On GMM Estimation of Distributions from Grouped Data

William Griffiths and Gholamreza Hajargasht

Department of Economics
University of Melbourne, Australia

November 24, 2014

Abstract

For estimating distributions from grouped data, setting up moment conditions in terms of group shares and group means leads to an optimal weight matrix and a GMM objective function that are considerably simpler than those from a previous specification. Minimisation is more efficient and convergence is more reliable.

Key words: Optimal Weight Matrix, Income Distributions, Share Data
JEL Classification: C13, C16, D31

Corresponding author:

William Griffiths
Department of Economics
University of Melbourne
Vic 3010
Australia
Email: wegrif@unimelb.edu.au
Telephone: +613 8344 3622
Fax: +613 8344 6899

On GMM Estimation of Distributions from Grouped Data

William Griffiths and Gholamreza Hajargasht

1. Introduction

Consider a sample of T observations $\left(y_1, y_2, \ldots, y_T\right)$, randomly drawn from a parametric distribution $f(y;\phi)$, $(y>0)$, and grouped into N classes defined by exogenously chosen class limits $(z_0,z_1),(z_1,z_2),\ldots,(z_{N-1},z_N)$, with $z_0=0$ and $z_N=\infty$. Let $g_i\left(y\right)$ be an indicator function such that $g_i\left(y\right)=1$ if $z_{i-1}<y\le z_i$, and 0 otherwise. Assume that the data available to the researcher are (a) the sample mean $\bar{y}$, (b) the proportion of observations in each class

$$c_i = \frac{1}{T}\sum_{t=1}^{T} g_i\left(y_t\right) = \frac{T_i}{T}, \tag{1}$$

and (c) the proportion of the total value of all observations in each class

$$s_i = \frac{1}{T\bar{y}}\sum_{t=1}^{T} y_t g_i\left(y_t\right). \tag{2}$$

Our problem is to estimate ϕ, and, if they are unknown, the class limits $z_1, z_2, \ldots, z_{N-1}$.

The motivation for this problem is the availability and use of grouped data on income or expenditure, typically provided in this form on the websites of the World Bank and the World Institute for Development Economics Research (WIDER). Data on population shares c_i, income shares s_i, and mean income $\bar{y}$ are available for estimating income distributions $f(y;\phi)$. Examples of where such data are used for estimation and measuring poverty and inequality are Chotikapanich et al. (2007, 2012) and Hajargasht et al. (2012). A method of moments estimator that utilises c_i, s_i, and $\bar{y}$ to estimate beta-2 distributions was proposed by Chotikapanich et al. (2007), and later used in a large scale study of changes in global and regional inequality by Chotikapanich et al. (2012). Hajargasht et al. (2012) refined this earlier work by deriving an optimal GMM estimator, and showing how it can be used to estimate parametric income distributions of any form.

These studies set up moment conditions for the proportions c_i, and for either mean income in each group

$$\bar{y}_i = \frac{1}{T_i}\sum_{t=1}^{T} y_t g_i\left(y_t\right) = \frac{s_i\bar{y}}{c_i}, \tag{3}$$

or for that part of total mean income in the *i*-th group

$$\tilde{y}_i = \frac{1}{T}\sum_{t=1}^{T} y_t g_i\left(y_t\right) = s_i\bar{y} = c_i\bar{y}_i. \tag{4}$$

Chotikapanich et al (2007, 2012) used moment conditions for $\bar{y}_i$, whereas, for deriving an optimal GMM estimator, Hajargasht et al (2012) found it easier to work with $\tilde{y}_i$. In this paper we derive an expression for the optimal GMM estimator when using the moment conditions for the group mean incomes $\bar{y}_i$. Although both approaches are asymptotically equivalent, specification of the moment conditions in terms of the group means $\bar{y}_i$ is more natural. More importantly, the resulting GMM objective function is more convenient computationally than its counterpart for $\tilde{y}_i$; the minimization problem is simpler and convergence is easier to achieve. A small Monte Carlo experiment is used to demonstrate the validity and practicality of the proposed estimator.

