THE UNIVERSITY OF
MELBOURNE

Department of Economics

Working Paper Series

GMM Estimation of Income Distributions from Grouped Data

Gholamreza Hajargsht
William E. Griffiths
Joseph Brice
D.S. Prasada Rao
Duangkamon Chotikapanich

Sep 2011

Research Paper Number 1129

ISSN: 0819-2642
ISBN: 978 0 7340 4482 2

Department of Economics
The University of Melbourne
Parkville VIC 3010
www.economics.unimelb.edu.au

GMM Estimation of Income Distributions from Grouped Data

Gholamreza Hajargasht
Department of Economics
University of Melbourne, Australia

William E. Griffiths
Department of Economics
University of Melbourne, Australia

Joseph Brice
Department of Economics
University of Queensland, Australia

D.S. Prasada Rao
Department of Economics
University of Queensland, Australia

Duangkamon Chotikapanich
Department of Econometrics and Business Statistics
Monash University, Melbourne, Australia

Aug 29, 2011

Corresponding author:

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

Abstract

We develop a GMM procedure for estimating income distributions from grouped data with unknown group bounds. The approach enables us to obtain standard errors for the estimated parameters and functions of the parameters, such as inequality and poverty measures, and to test the validity of an assumed distribution using a *J*-test. Using eight countries/regions for the year 2005, we show how the methodology can be applied to estimate the parameters of the generalized beta distribution of the second kind, and its special-case distributions, the beta-2, Singh-Maddala, Dagum, generalized gamma and lognormal distributions. This work extends earlier work (Chotikapanich et al., 2007, 2012) that did not specify a formal GMM framework, did not provide methodology for obtaining standard errors, and considered only the beta-2 distribution. The results show that generalized beta distribution fits the data well and outperforms other frequently used distributions.

***Keywords*:** GMM; generalized beta distribution; grouped data; inequality and poverty.

JEL classification numbers: C13, C16, D31

1. Introduction

The estimation of income distributions has played an important role in the measurement of inequality and poverty and, more generally, in welfare comparisons over time and space. Access to what is a vast literature on the modeling of income distributions, the characteristics of different specifications, and various estimation methods, is conveniently achieved through a volume by Kleiber and Kotz (2003), the collection of papers in Chotikapanich (2008), and papers by Bandourian et al. (2003) and McDonald and Xu (1995).

For carrying out large scale investigations that involve many countries, different time periods, and the estimation of regional and global income distributions (see, for example, Milanovic (2002) and Chotikapanich et al. (2012)), compilation of the necessary country-specific income distribution data is a major research problem. The data are generally drawn from household expenditure and income surveys that are conducted once in five years in most countries. Because compilation of data and data dissemination from these surveys are resource intensive, much of the raw data are not readily available for researchers. More regularly disseminated data take the form of summary statistics that include mean income, measures of inequality like the Gini coefficient, and grouped data in the form of income and population shares. Two sources of such data are the World Bank and the World Institute for Development Economics Research.[1] The focus of this paper is the estimation of country-level income distributions from limited data of this form. Specifically, our objective is to develop and apply a generalized method of moments (GMM) estimator for income distributions, using data that is in the form of population shares and group mean incomes, with unknown group limits. Group mean incomes for each group in a given country can be computed from the readily available data on the county's mean per capita income, and its income and

[1] http://go.worldbank.org/6F2DBUXBE0 and http://www.wider.unu.edu/research/Database/en_GB/wiid/

population shares. To achieve comparability over countries and time, mean incomes that have been adjusted for purchasing power parity are available from the World Bank and the Penn World Tables.[2]

Some of us have tackled this problem before. Chotikapanich et al. (2007) (hereafter CGR) suggest a GMM estimator for the beta-2 distribution, apply it to a sample of 8 countries in two time periods, and illustrate how the estimated distributions can be combined to derive a regional distribution, find Lorenz curves, and measure inequality. In a more extensive study, Chotikapanich et al. (2012) use the same technique to estimate the global income distribution as a mixture of GMM-estimated beta-2 distributions for 91 countries in 1993 and 2000. The GMM technique is a natural candidate because it can be implemented with aggregated data rather requiring individual income observations. Chotikapanich et al. (2007, 2012) describe the main features of their approach, and show that it works well, but their method was deficient in several respects. They did not set up their estimator within a formal GMM framework, they used an arbitrarily specified weight matrix, and, because of the lack of an asymptotic covariance matrix for the estimator, they did not provide any standard errors.

These deficiencies are remedied in this paper. We define a formal set of moment conditions and construct an optimal weight matrix, leading to an asymptotically efficient estimator. Deriving the optimal weight matrix for the estimator makes it possible to estimate the asymptotic covariance matrix of the estimator which in turn provides measures of reliability in the form of standard errors of the estimated parameters and any functions of the parameters used regularly in the area of income distributions, such as inequality and poverty measures. We also extend the CGR framework to one that can accommodate any income distribution, not just the beta-2 distribution. In our empirical work we focus on the

[2] http://pwt.econ.upenn.edu/

generalized beta distribution of the second kind (GB2 distribution) and its popular special cases, the beta-2, Singh-Maddala, Dagum, generalized gamma and lognormal distributions. We show how the adequacy of an income distribution can be assessed using (i) the *J*-test for the validity of excess moment restrictions, and (ii) a comparison of predicted and observed income shares. We also illustrate how estimates of the parameters can be used to plot income distributions and their confidence bounds, and compute inequality and poverty measures.

It is useful to note how our current and past work differs from related parallel work by Wu and Perloff (2005, 2007) who also consider GMM estimation of income distributions from grouped data. Wu and Perloff use a maximum entropy density to approximate the underlying income distribution and use simulation to estimate the optimal weight matrix that is used in a two-step estimator. In this paper, we show how the optimal weight matrix can be expressed in terms of the moments and moment distribution functions of any income distribution; then, in our empirical work, we estimate the GB2 as a general and flexible class of income distributions. Knowing the parametric specification means we are able to specify the optimal weight matrix as a function of the unknown parameters and obtain optimal GMM estimates in one step. Our past work (Chotikapanich et al. (2007, 2012)) used the more restrictive beta-2 distribution and a sub-optimal weight matrix. Other distinguishing features of our current work are our emphasis on estimation of group bounds, and the asymptotic covariance matrix that can be used to find standard errors for all estimated parameters (including the bounds) and functions of those parameters.

The paper is organized as follows. In Section 2 the GMM methodology is described in general for estimating the parameters of any income distribution. In Section 3 we provide the expressions that are needed for GMM estimation of the GB2, beta-2, Singh-Maddala, Dagum, generalized gamma and lognormal distributions. These expressions include the moments, distribution functions and first-moment distribution functions. We refer to an appendix where derivatives for computing the GMM asymptotic covariance matrix can be found. Expressions for inequality and poverty measures are also provided. Section 4 contains a description of the data used to illustrate the theoretical framework and the ability of the GMM technique to recover income densities. We selected eight countries/areas for the year 2005: China rural, China urban, India rural, India urban, Pakistan, Russia, Poland and Brazil. The results presented in Section 5 include parameter estimates and their standard errors, plots of income densities and their confidence bounds, goodness-of-fit assessment, and inequality and poverty measures. Concluding remarks are provided in Section 6.

2. The GMM estimator

Suppose that we have a sample of T observations $\left(y_1, y_2, \ldots, y_T\right)$ that are assumed to be randomly drawn from a parametric income distribution $f(y;\phi)$, and that these observations have been grouped into N income classes $(z_0, z_1), (z_1, z_2), \ldots, (z_{N-1}, z_N)$, with $z_0 = 0$ and $z_N = \infty$. Let the mean class incomes for the N classes be given by $\bar{y}_1, \bar{y}_2, \ldots, \bar{y}_N$; and let the proportions of observations in each class be given by $c_1, c_2, \ldots, c_N$. Given available data on the $\bar{y}_i$ and the c_i, but not the z_i, our problem is to estimate ϕ, along with the unknown class limits $z_1, z_2, \ldots, z_{N-1}$. To tackle this problem using GMM estimation we create a set of sample moment conditions

$$\mathbf{H}\left(\boldsymbol{\theta}\right) = \frac{1}{T}\sum_{t=1}^{T}\mathbf{h}\left(y_t, \boldsymbol{\theta}\right) \tag{1}$$

such that $\operatorname{plim} \mathbf{H}\left(\boldsymbol{\theta}\right) = \mathbf{0}$, where $\boldsymbol{\theta} = \left(z_1, z_2, \ldots, z_{N-1}, \phi'\right)'$. The GMM estimator $\hat{\boldsymbol{\theta}}$ is defined as

$$\hat{\boldsymbol{\theta}} = \arg\min_{\boldsymbol{\theta}} \mathbf{H}\left(\boldsymbol{\theta}\right)' \mathbf{W}\, \mathbf{H}\left(\boldsymbol{\theta}\right) \tag{2}$$

where $\mathbf{W}$ is a weight matrix. In what follows we consider first the moment conditions and then the weight matrix.

2.1 The moment conditions

To set up the moment conditions we begin by defining the population proportion for the *i*-th income class as:

$$k_i(\boldsymbol{\theta}) = \int_{z_{i-1}}^{z_i} f(y;\phi)dy \qquad i = 1,2,\ldots,N \tag{3}$$

Also, the population class mean income for income class *i* is defined as

$$\mu_i(\boldsymbol{\theta}) = \frac{\int_{z_{i-1}}^{z_i} y f(y;\phi)dy}{\int_{z_{i-1}}^{z_i} f(y;\phi)dy} \qquad i = 1,2,\ldots,N \tag{4}$$

Then, letting $g_i(y)$ be an indicator function such that

$$g_i(y) = \begin{cases} 1 & \text{if } z_{i-1} < y \le z_i \\ 0 & \text{otherwise} \end{cases}$$

we have

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

The corresponding sample moment is

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

where T_i is the number of observations in group *i*. Thus, for the proportion of observations in each group, we have $N-1$ moment conditions

$$\frac{1}{T}\sum_{t=1}^{T} g_i\left(y_t\right)-E\left[g_i\left(y\right)\right] = c_i-k_i\left(\boldsymbol{\theta}\right) \qquad i=1,2,\ldots,N-1 \tag{5}$$

The moment condition for $i=N$ 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 redundant.

To obtain the moment conditions for the class means, we note that

$$\begin{aligned}\frac{1}{k_i\left(\boldsymbol{\theta}\right)}E\left[yg_i\left(y\right)\right] &= \frac{1}{k_i\left(\boldsymbol{\theta}\right)}\int_0^{\infty} yg_i\left(y\right)f\left(y;\boldsymbol{\phi}\right)dy \\ &= \frac{1}{k_i\left(\boldsymbol{\theta}\right)}\int_{z_{i-1}}^{z_i} yf\left(y;\boldsymbol{\phi}\right)dy \\ &= \mu_i\left(\boldsymbol{\theta}\right)\end{aligned}$$

Now, since $\operatorname{plim} c_i = k_i\left(\boldsymbol{\theta}\right)$,

$$\operatorname{plim}\left(\frac{1}{c_i}\frac{1}{T}\sum_{t=1}^{T} y_t g_i\left(y_t\right)\right)=\mu_i\left(\boldsymbol{\theta}\right)$$

Then, noting that

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

we have,

$$\operatorname{plim}\frac{1}{c_i}\frac{1}{T}\sum_{t=1}^{T} y_t g_i\left(y_t\right)=\operatorname{plim}\ \bar{y}_i=\mu_i\left(\theta\right)$$

Thus, for the class means we have the N moment conditions

$$\frac{1}{c_i}\frac{1}{T}\sum_{t=1}^{T} y_t g_i\left(y_t\right) - \text{plim}\left[\frac{1}{c_i}\frac{1}{T}\sum_{t=1}^{T} y_t g_i\left(y_t\right)\right] = \bar{y}_i - \mu_i\left(\theta\right) \qquad i = 1,2,\ldots,N \tag{6}$$

Collecting all the terms we can write

$$\mathbf{h}\left(y_t,\boldsymbol{\theta}\right) = \begin{bmatrix} g_1\left(y_t\right) - k_1\left(\boldsymbol{\theta}\right) \\ \vdots \\ g_{N-1}\left(y_t\right) - k_{N-1}\left(\boldsymbol{\theta}\right) \\ \dfrac{y_t g_1\left(y_t\right)}{c_1} - \mu_1\left(\boldsymbol{\theta}\right) \\ \vdots \\ \dfrac{y_t g_N\left(y_t\right)}{c_N} - \mu_N\left(\boldsymbol{\theta}\right) \end{bmatrix} \tag{7}$$

and

$$\mathbf{H}\left(\boldsymbol{\theta}\right) \;=\; \frac{1}{T}\sum_{t=1}^{T}\mathbf{h}\left(y_t;\boldsymbol{\theta}\right) \;=\; \begin{bmatrix} c_1 - k_1\left(\boldsymbol{\theta}\right) \\ \vdots \\ c_{N-1} - k_{N-1}\left(\boldsymbol{\theta}\right) \\ \bar{y}_1 - \mu_1\left(\boldsymbol{\theta}\right) \\ \vdots \\ \bar{y}_N - \mu_N\left(\boldsymbol{\theta}\right) \end{bmatrix} \tag{8}$$

If K is the dimension of ϕ (the number of unknown parameters in the income density), then there are $2N-1$ moment conditions and $N+K-1$ unknown parameters.

For computational purposes, it is typically more convenient to express $k_i\left(\boldsymbol{\theta}\right)$ and $\mu_i\left(\boldsymbol{\theta}\right)$ in terms of distribution functions. If $\mu = E(y) = \int_0^\infty y f(y;\phi)dy$ is the mean of y, $F(y;\phi)$ is the distribution function, and

$$F_1(y;\phi) = \frac{\int_0^y t f(t;\phi)dt}{\mu} \tag{9}$$

is the first moment distribution function, then, from (3), (4) and (9),

$$k_i(\boldsymbol{\theta}) = F(z_i;\phi) - F(z_{i-1};\phi) \tag{10}$$

and

$$\mu_i(\boldsymbol{\theta}) = \frac{\mu\left(F_1(z_i;\phi) - F_1(z_{i-1};\phi)\right)}{F(z_i;\phi) - F(z_{i-1};\phi)} \tag{11}$$

In Section 3 we give explicit expressions for μ, $F(z_i;\phi)$ and $F_1(z_i;\phi)$ for the GB2 distribution, and its special cases, the beta-2, Singh-Maddala, Dagum, generalized gamma and lognormal distributions. Inserting these expressions into the moment conditions in (7) and (8), and including expressions for the elements of the weighting matrix that we consider in the next section, makes the GMM estimator operational.

