal distribution — A new parametric model for size distributions. *Communications in Statistics - Theory and Methods* **33**: 1733–1753.

Salem ABZ, Mount TD. 1974. A convenient descriptive model of income distribution: The gamma density. *Econometrica* **42**: 1115–1127.

Singh SK, Maddala GS. 1976. A function for size distribution of incomes. *Econometrica* **44**: 963–970.

Slottje DJ. 1984. A measure of income inequality in the U.S. for the years 1952–1980 based on the beta distribution of the second kind. *Economics Letters* **15**: 369–375.

Tachibanaki T, Suruga T, Atoda N. 1997. Estimations of income distribution parameters for individual observations by maximum likelihood method. *Journal of the Japan Statistical Society* **27**: 191–203.

van Dijk HK, Kloek T. 1980. Inferential procedures in stable distributions for class frequency data on incomes. *Econometrica* **48**: 1139–1148.

Wichitaksorn N, Tsurumi H. 2013. Comparison of MCMC algorithms for the estimation of Tobit model with non-normal error: The case of asymmetric Laplace distribution. *Computational Statistics & Data Analysis* **67**: 226–235.

Table 1: Simulated data: Posterior estimates

		RWMH				TaRBMH			
	true values	Mean	95%CI		IF	Mean	95%CI		IF
a	5.00	5.345	3.298	8.046	14025.644	5.565	3.678	8.235	151.021
b	30.00	29.005	26.751	31.985	10903.941	28.771	26.828	31.939	37.027
c	0.95	0.932	0.854	0.979	9934.476	0.938	0.869	0.981	122.364
p	0.50	0.481	0.283	0.819	11192.787	0.456	0.275	0.716	130.780
q	0.80	0.681	0.384	1.100	12645.934	0.637	0.366	1.097	118.345

Note: Posterior means (Mean), 95% credible intervals (95%CI), and inefficiency factors (IF) are displayed. The acceptance rates are as follows: around 20% in the RWMH algorithm, and around 90% in the TaRBMH algorithm.

Table 2: Simulated data: Marginal likelihood

	Log ML	SE
DA	-2.148	0.916
GB2	-0.649	0.136
GB (RWMH)	-0.860	0.316
GB (TaRBMH)	1.753	1.368

Note: The log of marginal likelihoods (Log ML) and the standard errors (SE), which are calculated using delta method, are displayed.

Table 3: Real data: Posterior estimates for U.S. data

U.S. Family Income Data in 1970									
	MLE	RWMH				TaRBMH			
		Mean	95%CI		IF	Mean	95%CI		IF
a	4.797	4.740	4.656	4.834	6112.176	4.828	4.381	5.293	73.503
b	44.29	44.331	43.646	44.834	4984.985	44.175	42.977	45.606	60.753
c	0.997	0.996	0.995	0.998	5178.765	0.997	0.995	0.999	2.393
p	0.316	0.321	0.311	0.330	5250.346	0.315	0.284	0.350	69.651
q	0.695	0.707	0.691	0.722	3726.839	0.692	0.600	0.806	74.333
U.S. Family Income Data in 1975									
	MLE	RWMH				TaRBMH			
		Mean	95%CI		IF	Mean	95%CI		IF
a	2.887	2.933	2.704	3.161	201.410	2.937	2.740	3.175	107.947
b	54.87	54.795	52.293	57.700	147.279	54.741	52.140	57.386	104.862
c	1.000	0.996	0.988	1.000	23.687	0.996	0.988	1.000	10.618
p	0.561	0.551	0.500	0.609	174.675	0.550	0.499	0.598	102.898
q	1.587	1.575	1.355	1.840	163.258	1.570	1.345	1.790	115.337
U.S. Family Income Data in 1980									
	MLE	RWMH				TaRBMH			
		Mean	95%CI		IF	Mean	95%CI		IF
a	2.587	2.759	2.544	2.978	813.741	2.792	2.567	3.053	94.066
b	64.48	59.331	54.884	64.448	563.933	58.799	54.199	63.728	80.060
c	1.000	0.977	0.957	0.997	114.425	0.976	0.958	0.997	12.082
p	0.599	0.558	0.509	0.614	689.830	0.550	0.493	0.608	89.899
q	1.961	1.605	1.313	1.969	605.153	1.564	1.248	1.912	90.826
U.S. Family Income Data in 1985									
	MLE	RWMH				TaRBMH			
		Mean	95%CI		IF	Mean	95%CI		IF
a	2.498	2.765	2.542	2.991	800.566	2.782	2.552	3.072	137.168
b	66.06	58.104	53.740	62.941	639.629	57.837	53.171	62.946	108.665
c	1.000	0.978	0.961	0.997	150.480	0.978	0.961	0.997	20.292
p	0.578	0.517	0.471	0.572	708.974	0.514	0.456	0.568	126.702
q	1.793	1.316	1.075	1.606	684.367	1.299	1.027	1.600	123.146
U.S. Family Income Data in 1990									
	MLE	RWMH				TaRBMH			
		Mean	95%CI		IF	Mean	95%CI		IF
a	2.731	3.039	2.771	3.267	2760.405	3.046	2.752	3.368	86.386
b	62.19	55.059	51.710	59.330	1845.727	55.069	51.180	59.889	76.529
c	1.000	0.978	0.967	0.993	198.495	0.978	0.967	0.995	20.885
p	0.519	0.463	0.424	0.514	2493.939	0.462	0.411	0.518	83.443
q	1.358	0.991	0.831	1.219	2229.423	0.991	0.793	1.245	86.619