Throughout, we treat the class limits $\left(z_1, z_2, \ldots, z_{N-1}\right)$ as unknown since doing so is more general than treating them as known, and they are not provided in the data source that motivated this study.[1] However, our results also hold for the case where $\left(z_1, z_2, \ldots, z_{N-1}\right)$ are known; there is simply a reduction in the number of parameters to be estimated. Also, since our GMM estimator makes a distributional assumption about the data generating process, it differs from the traditional GMM estimator which is based on a less restrictive set of

[1] The World Bank and WIDER are primarily interested in Lorenz curve estimation where class limits are not required.

assumptions. We utilize GMM in spite of the distributional assumption because derivation of the likelihood function that uses information on both the c_i and the $\bar{y}_i$ is not straightforward.

In Section 2 we review the moment conditions, optimal weight matrix and GMM objective function set up and derived by Hajargasht et al. (2012) for using data on $\left(c_i, \tilde{y}_i\right)$. In Section 3 we use these results to derive the optimal weight matrix and GMM objective function for the case where $\left(c_i, \bar{y}_i\right)$ are used to set up the moment conditions. Results from a Monte Carlo experiment are presented in Section 4.

2. Previous Results

Let the complete set of unknown parameters be given by $\boldsymbol{\theta} = \left(z_1, z_2, \ldots, z_{N-1}, \phi'\right)'$. Defining the population moments corresponding to c_i and $\tilde{y}_i$ as $k_i\left(\boldsymbol{\theta}\right)$ and $\tilde{\mu}_i\left(\boldsymbol{\theta}\right)$, respectively, we have, from equations (1) and (4),

$$k_i\left(\boldsymbol{\theta}\right) = E\left[g_i\left(y\right)\right] = \int_0^\infty g_i\left(y\right) f\left(y;\phi\right) dy = \int_{z_{i-1}}^{z_i} f(y;\phi) dy \tag{5}$$

and

$$\tilde{\mu}_i\left(\boldsymbol{\theta}\right) = E\left[y g_i\left(y\right)\right] = \int_0^\infty y g_i\left(y\right) f\left(y;\phi\right) dy = \int_{z_{i-1}}^{z_i} y f(y;\phi) dy \; . \tag{6}$$

Setting up corresponding moment conditions in matrix notation, we define

$$\mathbf{H}\left(\boldsymbol{\theta}\right) = \frac{1}{T}\sum_{t=1}^{T} \mathbf{h}\left(y_t;\boldsymbol{\theta}\right) = \frac{1}{T}\sum_{t=1}^{T} \begin{bmatrix} g_1\left(y_t\right) - k_1\left(\boldsymbol{\theta}\right) \\ \vdots \\ g_{N-1}\left(y_t\right) - k_{N-1}\left(\boldsymbol{\theta}\right) \\ y_t g_1\left(y_t\right) - \tilde{\mu}_1\left(\boldsymbol{\theta}\right) \\ \vdots \\ y_t g_N\left(y_t\right) - \tilde{\mu}_N\left(\boldsymbol{\theta}\right) \end{bmatrix} = \begin{bmatrix} c_1 - k_1\left(\boldsymbol{\theta}\right) \\ \vdots \\ c_{N-1} - k_{N-1}\left(\boldsymbol{\theta}\right) \\ \tilde{y}_1 - \tilde{\mu}_1\left(\boldsymbol{\theta}\right) \\ \vdots \\ \tilde{y}_N - \tilde{\mu}_N\left(\boldsymbol{\theta}\right) \end{bmatrix} = \begin{bmatrix} \mathbf{c}_{-N} - \mathbf{k}_{-N} \\ \tilde{\mathbf{y}} - \tilde{\boldsymbol{\mu}} \end{bmatrix} \tag{7}$$