2.2 The weighting matrix

The simplest weighting matrix is that where $\mathbf{W} = \mathbf{I}$. However, since the last N moment conditions involving the class means are of a much higher order of magnitude than the first $N-1$, which involve proportions, setting $\mathbf{W} = \mathbf{I}$ gives an undesirably large relative weight to the last N conditions. Under these circumstances, the last N conditions tend to dominate the estimation procedure and, as noted by CGR, this can lead to perverse outcomes where $\hat{z}_{i-1} > \hat{z}_i$ for some i. To ensure both sets of moment conditions were on a similar scale, CGR used a diagonal weighting matrix, with diagonal elements $(c_1^{-2}, c_2^{-2}, \ldots, c_{N-1}^{-2}, \bar{y}_1^{-2}, \bar{y}_2^{-2}, \ldots, \bar{y}_N^{-2})$. For future reference, we refer to this matrix as $\mathbf{W}_{CGR}$. The motivation behind this weight matrix was that it led to an estimator that minimized the sum of squares of percentage errors in the moment conditions. It is not an optimal weight matrix, however. Its diagonal elements are not equal to the inverses of the variances of the moment conditions, and it ignores correlations between moment conditions. Deriving an optimal weight matrix is crucial for deriving an asymptotically efficient estimator and facilitates derivation of standard errors for the parameter estimates.

The optimal weight matrix is given by:

$$\mathbf{W} = \left[\operatorname{plim} \frac{1}{T} \sum_{t=1}^{T} \mathbf{h}\left(y_t, \boldsymbol{\theta}\right) \mathbf{h}\left(y_t, \boldsymbol{\theta}\right)' \right]^{-1} \tag{12}$$

where $\mathbf{W}^{-1}$ is traditionally estimated from

$$\hat{\mathbf{W}}^{-1} = \frac{1}{T} \sum_{t=1}^{T} \mathbf{h}\left(y_t, \hat{\boldsymbol{\theta}}\right) \mathbf{h}\left(y_t, \hat{\boldsymbol{\theta}}\right)' \tag{13}$$

with $\hat{\boldsymbol{\theta}}$ being a first-step estimator obtained by minimizing $\mathbf{H}(\theta)'\mathbf{W}\mathbf{H}(\theta)$ for a pre-specified $\mathbf{W}$. The estimator $\hat{\mathbf{W}}$ depends on both the sample data and estimates of the parameters $\hat{\boldsymbol{\theta}}$. It turns out not to be suitable for our problem because it contains terms of the form $\sum_{t=T_{i-1}+1}^{T_i} y_t^2$ which are not available from the grouped data. However, instead of using (13), we can take the probability limit in (12) and obtain a result that depends only on the unknown parameters, not on the sample data. This result is derived in Appendix A.1 and is presented in the following equations. To ease the notation, we write k_i for $k_i\left(\boldsymbol{\theta}\right)$ and μ_i for $\mu_i\left(\boldsymbol{\theta}\right)$. Then, partitioning $\mathbf{W}^{-1}$ as

$$\mathbf{W}^{-1} = \begin{bmatrix} \mathbf{W}^{11} & \mathbf{W}^{12} \\ \mathbf{W}^{21} & \mathbf{W}^{22} \end{bmatrix} \tag{14}$$

we find that $\mathbf{W}^{11}$ is an $\left(N-1\right)\times\left(N-1\right)$ matrix with diagonal elements

$$w_{ii}^{11} = k_i\left(1-k_i\right) \qquad i = 1,2,\ldots,N-1$$

and off-diagonal elements

$$w_{ij}^{11} = -k_i k_j \qquad\qquad i \neq j, (i, j) = 1, 2, \ldots, N-1$$

$\mathbf{W}^{22}$ is an $N \times N$ matrix with diagonal elements

$$w_{ii}^{22} = \frac{\mu_i^{(2)}}{k_i} - \mu_i^2 \qquad\qquad i = 1, 2, \ldots, N$$

and off-diagonal elements

$$w_{ij}^{22} = -\mu_i \mu_j \qquad\qquad i \neq j, (i, j) = 1, 2, \ldots, N$$

where

$$\begin{aligned} \mu_i^{(2)} &= \operatorname{plim}\left(\overline{y_i^2}\right) = \operatorname{plim}\left(\frac{1}{T_i}\sum\nolimits_{t=T_{i-1}+1}^{T_i} y_t^2\right) \\ &= \frac{1}{k_i}\left[\mu^{(2)}\left\{F_2(z_i;\phi) - F_2(z_{i-1};\phi)\right\}\right], \end{aligned} \tag{15}$$

$F_2(z_i;\phi) = \int_0^{z_i} y^2 f(y;\phi)dy \Big/ \mu^{(2)}$ is the second moment distribution function, and $\mu^{(2)} = E(y^2)$ is the second moment for *y*. See Appendix A.2. The elements in the off-diagonal blocks of $\mathbf{W}^{-1}$ are

$$\mathbf{W}_{(N-1)\times N}^{12} = \mathbf{W}^{21\prime} = \begin{bmatrix} \mu_1(1-k_1) & -\mu_2 k_1 & \cdots & -\mu_{N-1}k_1 & -\mu_N k_1 \\ -\mu_1 k_2 & \mu_2(1-k_2) & \cdots & -\mu_{N-1}k_2 & -\mu_N k_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ -\mu_1 k_{N-1} & -\mu_2 k_{N-1} & \cdots & \mu_{N-1}(1-k_{N-1}) & -\mu_N k_{N-1} \end{bmatrix} \tag{16}$$

The matrix $\mathbf{W}^{-1}$ defined in this way is a convenient one because it depends only on the unknown parameters and not on the sample data. In traditional GMM estimation where no distributional assumptions are made, this matrix is not available and one must resort to using the matrix $\hat{\mathbf{W}}^{-1}$ specified in (13). In our case, because we will be assuming a specific parametric distribution for $f(y;\phi)$, taking the probability limit in (12) enables us to express the result in terms of unknown parameters.

For minimizing the GMM objective function, we require $\mathbf{W}$, not $\mathbf{W}^{-1}$. For given values (or estimates) of $\boldsymbol{\theta}$, $\mathbf{W}$ can be readily found by numerically inverting $\mathbf{W}^{-1}$. However, it is useful to provide an analytical expression for $\mathbf{W}$ to improve computational efficiency and to give insights into the minimization process. Working in this direction, we begin by expressing $\mathbf{W}^{-1}$ in matrix notation. We define the following N-dimensional vectors.

$$\mathbf{k} = (k_1, k_2, \ldots, k_N)' \qquad\qquad \boldsymbol{\mu} = (\mu_1, \mu_2, \ldots, \mu_N)'$$

$$\mathbf{a} = \left(\mu_1^{(2)}/k_1, \mu_2^{(2)}/k_2, \ldots, \mu_N^{(2)}/k_N\right)'$$

The $N-1$ dimensional vectors obtained by deleting the last element in each of the above vectors will be denoted by $\mathbf{k}_{-N}$, $\boldsymbol{\mu}_{-N}$ and $\mathbf{a}_{-N}$, respectively. Also, for any vector $\mathbf{x}$, we use the notation $\mathbf{D}(\mathbf{x})$ to denote a diagonal matrix with the elements of $\mathbf{x}$ on the diagonal. Then, we can write

$$\begin{aligned}
\mathbf{W}^{-1} &= \left[\begin{array}{c:c} \mathbf{D}(\mathbf{k}_{-N}) - \mathbf{k}_{-N}\mathbf{k}'_{-N} & \begin{bmatrix}\mathbf{D}(\boldsymbol{\mu}_{-N}) & \mathbf{0}_{N-1}\end{bmatrix} - \mathbf{k}_{-N}\boldsymbol{\mu}' \\ \hdashline \begin{bmatrix}\mathbf{D}(\boldsymbol{\mu}_{-N}) \\ \mathbf{0}'_{N-1}\end{bmatrix} - \boldsymbol{\mu}\mathbf{k}'_{-N} & \mathbf{D}(\mathbf{a}) - \boldsymbol{\mu}\boldsymbol{\mu}' \end{array}\right] \\
&= \left[\begin{array}{c:c} \mathbf{D}(\mathbf{k}_{-N}) & \begin{bmatrix}\mathbf{D}(\boldsymbol{\mu}_{-N}) & \mathbf{0}_{N-1}\end{bmatrix} \\ \hdashline \begin{bmatrix}\mathbf{D}(\boldsymbol{\mu}_{-N}) \\ \mathbf{0}'_{N-1}\end{bmatrix} & \mathbf{D}(\mathbf{a}) \end{array}\right] - \begin{bmatrix}\mathbf{k}_{-N} \\ \boldsymbol{\mu}\end{bmatrix}\begin{bmatrix}\mathbf{k}'_{-N} & \boldsymbol{\mu}'\end{bmatrix}
\end{aligned} \tag{17}$$

where $\mathbf{0}_{N-1}$ is an $(N-1)$ dimensional vector of zeros.

In Appendix A.3, we show how an analytical expression for $\mathbf{W}$ can be derived from (17). The result is given in (18) where we use the notation $v_i = \mu_i^{(2)} - \mu_i^2$ to denote the variance of the i-th group, $\mathbf{v} = (v_1, v_2, \ldots, v_N)'$, and $\boldsymbol{\iota}_{N-1}$ is an $(N-1)$-dimensional vector of ones. Also, with a slight abuse of matrix notation, we use $\mathbf{D}(\mathbf{x}/\mathbf{s})$ to denote a diagonal matrix with diagonal $(x_1/s_1, x_2/s_2, \ldots, x_N/s_N)$, for any two vectors $\mathbf{x}$ and $\mathbf{s}$.

$$\mathbf{W} = \left[\begin{array}{c:c} \mathbf{D}(\mathbf{a}_{-N}/\mathbf{v}_{-N}) + (a_N/v_N)\boldsymbol{\iota}_{N-1}\boldsymbol{\iota}'_{N-1} & \left[-\mathbf{D}(\boldsymbol{\mu}_{-N}/\mathbf{v}_{-N}) \quad (\mu_N/v_N)\boldsymbol{\iota}_{N-1}\right] \\ \hdashline \begin{bmatrix} -\mathbf{D}(\boldsymbol{\mu}_{-N}/\mathbf{v}_{-N}) \\ (\mu_N/v_N)\boldsymbol{\iota}'_{N-1} \end{bmatrix} & \mathbf{D}(\mathbf{k}/\mathbf{v}) \end{array}\right] \tag{18}$$

An advantage of expressing $\mathbf{W}$ in this form is that it contains a large number of zeros in all sub-matrix blocks except the upper left. The total number of zeros is $3N^2 - 7N + 4$. It may also help explain why the CGR estimator, which uses a diagonal $\mathbf{W}$, performed well (in the sense that it led to estimates comparable to those from optimal GMM) in our empirical work.

For any particular distribution, the extra information needed to compute $\mathbf{W}$ (that was not also needed to compute the moment conditions) is the second moment $\mu^{(2)}$ and the second moment distribution function $F_2(z_i;\phi)$. Expressions for these quantities for a number of distributions are provided in Section 3.

2.3 Empirical implementation of GMM estimation

We consider three GMM estimators: the CGR estimator $\hat{\boldsymbol{\theta}}_{CGR}$ that uses the weight matrix $\mathbf{W}_{CGR}$; a two-step estimator that uses $\hat{\boldsymbol{\theta}}_{CGR}$ to compute an estimate of the optimal weight matrix, $\mathbf{W}\left(\hat{\boldsymbol{\theta}}_{CGR}\right)$; and a one-step estimator obtained by minimizing the complete objective function with respect to $\boldsymbol{\theta}$ where both the moment conditions and $\mathbf{W}$ are functions of $\boldsymbol{\theta}$. Hansen et al. (1996) refer to the one-step estimator as the “continuous-updating estimator”.

The three estimators are given by

$$\hat{\boldsymbol{\theta}}_{CGR} = \arg\min_{\boldsymbol{\theta}} \mathbf{H}(\boldsymbol{\theta})'\mathbf{W}_{CGR}\mathbf{H}(\boldsymbol{\theta}) \tag{19}$$

$$\hat{\boldsymbol{\theta}}_{2-STEP} = \arg\min_{\boldsymbol{\theta}} \mathbf{H}(\boldsymbol{\theta})'\mathbf{W}\left(\hat{\boldsymbol{\theta}}_{CGR}\right)\mathbf{H}(\boldsymbol{\theta}) \tag{20}$$

$$\hat{\boldsymbol{\theta}}_{1-STEP} = \arg\min_{\boldsymbol{\theta}} \mathbf{H}(\boldsymbol{\theta})'\mathbf{W}(\boldsymbol{\theta})\mathbf{H}(\boldsymbol{\theta}) \tag{21}$$

2.4 Asymptotic covariance for the GMM estimator

To specify the asymptotic covariance matrices for the estimators, we first define the matrix of partial derivatives of the moment equations with respect to the parameters as

$$\mathbf{G}_{(2N-1)\times(N+K-1)} = \frac{\partial \mathbf{H}}{\partial \boldsymbol{\theta}'} \tag{22}$$

Then, an estimator for the asymptotic covariance matrix of the one and two-step estimators is given by (see, for example, Cameron and Trivedi, 2005 p. 176)

$$\mathbf{var}\left(\hat{\boldsymbol{\theta}}\right) = \frac{1}{T}\left(\mathbf{G}(\hat{\boldsymbol{\theta}})'\mathbf{W}(\hat{\boldsymbol{\theta}})\mathbf{G}(\hat{\boldsymbol{\theta}})\right)^{-1} \tag{23}$$

An estimator for the asymptotic covariance matrix of $\hat{\boldsymbol{\theta}}_{CGR}$ is given by

$$\mathbf{var}\left(\hat{\boldsymbol{\theta}}_{CGR}\right) = \frac{1}{T}\left(\hat{\mathbf{G}}'\mathbf{W}_{CGR}\hat{\mathbf{G}}\right)^{-1}\hat{\mathbf{G}}'\mathbf{W}_{CGR}\hat{\mathbf{W}}^{-1}\mathbf{W}_{CGR}\hat{\mathbf{G}}\left(\hat{\mathbf{G}}'\mathbf{W}_{CGR}\hat{\mathbf{G}}\right)^{-1} \tag{24}$$

where $\hat{\mathbf{G}}$ and $\hat{\mathbf{W}}^{-1}$ are equal to $\mathbf{G}(\hat{\boldsymbol{\theta}}_{CGR})$ and $\mathbf{W}^{-1}(\hat{\boldsymbol{\theta}}_{CGR})$, respectively.

