THE UNIVERSITY OF MELBOURNE

Department of Economics

Working Paper Series

Inference for Income Distributions Using Grouped Data

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

Feb 6, 2012

Research Paper Number 1140

ISSN: 0819-2642
ISBN: 978 0 7340 44938

This paper is substantially revised version of Working Paper No. 1129

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

Inference for Income Distributions Using 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

Feb 6, 2012

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 general approach to estimation and inference for income distributions using grouped or aggregate data that are typically available in the form of population shares and class mean incomes, with unknown group bounds. Generic moment conditions and an optimal weight matrix that can be used for GMM estimation of any parametric income distribution are derived. Our derivation of the weight matrix and its inverse allows us to express the seemingly complex GMM objective function in a relatively simple form that facilitates estimation. We show that our proposed approach, that incorporates information on class means as well as population proportions, is more efficient than maximum likelihood estimation of the multinomial distribution that uses only population proportions. In contrast to the earlier work of Chotikapanich et al. (2007, 2012), that did not specify a formal GMM framework, did not provide methodology for obtaining standard errors, and restricted the analysis to the beta-2 distribution, we provide standard errors for estimated parameters and relevant functions of them, such as inequality and poverty measures, and we provide methodology for all distributions. A test statistic for testing the adequacy of a distribution is proposed. 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. We test the adequacy of each distribution and compare predicted and actual income shares, where the number of groups used for prediction can differ from the number used in estimation. Estimates and standard errors for inequality and poverty measures are provided.

KEY WORDS: GMM; Generalized beta distribution; Inequality and poverty.

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, Bourguignon and Morrisson (2002), 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 provision of econometric methodology for fitting income distributions to limited country-level data of this form. Specifically, our objective is to develop a framework for generalized method of moments (GMM) estimation of income distributions, when the available data are in the form of population shares and group mean incomes, with unknown group limits. When data are available as population and income shares, group mean incomes for a country can be computed from readily available data on the country's mean per capita

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

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

A limited solution to this problem was offered by Chotikapanich et al. (2007) (hereafter CGR) who suggested a method-of-moments estimator for the beta-2 distribution, implemented through nonlinear least squares. They applied it to a sample of 8 countries in two time periods, and illustrated 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) used the same technique to estimate the global income distribution as a mixture of beta-2 distributions for 91 countries in 1993 and 2000. 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. Their estimator was not an optimal GMM estimator and was not set up 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. Also, their methodology was limited to fitting the beta-2 distribution.

This paper represents a major step forward from this previous work. We establish moment conditions in a general form, and construct an optimal weight matrix, leading to a GMM estimator that is asymptotically efficient. Formally incorporating the optimal weight matrix into GMM estimation makes it possible to find the asymptotic covariance matrix of the estimator, from which we can find standard errors for the estimated parameters and for functions of the parameters used regularly in the area of income distributions, such as inequality and poverty measures. Also, our framework can be used to fit any income distribution, not just the beta-2 distribution considered by CGR. Because the GMM estimator

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

uses information on both population shares and class means, we are able to show that the covariance matrix for the GMM estimator is less than that for the maximum likelihood estimator applied to a multinomial likelihood function that uses only the population shares. By deriving both the weight matrix and its inverse, we are able to show how the seemingly complex GMM objective function can be written in a relatively simple form that is potentially amenable for use in standard software. To check the adequacy of an income distribution, we propose a test based on the standard *J*-statistic used for testing the validity of excess moment conditions. In the empirical work reported in this paper, we focus on the 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 goodness-of-fit can be assessed using a comparison of predicted and observed income shares, where the number of groups used for prediction can differ from that used for estimation. We also illustrate how estimates of the parameters and their covariance matrix can be used to plot income distributions along with their confidence bounds, and to compute inequality and poverty measures along with their standard errors.

