CrossMark

Simulation Studies Comparing Dagum and Singh–Maddala Income Distributions

Kazuhiko Kakamu[1]

Abstract Dagum and Singh–Maddala distributions have been widely assumed as models for income distribution in empirical analyses. The properties of these distributions are well known and several estimation methods for these distributions from grouped data have been discussed widely. Moreover, previous studies argue that the Dagum distribution gives a better fit than the Singh–Maddala distribution in the empirical analyses. This study explores the reason why Dagum distribution is preferred to the Singh–Maddala distribution in terms of the akaike information criterion through Monte Carlo experiments. In addition, the properties of the Gini coefficients and the top income shares from these distributions are examined by means of root mean square errors. From the experiments, we confirm that the fit of the distributions depends on the relationships and magnitudes of the parameters. Furthermore, we confirm that the root mean square errors of the Gini coefficients and top income shares depend on the relationships of the parameters when the data-generating processes are a generalized beta distribution of the second kind.

Keywords Akaike information criterion (AIC) · Dagum distribution · Gini coefficients · Root mean square error (RMSE) · Singh–Maddala distribution · Top income share

✉ Kazuhiko Kakamu
kakamu@person.kobe-u.ac.jp

1 Graduate School of Business Administration, Kobe University, 2-1, Rokkodai, Nada, Kobe 657-8501, Japan

1 Introduction

Kleiber (1996) shows that Dagum and Singh–Maddala income distributions are closely related, although he states that the Dagum distribution should provide a better fit than the Singh–Maddala distribution. Moreover, Tadikamalla (1980) shows that the shape of the Dagum distribution is considerably more flexible than that of the Singh–Maddala distribution. Therefore, the Dagum distribution more often is assumed as the income distribution in empirical analyses. Accordingly, Kleiber (2008) reviews studies that examine income distributions of the Dagum type.

In empirical analyses, grouped data have been utilized widely and many works have examined several distributions from grouped data (e.g., McDonald and Ransom 1979a). Although McDonald and Mantrala (1995) show that the estimates are different if the number of groups is different, even if the same data source is used,[1] the properties of the estimators have been examined rarely. As far as we know, the properties of the estimators have been examined only by McDonald and Ransom (1979b), who assume a gamma distribution. However, their focus is on comparing several estimation methods. In addition, the properties of the model selection criteria have not been examined, in spite of the fact that several model selection criteria, including the akaike information criterion (AIC), have been utilized widely in empirical analyses. Therefore, we consider that the properties of the model selection criteria, which take into account the effect of group numbers, become important when we use these distributions in empirical analyses.

Moreover, the parameters are used not only for their own inferences, but also for the inference of the Gini coefficients (e.g., Atoda et al. 1988; Chotikapanich and Griffiths 2000; Nishino and Kakamu 2011). However, it is well known that the Gini coefficients are highly nonlinear functions of the parameters. As shown by Gelfand et al. (1990), the MCMC method can yield valid inferences on nonlinear functions of the parameters, such as the Gini coefficients. Therefore, it would be worthwhile to examine the properties of the Gini coefficients in a Bayesian framework.[2]

Recently, researchers of inequality have become concerned with not only the Gini coefficient but also the top income share. For example, Atkinson et al. (2011) examine the history of the top income shares and Brzezinski (2013) examines the properties of the top income share. While the Gini coefficient measures inequality among whole households or individuals, top income share is constructed as ratios (portions) of upper tails to total income of households or individuals. Therefore, top income share might be sensitive to the choice of the underlying hypothetical distributions. Thus, we analyze top income share as an alternative inequality measure to the Gini coefficient.

---

[1] Kakamu and Nishino (2014) show that the estimates of the generalized beta (GB) distribution do not change, even if the number of groups is different when a Markov chain Monte Carlo (MCMC) method is utilized. Therefore, we also utilize the MCMC method to estimate the parameters of the models.

[2] It is desirable to use the maximum likelihood estimates to calculate the Gini coefficients. However, as stated in the Appendix, it is sometimes difficult to find the mode of parameters by the maximum likelihood estimates when we assume the Dagum and Singh–Maddala distributions are the income distribution. This is another reason why we use the MCMC method.

This study has two objectives. One is to explore the reason why Dagum distribution is preferred to Singh–Maddala distribution in terms of AICs through Monte Carlo experiments, in which the data-generating processes (DGPs) are the generalized beta distribution of the second kind (GB2 distribution). This involves considering the Dagum and Singh–Maddala distributions as special cases. The other objective is to examine the properties of the Gini coefficients and top income shares for these two distributions by means of root mean square errors (RMSEs). From the experiments, we confirm that the fit of the distributions depends on the relationships and magnitudes of the parameters and that the RMSEs of the Gini coefficients and top income shares also depend on the relationships of the parameters.