In (7), the moment condition for c_N is omitted because the result $\sum_{i=1}^{N} k_i\left(\boldsymbol{\theta}\right) = \sum_{i=1}^{N} c_i = 1$ makes one of the N conditions for the c_i redundant. The notation **c**, **k**, $\tilde{\mathbf{y}}$, and $\tilde{\boldsymbol{\mu}}$ is used to denote N-dimensional vectors containing the elements $c_i, k_i, \tilde{y}_i$, and $\tilde{\mu}_i$, respectively, and the subscript "$-N$" denotes a corresponding $(N-1)$-dimensional vector with the last element excluded; the dependence on θ has been dropped from these vectors for notational ease. The GMM objective function is given by $\varphi = \mathbf{H}(\boldsymbol{\theta})'\mathbf{W}\,\mathbf{H}(\boldsymbol{\theta})$, and the GMM estimator by

$$\hat{\boldsymbol{\theta}} = \arg\min_{\boldsymbol{\theta}} \mathbf{H}(\boldsymbol{\theta})'\mathbf{W}\,\mathbf{H}(\boldsymbol{\theta}) \tag{8}$$

where $\mathbf{W}$ is a weight matrix. The optimal weight matrix is given by the inverse of the covariance matrix of the limiting distribution of $T^{1/2}\mathbf{H}(\boldsymbol{\theta})$ (see e.g., Cameron and Trivedi 2005, pp 174). It can be obtained by taking the probability limit

$$\text{var}\left[T^{1/2}\mathbf{H}(\boldsymbol{\theta})\right] = \text{plim}\frac{1}{T}\sum_{t=1}^{T}\mathbf{h}(y_t,\boldsymbol{\theta})\mathbf{h}(y_t,\boldsymbol{\theta})' \; . \tag{9}$$

Hajargasht et al. (2012) showed that

$$\text{var}\left[T^{1/2}\mathbf{H}(\boldsymbol{\theta})\right] = \left[\begin{array}{c:c} \mathbf{D}(\mathbf{k}_{-N}) & \left[\mathbf{D}(\tilde{\boldsymbol{\mu}}_{-N}) \quad \mathbf{0}_{N-1}\right] \\ \hdashline \begin{bmatrix} \mathbf{D}(\tilde{\boldsymbol{\mu}}_{-N}) \\ \mathbf{0}'_{N-1} \end{bmatrix} & \mathbf{D}(\tilde{\boldsymbol{\mu}}^{(2)}) \end{array}\right] - \begin{bmatrix} \mathbf{k}_{-N} \\ \tilde{\boldsymbol{\mu}} \end{bmatrix}\left[\mathbf{k}'_{-N} \quad \tilde{\boldsymbol{\mu}}'\right] \tag{10}$$

where $\mathbf{0}_{N-1}$ is an $(N-1)$ dimensional vector of zeroes, $\mathbf{D}(\mathbf{x})$ denotes a diagonal matrix with elements of a vector $\mathbf{x}$ on the diagonal, and $\tilde{\boldsymbol{\mu}}^{(2)} = \left(\tilde{\mu}_1^{(2)}, \tilde{\mu}_2^{(2)}, \ldots, \tilde{\mu}_N^{(2)}\right)'$, with

$$\tilde{\mu}_i^{(2)}(\boldsymbol{\theta}) = \int_{z_{i-1}}^{z_i} y^2 f(y;\phi)dy = \int_0^\infty y^2 g_i(y)\, f(y;\phi)\, dy = E\left[y^2 g_i(y)\right] . \tag{11}$$

They then go on to show that inverting (10), and simplifying $\varphi = \mathbf{H}(\boldsymbol{\theta})'\,\mathbf{W}\,\mathbf{H}(\boldsymbol{\theta})$, leads to the GMM objective function

$$\varphi = \sum_{i=1}^{N} w_{1i}(c_i - k_i)^2 + \sum_{i=1}^{N} w_{2i}(\tilde{y}_i - \tilde{\mu}_i)^2 - 2\sum_{i=1}^{N} w_{3i}(c_i - k_i)(\tilde{y}_i - \tilde{\mu}_i) \tag{12}$$

where $w_{1i} = \tilde{\mu}_i^{(2)}/v_i$, $w_{2i} = k_i/v_i$, and $w_{3i} = \tilde{\mu}_i/v_i$, with $v_i = k_i\tilde{\mu}_i^{(2)} - \tilde{\mu}_i^2$. What is important to note is that minimising this function involves minimising weighted sums of squares of

$(c_i - k_i)$ and $(\tilde{y}_i - \tilde{\mu}_i)$, as well as a weighted sum of the cross product terms $(c_i - k_i)(\tilde{y}_i - \tilde{\mu}_i)$. In the formulation that we consider in the next section, the weight matrix $\mathbf{W}$ turns out to be such that the need for cross product terms can be avoided.