In our empirical work we focused on standard errors for the one and two-step estimators and functions of them. We successfully used both analytical and numerical derivatives to calculate the elements of the $\mathbf{G}$ matrix. Some expressions for obtaining analytical derivatives for a variety of distributions are given in Appendix B. This appendix is best consulted after we consider specific income distributions in Section 3.

2.5 The J statistic

Under the null hypothesis that the moment conditions are correct $\left(\text{plim}\,\mathbf{H}(\boldsymbol{\theta})=\mathbf{0}\right)$, the *J* statistic

$$J = T\,\mathbf{H}\left(\hat{\boldsymbol{\theta}}\right)'\mathbf{W}\left(\hat{\boldsymbol{\theta}}\right)\mathbf{H}\left(\hat{\boldsymbol{\theta}}\right) \xrightarrow{d} \chi^2_{N-K} \tag{25}$$

In traditional GMM estimation this test statistic is used to assess whether excess moment conditions are valid. In our case, since we assume a particular form of parametric income distribution, and use its properties to construct the moment conditions and weight matrix, the *J* statistic can be used to test the validity of the assumed income distribution.

3. Income Distributions

A large number of probability density functions has been suggested in the literature for modelling income distributions. See, for example, McDonald and Ransom (1979), McDonald (1984), McDonald and Xu (1995), Creedy and Martin (1997), Bandourian et al. (2003) and Kleiber and Kotz (2003). One of the flexible distributions introduced by McDonald (1984) is the four parameter generalized beta distribution of the second kind (GB2). This distribution has analytical properties that make it well suited to the analysis of income distributions (Parker, 1999), and, as we will see, it provides a very good fit to the observed data (see also, Bordley et al. (1996) and Bandourian et al (2003)). In this section we describe the GB2 distribution and its characteristics needed for GMM estimation. We also present results needed for GMM estimation of the beta2, Dagum, Singh-Maddala, generalized gamma and lognormal distributions that can be obtained as special cases of the GB2 distribution.

3.1 *The generalized beta income distribution of the second kind*

The GB2 distribution whose parameters are $\phi = (a, b, p, q)'$ has probability density function (pdf)

$$f(y; a, b, p, q) = \frac{a y^{ap-1}}{b^{ap} B(p, q) \left(1 + \left(\frac{y}{b}\right)^{a}\right)^{p+q}} \qquad y > 0 \tag{26}$$

where y is income, $b > 0, p > 0, q > 0,\ a > 0$ and

$$B(p, q) \quad = \quad \frac{\Gamma(p)\ \Gamma(q)}{\Gamma(p+q)} \quad = \quad \int_0^1 t^{p-1}(1-t)^{q-1} dt$$

is the beta function. The GB2 distribution is a generalization of the standard beta distribution defined on the (0,1) interval. If U is a standard beta random variable with parameters (p, q), then $Y = b\left[U/(1-U)\right]^{1/a} \sim GB2(a, b, p, q)$. The inverse of this transformation is

$$U = \frac{\left(Y/b\right)^{a}}{1 + \left(Y/b\right)^{a}}$$

Thus, the cumulative distribution function (cdf) of the GB2 distribution is given by

$$F(y; a, b, p, q) = \frac{1}{B(p, q)} \int_0^u t^{p-1}(1-t)^{q-1} dt = B_u(p, q) \tag{27}$$

where $u = \left(y/b\right)^{a} \Big/ \left[1 + \left(y/b\right)^{a}\right]$ and the function $B_u(p, q)$ is the cdf for the standard beta distribution. Expressing the GB2 cdf in this form is convenient because $B_u(p, q)$ is readily computed by computer software.

The quantities required to compute the moment conditions and the weighting matrix for GMM estimation are the first and second moments μ and $\mu^{(2)}$, the distribution function

$F(y;\phi)$, and the first and second moment distribution functions $F_1(y;\phi)$ and $F_2(y;\phi)$. The moments of order k exist only if $-ap < k < aq$, and, if a moment exists, it is given by (Kleiber and Kotz, 2003, p.188)

$$\mu^{(k)} = \frac{b^k B\left(p + k/a\,,\, q - k/a\right)}{B\left(p,q\right)} \tag{28}$$

Also, the k-th moment distribution function $F_k(y;\phi) = \int_0^y t^k f(t;\phi)dt \Big/ \mu^{(k)}$ can be written as (Kleiber and Kotz, 2003, p.192)

$$F_k(y;\phi) = F\left(y;a,b,p+\frac{k}{a},q-\frac{k}{a}\right) = B_u\left(p+\frac{k}{a},q-\frac{k}{a}\right) \tag{29}$$

This expression is a computationally convenient one, because it allows us to compute the first and second moment distribution functions using a standard beta distribution function with different parameters.

Income distributions are often estimated to assess inequality and poverty. To illustrate, we consider two inequality measures, the Gini and Theil coefficients, and two poverty measures, the headcount ratio and the Foster-Greer-Thorbecke measure with an inequality aversion parameter of 2 ($FGT(2)$) (Foster, Greer and Thorbecke, 1984). General forms for the Gini and Theil coefficients, and their expressions in terms of the parameters of the GB2 distribution, are (McDonald, 1984; McDonald and Ransom, 2008)

$$\begin{aligned} G &= -1 + \frac{2}{\mu}\int_0^\infty y\,F(y;\phi)\,f(y;\phi)\,dy \\ &= \frac{2B\left(2p+1/a,\ 2q-1/a\right)}{B\left(p,q\right)B\left(p+1/a,\ q-1/a\right)}\left\{\frac{1}{p}\,{}_3F_2\left[1,p+q,2p+\frac{1}{a};\ p+1,2\left(p+q\right);\ 1\right]\right. \\ &\qquad \left. -\frac{1}{p+1/a}\,{}_3F_2\left[1,p+q,2p+\frac{1}{a};\ p+\frac{1}{a}+1,2\left(p+q\right);\ 1\right]\right\} \end{aligned} \tag{30}$$

$$\begin{aligned} T &= \int_0^\infty \left(\frac{y}{\mu}\right)\ln\left(\frac{y}{\mu}\right) f(y;\phi)\,dy \\ &= \frac{1}{a}\left[\psi\left(p+1/a\right)-\psi\left(q-1/a\right)\right]+\ln\left(b/\mu\right) \end{aligned} \tag{31}$$

where ${}_3F_2$ is a generalized hypergeometric function and $\psi(t) = d\log\Gamma(t)/dt$ is the digamma function. The hypergeometric function can be computed by Matlab, but we found it was more efficient and reliable to numerically integrate $G = -1+\left(2/\mu\right)\int_0^\infty y\,F(y;\phi)\,f(y;\phi)\,dy$.

For a given poverty line *x*, the headcount ratio is the proportion of population with incomes less than *x*, and so, for the GB2 distribution, it is simply given by

$$H_x = F(x;\phi) = B_u(p,q) \qquad \text{where} \qquad u = \frac{\left(x/b\right)^a}{1+\left(x/b\right)^a} \tag{32}$$

The measure $FGT_x(2)$ considers not just the proportion of poor, but also how far the poor are below the poverty line. It is defined as

$$FGT_x(2) = \int_0^x \left(\frac{x-y}{x}\right)^2 f(y)dy \tag{33}$$

Following Kakwani (1999), it is convenient to express $FGT_x(2)$ in terms of the mean and variance of the poor (μ_x and σ_x^2), the headcount ratio H_x , and the income gap ratio $g_x = \left(x-\mu_x\right)/x$. Definitions of these quantities, and expressions for them in terms of the parameters of the GB2 distribution and some of its special cases are given in Table 1.

3.2 Some commonly used distributions

A number of popular distributions are special cases of the GB2 distribution. Borrowing from McDonald (1984), we display those distributions in Figure 1. In our work we estimated the GB2, beta-2, Singh-Maddala, Dagum, generalized gamma and lognormal distributions. In the remainder of this section we provide or refer to the quantities μ, $\mu^{(2)}$, $F(y;\phi)$, $F_1(y;\phi)$ and $F_2(y;\phi)$, that are required for GMM estimation of these special-case distributions. For the inequality and poverty measures, we confined estimation to the GB2, beta-2, Singh-Maddala, and Dagum distributions. Expressions for the inequality and poverty measures in terms of the parameters of these distributions are given in Table 1.

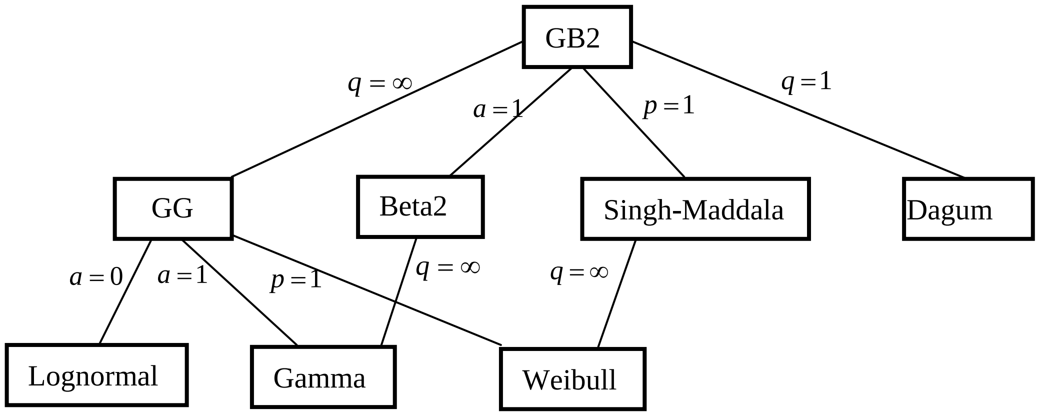

Figure1- Relationship of GB2 with other distributions (adopted from McDonald (1984))

The required first and second moments and distribution functions for the beta-2, Singh-Maddala, and Dagum distributions are readily obtained from the more general GB2 expressions in (28) and (29) by setting $a=1$, $p=1$ and $q=1$, respectively. Simplifications for the means of these distributions are given in Table 1. For the beta-2 distribution it is also worth noting that the second moment reduces to

$$\mu^{(2)} = \frac{b^2 p(p+1)}{(q-1)(q-2)}$$

There are also useful simplifications for the cdf's for the Singh-Maddala and Dagum distributions. For the Singh-Maddala distribution we have

$$F(y;a,b,1,q)=1-\left[1+\left(\frac{y}{b}\right)^{a}\right]^{-q}$$

and, for the Dagum distribution,

$$F(y;a,b,p,1)=1-\left[1+\left(\frac{y}{b}\right)^{-a}\right]^{-p}$$

3.3 Generalized gamma distribution

The generalized gamma pdf is given by

$$f(y;a,p,\beta)=\frac{a}{\beta^{ap}\Gamma(p)}y^{ap-1}\exp\left(-\left(\frac{y}{\beta}\right)^{a}\right) \tag{34}$$

It is obtained as a special case of the GB2 distribution in (26) by setting $b=\beta q^{1/a}$ and taking the limit as $q\to\infty$ (McDonald 1984). The standard gamma pdf

$$f(u)=\frac{1}{b^{p}\Gamma(p)}u^{p-1}\exp\left(-\frac{u}{b}\right) \tag{35}$$

can be obtained from (34) using the transformation $u=y^{a}$, and redefining *b* as $b=\beta^{a}$. Thus, values for the cdf of the generalized gamma distribution can be computed from

$$F(y;a,p,\beta)=G_{u}(p,b) \qquad \text{with } u=y^{a} \text{ and } b=\beta^{a}$$

where $G_{u}(p,b)=\int_{0}^{u/b}t^{p-1}e^{-t}dt\Big/\Gamma(p)$ is the cdf of the standard gamma distribution with parameters *p* and *b*.

The moments and moment distribution functions for the generalized gamma are given by (McDonald, 1984; Butler and McDonald, 1989)

$$\mu^{(k)} = \frac{\beta^k \Gamma\left(p + k/a\right)}{\Gamma(p)} \tag{36}$$

$$F_k(y;a,p,\beta) = F\left(y;a,p+k/a,\beta\right) = G_u\left(p+k/a,b\right) \tag{37}$$

These expressions complete what is needed for computing the moment conditions and the weight matrix.

3.4 Lognormal distribution

The lognormal pdf

$$f\left(y;\mu,\sigma^2\right) = \frac{1}{y\sqrt{2\pi}\sigma}\exp\left(-\frac{(\ln y - \mu)^2}{2\sigma^2}\right) \tag{38}$$

can be obtained as a special case of the generalized gamma distribution by setting $\beta^a = \sigma^2 a^2$ and $p = (a\mu + 1)/\beta^a$, and taking the limit as $a \to 0$ (McDonald, 1984). Its cdf is

$$F\left(y;\mu,\sigma^2\right) = \Phi\left(\frac{\ln(y) - \mu}{\sigma}\right) \tag{39}$$

where Φ is the standard normal cdf. Its moments are

$$\mu^{(k)} = \exp\left(k\mu + \frac{k^2\sigma^2}{2}\right)$$

and values for its moment distribution functions can be computed from

$$F_k\left(y;\mu,\sigma^2\right) = F\left(y;\mu + k\sigma^2,\sigma^2\right) = \Phi\left(\frac{\ln(y) - \mu - k\sigma^2}{\sigma}\right) \tag{40}$$

Details can be found in Aitchison and Brown (1957) or Keliber and Kotz (2003).