Note: MLE estimates are based on the results from Bordley et al. (1997). The acceptance rates are as follows: around 50% in the RWMH algorithm every year, around 70% in the TaRBMH in 1975, and around 90% in the TaRBMH algorithm in other years.

Table 4: Real data: Marginal likelihood for U.S. data

	DA		GB2		GB	
	Log ML	SE	Log ML	SE	Log ML	SE
1975	-31.380	0.272	-28.108	0.213	-25.627	0.650
1980	-129.833	0.698	-103.294	0.934	-105.203	2.438
1985	-70.291	0.226	-28.633	0.299	-32.452	0.583
1990	-32.287	0.363	-10.207	0.491	-17.680	0.901
1995	-8.993	0.295	-4.444	0.598	-7.941	1.316

Table 5: Real data: Marginal likelihood for FIES data in Japan

	Quintile		Decile	
	Log ML	SE	Log ML	SE
DA	15.167	0.440	14.200	0.318
GB2	16.254	0.146	32.365	0.285
GB	15.826	0.234	34.746	1.696

Table 6: Real data: Gini coefficients for FIES data in Japan

	Quintile			Decile		
	Mean	95%CI		Mean	95%CI	
DA	0.276	0.270	0.283	0.270	0.264	0.275
GB2	0.258	0.252	0.265	0.256	0.251	0.261
GB	0.247	0.223	0.265	0.244	0.236	0.252
Gastwirth (1972)		0.238	0.261		0.249	0.256

Note: Gastwirth (1972) gives the lower and the upper bounds of Gini coefficient (written in 95%CI columns).

Table 7: Real data: Posterior estimates for FIES data in Japan in 2009

	Quintile				Decile			
	Mean	95%CI		IF	Mean	95%CI		IF
a	2.940	1.534	5.414	493.659	2.117	1.343	3.153	503.148
b	0.634	0.463	0.821	49.460	0.515	0.352	0.684	70.550
c	0.952	0.887	0.996	162.361	0.932	0.895	0.961	261.024
p	1.658	0.544	3.929	431.401	2.893	1.336	5.947	547.177
q	1.529	0.436	3.669	329.379	1.751	0.848	3.552	450.568

Note: The acceptance rate is around 80% in the TaRBMH algorithm.