3. Optimal Weight Matrix and Objective Function for $(c_i, \bar{y}_i)$

Considering now the moment conditions for $(c_i, \bar{y}_i)$, we have $E\left[c_i - k_i(\boldsymbol{\theta})\right] = 0$ (as before) and $\text{plim}\left[\bar{y}_i - \mu_i(\boldsymbol{\theta})\right] = 0$, where $\mu_i(\boldsymbol{\theta}) = \tilde{\mu}_i(\boldsymbol{\theta})/k_i(\boldsymbol{\theta})$.[2] Collecting these terms into a vector, we define

$$\mathbf{L}(\boldsymbol{\theta}) = \begin{bmatrix} c_1 - k_1(\boldsymbol{\theta}) \\ \vdots \\ c_{N-1} - k_{N-1}(\boldsymbol{\theta}) \\ \bar{y}_1 - \mu_1(\boldsymbol{\theta}) \\ \vdots \\ \bar{y}_N - \mu_N(\boldsymbol{\theta}) \end{bmatrix} = \begin{bmatrix} \mathbf{c}_{-N} - \mathbf{k}_{-N} \\ \bar{\mathbf{y}} - \boldsymbol{\mu} \end{bmatrix}. \tag{13}$$

This time the GMM problem is to minimize $\varphi^* = \mathbf{L}(\theta)'\mathbf{W}^*\mathbf{L}(\theta)$ where for an optimal weight matrix $\mathbf{W}^*$, we require the inverse of the covariance matrix of the limiting distribution of $T^{1/2}\mathbf{L}(\boldsymbol{\theta})$. Working towards this covariance matrix, we define the stochastic components of $\mathbf{L}(\boldsymbol{\theta})$, and a 2*N*-dimensional version of $\mathbf{H}(\boldsymbol{\theta})$, as $\boldsymbol{\eta}_{-N}$ and $\tilde{\boldsymbol{\eta}}$, respectively, where

$$\boldsymbol{\eta}_{-N} = \begin{bmatrix} c_1 \\ \vdots \\ c_{N-1} \\ \tilde{y}_1/c_1 \\ \vdots \\ \tilde{y}_N/c_N \end{bmatrix} \quad \text{and} \quad \tilde{\boldsymbol{\eta}} = \begin{bmatrix} c_1 \\ \vdots \\ c_N \\ \tilde{y}_1 \\ \vdots \\ \tilde{y}_N \end{bmatrix} \quad \Rightarrow \quad \boldsymbol{\eta}_{-N} = \begin{bmatrix} \tilde{\eta}_1 \\ \vdots \\ \tilde{\eta}_{N-1} \\ \tilde{\eta}_{N+1}/\tilde{\eta}_1 \\ \vdots \\ \tilde{\eta}_{2N}/\tilde{\eta}_N \end{bmatrix} \tag{14}$$

[2] In this case we use probability limits instead of expectations because $E(\bar{y}_i) \neq E(\tilde{y}_i)/E(c_i) = \tilde{\mu}_i(\boldsymbol{\theta})/k_i(\boldsymbol{\theta})$. Using probability limits may mean it is better to call the estimator a "minimum distance estimator". The asymptotic distribution is the same, however. See, for example, Greene (2012, Ch.13).

Then, using the delta method, the covariance matrix of the limiting distribution can be written as

$$\mathrm{var}\left[T^{1/2}\mathbf{L}(\boldsymbol{\theta})\right] = \mathrm{plim}\left(\frac{\partial \boldsymbol{\eta}_{-N}}{\partial \tilde{\boldsymbol{\eta}}'}\right)\mathrm{var}\left(T^{1/2}\tilde{\boldsymbol{\eta}}\right)\mathrm{plim}\left(\frac{\partial \boldsymbol{\eta}'_{-N}}{\partial \tilde{\boldsymbol{\eta}}}\right) \tag{15}$$