The rest of this paper is organized as follows. In Sect. 2, we briefly discuss the relationship between the Dagum and Singh–Maddala distributions and set forth the framework for Bayesian inference. Section 3 presents the results obtained from the Monte Carlo simulations. Finally, brief conclusions and remaining issues are presented in Sect. 4.

2 Dagum and Singh–Maddala Income Distributions

Dagum (1977) and Singh and Maddala (1976) distributions are special cases of the GB2 distribution proposed by McDonald (1984). The probability density function (PDF) of the GB2 distribution with parameters $a > 0$, $b > 0$, $p > 0$, and $q > 0$ is defined as

$$f_{GB2}(x) = \frac{ax^{ap-1}}{b^{ap} B(p, q) \left[1 + \left(\frac{x}{b} \right)^a \right]^{p+q}}, \ x > 0,$$

where $B(\cdot, \cdot)$ is a beta function, b is a scale, and a, p, and q are shape parameters. In particular, parameter p makes the shape of the upper tail change and parameter q makes the shape of the lower tail change. The distribution proposed by Majumder and Chakravarty (1990) is simply a reparameterization of the GB2 distribution.

This four-parameter family nests most of the previously considered income distributions as special or limiting cases. Our notations are as follows: $GB2(a, b, p, 1) = \mathcal{DA}(a, b, p)$ and $GB2(a, b, 1, q) = \mathcal{SM}(a, b, q)$, where $GB2(a, b, p, q)$, $\mathcal{DA}(a, b, p)$, and $\mathcal{SM}(a, b, q)$ are GB2 distribution with parameters a, b, p, and q; the Dagum distribution has parameters a, b, and p; and the Singh–Maddala distribution has parameters a, b, and q. These relationships and related distributions are summarized in Kleiber and Kotz (2003) in detail. In addition, Kleiber (1996) shows that $X \sim \mathcal{SM}(a, b, r) \Leftrightarrow \frac{1}{X} \sim \mathcal{DA}\left(a, \frac{1}{b}, r \right)$. Therefore, we can confirm that the Dagum and Singh–Maddala distributions are related closely.

The PDF and cumulative density function (CDF) of the Dagum distribution are expressed as

$$f_{DA}(x) = \frac{apx^{ap-1}}{b^{ap} \left[1 + \left(\frac{x}{b} \right)^a \right]^{p+1}}, \ x > 0, \tag{1}$$

$$F_{DA}(x) = \left[1 + \left(\frac{x}{b}\right)^{-a}\right]^{-p}, \quad x > 0, \tag{2}$$

respectively. On the other hand, the PDF and CDF of the Singh–Maddala distribution are given as

$$f_{SM}(x) = \frac{aqx^{a-1}}{b^a \left[1 + \left(\frac{x}{b}\right)^a\right]^{1+q}}, \quad x > 0, \tag{3}$$

$$F_{SM}(x) = 1 - \left[1 + \left(\frac{x}{b}\right)^a\right]^{-q}, \quad x > 0, \tag{4}$$

respectively. As shown above, these two distributions have closed-form expressions in their CDFs, unlike the GB2 distribution; furthermore, their CDFs are invertible. Therefore, these two distributions are tractable and it is possible that the statistical inference is implemented using these PDFs and CDFs.[3]

To estimate the parameters of the income distribution from grouped data, Kakamu and Nishino (2014) show that the Bayesian approach is preferred to the maximum likelihood method. Therefore, we take the Bayesian approach and introduce the likelihood function, following Nishino and Kakamu (2011). Let $\boldsymbol{\theta}$ be the vector of parameters of the relevant income distribution and let $\mathbf{x} = (x_1, x_2, \ldots, x_k)$ be the vector of observed income in the form of grouped data. Then, x_is are strictly known. In the general form, the likelihood function given by the PDF and CDF is expressed as follows

$$
\begin{aligned}
L(\mathbf{x}|\boldsymbol{\theta}) \\
= n! \frac{F(x_1)^{n_1-1}}{(n_1-1)!} f(x_1) \left\{ \prod_{i=2}^{k} \frac{(F(x_i) - F(x_{i-1}))^{n_i - n_{i-1} - 1}}{(n_i - n_{i-1} - 1)!} f(x_i) \right\} \\
\times \frac{(1 - F(x_k))^{n-n_k}}{(n - n_k)!},
\end{aligned}
\tag{5}
$$