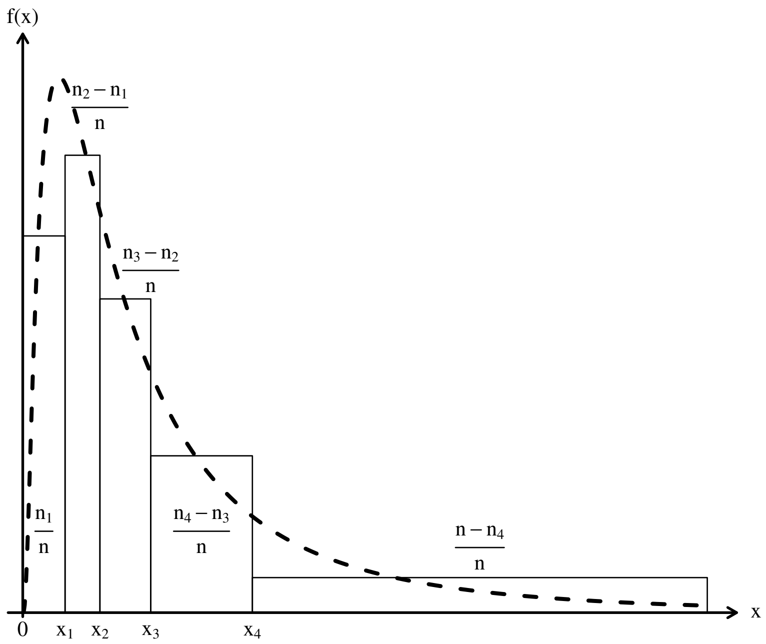

Figure 1: Quintile data

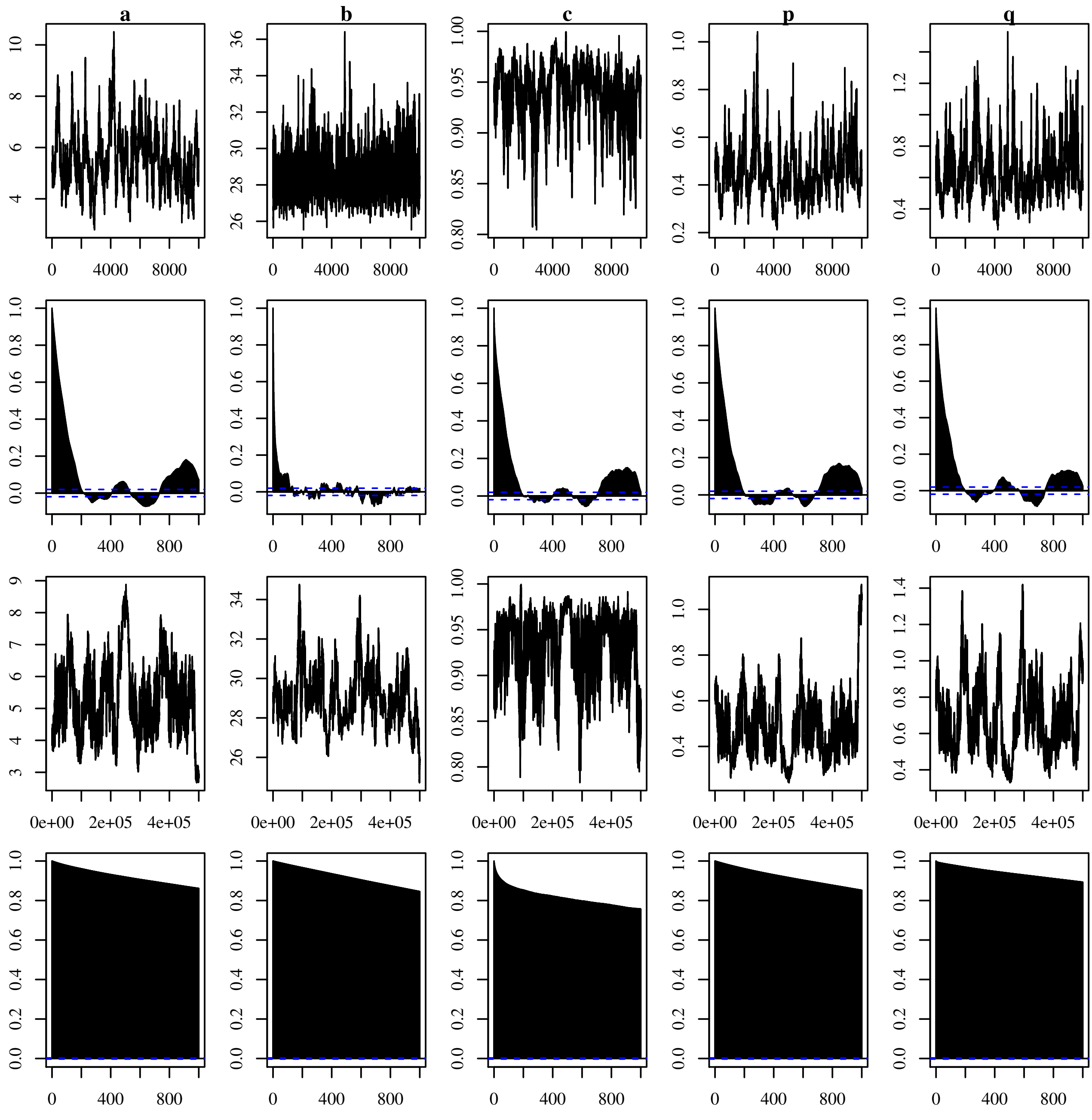

Figure 2: Sampling results from the simulated data: Time-series plots of the draws from the posterior and corresponding ACFs using the TaRBMH (top panel) and the RWMH (bottom panel)