Differentiating and taking probability limits, yields

$$\mathrm{plim}\left(\frac{\partial \boldsymbol{\eta}_{-N}}{\partial \tilde{\boldsymbol{\eta}}'}\right) = \begin{pmatrix} [\mathbf{I}_{N-1}, \mathbf{0}_{(N-1)\times 1}] & \mathbf{0}_{(N-1)\times N} \\ \mathbf{D}\left(-\tilde{\boldsymbol{\mu}}/\mathbf{k}^2\right) & \mathbf{D}(\mathbf{1}/\mathbf{k}) \end{pmatrix} \tag{16}$$

where, with a slight abuse of matrix algebra notation, we are using $\mathbf{D}\left(-\tilde{\boldsymbol{\mu}}/\mathbf{k}^2\right)$ to denote a diagonal matrix with elements $\left(-\mu_i/k_i^2\right)$ on the diagonal. Similar definitions apply to $\mathbf{D}(\mathbf{1}/\mathbf{k})$ and other diagonal matrices that follow. Using a 2*N*-dimensional version of equation (10), we can show that

$$\mathrm{var}\left(T^{1/2}\tilde{\boldsymbol{\eta}}\right) = \begin{pmatrix} \mathbf{D}(\mathbf{k}) - \mathbf{k}\mathbf{k}' & \mathbf{D}(\tilde{\boldsymbol{\mu}}) - \mathbf{k}\tilde{\boldsymbol{\mu}}' \\ \mathbf{D}(\tilde{\boldsymbol{\mu}}) - \tilde{\boldsymbol{\mu}}\mathbf{k}' & \mathbf{D}\left(\tilde{\boldsymbol{\mu}}^{(2)}\right) - \tilde{\boldsymbol{\mu}}\tilde{\boldsymbol{\mu}}' \end{pmatrix} \tag{17}$$

From (15), (16) and (17), tedious matrix manipulation yields

$$\begin{aligned} \mathrm{var}\left[T^{1/2}\mathbf{L}(\boldsymbol{\theta})\right] &= \begin{pmatrix} \mathbf{D}(\mathbf{k}_{-N}) - \mathbf{k}_{-N}\mathbf{k}'_{-N} & \mathbf{0} \\ \mathbf{0} & \mathbf{D}\left(\tilde{\boldsymbol{\mu}}^{(2)}/\mathbf{k}^2 - \tilde{\boldsymbol{\mu}}^2/\mathbf{k}^3\right) \end{pmatrix} \\ &= \begin{pmatrix} \mathbf{D}(\mathbf{k}_{-N}) - \mathbf{k}_{-N}\mathbf{k}'_{-N} & \mathbf{0} \\ \mathbf{0} & \mathbf{D}(\boldsymbol{\upsilon}/\mathbf{k}) \end{pmatrix} \end{aligned} \tag{18}$$

where $\upsilon_i = \tilde{\mu}_i^{(2)}/k_i - \tilde{\mu}_i^2/k_i^2$. Then, using $\left[\mathbf{D}(\mathbf{k}_{-N}) - \mathbf{k}_{-N}\mathbf{k}'_{-N}\right]^{-1} = \mathbf{D}(\mathbf{1}/\mathbf{k}_{-N}) + \mathbf{1}\mathbf{1}'/k_N$, where $\mathbf{1}$ is a vector of ones, we have

$$\mathbf{W}^* = \left(\mathrm{var}\left[T^{1/2}\mathbf{L}(\boldsymbol{\theta})\right]\right)^{-1} = \begin{pmatrix} \mathbf{D}(\mathbf{1}/\mathbf{k}_{-N}) + \mathbf{1}\mathbf{1}'/k_N & \mathbf{0} \\ \mathbf{0} & \mathbf{D}(\mathbf{k}/\boldsymbol{\upsilon}) \end{pmatrix} \tag{19}$$

Now,

$$\left(\mathbf{c}_{-N}-\mathbf{k}_{-N}\right)'\left[\mathbf{D}\left(\mathbf{1}/\mathbf{k}_{-N}\right)+\mathbf{11}'/k_N\right]\left(\mathbf{c}_{-N}-\mathbf{k}_{-N}\right)$$
$$=\left(\mathbf{c}_{-N}-\mathbf{k}_{-N}\right)'\mathbf{D}(\mathbf{1}/\mathbf{k}_{-N})\left(\mathbf{c}_{-N}-\mathbf{k}_{-N}\right)+\left(\sum_{i=1}^{N-1}\left(c_i-k_i\right)\right)^2\Big/k_N \quad (20)$$
$$=\sum_{i=1}^{N}(c_i-k_i)^2/k_i$$

with the last equality coming from $\sum_{i=1}^{N-1}\left(c_i-k_i\right)=k_N-c_N$.