4. Description of data and sources

To illustrate the methodology described in Sections 2 and 3, we use income distribution data from the PovcalNet website developed by the World Bank poverty research group. This database is set up for the purpose of poverty assessment for individual countries, regions and globally. The data are provided in grouped form and can be downloaded from http://go.worldbank.org/WE8P1I8250. They are available for developing countries for a number of years ranging from 1981 to 2005. The latest version of the data was updated in August 2008 to incorporate 2005 purchasing power parity estimated by the World Bank International Comparison Program. To use a reasonably diverse cross section of countries to test the performance of the estimator, we chose as examples Brazil, China, India, Pakistan, Russia and Poland for the year 2005. Separate data are available for rural and urban regions in India and China, making a total 8 different data sets. We will refer to each data set as coming from a region, where a region can be a country, or rural or urban China or India.

For most of the chosen regions, population shares c_i and the corresponding income shares g_i were available for 20 groups. Exceptions were India rural and urban which each had 12 groups, and China rural which had 17 groups. In line with India rural and urban, we aggregated the data from the other regions into 12 groups. Having 12 regions for all regions has the advantage of uniformity for estimation, and it provides an opportunity for checking the ability of the estimated model to predict income shares for groups other than those used for estimation, a procedure that we consider in Section 5. The population proportions in each region were not identical, but in most cases they were approximately 0.05 for the first and last two groups and 0.1 for the remaining groups. China rural was an exception where there was a much smaller proportion in the earlier groups.

Also available from the World Bank website is each region's mean income $\bar{y}$, found from surveys and then converted using a 2005 purchasing-power-parity exchange rate. To use the methodology described in Sections 2 and 3, we need the data on class mean incomes $\bar{y}_i$. They are obtained as $\bar{y}_i = g_i\,\bar{y}/c_i$. For computing standard errors and the J statistic, we also need the sample sizes T for each of the surveys. Unfortunately, although the website provides data on the population size of each region, it does not have comprehensive data on the sample sizes T for each of the surveys. For our calculations we use $T = 20{,}000$. This is a conservative value since most of surveys have sample sizes which are much larger. If standard errors or *J*-statistics for other sample sizes are of interest, they can be obtained from our results by multiplying by the appropriate scaling factor.

5. Empirical analysis

Our presentation and discussion of the results begins in Section 5.1 with consideration of the estimated income distributions for the eight regions. Goodness-of-fit of the distributions is assessed in Section 5.2, using *J*-statistics and a comparison of predicted and observed income shares. Levels of inequality and poverty obtained from different distributions are reported in Section 5.3.

5.1 *Country-specific income distributions*

Table 2 contains the estimated class limits and parameters of the GB2 distribution obtained using the GMM estimation procedure outlined in Sections 2 and 3. For each region, we report three different sets of estimates – those from the CGR, the two-step, and the one-step estimations. Standard errors for both the two-step and one-step estimates are also reported. There are no dramatic differences between the estimates from the three different estimators. The two-step and the one-step estimates are almost identical, and the CGR estimates are only slightly different. Similarly, the standard errors obtained using two-step estimation are very close to those obtained from one-step estimation. The magnitudes of the standard errors for the class limits are very small suggesting we are estimating their values precisely. However, standard errors for some of the estimates $\left(\hat{a},\hat{b},\hat{p},\hat{q}\right)$ are relatively large, implying wide confidence intervals around the corresponding parameters. In most regions, separate hypothesis tests for $H_0 : p = 1$ (Singh-Maddala), $H_0 : q = 1$ (Dagum), and $H_0 : a = 1$ (beta-2) would not be rejected. The situation may change, of course, if we use a larger sample size, but one of the 3-parameter distributions may be an adequate representation for some cases. More light is shed on this issue when we examine goodness of fit.

To save space we have not reported estimates and standard errors for distributions other than the GB2; they are available from the authors on request. Estimates of the z_i and their standard errors were similar for all distributions. Standard errors for the estimated parameters of the beta-2, Singh-Maddala and Dagum distributions (which have parameters in common with the GB2) were much smaller than those for the GB2, reflecting the drop from 4 to 3 parameters.

In all cases we computed standard errors using both numerical derivatives and analytical derivatives, and where both sets were computable, they produced identical results. There were a few cases where the computation of analytical derivatives failed – beta-2 estimates for Pakistan and India (rural and urban). These failures corresponded to solutions where *p* was very large and *b* was very small; estimation was unstable, with different starting values leading to different local minima. Analytical standard errors could not be found because the hypergeometric function in Matlab broke down. Numerical standard errors could still be found, however. This problem did not arise with the GB2 and other distributions.

Figure 2 contains graphs of the estimated GB2 pdfs for China rural and urban, India rural and urban, Brazil and Poland, along with 95% confidence bounds for these distributions. To find the confidence bounds, standard errors were computed for the estimated pdfs at a number of income levels using the covariance matrix of the parameter estimates, the delta rule, and numerical derivatives. The narrowness of the confidence bounds suggests we are accurately estimating the pdfs, despite relatively large confidence intervals for some of the parameters. A comparison of the urban and rural pdfs for India and China shows clearly the larger incomes of the urban populations. In India, it is interesting that the rural and urban modes are similar, but the urban pdf has a much fatter tail. Using a population weighted mixture of the rural and urban components, in Figure 3 we have graphed the pdf and cdf for all of China, alongside those of the rural and urban subpopulations. If more countries are considered, similar mixtures can be obtained for larger regions such as continents or the whole world. See, for example, Chotikapanich et al. (2012).

A potential estimation problem for all distributions other than the generalized gamma and lognormal, is the non-existence of the second moment. For the existence of the *k*-th moment, the GB2 distribution requires $aq > k$. This condition is the same for the moments of the Singh-Maddala distribution, it becomes $q > k$ for the beta-2 distribution and $a > k$ for the Dagum distribution. Since the optimal weighting matrix requires the existence of second order moments, if the CGR estimates violate one of these inequalities, we cannot proceed with two-step estimation of the offending distribution. Also, our experience suggests one-step estimation breaks down. (CGR estimation is still feasible.) We encountered this problem with Brazil, a country with relatively high inequality, for estimations with the GB2, Singh-Maddala and Dagum distributions, but not the beta-2 distribution. In the results reported in Table 2 we overcame the problem by minimizing the objective function subject to the constraint $aq > 2$. This solution may not be entirely satisfactory. The underlying income distribution may indeed not have second moments, the standard errors for the boundary solutions that result may not be valid, and inequality appeared to be underestimated relative to values reported by the World Bank.

We also found that the generalized gamma distribution can be difficult to estimate. Sometimes estimation would break down, particularly when there was a tendency for the estimate for *a* to become small. We suspect that small values of *a* were making calculation of $\Gamma\left(p+1/a\right)$ troublesome. We tried different parameterizations and different starting values, and in all cases managed to get convergence. However, we are not confident that all our solutions correspond to global minima.

5.2 *Goodness-of-fit analysis*

In this section we assess the adequacy of the various distributions for modelling the observed population and income shares (or income shares converted to class mean incomes). Two criteria are used: (i) the *J* test to test whether the moment conditions are valid for each of the distributions, and (ii) a comparison of observed and predicted income shares.

Table 3 presents the *p*-values for the *J* statistics calculated for all distributions considered and for all example regions. Under the null hypothesis that the moment conditions are correct, the *J* statistic has a χ^2 distribution with degrees of freedom equal to the number of excess moment conditions. In the case of the GB2 distribution, we have 23 moment conditions and 15 parameters giving degrees of freedom of 8. For the 3-parameter distributions the degrees of freedom is 9, and for the log normal it is 10. The very large *p*-values for the GB2 distribution show that its moment conditions are compatible with the sample moment conditions. The smaller, although still large, *p*-value for Brazil could be attributable to minimizing a restricted version of the objective function. Results for the other distributions suggest the beta-2 distribution is adequate for all regions except India rural and Pakistan – two of the three regions where estimation was unstable with large $\hat{p}$ and small $\hat{b}$. The Singh-Maddala distribution is unsuitable for China rural, India urban, Russia, Poland and Brazil. The same is true for the Dagum distribution for China rural, Russia and Brazil, and to a lesser extent Poland. At a 5% level of significance, the generalized gamma distribution is rejected for all regions except China urban, Russia and Poland; the lognormal is rejected for all regions except Russia. Based on these results, we conclude that the GB2 distribution is both the best-fitting and an adequate model, the beta-2 and Dagum distributions show some promise, and the Singh-Maddala, generalized gamma and lognormal are inadequate in most cases.

We also assess the goodness-of-fit of the distributions by comparing the observed income shares g_i with the predicted income shares derived from the estimated distributions. The distributions were estimated using 12 groups, obtained by aggregating 20 original groups in all regions except India rural and urban and China rural. No aggregation was carried out on the original 12 groups available for India rural and urban, and, in the case of China rural, 17 groups were aggregated to 12. To assess goodness-of-fit, we examined the ability of the models to predict the income shares in the original groups (20 in most cases) from the distributions estimated from 12 groups.

The income shares were predicted in the following way. Beginning with the original population shares c_i, and corresponding cumulative proportions $\pi_i = \sum_{j=1}^{i} c_j$, we found class limits z_i (not necessarily the same as the previously-estimated class limits) by solving the equations

$$F\left(z_i;\hat{\phi}\right) = \pi_i$$

Then, predicted cumulative income shares $\hat{\eta}_i$ were found from the first moment distribution function

$$\hat{\eta}_i = F_1\left(z_i;\hat{\phi}\right)$$

giving the predicted income shares $\hat{g}_i = \hat{\eta}_i - \hat{\eta}_{i-1}$.

Note that when the number of groups used for estimation differs from the number used for predicting the income shares, the class limits (z_i) in the above two equations will, by necessity, be different from the estimated z_i. When the same number of groups is used for estimation and prediction, we have two alternatives for predicting the income shares. We can use the above two equations as already described, or we can simply use $\hat{\eta}_i = F_1\left(\hat{z}_i;\hat{\phi}\right)$ where $\hat{z}_i$ are the original estimates of the class limits. We used the former approach in all cases. Since it uses less information from GMM minimization, it is likely to be a more stringent test of predictive ability.

We present a comparison of the predicted and actual income shares (in percentage form) for the GB2 distribution in Table 4. Table 5 contains the root-mean-squared errors, $\sqrt{N^{-1}\sum_{i=1}^{N}\left[100\left(\hat{g}_i - g_i\right)\right]^2}$, for all distributions. In Table 4 the observed and predicted income shares are remarkably similar for all regions, giving strong support for the GB2 distribution. This outcome is very encouraging given that the parameters of the distributions have been estimated from limited data, the predictions are partially "out-of-sample" for most countries, and the class limits z_i implied by the estimated parameters, not the z_i giving the "best fit", were used to compute the predicted income proportions.

In Table 5, the GB2 distribution performs the best in five out of the eight cases; the exceptions were India rural, Russia and Brazil where the preferred distributions (Singh-Maddala in the case of India rural and beta-2 in the other cases) were marginally better than GB2. With the exception of Russia, the lognormal and generalized gamma distributions performed very poorly in all cases. A possibly counterintuitive result is that the lognormal distribution outperformed the generalized gamma distribution. Since the lognormal can be viewed as a 2-parameter special case of the 3-parameter generalized gamma, we would expect the generalized gamma to do better. The problem with the generalized gamma seemed to lie in predicting the share of the last group. If this group was omitted, the predictions from the lognormal were worse. We speculated earlier that, with the generalized gamma, we may not have always reached a global minimum. That could be the reason for poor prediction of the last share.

5.3 *Inequality and poverty*

In this subsection we illustrate how the parameter estimates can be used to estimate inequality and poverty. The Gini and Theil coefficients were calculated using the expressions given in (30) and (31) for the GB2 distributions and using those in Table 1 for the beta-2, Singh-Maddala and Dagum distributions. Standard errors were computed numerically using the delta rule and the covariance matrix of the parameter estimates. Table 6 reports the estimated Gini and Theil coefficients and their corresponding standard errors. It is found that GB2 and beta-2 give similar results for the Gini and Theil coefficients while Singh-Maddala and Dagum give slightly different results. In terms of the standard errors, those from beta-2 seem to be the smallest. However, in all cases the standard errors are relatively small compared to the estimated coefficients. Inequality is highest in Brazil followed by India urban; it is lowest in India rural.

Table 7 reports poverty incidence using the headcount ratio (HCR) and the FGT(2) measure, both expressed as percentages, using a poverty line of \$1.25 per day, or, in the monthly units used in estimation, \$38. Values are calculated from the expressions in Table 1 for the beta-2, Singh-Maddala, Dagum and GB2 distributions. The corresponding standard errors are also reported. Poverty is greatest in rural and urban India, followed by rural China and Pakistan, then Brazil. There is much less poverty in urban China, Russia and Poland. The estimates can be sensitive to the chosen distribution, particularly when we are in the tail of the distribution where the level of poverty is low; see, for example, Russian and Poland.

6. **Summary and conclusions**

Estimation of income distributions is critical for monitoring inequality and poverty at both national and international levels. Studies which attempt to estimate the global income distribution taking into account both within-country and between-country inequality typically utilize data provided in aggregated form by either the World Institute for Development Economics Research (WIDER) or the World Bank. See, for example, Milanovic (2002) and Chotikapanich et al (2012). Previous work by Chotikapanich et al (2012) used a method-of-moments estimator to estimate beta-2 income distributions from this data. In this paper we have extended their work by providing moment conditions and the optimal weight matrix that can be used for GMM estimation of any class of income distributions. Specific expressions for the moment conditions and the optimal weight matrix were provided for the more general GB2 distribution, its obvious special cases the beta-2, Dagum and Singh-Maddala distributions, and its less obvious special cases, the generalized gamma and lognormal distributions. We also show how to get standard errors for the optimal GMM estimates. Once the parameters have been estimated, along with the covariance matrix of the estimator, they can be used in a variety of ways. Values for the density function, distribution function and

Lorenz curve and their confidence bounds can be found at a number of income values and then graphed. Distributions for larger regions can be obtained as population weighted mixtures of individual countries. Inequality and poverty measures and their standard errors can also be computed. We have illustrated the methodology and how a number of economically relevant quantities can be estimated from it, using data on 6 selected countries that included 8 different regions. We found that the methodology can be readily implemented and that the GB2 distribution provides a good fit in terms of the validity of its moment conditions, and the accuracy of predicted income shares from the estimated distributions.