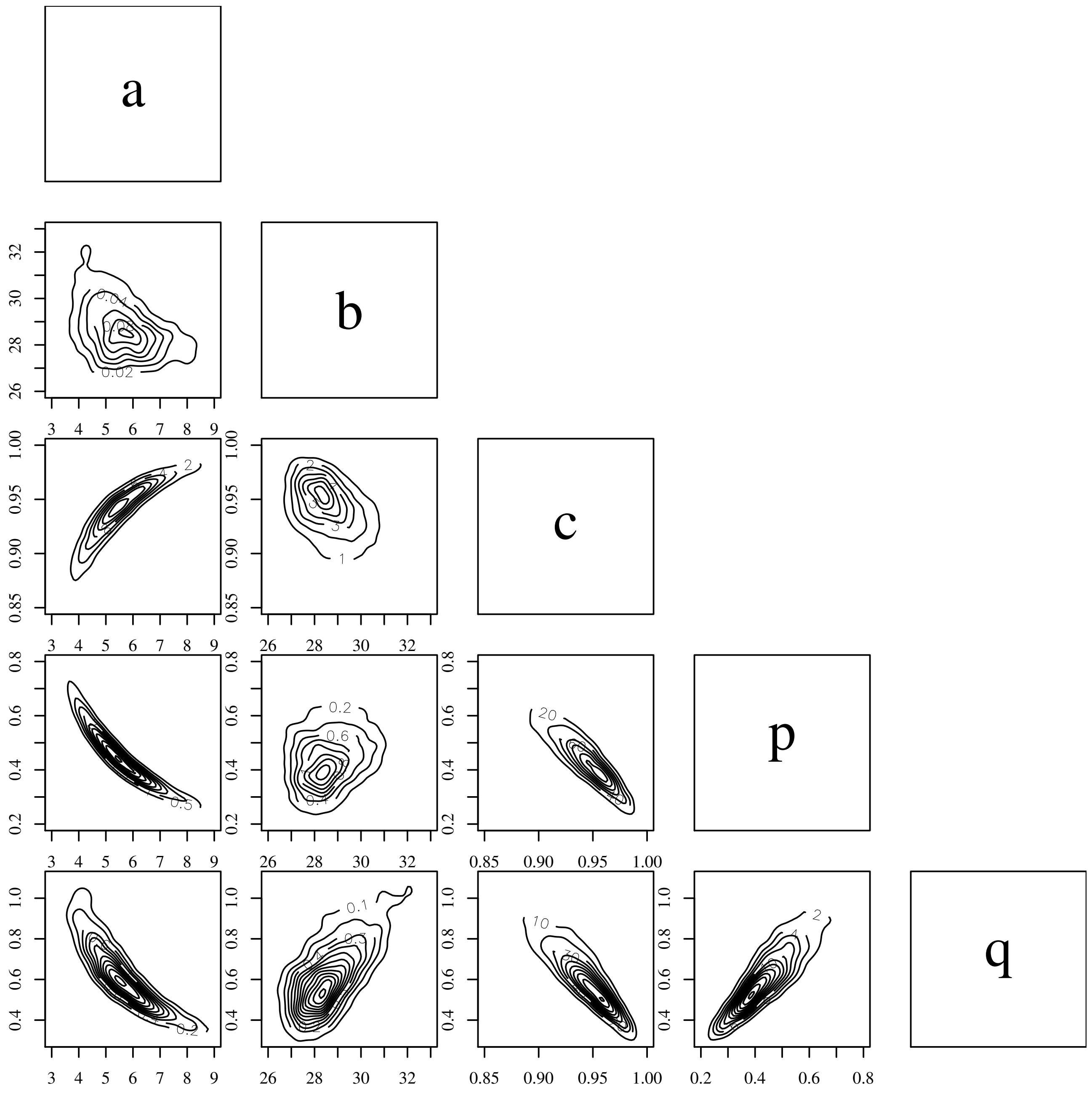

Figure 3: Joint posterior distributions from the simulated data

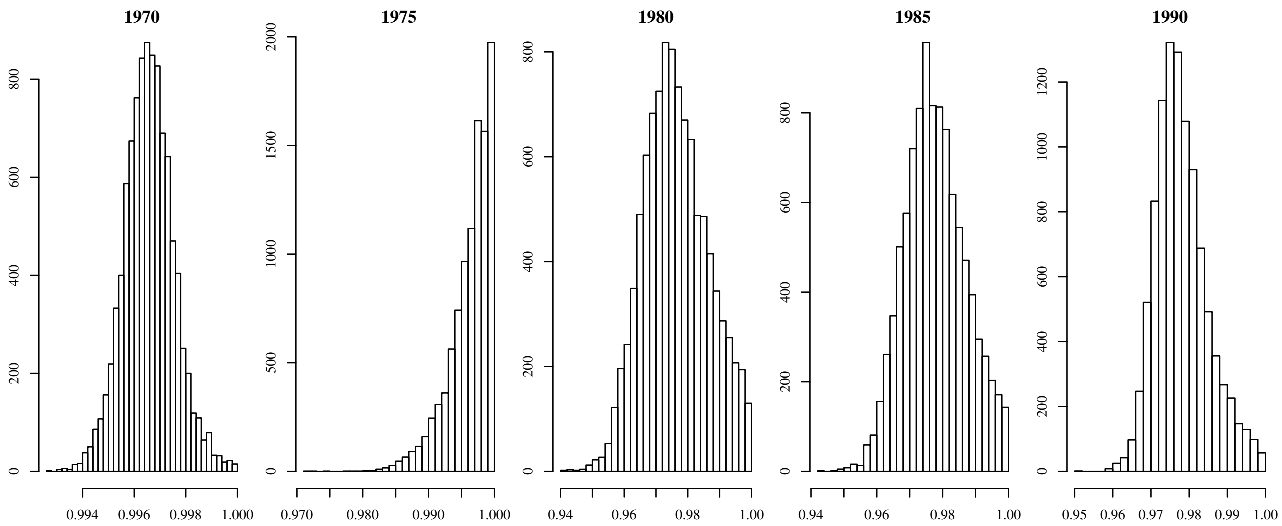

Figure 4: Marginal posterior