Our work differs from related work by Wu and Perloff (2005, 2007) who also consider estimation of income distributions from grouped data. Wu and Perloff approximate the underlying income distribution with a maximum entropy density, estimated using a two-step GMM estimator with a simulated weight matrix. 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. Other distinguishing features of our work are our emphasis on estimation of class limits, and the asymptotic covariance matrix that can be used to find standard errors for all estimated parameters (including the class limits) 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. The moment conditions are specified and we refer to an appendix where the weight matrix and its inverse are derived. In Section 3 we summarize 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. 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. GMM ESTIMATION

Consider a sample of T observations $\left(y_1, y_2, \ldots, y_T\right)$, randomly drawn from a parametric income distribution $f(y;\phi)$, and grouped into N income classes $(z_0, z_1), (z_1, z_2), \ \ldots, (z_{N-1}, z_N)$, with $z_0 = 0$ and $z_N = \infty$. These grouped data are generally available in one of two forms: (i) as population proportions $\mathbf{c}' = \left(c_1, c_2, \ldots, c_N\right)$, income shares $\left(s_1, s_2, \ldots, s_N\right)$, and mean income $\bar{y}$, or (ii) as population proportions $\mathbf{c}$, and class mean incomes $\left(\bar{y}_1, \bar{y}_2, \ldots, \bar{y}_N\right)$. We work with the population proportions and class mean incomes. This form of the data can be obtained from the population and income shares using $\bar{y}_i = s_i \bar{y} / c_i$. 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

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

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

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

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

2.1 Specification of the Moment Conditions

To set up the moment conditions corresponding to the sample proportions c_i, we note that the corresponding population proportions, that we denote by $k_i(\boldsymbol{\theta})$, are given by

$$k_i(\boldsymbol{\theta}) = \int_{z_{i-1}}^{z_i} f(y;\phi)dy = \int_0^{\infty} g_i(y)\, f(y;\phi)dy = E\left[g_i(y)\right] \qquad i = 1, 2, \ldots, N \tag{3}$$

where $g_i(y)$ is an indicator function such that

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

The sample moment corresponding to $E\left[g_i(y)\right] = k_i(\boldsymbol{\theta})$ 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

$$\frac{1}{T}\sum_{t=1}^{T} g_i(y_t) - E\left[g_i(y)\right] \;=\; c_i - k_i(\boldsymbol{\theta}) \qquad i = 1, 2, \ldots, N-1 \tag{4}$$

The moment condition for $i = N$ is omitted because the result $\sum_{i=1}^{N} k_i(\boldsymbol{\theta}) = \sum_{i=1}^{N} c_i = 1$ makes one of the N conditions redundant.

To obtain moment conditions that use the information in the class mean incomes $\bar{y}_i$, it turns out to be convenient to consider $\tilde{y}_i$ defined as

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

Since $\sum_{i=1}^{N} \tilde{y}_i = \sum_{i=1}^{N} c_i \bar{y}_i = \bar{y}$, where $\bar{y}$ is sample mean income, we interpret $\tilde{y}_i$ as that part of sample mean income that comes from the i-th income class. Its corresponding population quantity is

$$\mu_i(\boldsymbol{\theta}) = \int_{z_{i-1}}^{z_i} y f(y;\phi) dy = \int_{0}^{\infty} y g_i(y) f(y;\phi) dy = E\left[y g_i(y) \right] \qquad i = 1, 2, \ldots, N \qquad (5)$$

Thus, our second set of moment conditions are

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

These moment conditions are different from those used by CGR. They used $\bar{y}_i - \mu_i(\theta)/k_i(\theta)$ instead of $\tilde{y}_i - \mu_i(\theta)$. The problem with $\bar{y}_i - \mu_i(\theta)/k_i(\theta)$ is that

$$E(\bar{y}_i) = E\left[\frac{T^{-1}\sum_{t=1}^{T} y_t g_i(y_t)}{T^{-1}\sum_{t=1}^{T} g_i(y_t)} \right] \neq \frac{\mu_i(\theta)}{k_i(\theta)}$$

although it is true that $\text{plim}(\bar{y}_i) = \mu_i(\theta)/k_i(\theta)$. Also, derivation of the corresponding weight matrix for $\tilde{y}_i - \mu_i(\theta)$ is straightforward, whereas the weight matrix for $\bar{y}_i - \mu_i(\theta)/k_i(\theta)$ is more difficult to derive.