---

[3] Once the parameters are estimated, we could calculate the Gini coefficients from the parameters. The Gini coefficients from the Dagum and Singh–Maddala distributions are expressed by

$$G_{DA} = \frac{\Gamma(p)\Gamma(2p + 1/a)}{\Gamma(2p)\Gamma(p + 1/a)} - 1,$$

$$G_{SM} = 1 - \frac{\Gamma(q)\Gamma(2q - 1/a)}{\Gamma(q - 1)\Gamma(2q)},$$

where $\Gamma(\cdot)$ is a gamma function.
Moreover, $\chi\%$ top income share from these distributions, where $\chi = 100 \times (1 - z)$, are expressed by

$$T_{DA,\chi} = 1 - I_t(p + 1/a, 1 - 1/a), \text{ where } t = z^{1/p}$$

$$T_{SM,\chi} = 1 - I_t(1 + 1/a, q - 1/a), \text{ where } t = 1 - (1 - z)^{1/q},$$

where $I_x(a, b)$ is the incomplete beta function ratio. In this study, we focus on 10 % ($z = 0.9$) top income shares.

where n_i is the number of observations when income is less than x_i and n is the number of total observations. If we substitute (1) and (2) into (5), then we could derive an inference for the Dagum distribution. Similarly, if we use (3) and (4), we could draw an inference for the Singh–Maddala distribution.

To proceed with the Bayesian analysis, we need to assume the prior distribution as $\pi(\boldsymbol{\theta})$. Given the likelihood function (5) and the prior distribution $\pi(\boldsymbol{\theta})$, the posterior distribution is expressed as

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

Using this posterior distribution, posterior inference via the MCMC method proposed by Chotikapanich and Griffiths (2000) is implemented.[4]

Finally, we specify the prior distribution as $\pi(\boldsymbol{\theta}_{SM}) \propto b^{-1}$, which is the same assumption as in Chotikapanich and Griffiths (2000). In addition, we assume that the prior distribution for the Dagum distribution is $\pi(\boldsymbol{\theta}_{DA}) \propto b^{-1}$.

3 Monte Carlo Experiments

We now explain the setup for the Monte Carlo simulations. First, we set the number of observations as $n = 10{,}000$ and $100{,}000$ to evaluate the effect of the number of observations. In addition, we assume the number of groups as quintile ($k = 4$) and decile ($k = 9$) to evaluate the effects of the number of groups.

Given n and k, we assume the true DGP to be a GB2 distribution, and L samples of x_i for $i = 1, 2, \ldots, k$ are generated. That is, we perform L simulation runs for the Dagum and Singh–Maddala distributions; in this section, $L = 500$.

The simulation procedure in this section is as follows:

(i) Given $a = 3.0$; $b = 6.0$; $p = 0.5, 0.6, 0.7, \ldots, 2.9, 3.0$; and $q = 0.5, 0.6, 0.7, \ldots, 2.9, 3.0$, we generate random numbers x_js, $j = 1, 2, \ldots, n$ from the GB2 distribution.[5]

---

[4] Some modifications are required to estimate the parameters. Therefore, the details of the modified algorithm are summarized in the Appendix.

[5] To generate a random number from GB2 with parameters a, b, p, and q, we first generate Z from a beta distribution with parameters p and q. Then, $X = b\left(\dfrac{Z}{1-Z}\right)^{\frac{1}{a}}$ becomes the random number from the GB2 distribution. In addition, the Gini coefficient from the GB2 distribution is as follows

$$
\begin{aligned}
G_{GB2} = {} & \frac{B(2p + 1/a, 2q - 1/a)}{B(p, q)B(p + 1/a, q - 1/a)} \\
& \times \left[\frac{1}{p} {}_3F_2(1, p + q, 2p + 1/a; p + 1 + 1/a, 2(p + q); 1) \right. \\
& \left. - \frac{1}{p + 1/a} {}_3F_2(1, p + q, 2p + 1/a; p + 1 + 1/a, 2(p + q); 1) \right],
\end{aligned}
$$

where ${}_3F_2$ is a hypergeometric function. However, we calculate the Gini coefficient as well as the top income share by the numerical integration for the same reason as Hajargasht et al. (2012). Moreover, we examine several different values of a and b. However, as similar tendencies can be found, we focus on the cases of $a = 3.0$ and $b = 6.0$.

(ii) Sort the random numbers in ascending order and pick up x_i, which corresponds to the n_i-th observation, $i = 1, 2, \ldots, k$.

(iii) Given $\mathbf{x} = (x_1, x_2, \ldots, x_k)$, we obtain the estimates using the Dagum and Singh–Maddala distributions. In the MCMC procedure, we run the MCMC algorithm, with 7000 iterations excluding the first $1, 000$ iterations.