Acknowledgements

This project has been supported by Australian Research Council Grant DP1094632.

Appendix 1 Optimal Weighting Matrix

A.1 Finding $\mathbf{W}^{-1}$ as a probability limit

We require

$$\mathbf{W}^{-1} = \text{plim}\left[\frac{1}{T}\sum_{t=1}^{T}\mathbf{h}\left(y_t,\boldsymbol{\theta}\right)\mathbf{h}\left(y_t,\boldsymbol{\theta}\right)'\right] = \text{plim}\begin{pmatrix} \mathbf{P}_{(N-1)\times(N-1)} & \mathbf{Q}_{(N-1)\times N} \\ \mathbf{Q}'_{N\times(N-1)} & \mathbf{M}_{N\times N} \end{pmatrix}$$

It is straightforward to show that the elements of $\mathbf{P}$, $\mathbf{Q}$ and $\mathbf{M}$ are

$$p_{ii} = \frac{1}{T}\sum_{t=1}^{T}\left[g_i(y_t)-k_i\right]^2 = \frac{1}{T}\sum_{t=1}^{T}\left[g_i(y_t)^2 + k_i^2 - 2g_i(y_t)k_i\right] = c_i + k_i^2 - 2c_ik_i$$

$$p_{ij} = \frac{1}{T}\sum_{t=1}^{T}\left[g_i(y_t)-k_i\right]\left[g_j(y_t)-k_j\right] = -c_ik_j - c_jk_i + k_ik_j \qquad \left(i \neq j\right)$$

$$q_{ii} = \frac{1}{T}\sum_{t=1}^{T}\left[g_i(y_t)-k_i\right]\left[\frac{y_t}{c_i}g_i(y_t)-\mu_i\right] = \bar{y}_i - c_i\mu_i - k_i\,\bar{y}_i + k_i\mu_i$$

$$q_{ij} = \frac{1}{T}\sum_{t=1}^{T}\left[g_i(y_t)-k_i\right]\left[\frac{y_t}{c_j}g_j(y_t)-\mu_j\right] = -c_i\mu_j - k_i\,\bar{y}_j + k_i\mu_j \qquad \left(i \neq j\right)$$

$$m_{ii}=\frac{1}{T}\sum_{t=1}^{T}\left[\frac{1}{c_i}g_i(y_t)y_t-\mu_i\right]^2=\frac{\sum_{t=T_{i-1}+1}^{T_i}y_t^2}{T_ic_i}+\mu_i^2-2\bar{y}_i\mu_i$$

$$m_{ij}=\frac{1}{T}\sum_{t=1}^{T}\left[\frac{y_tg_i(y_t)}{c_i}-\mu_i\right]\left[\frac{y_tg_j(y_t)}{c_j}-\mu_j\right]=-\bar{y}_i\mu_j-\mu_i\bar{y}_j+\mu_i\mu_j \qquad (i\neq j)$$

Now, let $\overline{y_i^2}=\frac{1}{T_i}\sum_{t=T_{i-1}+1}^{T_i}y_t^2$ be the average of the squared observations in each group, and let $\mu_i^{(2)}(\theta)=\mu_i^{(2)}=\text{plim}\left(\overline{y_i^2}\right)$. Using this result and also that $\text{plim}\,c_i=k_i$ and $\text{plim}\,\bar{y}_i=\mu_i$, we have for the elements of the matrix $\mathbf{W}^{-1}$ (see equation (17)),

$$\text{plim}\,p_{ii}=k_i\left(1-k_i\right)$$

$$\text{plim}\,p_{ij}=-k_ik_j$$

$$\text{plim}\,q_{ii}=\mu_i\left(1-k_i\right)$$

$$\text{plim}\,q_{ij}=-k_i\mu_j$$

$$\text{plim}\,m_{ii}=\mu_i^{(2)}/k_i-\mu_i^2$$

$$\text{plim}\,m_{ij}=-\mu_i\mu_j$$

A.2 Second moment for the i-th group

It is convenient to write $\mu_i^{(2)}$ in terms of the second moment distribution function

$$F_2(y;\phi)=\frac{\int_0^y t^2f(t;\phi)dt}{E(y^2)}$$

Working in this direction, we have

$$\mu_i^{(2)} = \text{plim}\left(\overline{y_i^2}\right) = \text{plim}\left(\frac{1}{T_i}\sum_{t=T_{i-1}+1}^{T_i} y_t^2\right)$$

$$= \text{plim}\left(\frac{T}{T_i}\right)\text{plim}\left(\frac{1}{T}\sum_{t=1}^{T} y_t^2 g_i(y_t)\right)$$

$$= \frac{1}{k_i}\int_0^\infty y^2 g_i(y) f(y;\phi)dy$$

$$= \frac{1}{k_i}\int_{z_{i-1}}^{z_i} y^2 f(y;\phi)dy$$

$$= \frac{1}{k_i}\left[\mu^{(2)}\left\{F_2(z_i;\phi) - F_2(z_{i-1};\phi)\right\}\right]$$

where $\mu^{(2)} = E(y^2)$.

A.3 Deriving $\mathbf{W}$ *from* $\mathbf{W}^{-1}$

With obvious definitions for $\mathbf{A}$ and $\mathbf{c}$, we write $\mathbf{W}^{-1}$ as

$$\mathbf{W}^{-1} = \begin{bmatrix} \mathbf{D}(\mathbf{k}_{-N}) & \begin{bmatrix}\mathbf{D}(\boldsymbol{\mu}_{-N}) & \mathbf{0}_{N-1}\end{bmatrix} \\ \begin{bmatrix}\mathbf{D}(\boldsymbol{\mu}_{-N}) \\ \mathbf{0}'_{N-1}\end{bmatrix} & \mathbf{D}(\mathbf{a}) \end{bmatrix} - \begin{bmatrix}\mathbf{k}_{-N} \\ \boldsymbol{\mu}\end{bmatrix}\begin{bmatrix}\mathbf{k}'_{-N} & \boldsymbol{\mu}'\end{bmatrix} = \mathbf{A} - \mathbf{cc}'$$

To invert this matrix, we use the result

$$\left(\mathbf{A} - \mathbf{cc}'\right)^{-1} = \mathbf{A}^{-1} + \frac{1}{1-\mathbf{c}'\mathbf{A}^{-1}\mathbf{c}}\mathbf{A}^{-1}\mathbf{cc}'\mathbf{A}^{-1}$$

First, we need to obtain $\mathbf{A}^{-1}$ which we partition as

$$\mathbf{A}^{-1} = \begin{bmatrix}\mathbf{A}^{11} & \mathbf{A}^{12} \\ \mathbf{A}^{21} & \mathbf{A}^{22}\end{bmatrix}$$

Using results on the partitioned inverse of a matrix, we have

$$\mathbf{A}^{11} = \left[\mathbf{D}\left(\mathbf{k}_{-N}\right) - \left[\mathbf{D}\left(\boldsymbol{\mu}_{-N}\right) \quad \mathbf{0}_{N-1}\right]\mathbf{D}\left(\mathbf{a}\right)^{-1}\begin{bmatrix}\mathbf{D}\left(\boldsymbol{\mu}_{-N}\right) \\ \mathbf{0}'_{N-1}\end{bmatrix}\right]^{-1}$$

This is a diagonal matrix of dimension $\left(N-1\right)$ with diagonal elements

$$\left(k_i - \frac{\mu_i^2}{a_i}\right)^{-1} = \left(\frac{\mu_i^{(2)} - \mu_i^2}{\mu_i^{(2)}/k_i}\right)^{-1} = \frac{a_i}{v_i}$$

Thus, we have $\mathbf{A}^{11} = \mathbf{D}\left(\mathbf{a}_{-N}/\mathbf{v}_{-N}\right)$

Also,

$$\mathbf{A}^{22} = \left[\mathbf{D}(\mathbf{a}) - \begin{bmatrix}\mathbf{D}(\boldsymbol{\mu}_{-N}) \\ \mathbf{0}'_{N-1}\end{bmatrix}\mathbf{D}(\mathbf{k}_{-N})^{-1} \quad \begin{bmatrix}\mathbf{D}(\boldsymbol{\mu}_{-N}) & \mathbf{0}_{N-1}\end{bmatrix}\right]^{-1} = \begin{bmatrix}\mathbf{D}(\mathbf{k}_{-N}/\mathbf{v}_{-N}) & \mathbf{0}_{N-1} \\ \mathbf{0}'_{N-1} & k_N/\mu_N^{(2)}\end{bmatrix}$$

and

$$\mathbf{A}^{12} = \mathbf{A}^{21\prime} = -\mathbf{D}(\mathbf{k}_{-N})^{-1}\begin{bmatrix}\mathbf{D}(\boldsymbol{\mu}_{-N}) & \mathbf{0}_{N-1}\end{bmatrix}\begin{bmatrix}\mathbf{D}(\mathbf{k}_{-N}/\mathbf{v}_{-N}) & \mathbf{0}_{N-1} \\ \mathbf{0}'_{N-1} & k_N/\mu_N^{(2)}\end{bmatrix} = \begin{bmatrix}-\mathbf{D}(\boldsymbol{\mu}_{-N}/\mathbf{v}_{-N}) & \mathbf{0}_{N-1}\end{bmatrix}$$

Then,

$$\mathbf{A}^{-1}\mathbf{c} = \begin{bmatrix}\mathbf{D}(\mathbf{a}_{-N}/\mathbf{v}_{-N}) & -\mathbf{D}(\boldsymbol{\mu}_{-N}/\mathbf{v}_{-N}) & \mathbf{0}_{N-1} \\ -\mathbf{D}(\boldsymbol{\mu}_{-N}/\mathbf{v}_{-N}) & \mathbf{D}(\mathbf{k}_{-N}/\mathbf{v}_{-N}) & \mathbf{0}_{N-1} \\ \mathbf{0}'_{N-1} & \mathbf{0}'_{N-1} & k_N/\mu_N^{(2)}\end{bmatrix}\begin{bmatrix}\mathbf{k}_{-N} \\ \boldsymbol{\mu}_{-N} \\ \mu_N\end{bmatrix}$$

$$= \begin{bmatrix}\boldsymbol{\iota}_{N-1} \\ \mathbf{0}_{N-1} \\ k_N\mu_N/\mu_N^{(2)}\end{bmatrix}$$

and

$$1 - \mathbf{c}'\mathbf{A}^{-1}\mathbf{c} = 1 - \sum_{i=1}^{N-1} k_i - \frac{k_N\mu_N^2}{\mu_N^{(2)}} = \frac{v_N}{a_N}$$

Thus,

$$\frac{\mathbf{A}^{-1}\mathbf{c}\mathbf{c}'\mathbf{A}^{-1}}{1 - \mathbf{c}'\mathbf{A}^{-1}\mathbf{c}} = \frac{a_N}{v_N}\begin{bmatrix}\boldsymbol{\iota}_{N-1}\boldsymbol{\iota}'_{N-1} & \mathbf{0} & \frac{k_N\mu_N}{\mu_N^{(2)}}\boldsymbol{\iota}_{N-1} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \frac{k_N\mu_N}{\mu_N^{(2)}}\boldsymbol{\iota}'_{N-1} & \mathbf{0} & \left[\frac{k_N\mu_N}{\mu_N^{(2)}}\right]^2\end{bmatrix} = \begin{bmatrix}\boldsymbol{\iota}_{N-1}\boldsymbol{\iota}'_{N-1}\frac{a_N}{v_N} & \mathbf{0} & \frac{\mu_N}{v_N}\boldsymbol{\iota}_{N-1} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \frac{\mu_N}{v_N}\boldsymbol{\iota}'_{N-1} & \mathbf{0} & \frac{k_N\mu_N^2}{v_N}\end{bmatrix}$$

Finally,

$$\mathbf{W} = \mathbf{A}^{-1} + \frac{\mathbf{1}}{\mathbf{1 - c'A^{-1}c}} \mathbf{A}^{-1}\mathbf{cc'A}^{-1}$$

$$= \begin{bmatrix} \mathbf{D}(\mathbf{a}_{-N}/\mathbf{v}_{-N}) & -\mathbf{D}(\boldsymbol{\mu}_{-N}/\mathbf{v}_{-N}) & \mathbf{0}_{N-1} \\ -\mathbf{D}(\boldsymbol{\mu}_{-N}/\mathbf{v}_{-N}) & \mathbf{D}(\mathbf{k}_{-N}/\mathbf{v}_{-N}) & \mathbf{0}_{N-1} \\ \mathbf{0}'_{N-1} & \mathbf{0}'_{N-1} & k_N/\mu_N^{(2)} \end{bmatrix} + \begin{bmatrix} \boldsymbol{\iota}_{N-1}\boldsymbol{\iota}'_{N-1}\dfrac{a_N}{v_N} & \mathbf{0} & \dfrac{\mu_N}{v_N}\boldsymbol{\iota}_{N-1} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \dfrac{\mu_N}{v_N}\boldsymbol{\iota}'_{N-1} & \mathbf{0} & \dfrac{k_N\mu_N^2}{v_N} \end{bmatrix}$$

$$= \begin{bmatrix} \mathbf{D}(\mathbf{a}_{-N}/\mathbf{v}_{-N}) + (a_N/v_N)\boldsymbol{\iota}_{N-1}\boldsymbol{\iota}'_{N-1} & -\mathbf{D}(\boldsymbol{\mu}_{-N}/\mathbf{v}_{-N}) & (\mu_N/v_N)\boldsymbol{\iota}_{N-1} \\ -\mathbf{D}(\boldsymbol{\mu}_{-N}/\mathbf{v}_{-N}) & \mathbf{D}(\mathbf{k}_{-N}/\mathbf{v}_{-N}) & \mathbf{0} \\ (\mu_N/v_N)\boldsymbol{\iota}'_{N-1} & \mathbf{0} & k_N/v_N \end{bmatrix}$$