Collecting all the terms in (4) and (6), we can write

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

and

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

where $\mathbf{c}$, $\mathbf{k}$, $\tilde{\mathbf{y}}$, and $\boldsymbol{\mu}$ are N-dimensional vectors containing the elements $c_i, k_i, \tilde{y}_i$, and μ_i, respectively, and the subscript "$-N$" denotes a corresponding $(N-1)$-dimensional vector with the last element excluded. Also, we drop the argument θ from these vectors to ease notation. 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 $k_i(\boldsymbol{\theta})$ and $\mu_i(\boldsymbol{\theta})$ 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 its distribution function, and

$$F_1(y;\phi) = \frac{1}{\mu}\int_0^y t f(t;\phi)dt \tag{9}$$

is its first moment distribution function, then, from (3), (5) and (9),

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

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

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 Derivation of Weight 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 for their moment conditions, with diagonal elements $(c_1^{-2}, c_2^{-2}, \ldots, c_{N-1}^{-2},\ \bar{y}_1^{-2}, \bar{y}_2^{-2}, \ldots, \bar{y}_N^{-2})$. The estimator from this weight matrix minimizes the sum of squares of percentage errors in the moment conditions, but it is not an optimal weight matrix. 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 $\sum_{t=1}^{T} y_t^2 g_i(y_t)$ 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. From Appendix A.1, we find

$$\begin{aligned} \mathbf{W}^{-1} &= \text{plim} \frac{1}{T} \sum_{t=1}^{T} \mathbf{h}\left(y_t, \boldsymbol{\theta}\right) \mathbf{h}\left(y_t, \boldsymbol{\theta}\right)' \\ &= \left[\begin{array}{c:c} \mathbf{D}(\mathbf{k}_{-N}) & \left[\mathbf{D}(\boldsymbol{\mu}_{-N}) \quad \mathbf{0}_{N-1} \right] \\ \hdashline \begin{bmatrix} \mathbf{D}(\boldsymbol{\mu}_{-N}) \\ \mathbf{0}'_{N-1} \end{bmatrix} & \mathbf{D}(\boldsymbol{\mu}^{(2)}) \end{array} \right] - \begin{bmatrix} \mathbf{k}_{-N} \\ \boldsymbol{\mu} \end{bmatrix} \left[\mathbf{k}'_{-N} \quad \boldsymbol{\mu}' \right] \end{aligned} \tag{14}$$

where $\mathbf{0}_{N-1}$ is an $(N-1)$ dimensional vector of zeros, $\mathbf{D}(\mathbf{x})$ denotes a diagonal matrix with elements of a vector $\mathbf{x}$ on the diagonal, and $\boldsymbol{\mu}^{(2)} = \left(\mu_1^{(2)}, \mu_2^{(2)}, \ldots, \mu_N^{(2)}\right)'$, with

$$\mu_i^{(2)}(\boldsymbol{\theta}) = \int_{z_{i-1}}^{z_i} y^2 f(y;\phi)dy = \int_0^{\infty} y^2 g_i(y)\, f(y;\phi)\, dy = E\left[y^2 g_i(y) \right] \tag{15}$$

For minimizing the GMM objective function in (2), 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, an analytical expression for $\mathbf{W}$ is useful because it allows us to simplify the GMM objective function, gain insights into the minimization process, and improve computational efficiency. In Appendix A.2, we derive and present an analytical expression for $\mathbf{W}$. Then, using direct matrix multiplication, we can show that the GMM objective function simplifies to