(iv) We repeat (i)–(iii) L times, where $L = 500$, as mentioned above.

(v) From L estimates, we compute the AIC and RMSE for Gini coefficients and top income shares, as follows

$$\text{AIC} = -2L(\mathbf{x}|\boldsymbol{\theta}^{(l)}) + 2p,$$

$$\text{RMSE}_{\text{G}} = \left(\frac{1}{L} \sum_{l=1}^{L} (G^{(l)} - G)^2 \right)^{\frac{1}{2}},$$

$$\text{RMSE}_{\text{T}} = \left(\frac{1}{L} \sum_{l=1}^{L} (T^{(l)} - T)^2 \right)^{\frac{1}{2}},$$

where $\boldsymbol{\theta}^{(l)}$, $G^{(l)}$, and $T^{(l)}$ represent the estimators of $\boldsymbol{\theta}$, the Gini coefficient and the 10 % top income share in the l-th simulation run, respectively, and p denotes the number of parameters.

We compare the fit of the distributions by the AICs and the accuracy of the estimators of the Gini coefficient and 10 % top income share from the Dagum and Singh–Maddala distributions by RMSEs through the Monte Carlo studies. All the results reported here are generated using Ox version 7.00 (OS_X_64/U) (see Doornik 2009).

Figure 1 shows the selected frequencies of the Dagum distribution compared with the Singh–Maddala distribution, where the AICs are smaller.[6] From the figure, the standard results can be confirmed: the selected frequency increases as the number of observations or the number of groups increase, and when the DGPs are true. Moreover, we can establish that the selected frequencies are almost the same if $p = q$. On the other hand, the Singh–Maddala distributions are selected if $p < q$ and $p > 1.0$, while the Dagum distributions are selected if $p > q$ and $q > 1.0$. However, if $p < 1.0$ or $q < 1.0$, the opposite results can be confirmed in some regions, that is, the Singh–Maddala distributions are likely to be selected if $p > q$, while the Dagum distributions are likely to be selected if $p < q$ and the frequencies depend on the relationship between the parameters p and q. Therefore, we can conclude that the fit of the distributions seems to depend on the relationships and magnitudes of the parameters p and q if we compare the models by AIC. In fact, for example, McDonald and Mantrala (1995) estimate the parameters of Dagum, Singh–Maddala, and GB2 distributions and show that the Dagum distribution is preferred to the Singh–Maddala distribution when the parameters of GB2 distribution are $p = 0.5186$ and $q = 1.358$. From the figure, we can confirm that the Dagum distribution is preferred to the Singh–Maddala distribution at this point.