Appendix B Analytic expressions for derivatives of moment conditions

To find the asymptotic covariance matrix of the estimators we need $\mathbf{G}$, the matrix of derivatives of the moment conditions with respect to all the parameters. We can calculate the elements of $\mathbf{G}$ using numerical derivatives. However, it is also possible to calculate them analytically. To do so we first note that this matrix has the following structure:

$$\mathbf{G} = \begin{bmatrix} \dfrac{\partial \mathbf{H}^k}{\partial \mathbf{z}'} & \dfrac{\partial \mathbf{H}^k}{\partial \boldsymbol{\phi}'} \\ \dfrac{\partial \mathbf{H}^\mu}{\partial \mathbf{z}'} & \dfrac{\partial \mathbf{H}^\mu}{\partial \boldsymbol{\phi}'} \end{bmatrix}$$

where $\mathbf{H}$ is partitioned as $\mathbf{H}' = \left(\mathbf{H}^{k\prime}, \mathbf{H}^{\mu\prime}\right)$ with $\mathbf{H}^k$ denoting the moment conditions for the class proportions and $\mathbf{H}^\mu$ denoting the moment conditions for the class means. Also, we partition $\boldsymbol{\theta}$ as $\boldsymbol{\theta}' = (\mathbf{z}', \boldsymbol{\phi}')$ where $\mathbf{z}' = (z_1, z_2, \ldots, z_{N-1})$ and $\boldsymbol{\phi}$ is the vector of parameters in the income distribution. The elements in $\mathbf{H}^k$ and $\mathbf{H}^\mu$ for which we require derivatives are, respectively, $k_i(\boldsymbol{\theta}) = F(z_i;\boldsymbol{\phi}) - F(z_{i-1};\boldsymbol{\phi})$, and

$$\mu_i(\boldsymbol{\theta}) = \frac{\mu\left(F_1(z_i;\boldsymbol{\phi}) - F_1(z_{i-1};\boldsymbol{\phi})\right)}{F(z_i;\boldsymbol{\phi}) - F(z_{i-1};\boldsymbol{\phi})}$$

In this appendix, we focus on the derivatives of $F(z_i;\boldsymbol{\phi})$ and $F_1(z_i;\boldsymbol{\phi})$ with respect to z_i and the elements in $\boldsymbol{\phi}$. Finding the derivatives of μ and combining these derivatives with those of $F(z_i;\boldsymbol{\phi})$ and $F_1(z_i;\boldsymbol{\phi})$ to find the required derivatives of $\mu_i\left(\boldsymbol{\theta}\right)$ is straightforward, although tedious. The basic tool used to find expressions for the derivatives of $F(z_i;\boldsymbol{\phi})$ and $F_1(z_i;\boldsymbol{\phi})$ is the following standard result from calculus:

$$\frac{d}{d\theta}\int_{u_1(\theta)}^{u_2(\theta)} f(x,\theta)dx = \int_{u_2(\theta)}^{u_2(\theta)} \frac{\partial f(x,\theta)}{\partial\theta}dx + f(u_2,\theta)\frac{du_2}{d\theta} - f(u_1,\theta)\frac{du_1}{d\theta}$$

B.1 Derivatives for the GB2 distribution

Let $u_i = \left(z_i/b\right)^a \Big/ \left[1+\left(z_i/b\right)^a\right] = z_i^a \Big/ \left(b^a + z_i^a\right)$. Derivatives for $B_{u_i}\left(p+k/a, q-k/a\right)$ are provided. Setting $k=0$ gives the required expressions for $F(z_i;\boldsymbol{\phi})$; setting $k=1$ will give the required expressions for $F_1(z_i;\boldsymbol{\phi})$. Derivatives for the beta-2, Singh-Maddala and Dagum distributions, can be obtained by setting $a=1$, $p=1$ and $q=1$, respectively.

$$\frac{\partial\left(B_{u_i}\left(p+k/a, q-k/a\right)\right)}{\partial z_i} = \frac{\partial}{\partial z_i}\int_0^{u_i} \frac{x^{p+k/a-1}(1-x)^{q-k/a-1}}{B(p,q)}dx = \frac{\left(u_i\right)^{p+k/a-1}\left(1-u_i\right)^{q-k/a-1}}{B\left(p+k/a, q-k/a\right)}\frac{\partial u_i}{\partial z_i}$$

$$= \frac{a z_i^{ap+k-1} b^{aq-k}}{\left(b^a + z_i^a\right)^{p+q} B\left(p+k/a, q-k/a\right)}$$

$$\frac{\partial\left(B_{u_i}\left(p+k/a, q-k/a\right)\right)}{\partial z_j} = 0 \qquad i \neq j$$

$$\frac{\partial\left(B_{u_i}\left(p+k/a, q-k/a\right)\right)}{\partial b} = \frac{\left(u_i\right)^{p+k/a-1}\left(1-u_i\right)^{q-k/a-1}}{B\left(p+k/a, q-k/a\right)}\frac{\partial u_i}{\partial b} = \frac{-a z_i^{ap+k} b^{aq-k-1}}{\left(b^a + z_i^a\right)^{p+q} B\left(p+k/a, q-k/a\right)}$$

$$\frac{\partial\left(B_{u_i}\left(p+k/a,q-k/a\right)\right)}{\partial p}=\left[\psi(p+q)-\psi\left(p+\frac{k}{a}\right)\right]B_{u_i}\left(p+\frac{k}{a},q-\frac{k}{a}\right)$$

$$+\int_0^{u_i}\frac{x^{p+k/a-1}(1-x)^{q-k/a-1}}{B(p+k/a,q-k/a)}\ln(x)dx$$

$$=\left[\psi(p+q)-\psi(p+k/a)+\ln\left(u_i\right)\right]B_{u_i}\left(p+k/a,q-k/a\right)$$

$$-\frac{x^{p+k/a}}{\left(p+k/a\right)^2 B\left(p+k/a,q-k/a\right)}$$

$$\times\, {}_3F_2\left(p+k/a,p+k/a,1+k/a-q;p+k/a+1,p+k/a+1;u_i\right)$$

$$\frac{\partial\left(B_{u_i}\left(p+k/a,q-k/a\right)\right)}{\partial q}=\left[\psi(p+q)-\psi\left(q-\frac{k}{a}\right)\right]B_{u_i}\left(p+\frac{k}{a},q-\frac{k}{a}\right)$$

$$+\int_0^{u_i}\frac{x^{p+k/a-1}(1-x)^{q-k/a-1}}{B(p+k/a,q-k/a)}\ln(1-x)dx$$

$$=-\left[\psi(p+q)-\psi(q-k/a)+\ln\left(1-u_i\right)\right]B_{u_i}\left(p+k/a,q-k/a\right)$$

$$+\frac{(1-x)^{q-k/a}}{\left(q-k/a\right)^2 B\left(p+k/a,q-k/a\right)}$$

$$\times\, {}_3F_2\left(q-k/a,q-k/a,1-k/a-p;q-k/a+1,q-k/a+1;1-u_i\right)$$

$$\frac{\partial\left(B_{u_i}\left(p+k/a,q-k/a\right)\right)}{\partial a}=\frac{k}{a^2}\left(-\frac{\partial}{\partial p}\left(B_{u_i}\left(p+\frac{k}{a},q-\frac{k}{a}\right)\right)+\frac{\partial}{\partial q}\left(B_{u_i}\left(p+\frac{k}{a},q-\frac{k}{a}\right)\right)\right)$$

$$+\frac{a_i^{ap+k}b^{aq-k}(\ln a_i-\ln b)}{\left(b^a+a_i^a\right)^{p+q}B\left(p+k/a,q-k/a\right)}$$

In the derivatives with respect to p and q, ψ is the derivative of the log of the gamma function and ${}_3F_2$ represents the generalized hypergeometric function.

B.2 Derivatives for the generalized gamma distribution

For the generalized gamma distribution, we need the derivatives of $G_{u_i}\left(p+k/a,b\right)$ for $k=0$ and $k=1$ where $u_i=z_i^a$. They are

$$\frac{\partial G_{u_i}\left(p+k/a,q\right)}{\partial z_i}=\frac{\partial}{\partial z_i}\int_0^{u_i}\frac{x^{p+k/a-1}\exp\left(-x/b\right)}{b^{p+k/a}\Gamma\left(p+k/a\right)}dx=\frac{z_i^{a(p+k/a-1)}\exp\left(-u_i/b\right)}{b^{p+k/a}\Gamma\left(p+k/a\right)}\frac{\partial u_i}{\partial z_i}$$

$$=\frac{az_i^{ap+k/a-1}\exp\left(-u_i/b\right)}{b^{p+k/a}\Gamma\left(p+k/a\right)}$$

$$\frac{\partial G_{u_i}\left(p+k/a,q\right)}{\partial z_j}=0 \qquad (i\neq j)$$

$$\frac{\partial G_{u_i}\left(p+k/a,b\right)}{\partial b}=\int_0^{u_i}\frac{\partial}{\partial b}\left(\frac{x^{p+k/a-1}\exp\left(-x/b\right)}{b^{p+k/a}\Gamma\left(p+k/a\right)}\right)dx=-\frac{z_i^{ap+k}\exp\left(-u_i/b\right)}{b^{p+1+k/a}\Gamma\left(p+k/a\right)}$$

$$\frac{\partial G_{u_i}\left(p+k/a,b\right)}{\partial p}=\int_0^{u_i}\ln(x)\frac{x^{p+k/a-1}\exp\left(-x/b\right)}{b^{p+k/a}\Gamma(p+k/a)}dx-\left(\ln(b)+\frac{\psi(p+k/a)}{\Gamma(p+k/a)}\right)G_{u_i}\left(p+\frac{k}{a},b\right)$$

$$\frac{\partial G_{u_i}\left(p+k/a,b\right)}{\partial a}=\frac{z_i^{a(p+k/a)}\exp\left(-u_i/b\right)\ln z_i}{b^{p+k/a}\Gamma(p+k/a)}-\frac{k}{a^2}\frac{\partial G_{u_i}(p+k/a,q)}{\partial p}$$

The integral in the derivative with respect to p can be evaluated numerically.

B.3 Derivatives for the lognormal distribution

For the lognormal distribution, we need the derivatives of $\Phi\left(\left[\ln(z_i)-\mu-k\sigma^2\right]/\sigma\right)$ for $k=0$ and $k=1$. They are

$$\frac{\partial}{\partial z_i}\Phi\left(\frac{\ln(z_i)-\mu-k\sigma^2}{\sigma}\right)=\frac{1}{\sigma z_i}\phi\left(\frac{\ln z_i-\mu-k\sigma^2}{\sigma}\right)$$

$$\frac{\partial}{\partial z_j}\Phi\left(\frac{\ln(z_i)-\mu-k\sigma^2}{\sigma}\right)=0 \qquad \text{for } i\neq j$$

$$\frac{\partial}{\partial \mu}\Phi\left(\frac{\ln\left(z_i\right)-\mu-k\sigma^2}{\sigma}\right)=\frac{-1}{\sigma}\phi\left(\frac{\ln z_i-\mu-k\sigma^2}{\sigma}\right)$$

$$\frac{\partial}{\partial\sigma}\Phi\left(\frac{\ln(z_i)-\mu-k\sigma^2}{\sigma}\right)=-\frac{\ln(z)-\mu+k\sigma^2}{\sigma^2}\phi\left(\frac{\ln(z)-\mu-k\sigma^2}{\sigma}\right)$$

where $\phi(.)$ denotes the standard normal pdf.

References

Aitchison, J. and Brown, J.A.C. (1957), *The Lognormal Distribution*, Cambridge University Press, Cambridge.

Bandourian, R., McDonald, J. B., and Turley, R. S. (2003), “Income Distributions: An Inter-temporal Comparison over Countries," *Estadistica*, 55, 135-152.

Bordley, R.F., McDonald, J.B., and A. Mantrala (1996), “Something New, Something Old: Parametric Models for the Size Distribution of Income,” *Journal of Income Distribution*, 6, 91 - 103.

Butler, R.J. and McDonald, J.B. (1989), “Using Incomplete Moments to Measure Inequality”, *Journal of Econometrics*, 42, 109-119.

Cameron, A. C. and P.K. Trivedi (2005), *Microeconometrics: Method and Application*, Cambridge University Press, New York.

Chotikapanich, D., editor (2008), *Modeling Income Distributions and Lorenz Curves*, Springer, New York.

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

Chotikapanich, D., W.E. Griffiths, D. S. Prasada Rao, and V. Valencia (2012), “Global Income Distributions and Inequality, 1993 and 2000: Incorporating Country-Level

Inequality Modelled with Beta Distributions," *The Review of Economics and Statistics*, forthcoming.

Creedy, J., and Martin, V. L. (eds.) (1997), *Nonlinear Economic Models: Cross-sectional, Time Series and Neural Network Applications*, Cheltenham: Edward Elgar.

Foster, J., J. Greer, and E. Thorbecke (1984), "A Class of Decomposable Poverty Measures," *Econometrica*, 52, 761-766.

Hansen, L.P., J. Heaton and A. Yaron (1996), "Finite-Sample Properties of Some Alternative GMM Estimators", *Journal of Business and Economic Statistics* 14(3), 262-280.

Kakwani, N. (1999), "Inequality, Welfare and Poverty: Three Interrelated Phenomena," in J. Silber (ed), *Handbook on Income Inequality Measurement*, Kluwer: The Netherlands, 599-628.

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

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

McDonald, J. B. and Ransom, M. R. (1979), "Functional Forms, Estimation Techniques and the Distribution of Income," *Econometrica*, 6, 1513 – 1525.

McDonald, James. B. and Michael R. Ransom (2008), "The Generalized Beta Distribution as a Model for the Distribution of Income: Estimation of Related Measures of Inequality," (pp. 147-166), in Duangkamon Chotikapanich (Ed.), *Modeling Income Distributions and Lorenz Curves* (New York: Springer).