$$\begin{aligned} GMM &= \mathbf{H}'\mathbf{W}\mathbf{H} = \begin{bmatrix} \mathbf{c}_{-N} - \mathbf{k}_{-N} \\ \tilde{\mathbf{y}} - \boldsymbol{\mu} \end{bmatrix}' \mathbf{W} \begin{bmatrix} \mathbf{c}_{-N} - \mathbf{k}_{-N} \\ \tilde{\mathbf{y}} - \boldsymbol{\mu} \end{bmatrix} \\ &= \sum_{i=1}^{N} w_{1i}(c_i - k_i)^2 + \sum_{i=1}^{N} w_{2i}(\tilde{y}_i - \mu_i)^2 - 2\sum_{i=1}^{N} w_{3i}(c_i - k_i)(\tilde{y}_i - \mu_i) \end{aligned} \tag{16}$$

where $w_{1i} = \mu_i^{(2)} / v_i$, $w_{2i} = k_i / v_i$, $w_{3i} = \mu_i / v_i$, and $v_i = k_i \mu_i^{(2)} - \mu_i^2$. The advantages of expressing the GMM objective function in this way are its potential suitability for use in standard software, and that it can be conveniently written as a function of $2N$ rather than $2N-1$ moment conditions, with a weight matrix that consists of four diagonal blocks. Specifically, we can write

$$GMM = \mathbf{H}'_{(2N)} \mathbf{W}_{(2N)} \mathbf{H}_{(2N)} = \begin{bmatrix} \mathbf{c} - \mathbf{k} \\ \tilde{\mathbf{y}} - \boldsymbol{\mu} \end{bmatrix}' \begin{bmatrix} D(\mathbf{w}_1) & -D(\mathbf{w}_3) \\ -D(\mathbf{w}_3) & D(\mathbf{w}_2) \end{bmatrix} \begin{bmatrix} \mathbf{c} - \mathbf{k} \\ \tilde{\mathbf{y}} - \boldsymbol{\mu} \end{bmatrix} \tag{17}$$

where $\mathbf{w}_1, \mathbf{w}_2$ and $\mathbf{w}_3$ are *N*-dimensional vectors containing the elements w_{1i}, w_{2i} and w_{3i}, respectively; the subscript "$(2N)$" has been used to describe this alternative specification, where it is no longer necessary to drop a redundant moment condition. In this formulation, there are no cross-product terms involving moment conditions from different groups, but there is a cross-product term between the proportion and mean moment conditions for the *i*-th group.

The extra quantities needed to compute the weights (that were not also needed to compute the moment conditions) are the $\mu_i^{(2)}$. It is convenient to write

$$\mu_i^{(2)} = \mu^{(2)} \left[F_2(z_i;\phi) - F_2(z_{i-1};\phi) \right] \tag{18}$$

where $F_2(z_i;\phi) = \int_0^{z_i} y^2 f(y;\phi) dy \Big/ \mu^{(2)}$ is the second moment distribution function, and $\mu^{(2)} = E(y^2)$ is the second moment for *y*. Expressions for $F_2(z_i;\phi)$ and $\mu^{(2)}$, for a number of distributions, are provided in Section 3.

2.3 Empirical Implementation of GMM Estimation

We consider three GMM estimators. The first, $\hat{\boldsymbol{\theta}}_{CGR}$, is the CGR estimator modified to suit our specification of the moment conditions. It minimizes (16), but with weights $w_{1i} = 1/c_i^2$, $w_{2i} = 1/\tilde{y}_i^2$, and $w_{3i} = 0$. We use $\mathbf{W}_{(2N)}^{CGR}$ to denote the corresponding weight matrix. In this case the GMM estimator minimizes the sum of squares of the percentage errors in the moment conditions. The second estimator is a two-step estimator that uses $\hat{\boldsymbol{\theta}}_{CGR}$ to compute an estimate of the optimal weight matrix, $\mathbf{W}_{(2N)}\left(\hat{\boldsymbol{\theta}}_{CGR}\right)$. In our empirical work we used an iterative version of this estimator, updating $\mathbf{W}_{(2N)}$ after each iteration, and iterating until convergence, with convergence being achieved within 10-20 iterations. The third estimator is a one-step estimator obtained by minimizing the complete objective function with respect to $\boldsymbol{\theta}$ where both the moment conditions and $\mathbf{W}_{(2N)}$ are functions of $\boldsymbol{\theta}$. Hansen et al. (1996) refer to the one-step estimator as the "continuous-updating estimator". Although the iterative and continuously updating estimators have the same asymptotic distribution, they are typically different in finite samples (Hall, 2005, p.103).