---

[6]   The detailed tables, which are the source of the figures, are available upon request from the author.
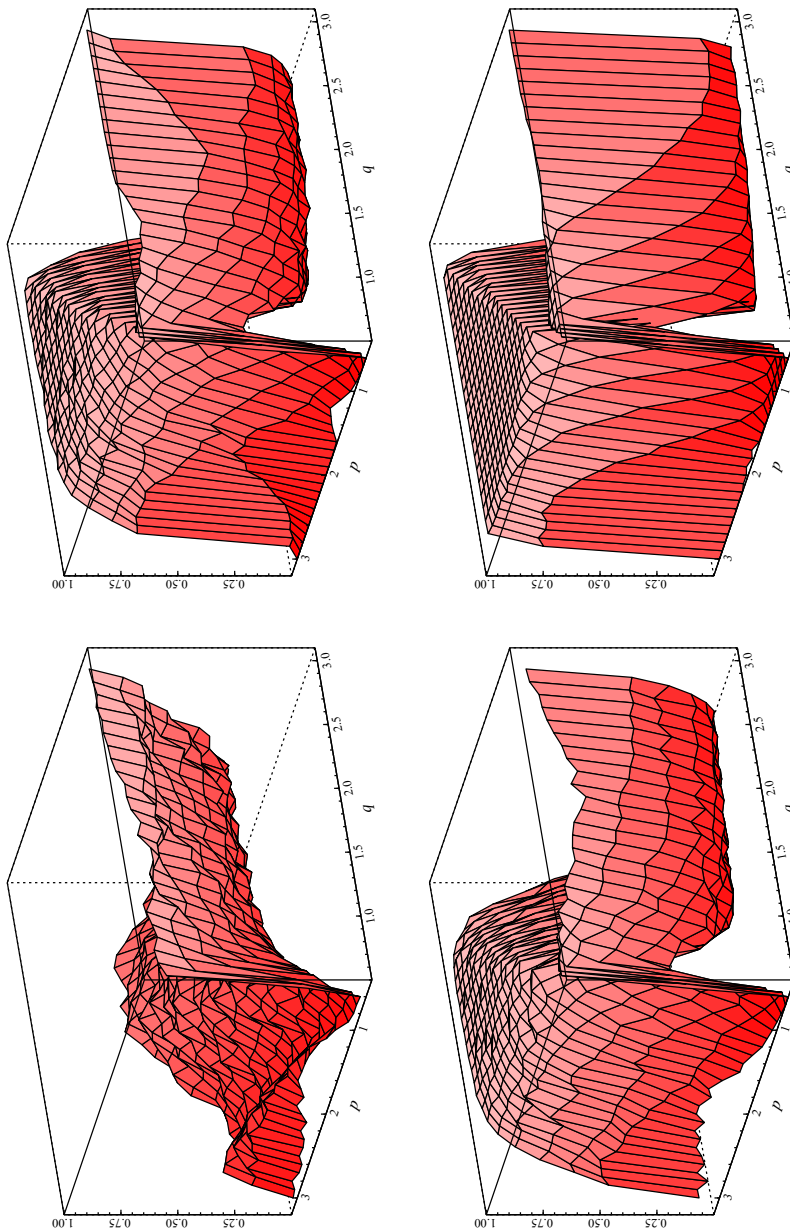
Fig. 1 Selected frequency from the AIC. *Note* The ratios with which the Dagum distributions are selected by AIC are displayed. The *top left* side shows the results from quintile data and $n = 10,000$, the *top right* side shows those from decile data and $n = 10,000$, the *bottom left* side shows those from quintile data and $n = 100,000$, and the *bottom right* side shows those from decile data and $n = 100,000$, respectively
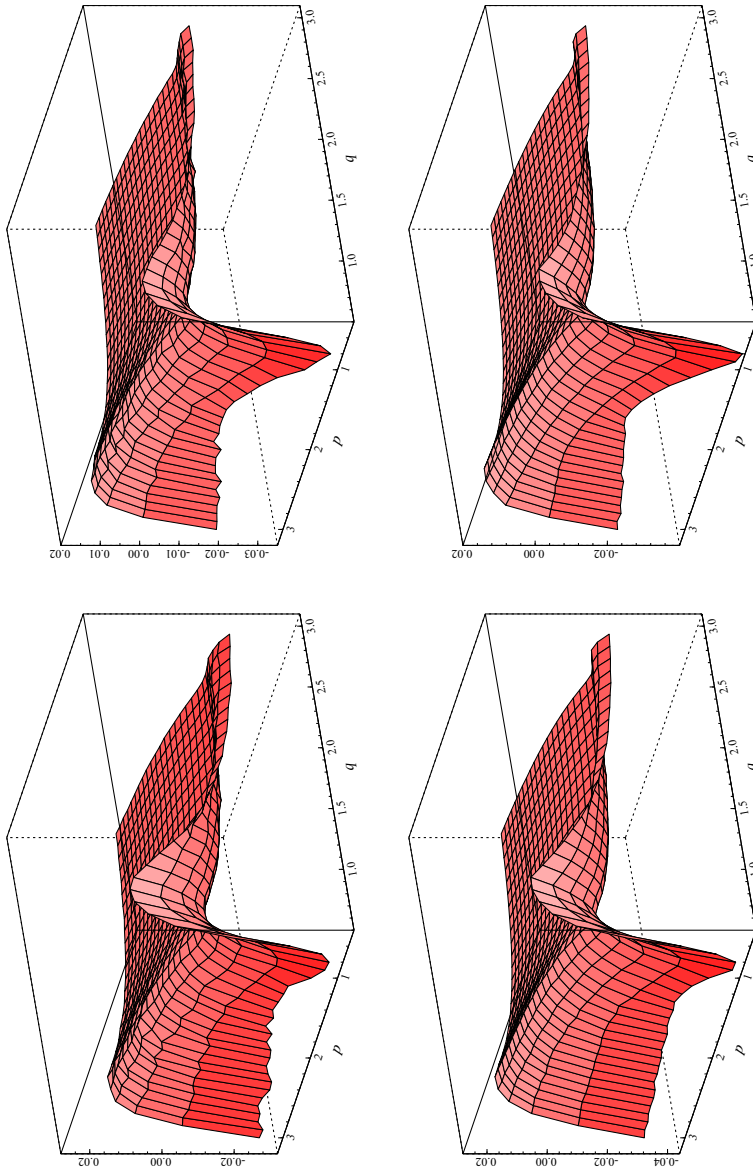
Fig. 2 Differences of RMSEs of Gini coefficients. *Note* The differences of the RMSEs from Singh–Maddala and Dagum distributions are displayed. Therefore, if the differences are larger than zero, the RMSEs from Dagum distribuion are smaller than those from Singh–Maddala distribution and vice versa. The *top