Then the GMM objective function can be written simply as

$$\varphi^*=\mathbf{L}(\theta)'\mathbf{W}^*\mathbf{L}(\theta)$$
$$=\begin{bmatrix}\mathbf{c}_{-N}-\mathbf{k}_{-N}\\ \bar{\mathbf{y}}-\boldsymbol{\mu}\end{bmatrix}'\begin{pmatrix}\mathbf{D}\left(\mathbf{1}/\mathbf{k}_{-N}\right)+\mathbf{11}'/k_N & \mathbf{0}\\ \mathbf{0} & \mathbf{D}\left(\mathbf{k}/\boldsymbol{\upsilon}\right)\end{pmatrix}\begin{bmatrix}\mathbf{c}_{-N}-\mathbf{k}_{-N}\\ \bar{\mathbf{y}}-\boldsymbol{\mu}\end{bmatrix} \quad (21)$$
$$=\sum_{i=1}^{N}\left(c_i-k_i\right)^2\Big/k_i+\sum_{i=1}^{N}k_i\left(\bar{y}_i-\mu_i\right)^2\Big/\upsilon_i$$

Thus, in this formulation the objective function can be written as a weighted sum of squares of the deviations $\left(c_i-k_i\right)$ and $\left(\bar{y}_i-\mu_i\right)$. In contrast to the earlier formulation written in terms of the deviations (c_i-k_i) and $(\tilde{y}_i-\tilde{\mu}_i)$, there is no cross product term, making the minimization problem simpler and convergence easier to obtain. Also of interest is the large sample covariance matrix of the estimator $\hat{\boldsymbol{\theta}}^*=\arg\min_{\boldsymbol{\theta}}\mathbf{L}(\theta)'\mathbf{W}^*\mathbf{L}(\theta)$ which turns out to be

$$\operatorname{var}\left(\hat{\boldsymbol{\theta}}^*\right)=\frac{1}{T}\left(\begin{bmatrix}\frac{\partial\mathbf{k}'}{\partial\boldsymbol{\theta}} & \frac{\partial\boldsymbol{\mu}'}{\partial\boldsymbol{\theta}}\end{bmatrix}\begin{bmatrix}\mathbf{D}(\mathbf{1}/\mathbf{k}) & \mathbf{0}\\ \mathbf{0} & \mathbf{D}(\mathbf{k}/\boldsymbol{\upsilon})\end{bmatrix}\begin{bmatrix}\partial\mathbf{k}/\partial\boldsymbol{\theta}'\\ \partial\boldsymbol{\mu}/\partial\boldsymbol{\theta}'\end{bmatrix}\right)^{-1}. \quad (22)$$

In the next section we carry out a Monte Carlo experiment to assess whether the estimator works well in practice, including whether equation (22) is an accurate reflection of the variance of the estimator in repeated samples.

4. A Monte Carlo Analysis

The design of the Monte Carlo experiment is as follows: We generate data from a generalized beta distribution of the second kind (GB2). This distribution is a flexible and popular candidate for estimating income distributions (see e.g., McDonald 1984, McDonald and Ransom 2008, or Kleiber and Kotz 2003). Its density function is defined as $f(y;\phi) = ay^{ap-1} \big/ \left(b^{ap}[1+(y/b)^a]^{p+q} \right)$, with parameters $\phi = (b, p, q, a)'$. The settings in the experiment were $b = 100$, $p = 1$, $q = 1.5$ and $a = 1.5$, implying a relatively heavy-tailed Singh-Maddala distribution with a Gini coefficient of 0.53. The Singh-Maddala is a special case of the GB2 with $p = 1$. Although the data were generated from the Singh-Maddala distribution, the more general GB2 distribution was estimated. We considered a heavy-tailed case because inference in such cases can be more challenging. In each of 1000 samples, we generated 10000 observations and created ten groups where the group bounds were the theoretical deciles (z_i's) of the distribution. The shares of observations for each group and the means for each group were computed and formed the data used in estimation. Computable expressions for $\mathbf{k}$, $\tilde{\boldsymbol{\mu}}$ and $\tilde{\boldsymbol{\mu}}^{(2)}$ in terms of the GB2 parameters are given in Hajargasht et al. (2012). To minimize φ^*, we used an iterated two-stage GMM estimator which begins with a non-optimal weight matrix in the first stage; in the second stage we minimize φ^* with the optimal weight matrix computed using the estimates from the first stage, and the process continues until convergence. Other details of GMM estimation in this context, but with different moment conditions and a different weight matrix, can be found in Hajargasht et al. (2012).