The three estimators are given by

$$\hat{\boldsymbol{\theta}}_{CGR} = \arg\min_{\boldsymbol{\theta}} \mathbf{H}_{(2N)}\left(\boldsymbol{\theta}\right)' \mathbf{W}_{(2N)}^{CGR} \mathbf{H}_{(2N)}\left(\boldsymbol{\theta}\right)$$

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

$$\hat{\boldsymbol{\theta}}_{1-STEP} = \arg\min_{\boldsymbol{\theta}} \mathbf{H}_{(2N)}\left(\boldsymbol{\theta}\right)' \mathbf{W}_{(2N)}\left(\boldsymbol{\theta}\right) \mathbf{H}_{(2N)}\left(\boldsymbol{\theta}\right)$$

2.4 Asymptotic Covariance Matrix and Relative Efficiency

The asymptotic covariance matrix of the one and two-step estimators is (see, for example, Cameron and Trivedi, 2005, p.176) $\mathbf{var}\left(\hat{\boldsymbol{\theta}}\right) = T^{-1}\left(\mathbf{G}(\boldsymbol{\theta})'\mathbf{W}(\boldsymbol{\theta})\mathbf{G}(\boldsymbol{\theta})\right)^{-1}$, where $\mathbf{G} = \partial\mathbf{H}/\partial\boldsymbol{\theta}'$. It is straightforward, although tedious, to show that

$$\left(\frac{\partial\mathbf{H}}{\partial\boldsymbol{\theta}'}\right)' \mathbf{W}\left(\frac{\partial\mathbf{H}}{\partial\boldsymbol{\theta}'}\right) = \left(\frac{\partial\mathbf{H}_{(2N)}}{\partial\boldsymbol{\theta}'}\right)' \mathbf{W}_{(2N)}\left(\frac{\partial\mathbf{H}_{(2N)}}{\partial\boldsymbol{\theta}'}\right)$$

Thus, the covariance matrix can also be written in terms of the complete set of $2N$ moment conditions and corresponding weight matrix as

$$\mathbf{var}\left(\hat{\boldsymbol{\theta}}\right) = \frac{1}{T}\left(\mathbf{G}_{(2N)}(\boldsymbol{\theta})'\,\mathbf{W}_{(2N)}(\boldsymbol{\theta})\mathbf{G}_{(2N)}(\boldsymbol{\theta})\right)^{-1} \tag{19}$$

where $\mathbf{G}_{(2N)} = \partial\mathbf{H}_{(2N)}/\partial\boldsymbol{\theta}'$. An estimate of $\mathbf{var}\left(\hat{\boldsymbol{\theta}}\right)$ is obtained by replacing θ by $\hat{\theta}$. In our empirical work we successfully used both analytical and numerical derivatives to calculate the elements of the $\mathbf{G}_{(2N)}$ 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.

One common way of estimating income distributions is via maximum likelihood (ML) estimation using a multinomial likelihood function, applied only to the population proportions. See, for example, McDonald (1984) and Bandourian et al. (2003). To show that this estimator is less efficient than the GMM estimator that use information on both population proportions and class mean incomes, we rewrite (19) as

$$\text{var}\left(\hat{\boldsymbol{\theta}}\right) = \frac{1}{T}\left(\begin{bmatrix}\dfrac{\partial\mathbf{k}'}{\partial\boldsymbol{\theta}} & \dfrac{\partial\boldsymbol{\mu}'}{\partial\boldsymbol{\theta}}\end{bmatrix}\begin{bmatrix}\mathbf{D}\left(\boldsymbol{\mu}^{(2)}/\mathbf{v}\right) & -\mathbf{D}\left(\boldsymbol{\mu}/\mathbf{v}\right)\\ -\mathbf{D}\left(\boldsymbol{\mu}/\mathbf{v}\right) & \mathbf{D}\left(\mathbf{k}/\mathbf{v}\right)\end{bmatrix}\begin{bmatrix}\partial\mathbf{k}/\partial\boldsymbol{\theta}'\\ \partial\boldsymbol{\mu}/\partial\boldsymbol{\theta}'\end{bmatrix}\right)^{-1} \tag{20}$$