left* side shows the results from quintile data and *n* = 10,000, the *top right* side shows those from decile data and *n* = 10,000, the *bottom left* side shows those from quintile data and *n* = 100,000, and the *bottom right* side shows those from decile data and *n* = 100,000, respectively
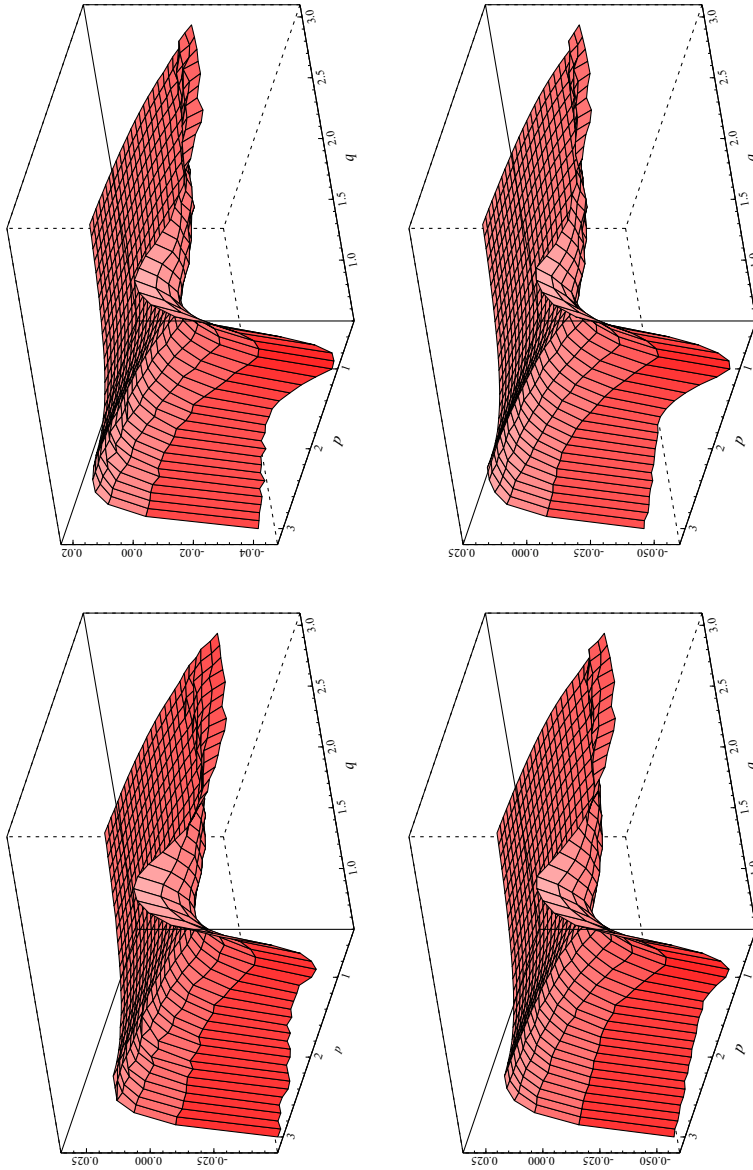
Fig. 3 Differences of RMSEs of top income shares. *Note* The differences of the RMSEs from Singh–Maddala and Dagum distributions are displayed. Therefore, if the differences are larger than zero, it means that the RMSEs from Dagum distribution are smaller than those from Singh–Maddala distribution and vice versa. The *top left* side shows the results from quintile data and $n = 10,000$, the *top right* side shows those from decile data and $n = 10,000$, the *bottom left* side shows those from quintile data and $n = 100,000$, and the *bottom right* side shows those from decile data and $n = 100,000$