distributions of c from the U.S. income data

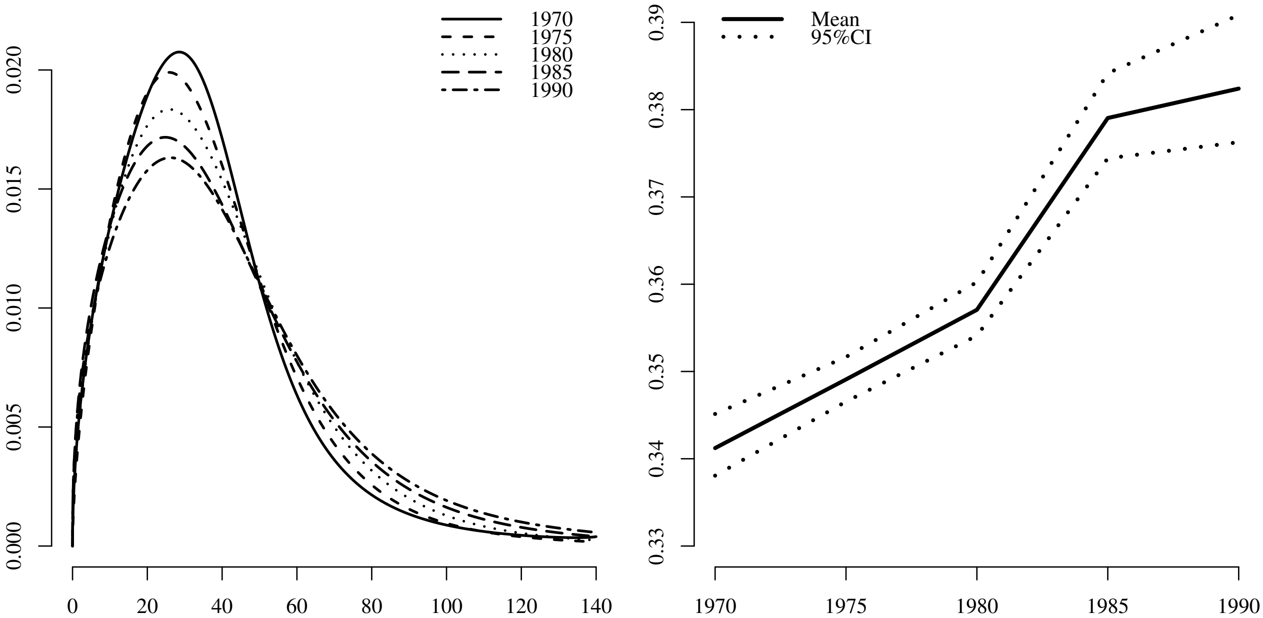

Figure 5: Changes in U.S. income distributions (left) and trends of the Gini coefficients (right)