McDonald, J. B. and Xu, Y. J. (1995), "A Generalization of the Beta Distribution with Applications," *Journal of Econometrics* 66, 133-152. Errata 69 (1995), 427-428.

Milanovic, B. (2002), “True World Income Distribution, 1988 and 1993: First Calculations based on Household Surveys Alone," *The Economic Journal*, 112, 51-92.

Parker, S. C. (1999), “The Generalized Beta as a Model for the Distribution of Earnings,” *Economics Letters*, 62, 197-200.

Wu, X. and Perloff, J. (2005), “China’s Income Distribution: 1985-2001”, *The Review of Economics and Statistics*, 87, 763-775.

Wu, X. and Perloff, J. (2007), “GMM Estimation of a Maximum Entropy Distribution with Interval Data”, *Journal of Econometrics*, 138, 532-546.

Table 1: Expressions for the Inequality and Poverty Measures

	General form	GB2	Beta-2 ($a=1$)	Singh-Maddala ($p=1$)	Dagum ($q=1$)
The mean μ	$\mu=\int_0^\infty y\,f(y)dy$	$\mu=b\frac{\Gamma(p+1/a)\Gamma(q-1/a)}{\Gamma(p)\Gamma(q)}$	$\mu=\frac{bp}{(q-1)}$	$\mu=\frac{b\Gamma(1/a)\Gamma(q-1/a)}{a\Gamma(q)}$	$\mu=\frac{b\Gamma(1/a)\Gamma(p+1/a)}{a\Gamma(p)}$
The Gini	$G=-1+\frac{2}{\mu}\int_0^\infty yF(y)\,f(y)dy$	Equation (30)	$G=\frac{2B(2p,2q-1)}{pB^2(p,q)}$	$G=1-\frac{\Gamma(q)\Gamma(2q-1/a)}{\Gamma(q-1/a)\Gamma(2q)}$	$G=\frac{\Gamma(p)\Gamma(2p+1/a)}{\Gamma(p+1/a)\Gamma(2p)}-1$
Theil index	$\int_0^\infty\left(\frac{y}{\mu}\right)\ln\left(\frac{y}{\mu}\right)f(y)dy$	$T=\frac{1}{a}\left[\psi(p+1/a)-\psi(q-1/a)\right]$ $+\ln(b/\mu)$	$T=\left[\psi(p+1)-\psi(q-1)\right]$ $+\ln(b/\mu)$	$T=\frac{1}{a}\left[\psi(1+1/a)-\psi(q-1/a)\right]$ $+\ln(b/\mu)$	$T=\frac{1}{a}\left[\psi(p+1/a)-\psi(1-1/a)\right]$ $+\ln(b/\mu)$
Head count ratio	$H_x=F(x)$	$H_x=B_u(p,q)$ where $u=\frac{(x/b)^a}{1+(x/b)^a}$	$H_x=B_u(p,q)$ where $u=\frac{x/b}{1+x/b}$	$H_x=B_u(p,q)$ where $u=\frac{(x/b)^a}{1+(x/b)^a}$	$H_x=B_u(p,q)$ where $u=\frac{(x/b)^a}{1+(x/b)^a}$
Mean income of the poor	$\mu_x=\frac{\int_0^x y\,f(y)dy}{F(x)}$	$\mu_x=\mu\frac{B_u(p+1/a,\,q-1/a)}{B_u(p,q)}$	$\mu_x=\mu\frac{B_u(p+1,q-1)}{B_u(p,q)}$	$\mu_x=\mu\frac{B_u(1+1/a,\,q-1/a)}{B_u(p,q)}$	$\mu_x=\mu\frac{B_u(p+1/a,\,1-1/a)}{B_u(p,q)}$
Variance of the income of the poor	$\sigma_x^2=\mu_x^{(2)}-\mu_x^2$ where $\mu_x^{(2)}$ is the second moment	$\mu_x^{(2)}=\frac{b^2B_u(p+2/a,\,q-2/a)}{H_x}\times$ $\frac{\Gamma(p+2/a)\Gamma(q-2/a)}{\Gamma(p)\Gamma(q)}$	$\mu_x^{(2)}=\frac{\mu b(p+1)}{H_x(q-2)}\times$ $B_u(p+2,q-2)$	$\mu_x^{(2)}=\frac{b^2B_u(1+2/a,q-2/a)}{H_x}\times$ $\frac{\Gamma(1+2/a)\Gamma(q-2/a)}{\Gamma(q)}$	$\mu_x^{(2)}=\frac{b^2B_u(p+2/a,1-2/a)}{H_x}\times$ $\frac{\Gamma(p+2/a)\Gamma(1-2/a)}{\Gamma(p)}$
$FGT_x(2)$	$\int_0^x\left(\frac{x-y}{x}\right)^2 f(y)dy$	$FGT_x(2)=H_x\left[g_x^2+(1-g_x)^2\frac{\sigma_x^2}{\mu_x^2}\right]$ where $g_x=\frac{x-\mu_x}{x}$ is known as the income gap ratio and x is the poverty line.			

Table 2: Estimated Coefficients from GB2 Distributions

	China Rural					China Urban				
	CGR	Two-Stage	SE	One-Stage	SE	CGR	Two-Stage	SE	One-Stage	SE
z_1	11.367	10.891	0.887	10.905	0.885	53.020	52.877	0.807	52.870	0.807
z_2	13.777	13.612	0.705	13.616	0.704	65.901	65.840	0.687	65.835	0.687
z_3	16.747	16.650	0.538	16.650	0.537	84.935	84.964	0.765	84.959	0.765
z_4	22.626	22.563	0.350	22.560	0.350	101.457	101.586	0.783	101.581	0.783
z_5	33.617	33.572	0.308	33.570	0.308	117.928	118.142	0.742	118.136	0.742
z_6	41.763	41.776	0.298	41.774	0.298	135.872	136.121	0.733	136.115	0.733
z_7	55.340	55.455	0.280	55.455	0.280	156.932	157.187	0.907	157.182	0.907
z_8	83.104	83.514	0.689	83.516	0.689	183.845	184.071	1.368	184.066	1.368
z_9	111.511	112.271	1.330	112.272	1.330	223.373	223.485	2.205	223.477	2.205
z_{10}	126.329	127.313	1.699	127.314	1.699	298.630	299.124	3.885	299.113	3.884
z_{11}	140.582	141.847	2.091	141.849	2.091	387.675	389.653	6.333	389.673	6.333
b	27.743	22.735	10.878	22.738	10.886	109.276	108.431	9.842	108.440	9.839
p	4.918	6.900	4.945	6.903	4.949	1.930	2.100	0.751	2.099	0.751
q	1.884	2.264	0.841	2.267	0.844	1.439	1.561	0.451	1.561	0.451
a	1.565	1.370	0.358	1.369	0.358	2.102	1.988	0.399	1.988	0.399
	India Rural					**India Urban**				
	CGR	**Two-Stage**	**SE**	**One-Stage**	**SE**	**CGR**	**Two-Stage**	**SE**	**One-Stage**	**SE**
z_1	20.971	20.910	0.347	20.907	0.347	19.729	19.637	0.321	19.635	0.321
z_2	24.340	24.274	0.279	24.271	0.279	23.479	23.385	0.250	23.382	0.250
z_3	28.770	28.736	0.285	28.734	0.285	29.128	29.100	0.254	29.097	0.254
z_4	32.686	32.707	0.263	32.705	0.263	34.664	34.699	0.270	34.697	0.270
z_5	36.457	36.519	0.224	36.519	0.224	40.174	40.267	0.269	40.266	0.269
z_6	40.254	40.347	0.212	40.346	0.212	46.570	46.694	0.264	46.692	0.264
z_7	44.988	45.095	0.264	45.095	0.264	54.726	54.826	0.296	54.825	0.296
z_8	51.394	51.495	0.390	51.496	0.390	64.941	64.974	0.427	64.974	0.427
z_9	61.431	61.491	0.583	61.490	0.582	81.217	81.181	0.750	81.181	0.750
z_{10}	80.736	80.756	0.917	80.751	0.917	112.682	112.584	1.503	112.581	1.503
z_{11}	105.900	106.302	1.621	106.304	1.621	153.378	153.802	2.737	153.803	2.736
b	30.296	29.744	2.942	29.745	2.943	6.776	6.761	14.811	6.760	14.812
p	1.417	1.580	0.666	1.581	0.667	21.052	21.058	48.051	21.058	48.043
q	0.609	0.661	0.172	0.662	0.173	2.366	2.387	1.178	2.389	1.180
a	4.244	3.963	0.877	3.960	0.876	1.201	1.194	0.444	1.193	0.444

Table 2: Estimated Coefficients from GB2 Distributions (cont.)

	Pakistan					Russia				
	CGR	Two-Stage	SE	One-Stage	SE	CGR	Two-Stage	SE	One-Stage	SE
z_1	26.652	26.689	0.429	26.689	0.429	79.750	79.910	1.348	79.908	1.348
z_2	30.975	30.975	0.349	30.975	0.349	101.364	101.428	1.122	101.427	1.122
z_3	37.077	37.044	0.358	37.043	0.358	135.210	135.094	1.226	135.094	1.226
z_4	42.285	42.224	0.342	42.224	0.342	166.344	166.169	1.328	166.167	1.328
z_5	47.476	47.407	0.302	47.406	0.302	198.668	198.505	1.329	198.502	1.329
z_6	53.161	53.105	0.281	53.104	0.281	234.775	234.660	1.317	234.657	1.317
z_7	59.917	59.909	0.342	59.908	0.342	277.852	277.703	1.535	277.699	1.535
z_8	68.735	68.803	0.510	68.802	0.510	333.446	333.146	2.296	333.138	2.296
z_9	82.008	82.104	0.782	82.102	0.782	414.316	414.016	3.903	414.003	3.902
z_{10}	108.371	108.395	1.284	108.392	1.284	564.702	564.324	7.318	564.298	7.318
z_{11}	141.530	141.576	2.171	141.575	2.171	735.808	733.461	11.722	733.447	11.721
b	37.832	37.689	4.570	37.688	4.570	170.347	171.246	49.656	171.226	49.660
p	1.904	1.903	0.838	1.903	0.838	6.152	5.643	4.261	5.643	4.261
q	0.813	0.802	0.211	0.803	0.211	4.619	4.201	2.622	4.201	2.622
a	3.297	3.320	0.719	3.320	0.719	0.952	1.004	0.375	1.004	0.375
	Poland					**Brazil**				
	CGR	**Two-Stage**	**SE**	**One-Stage**	**SE**	**CGR**	**Two-Stage**	**SE**	**One-Stage**	**SE**
z_1	95.853	95.527	1.591	95.517	1.591	30.022	30.289	0.564	30.318	0.564
z_2	117.507	117.131	1.285	117.120	1.284	44.736	44.816	0.554	44.820	0.560
z_3	150.310	150.100	1.353	150.089	1.353	70.527	70.698	0.745	70.631	0.763
z_4	179.794	179.867	1.424	179.857	1.424	96.635	96.729	0.851	96.612	0.859
z_5	209.978	210.315	1.390	210.305	1.390	125.954	125.798	0.880	125.655	0.879
z_6	243.370	243.905	1.352	243.893	1.352	160.347	159.434	1.035	159.290	1.052
z_7	282.956	283.603	1.552	283.591	1.552	204.023	201.667	1.532	201.568	1.598
z_8	333.865	334.264	2.285	334.256	2.285	264.539	261.845	2.439	261.900	2.551
z_9	407.922	407.971	3.814	407.963	3.814	361.476	360.107	3.745	360.607	3.829
z_{10}	546.846	547.417	7.068	547.397	7.068	563.605	569.661	9.003	571.962	9.200
z_{11}	708.788	712.313	11.450	712.337	11.450	826.852	845.195	35.083	853.705	38.098
b	148.476	140.202	54.804	140.218	54.801	179.415	160.600	14.908	159.153	15.532
p	4.761	5.428	4.092	5.428	4.092	1.935	1.489	0.386	1.514	0.419
q	2.574	2.826	1.403	2.827	1.403	2.148	1.501	0.481	1.512	0.520
a	1.373	1.290	0.425	1.290	0.425	1.097	1.335	0.243	1.323	0.257

Table 3: *p*-Values from *J*-Statistics

	GB2	B2	SM	Dagum	GGamma	LogN
China Rural	0.9998	0.9972	0.0000	0.0111	0.0035	0.0000
China Urban	0.9931	0.6083	0.4344	0.8222	0.8798	0.0009
India Rural	0.9955	0.0000	0.9696	0.9625	0.0000	0.0000
India Urban	0.9985	0.9994	0.0000	0.3073	0.0000	0.0000
Pakistan	0.9998	0.0080	0.9449	0.9991	0.0000	0.0000
Russia	0.9998	0.9999	0.0005	0.0017	0.5375	0.6798
Poland	0.9978	0.9977	0.0024	0.1515	0.1830	0.0016
Brazil	0.7866	0.3908	0.0060	0.0000	0.0039	0.0000

Table 4: Observed and Estimated Percentage Shares of Income based on GB2

China Rural		China Urban		India Rural		India Urban	
Estimated	Observed	Estimated	Observed	Estimated	Observed	Estimated	Observed
0.013	0.014	1.541	1.553	1.702	1.709	1.325	1.332
0.053	0.054	2.080	2.052	2.319	2.310	1.771	1.788
0.172	0.174	2.441	2.416	5.270	5.240	4.140	4.115
1.010	1.002	2.675	2.665	6.453	6.416	5.264	5.225
2.103	2.091	2.934	2.925	7.052	7.057	5.813	5.776
3.326	3.317	3.128	3.124	7.213	7.246	6.871	6.877
2.054	2.051	3.331	3.332	8.443	8.506	8.338	8.407
4.588	4.591	3.571	3.587	9.829	9.888	9.265	9.360
5.038	5.059	3.779	3.801	11.690	11.699	11.845	11.901
7.958	8.001	4.018	4.052	13.703	13.552	15.043	15.017
12.692	12.741	4.280	4.316	9.157	8.921	10.650	10.477
10.770	10.781	4.526	4.561	17.169	17.456	19.673	19.726
8.867	8.836	4.859	4.895				
6.903	6.852	5.163	5.188				
5.689	5.624	5.641	5.663				
4.322	4.258	6.139	6.138				
24.437	24.554	6.792	6.727				
		7.754	7.632				
		9.290	9.091				
		16.058	16.280				