Figure 2 shows the differences of the RMSEs of the Gini coefficients. From the figure, similar tendencies can be found, even if the number of observations n and/or the number of groups are changed. First, we find that the differences of the RSMEs are almost zero if $p = q$, $p > 1$, and $q > 1$. Second, the RMSEs from the Singh–Maddala distributions are smaller than those from the Dagum distributions if $p < q$ and $p > 1$, while the RMSEs from the Dagum distributions are smaller than those from the Singh–Maddala distributions if $p < q$ and $q > 1$. Moreover, the difference increases when parameters p and q move far away from $p = q$ and the differences become largest when p or q approaches 1. In other words, the RMSEs from the true distribution are smaller than those from the competing distribution. However, exceptions appear in the case of $p < 1$ or $q < 1$. In the case of $p < 1$, we find the parameter region where $q < 1.5$, while in the case of $q < 1$, the differences become smaller and cross zero, that is, the RMSEs from the Singh–Maddala distributions are smaller than those from the Dagum distributions. Therefore, we can conclude that the accuracy of the distribution seems to depend on the RMSE relationship of the parameters p and q, which is the same as the results from the AIC.

Figure 3 shows the differences of the RMSEs of the top income shares. From the figure, similar tendencies can be found, even if the number of observations n and/or the number of groups are changed and the tendencies are almost the same as the results from the Gini coefficients. First, we find that the differences of the RSMEs are almost zero if $p = q$, $p > 1$, and $q > 1$. Second, the RMSEs from the Singh–Maddala distributions are smaller than those from the Dagum distributions if $p < q$ and $p > 1$, while the RMSEs from the Dagum distributions are smaller than those from the Singh–Maddala distributions if $p < q$ and $q > 1$. Moreover, the differences increase when parameters p and q move far away from $p = q$ and the differences becomes largest when p or q approaches 1. In other words, the RMSEs from the true distribution are smaller than those from the competing distribution. However, exceptions appear in the case of $p < 1$ or $q < 1$. In the case of $p < 1$, we find the parameter region where $q < 1.5$, while in the case of $q < 1$, the differences become smaller and cross zero, that is, the RMSEs from the Singh–Maddala distributions are smaller than those from the Dagum distributions. Therefore, we can conclude that the accuracy of the distribution seems to depend on the RMSE relationship of the parameters p and q, which is the same as the results from the AIC and Gini coefficients.

Summarizing the results, if the DGPs are GB2 distribution, the case in which $p = 1$ and $q = 1$ plays an important role and the selected frequencies or RMSEs from the Gini coefficients or the top income shares appear as a type of saddle point; moreover, around the point, the performance of the AIC, Gini coefficients, and top income shares change. Therefore, it is important to check whether parameters p and/or q in the GB2 distribution are one, even when we choose the Dagum and Singh–Maddala distributions as income distribution.

4 Conclusion

This study compared the fit of the distributions in terms of AIC and the accuracy of the estimators of the Gini coefficients and top income shares from the Dagum and

Singh–Maddala distributions by RMSEs through Monte Carlo experiments. From the experiments, we confirmed that the fit of the distributions depends on the relationships and magnitudes of the parameters p and q and that the RMSEs of the Gini coefficients and top income shares depend on the relationships of the parameters when the DGPs are GB2 distribution. Therefore, a possible reason why the Dagum distribution is preferred to the Singh–Maddala distribution is that the estimated parameters lie in the desirable regions where the Dagum distribution is preferred when the GB2 distribution is estimated.

Finally, we mention issues that remain to be analyzed. We examined the fit of the distributions only by AIC. However, there are several other model selection procedures or tests. Therefore, it is worthwhile examining other methods, including, for example, the marginal likelihood by Chib and Jeliazkov (2001) and the Bayesian goodness-of-fit test by Johnson (2004). Moreover, we assume GB2 distribution as the DGP. However, the true income data might not be generated from the GB2 distribution but from the double Pareto-lognormal distribution by Reed and Jorgensen (2013) or some mixture of distributions. Therefore, it would be worthwhile to examine the performance of the Dagum and Singh–Maddala distribution when the true GDP is not GB2. However, we believe that our findings from the Monte Carlo experiments represent an interesting first step.

Appendix: Markov chain Monte Carlo Methods

In this appendix, we introduce a Markov chain Monte Carlo (MCMC) method to estimate the parameters of the distributions, which is proposed by Chotikapanich and Griffiths (2000) for the Singh–Maddala distribution. As the estimation procedure of the Dagum distribution is the same as that of the Singh–Maddala distribution, we continue the explanation focusing on the Singh–Maddala distribution.

Chotikapanich and Griffiths (2000) first proposed the MCMC method to estimate the parameters of the distribution from grouped data using a random walk Metropolis–Hastings (RWMH) algorithm. However, it is sometimes difficult to estimate the parameters using their algorithm. Therefore, some modifications are required. Their algorithm with some modifications is as follows.