where, in line with Appendix A.2, we use $\mathbf{D}\left(\mathbf{x}/\mathbf{s}\right)$ to denote a diagonal matrix with diagonal $\left(x_1/s_1, x_2/s_2, \ldots, x_N/s_N\right)$, for any two vectors $\mathbf{x}$ and $\mathbf{s}$. Expanding (20) yields

$$\text{var}\left(\hat{\boldsymbol{\theta}}\right) = \frac{1}{T}\left(\frac{\partial\mathbf{k}'}{\partial\boldsymbol{\theta}}\mathbf{D}\left(\boldsymbol{\mu}^{(2)}/\mathbf{v}\right)\frac{\partial\mathbf{k}}{\partial\boldsymbol{\theta}'} - 2\frac{\partial\mathbf{k}'}{\partial\boldsymbol{\theta}}\mathbf{D}\left(\boldsymbol{\mu}/\mathbf{v}\right)\frac{\partial\boldsymbol{\mu}}{\partial\boldsymbol{\theta}'} + \frac{\partial\boldsymbol{\mu}'}{\partial\boldsymbol{\theta}}\mathbf{D}\left(\mathbf{k}/\mathbf{v}\right)\frac{\partial\boldsymbol{\mu}}{\partial\boldsymbol{\theta}}\right)^{-1} \tag{21}$$

Using the definition of $\mathbf{v}$, and some algebraic manipulation, this expression can be written as

$$\text{var}\left(\hat{\boldsymbol{\theta}}\right) = \frac{1}{T}\left\{\frac{\partial\mathbf{k}'}{\partial\boldsymbol{\theta}}\mathbf{D}\left(1/\mathbf{k}\right)\frac{\partial\mathbf{k}}{\partial\boldsymbol{\theta}'} + \left(\frac{\partial\boldsymbol{\mu}'}{\partial\boldsymbol{\theta}} - \frac{\partial\mathbf{k}'}{\partial\boldsymbol{\theta}}\mathbf{D}\left(\boldsymbol{\mu}/\mathbf{k}\right)\right)\mathbf{D}\left(\mathbf{k}/\mathbf{v}\right)\left(\frac{\partial\boldsymbol{\mu}}{\partial\boldsymbol{\theta}'} - \mathbf{D}\left(\boldsymbol{\mu}/\mathbf{k}\right)\frac{\partial\mathbf{k}}{\partial\boldsymbol{\theta}'}\right)\right\}^{-1} \tag{22}$$

By setting up the multinomial likelihood function, and deriving the information matrix, it can be shown that the covariance matrix for the ML estimator based on only proportions is (see, for example, Cox (1984))

$$\mathrm{var}\left(\hat{\boldsymbol{\theta}}_{ML}\right) = \frac{1}{T}\left\{\frac{\partial \mathbf{k}'}{\partial \boldsymbol{\theta}}\mathbf{D}\left(1/\mathbf{k}\right)\frac{\partial \mathbf{k}}{\partial \boldsymbol{\theta}'}\right\}^{-1} \tag{23}$$

Since the second term in (22) is positive definite, it follows that $\mathrm{var}\left(\hat{\boldsymbol{\theta}}_{ML}\right) - \mathrm{var}\left(\hat{\boldsymbol{\theta}}\right)$ is positive definite and hence the GMM estimator that uses both population shares and class mean incomes is more efficient than the ML estimator that uses only population shares.