The results of the experiment are summarized in Table 1. The first column gives the true values of the group bounds (z_i), the parameters (b, p, q, a), and the Gini index computed from the values of the parameters using numerical integration. Column two contains the averages of the estimated parameters from the Monte Carlo samples. The third column provides the estimator variances predicted from GMM theory, obtained by inserting the true values of the parameters in the GMM variance formula in (22). Column four contains the variances of the estimated parameters from the Monte Carlo samples. For all the parameters, the Monte Carlo means and variances compare favourably with their theoretical counterparts. In Figure 1 we go further for a selection of the parameters, comparing the complete Monte Carlo distributions with the normal distributions prescribed by theory. The full lines represent the kernel densities of the estimated parameters from the Monte Carlo experiment; the dashed lines show the asymptotic distributions predicted by GMM theory. They are relatively close.

5. Conclusion

With the estimation of income distributions as our motivation, we have derived the moment conditions and optimal weight matrix for GMM estimation of distributions with grouped data when the moment conditions are based on data in the form of group population shares and group means. Importantly, we showed that in this case the optimal weight matrix is considerably simpler than that used in a previous formulation, leading to a computationally simpler objective function for GMM estimation. A limited Monte Carlo analysis showed that the resulting estimator works well.

References

Cameron, C. and P. Trivedi. (2005) *Microeconometrics: Methods and Applications*, New York, Cambridge University Press.

Chotikapanich, D., Griffiths, W.E., and Rao, D.S.P. (2007), "Estimating and Combining National Income Distributions using Limited Data", *Journal of Business and Economic Statistics*, 25, 97-109.

Chotikapanich, D., Griffiths, W., Prasada Rao, D.S. and Valencia, V. (2012), "Global income Distributions and Inequality, 1993 and 2000: Incorporating Country-Level Inequality Modelled with Beta Distributions", *Review of Economics and Statistics* 94, 52-73.

Greene, W. H. (2012) *Econometric Analysis*, 7th edition, New York, Prentice Hall.

Hajargasht, G., W. Griffiths, J. Brice, D. S. P. Rao, and D. Chotikapanich (2012) "Inference for Income Distributions using Grouped Data", *Journal of Business of Economic Statistics*, 30(4), 563-576.

Kleiber, C., Kotz, S. (2003) *Statistical Size distributions in Economics and Actuarial Sciences*, New York, John Wiley.

McDonald, J.B. (1984). "Some Generalized Functions of the Size Distribution of Income", *Econometrica*, 52, 647-663.

McDonald, J. and M. Ransom (2008) "The Generalized Beta Distribution as a Model for the Distribution of Income: Estimation of Related Measures of Inequality", in D. Chotikapanich (ed.), *Modeling Income Distributions and Lorenz Curves*, Springer, New York, 147-166.

Table 1 Monte Carlo Results

Par	True	Average Monte Carlo	Variance Theory	Variance Monte Carlo
z_1	17.430	17.440	0.0238	0.0260
z_2	29.521	29.519	0.0207	0.0207
z_3	41.613	41.607	0.0227	0.0223
z_4	54.805	54.804	0.0288	0.0300
z_5	70.139	70.144	0.0413	0.0427
z_6	89.169	89.160	0.0686	0.0633
z_7	114.890	114.889	0.1399	0.1360
z_8	154.690	154.655	0.4157	0.4247
z_9	236.700	236.703	3.3236	3.3389
b	100.000	100.402	38.1090	36.9999
p	1.000	1.007	0.0142	0.0140
q	1.500	1.516	0.0487	0.0459
a	1.500	1.506	0.0163	0.0156
Gini	0.533	0.533	0.000064	0.000063