Pakistan		Russia		Poland		Brazil	
Estimated	Observed	Estimated	Observed	Estimated	Observed	Estimated	Observed
1.708	1.700	1.024	1.017	1.259	1.264	0.344	0.343
2.196	2.212	1.513	1.528	1.752	1.755	0.668	0.698
2.487	2.495	1.835	1.851	2.064	2.057	0.918	0.936
2.705	2.720	2.115	2.125	2.330	2.320	1.148	1.157
2.917	2.929	2.378	2.378	2.578	2.565	1.375	1.372
3.121	3.134	2.636	2.635	2.820	2.800	1.608	1.621
3.309	3.312	2.896	2.887	3.061	3.044	1.896	1.894
3.505	3.501	3.163	3.146	3.309	3.306	2.083	2.089
3.709	3.699	3.443	3.424	3.567	3.570	2.508	2.529
3.924	3.905	3.742	3.719	3.841	3.850	2.512	2.547
4.157	4.133	4.064	4.051	4.136	4.154	2.996	3.064
4.413	4.378	4.420	4.408	4.461	4.490	3.381	3.443
4.691	4.659	4.818	4.818	4.825	4.878	4.124	4.114
5.034	5.017	5.276	5.277	5.242	5.305	4.061	4.037
5.433	5.443	5.817	5.818	5.735	5.792	5.114	5.080
5.944	5.991	6.480	6.470	6.340	6.361	5.813	5.792
6.583	6.664	7.339	7.377	7.125	7.108	6.970	7.000
7.571	7.626	8.552	8.572	8.240	8.143	8.703	8.890
9.312	9.332	10.586	10.732	10.124	9.915	12.104	12.755
17.280	17.150	17.902	17.767	17.192	17.322	31.674	30.639

Table 5: Root-Mean-Square Errors

	China Rural	China Urban	India Rural	India Urban	Pakistan	Russia	Poland	Brazil
GB2	0.0424	0.0763	0.1198	0.0673	0.0412	0.0468	0.0644	0.2779
B2	0.0812	0.2546	1.0566	0.0880	0.3840	0.0460	0.1025	0.2538
SM	0.9524	0.1559	0.0930	0.8243	0.1704	0.4705	0.3401	0.3026
Dagum	0.5686	0.1418	0.2891	0.4898	0.0535	0.5632	0.3313	0.5118
GGamma	1.4376	0.8576	2.0215	1.9490	1.3633	0.5218	0.8032	1.6572
LogN	1.0282	0.6312	1.6683	1.5764	1.0951	0.2915	0.5702	0.9261

Table 6: Gini and Theil Coefficients and Their Standard Errors

	B2		SM		Dagum		GB2	
	Gini	SE	Gini	SE	Gini	SE	Gini	SE
China Rural	0.358	0.006	0.371	0.011	0.372	0.007	0.359	0.006
China Urban	0.347	0.006	0.348	0.007	0.351	0.007	0.348	0.012
India Rural	0.299	0.004	0.305	0.007	0.302	0.007	0.304	0.017
India Urban	0.377	0.007	0.385	0.010	0.384	0.008	0.376	0.007
Pakistan	0.308	0.007	0.313	0.007	0.311	0.007	0.312	0.013
Russia	0.375	0.006	0.376	0.007	0.384	0.007	0.375	0.007
Poland	0.349	0.006	0.350	0.007	0.355	0.007	0.349	0.007
Brazil	0.543	0.008	0.531	0.014	0.517	0.010	0.553	0.013
	Theil	SE	Theil	SE	Theil	SE	Theil	SE
China Rural	0.236	0.011	0.323	0.027	0.290	0.015	0.241	0.013
China Urban	0.210	0.009	0.236	0.013	0.236	0.011	0.223	0.012
India Rural	0.159	0.005	0.197	0.013	0.180	0.010	0.191	0.013
India Urban	0.271	0.014	0.335	0.026	0.311	0.017	0.275	0.017
Pakistan	0.178	0.014	0.207	0.014	0.191	0.010	0.197	0.013
Russia	0.247	0.010	0.277	0.015	0.290	0.014	0.248	0.011
Poland	0.216	0.009	0.243	0.014	0.246	0.012	0.219	0.011
Brazil	0.603	0.029	0.614	0.056	0.581	0.034	0.672	0.113

Table 7: Poverty Measures (%) and Their Standard Errors

	B2		SM		Dagum		GB2	
	estimates	se	estimates	se	estimates	se	estimates	se
China R								
HCR	24.742	0.242	23.138	3.293	24.212	0.312	24.445	0.391
FGT2	2.162	0.064	1.639	1.222	1.922	0.061	2.152	0.065
China U								
HCR	1.548	0.143	2.078	0.901	1.969	0.130	1.827	0.165
FGT2	0.074	0.014	0.197	0.134	0.169	0.021	0.130	0.030
India R								
HCR	41.786	0.102	42.819	2.860	43.158	0.535	42.999	0.558
FGT2	3.672	0.091	3.371	1.356	3.466	0.078	3.418	0.085
India U								
HCR	35.508	0.305	33.498	3.408	34.739	0.417	35.337	0.463
FGT2	3.649	0.091	3.425	1.399	3.570	0.069	3.650	0.074
Pakistan								
HCR	22.054	0.464	19.796	3.770	21.210	0.332	20.790	0.603
FGT2	1.284	0.092	1.250	0.948	1.203	0.070	1.217	0.070
Russia								
HCR	0.302	0.045	0.596	0.258	0.539	0.053	0.296	0.080
FGT2	0.017	0.004	0.066	0.038	0.055	0.008	0.017	0.009
Poland								
HCR	0.042	0.015	0.245	0.269	0.180	0.033	0.071	0.040
FGT2	0.001	0.001	0.022	0.039	0.013	0.003	0.003	0.003
Brazil								
HCR	7.530	0.071	7.226	1.163	7.054	0.071	7.411	0.111
FGT2	1.325	0.037	1.488	N/A	1.499	0.036	1.393	0.047

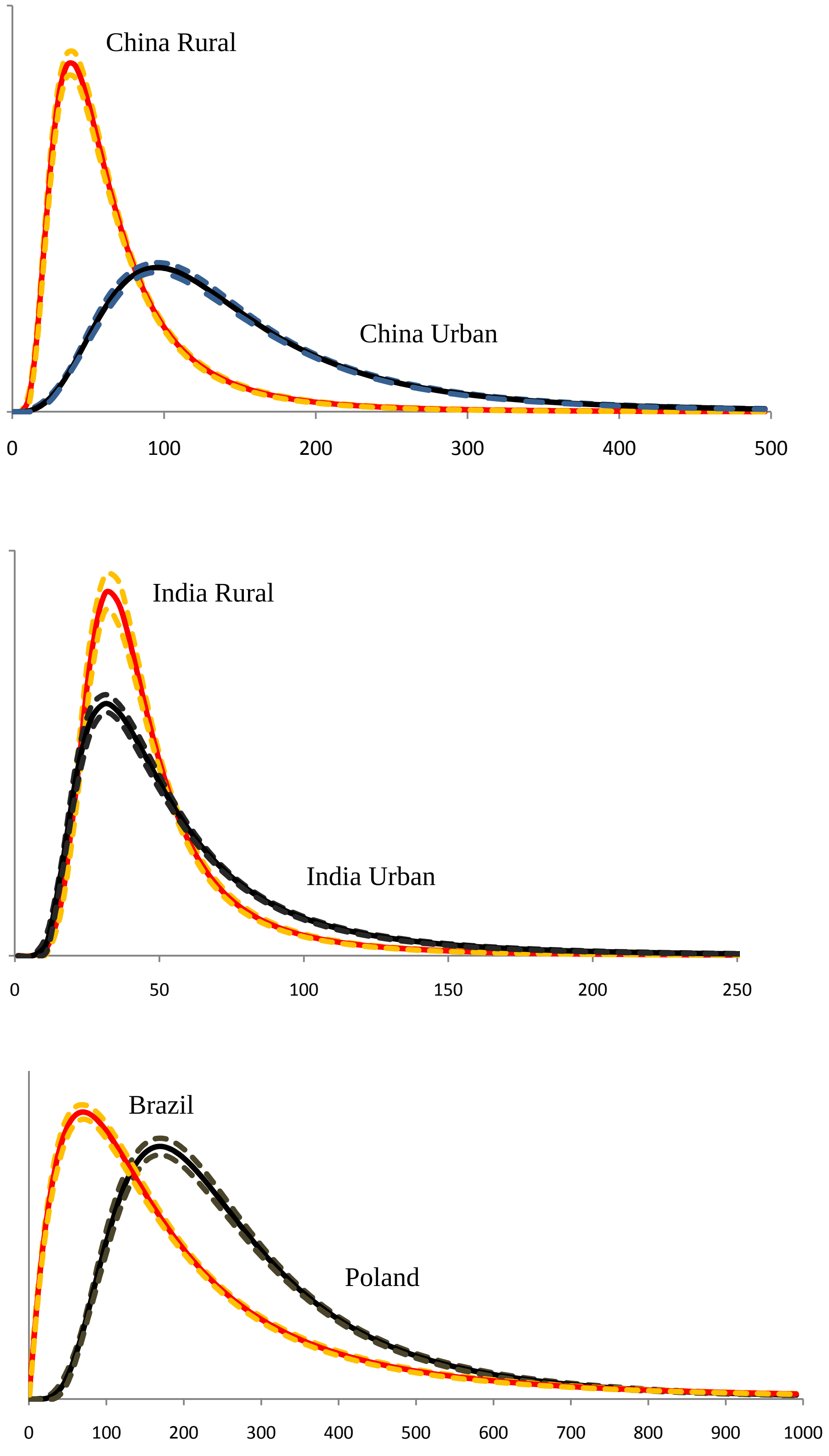

Figure 2. GB2-estimated pdfs for selected regions

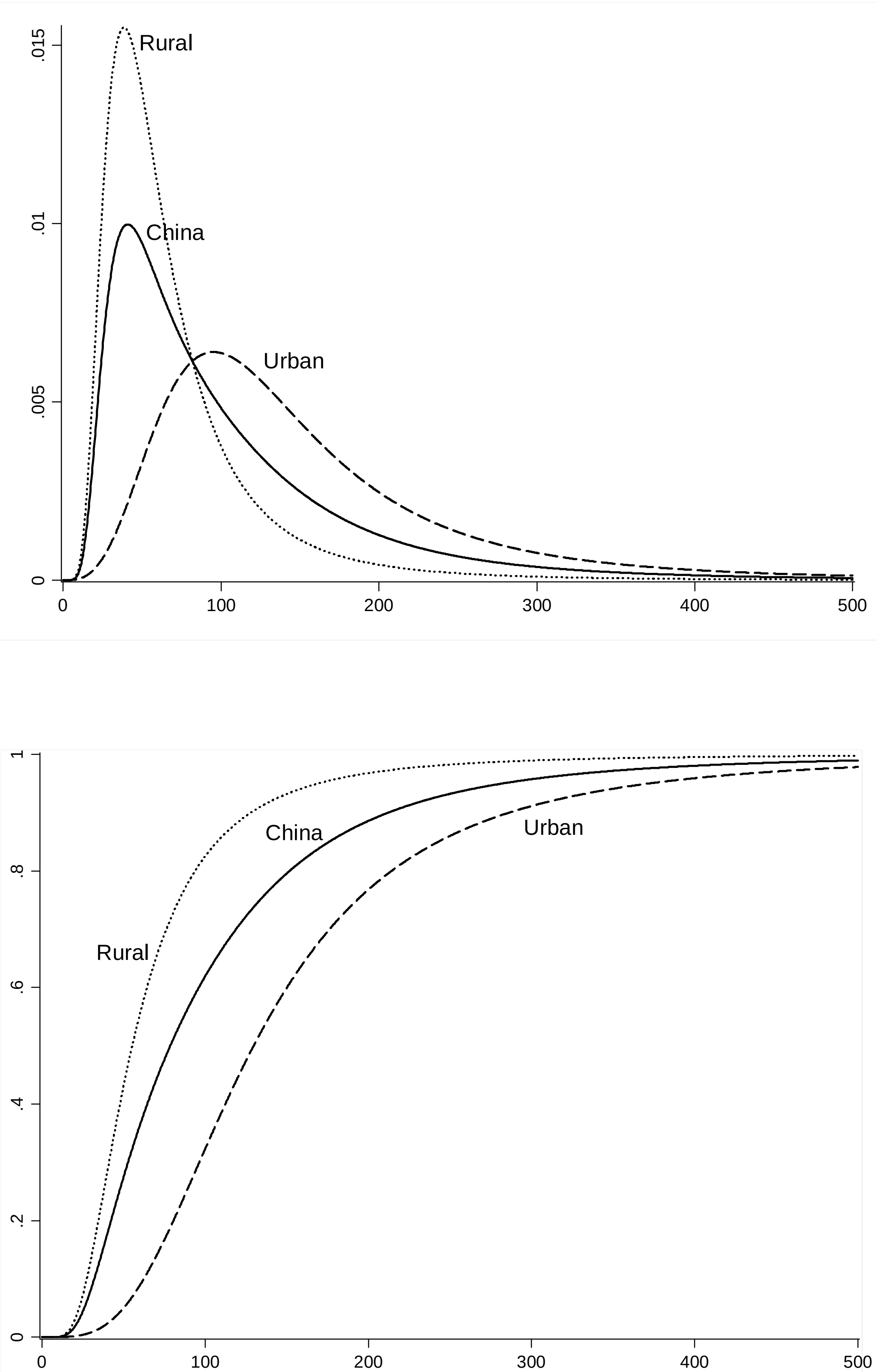

Figure 3 GB2-estimated pdf and cdf for all of China