2.5 Testing the Validity of an Income Distribution

Under the null hypothesis that the moment conditions are correct $\mathrm{plim}\,\mathbf{H}\left(\hat{\boldsymbol{\theta}}\right) = \mathbf{0}$ and

$$J = T\,\mathbf{H}\left(\hat{\boldsymbol{\theta}}\right)'\mathbf{W}\left(\hat{\boldsymbol{\theta}}\right)\mathbf{H}\left(\hat{\boldsymbol{\theta}}\right) \xrightarrow{d} \chi^2_{N-K} \tag{24}$$

In traditional GMM estimation this test statistic is used to assess whether excess moment conditions are valid. In our case, we assume a particular form of parametric income distribution and use its properties to construct both the moment conditions and the weight matrix. If the assumed distribution is invalid, then $\mathrm{plim}\,\mathbf{H}\left(\hat{\boldsymbol{\theta}}\right) \neq \mathbf{0}$ and $\mathrm{plim}\,\mathbf{W}\left(\hat{\boldsymbol{\theta}}\right) \neq \mathbf{W}\left(\boldsymbol{\theta}\right)$. Thus, the *J* statistic can be used to test the validity of the assumed income distribution.

3. APPLICATION TO MAJOR INCOME DISTRIBUTIONS

A large number of probability density functions (pdfs) has been suggested in the literature for modelling income distributions. See, for example, McDonald and Ransom (1979), McDonald (1984), McDonald and Xu (1995), Bandourian et al. (2003) and Kleiber and Kotz (2003). To illustrate the methodology, we have chosen the four-parameter generalized beta distribution of the second kind (GB2) introduced by McDonald (1984), and some of its popular special case distributions – the beta2, Dagum, Singh-Maddala, generalized gamma and lognormal distributions. The GB2 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)).

3.1 Moments and Moment Distribution Functions

To specify the moment conditions and weight matrix for each of the distributions, we need their first and second moments, their cumulative distributions functions (cdfs), and their first and second moment distribution functions, all expressed as computable functions of the parameters. These required quantities are summarized in Table 1. They can be found in various places; two convenient sources are McDonald (1984) and Kleiber and Kotz (2003).

The beta-2, Singh-Maddala, and Dagum distributions are readily obtained from the more general GB2 distribution by setting $a=1$, $p=1$ and $q=1$, respectively. The generalized gamma distribution can be obtained as a special case of the GB2 distribution by setting $b=\beta q^{1/a}$ and taking the limit as $q\to\infty$; the lognormal distribution can be obtained as a special case of the generalized gamma distribution by setting $\beta^a=\sigma^2 a^2$, $p=(a\mu+1)\big/\beta^a$, and taking the limit as $a\to 0$ (McDonald 1984). In Table 1, $B(\cdot,\cdot)$ is the beta function, $\Gamma(\cdot)$ is the gamma function, $B_u(p,q)=\int_0^u t^{p-1}(1-t)^{q-1}dt\Big/B(p,q)$ is the cdf for a standard beta random variable defined on the (0,1) interval, $G_u(p,b)=\int_0^{u/b} t^{p-1}e^{-t}dt\Big/\Gamma(p)$ is the cdf for a standard gamma random variable with parameters *p* and *b*, and $\Phi(\cdot)$ is the cdf for a standard normal random variable.

3.2 Inequality and Poverty Measures

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). In our empirical work these quantities were computed for the GB2, beta-2, Singh-Maddala and Dagum distributions.

The Gini coefficient is given by

$$G = -1 + \frac{2}{\mu}\int_0^\infty y\,F(y;\phi)\,f(y;\phi)\,dy \tag{25}$$

Expressions for G in terms of the parameters are given in Table 1 for all distributions except the GB2 and generalized gamma. Lengthy expressions written in terms of hypergeometric functions are available for these latter two distributions (McDonald, 1984; McDonald and Ransom, 2008). However, using MATLAB, we found that numerical evaluation of the integral in (25) was more efficient and reliable than computing values for the hypergeometric functions.

The Theil inequality coefficient is given by

$$T = \int_0^\infty \left(\frac{y}{\mu}\right)\ln\left(\frac{y}{\mu}\right) f(y;\phi)\,dy \tag{26}$$

and, for the GB2 distribution, it can be written as (McDonald and Ransom, 2008)

$$T = \frac{1}{a}\left[\psi\left(p + 1/a\right) - \psi\left(q - 1/a\right)\right] + \ln\left(b/\mu\right) \tag{27}$$

where $\psi(t) = d\log\Gamma(t)/dt$ is the digamma function. Values of T for the beta-2, Singh-Maddala and Dagum distributions are obtained from (27) by setting $a = 1$, $p = 1$ and $q = 1$, respectively.