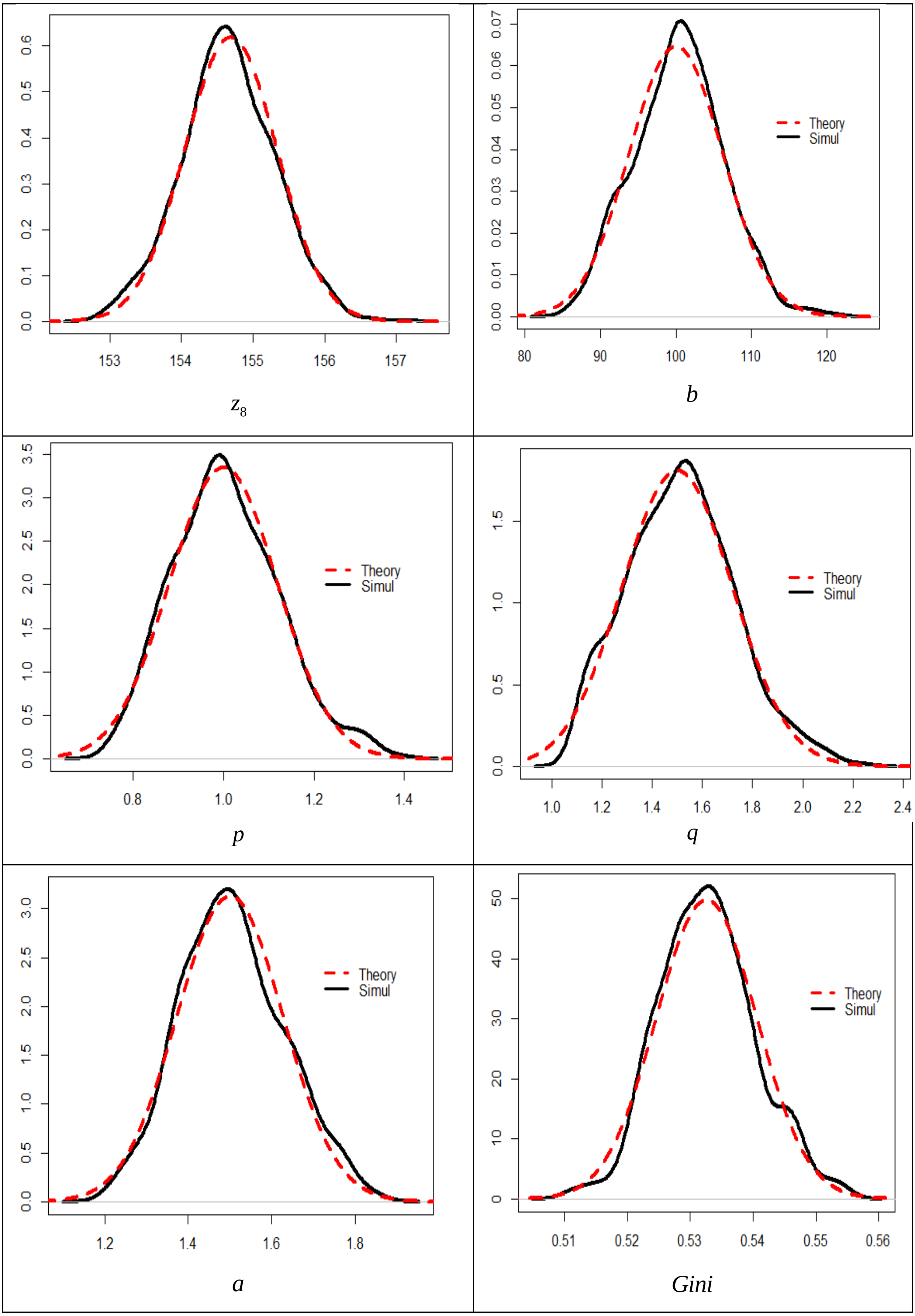

Figure 1: Distributions of the Estimates: Theoretical (dashed line) vs Monte Carlo