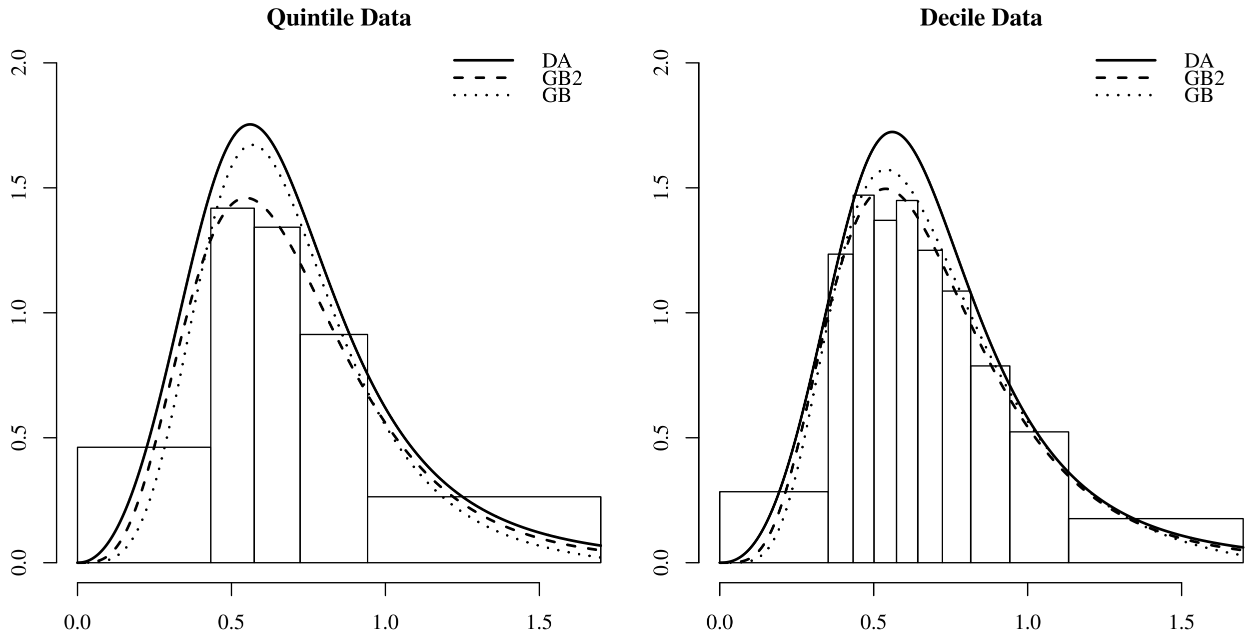

Figure 6: Distributions fitted to the representative sample in Japan

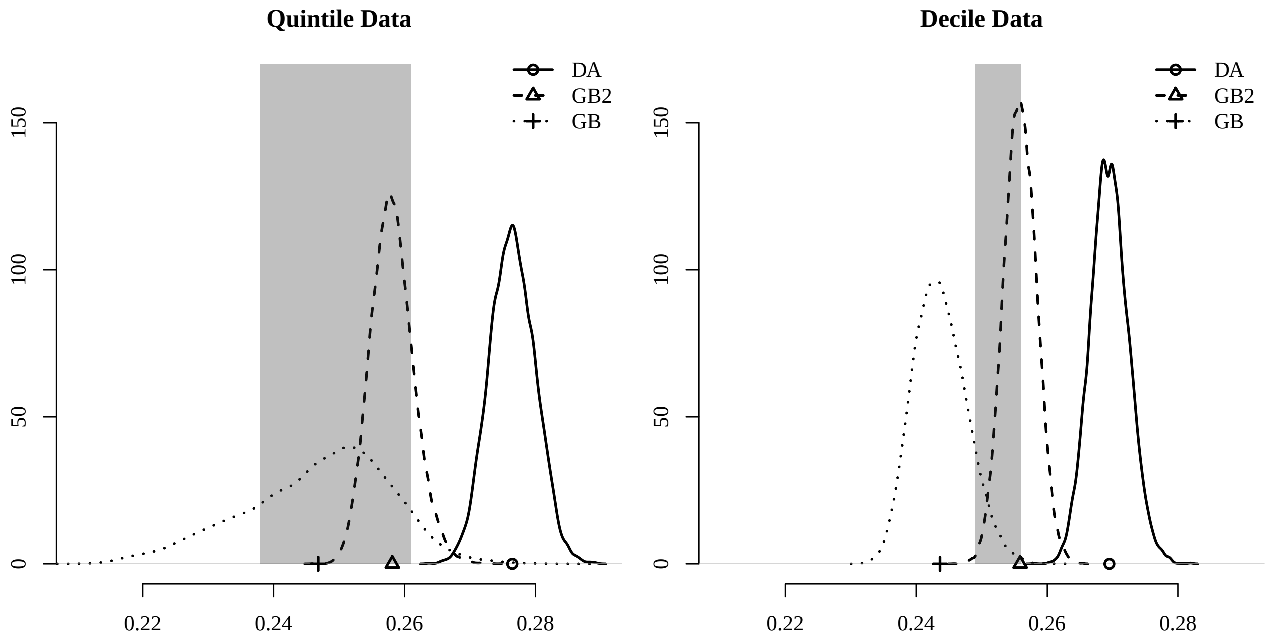

Figure 7: Posterior distributions of the Gini coefficients in Japan