1. Generate a candidate value $\boldsymbol{\theta}_{SM}^{new}$ from $\mathcal{N}(\boldsymbol{\theta}_{SM}^{(m-1)}, c^2\boldsymbol{\Sigma})$, where c is a tuning parameter and $\boldsymbol{\Sigma}$ is the maximum likelihood covariance estimate.[7]
2. Compute

---

[7] It is sometimes difficult to find the mode of the parameters by maximum likelihood method. Thus, we implement the simulated annealing of Goffe et al. (1994). Moreover, if the Cholesky decomposition of $\boldsymbol{\Sigma}$ fails, the modified Cholesky of Nocedal and Wright (2000) is utilized.

$$\alpha\left(\boldsymbol{\theta}_{SM}^{(m-1)}, \boldsymbol{\theta}_{SM}^{new}\right) = \min\left\{1, \frac{\pi(\boldsymbol{\theta}_{SM}^{new}|\mathbf{x})}{\pi(\boldsymbol{\theta}_{SM}^{(m-1)}|\mathbf{x})}\right\},$$

and if any of the elements of $\boldsymbol{\theta}_{SM}^{new}$ fall outside the feasible parameter region, then $\alpha\left(\boldsymbol{\theta}_{SM}^{(m-1)}, \boldsymbol{\theta}_{SM}^{new}\right) = 0$.

3. Generate a value u from $\mathcal{U}(0, 1)$, where $\mathcal{U}(a, b)$ is a uniform distribution on the interval (a, b).
4. If $u \leq \alpha\left(\boldsymbol{\theta}_{SM}^{(m-1)}, \boldsymbol{\theta}_{SM}^{new}\right)$, set $\boldsymbol{\theta}_{SM}^{(m)} = \boldsymbol{\theta}_{SM}^{new}$, otherwise $\boldsymbol{\theta}_{SM}^{(m)} = \boldsymbol{\theta}_{SM}^{(m-1)}$.
5. Return to step 1, with m set to $m + 1$.

References

Atkinson, A. B., Piketty, T., & Saez, E. (2011). Top incomes in the long run of history. *Journal of Economic Literature*, *49*, 3–71.

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

Brzezinski, M. (2013). Asymptotic and bootstrap inference for top income share. *Economics Letters*, *120*, 10–13.

Chib, S., & Jeliazkov, I. (2001). Marginal likelihood from the Metropolis–Hastings output. *Journal of the American Statistical Association*, *96*, 270–281.

Chotikapanich, D., & Griffiths, W. E. (2000). Posterior distributions for the Gini coefficient using grouped data. *Australian and New Zealand Journal of Statistics*, *42*, 383–392.

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

Doornik, J. A. (2009). *Ox: An object oriented matrix programming language*. London: Timberlake.

Gelfand, A. E., Hills, S. E., Racine-Poon, A., & Smith, A. F. M. (1990). Illustration of Bayesian inference in normal data models using Gibbs sampling. *Journal of the American Statistical Association*, *85*, 972–985.

Goffe, W. L., Ferrier, G. D., & Rogers, J. (1994). Global optimization of statistical functions with simulated annealing. *Journal of Econometrics*, *60*, 65–99.

Hajargasht, G., Griffiths, W. E., Brice, J., Rao, D. S. P., & Chotikapanich, D. (2012). Inference for income distributions using grouped data. *Journal of Business & Economic Statistics*, *30*, 563–575.

Johnson, V. (2004). A Bayesian χ^2 test for goodness-of-fit. *Annals of Statistics*, *32*, 2361–2384.

Kakamu, K., & Nishino, H. (2014). Bayesian estimation of beta type distribution parameters based on grouped data. mimeo.

Kleiber, C. (1996). Dagum vs. Singh–Maddala income distributions. *Economics Letters*, *53*, 265–268.

Kleiber, C. (2008). A guide to the Dagum distributions. In D. Chotikapanich (Ed.), *Modelling income distributions and Lorenz curves* (pp. 97–117). New York: Springer.

Kleiber, C., & Kotz, S. (2003). *Statistical size distributions in economics and actuarial science*. New York: Wiley.

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

McDonald, J. B. (1984). Some generalized functions for the size distribution of income. *Econometrica*, *52*, 647–663.

McDonald, J. B., & Mantrala, A. (1995). The distribution of personal income: Revisited. *Journal of Applied Econometrics*, *10*, 201–204.

McDonald, J. B., & Ransom, M. R. (1979a). Functional forms, estimation techniques and the distribution of income. *Econometrica*, *47*, 1513–1525.

McDonald, J. B., & Ransom, M. R. (1979b). Alternative parameter estimators based upon grouped data. *Communications in Statistics*, *A8*, 899–917.

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

Nocedal, J., & Wright, S. J. (2000). *Numerical optimization* (2nd ed.). New York: Springer.

Reed, W. J., & Jorgensen, M. (2013). The double Pareto-lognormal distribution–a new parametric model for size distributions. *Communications in Statistics, Theory and Methods*, *33*, 1733–1753.

Singh, S. K., & Maddala, G. S. (1976). A function for size distribution of income. *Econometrica*, *44*, 963–970.

Tadikamalla, P. R. (1980). A look at the Burr and related distributions. *International Statistical Review*, *48*, 337–344.