For a given poverty line x, the headcount ratio is the proportion of population with incomes less than x, and so, for all distributions, it is simply given by $F(x;\phi)$. 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

$$\begin{aligned} FGT_x(2) &= \int_0^x \left(\frac{x-y}{x}\right)^2 f(y)dy \\ &= F(x;\phi) - \frac{2\mu}{x}F_1(x;\phi) + \frac{\mu^{(2)}}{x^2}F_2(x;\phi) \end{aligned} \tag{28}$$

The quantities necessary for computing (28) are given in Table 1.

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 of 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 s_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 groups 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.

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 $\tilde{y}_i$, defined as

$\tilde{y}_i = c_i \bar{y}_i = s_i \bar{y}$. 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 $T = 10,000$. 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 where we consider the estimated income distributions for the eight regions. Goodness-of-fit 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 the 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, two-stage, and one-stage estimations. Standard errors for both the two-stage and one-stage estimates are also reported. There are no dramatic differences between the estimates from the three different estimators. The two-stage and one-stage estimates are almost identical, particularly for the class limits, and the CGR estimates are only slightly different. Similarly, the standard errors obtained using two-stage estimation are very close to those obtained from one-stage 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 estimates and standard errors for distributions other than the GB2 are not reported; 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)) because the hypergeometric function in MATLAB broke down. 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. Numerical standard errors could still be found, however. This problem did not arise with the GB2 and other distributions.

Figure 1 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 2 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. 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, obtained using the 2-stage estimates. 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 lognormal it is 10. At a 5% significance level, the moment conditions for the GB2 distribution are not rejected for 5 out of 8 of the regions. All the other distributions are rejected for all other countries with the exception of the beta-2 distribution for China rural and Russia. These test outcomes are not as pleasing as one might hope. It would have been more satisfactory if the GB2 distribution was accepted for all regions, and some of the other distributions were accepted for a wider selection of countries. However, acceptance of a particular distribution is likely to be difficult with a large sample size, and, despite the test outcomes, the ability of the GB2 distribution to predict income shares was very good.

Goodness-of-fit in terms of predicting income shares was carried out by comparing the observed income shares s_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 $\hat{s}_i = \hat{\eta}_i - \hat{\eta}_{i-1}$.

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, $\sqrt{N^{-1}\sum_{i=1}^{N}\left[100\left(\hat{s}_i - s_i\right)\right]^2}$, 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 or close to the best for all regions except India rural and Brazil. The generalized gamma and lognormal distributions generally performed poorly relative to the other distributions. The other special cases of the GB2 distribution did well for some regions and not so well for others. 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 in Table 1 and equations (25)-(27), for the GB2, 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 generally provides a good fit in terms of the validity of its moment conditions, and the accuracy of predicted income shares from the estimated distributions.

APPENDIX A: OPTIMAL WEIGHTING MATRIX

A.1 Derivation of $\mathbf{W}^{-1}$

From (7), we partition $\mathbf{h}(y_t,\boldsymbol{\theta})$ as $\left[\mathbf{h}_1' \quad \mathbf{h}_2'\right]'$, with $\mathbf{h}_1$ denoting the conditions relating to the proportions and $\mathbf{h}_2$ denoting the conditions relating to the class means. Then,

$$\mathbf{W}^{-1} = \text{plim}\left[\frac{1}{T}\sum_{t=1}^{T}\mathbf{h}(y_t,\boldsymbol{\theta})\mathbf{h}(y_t,\boldsymbol{\theta})'\right] = \text{plim}\begin{pmatrix} \frac{1}{T}\sum_{t=1}^{T}\mathbf{h}_1\mathbf{h}_1' & \frac{1}{T}\sum_{t=1}^{T}\mathbf{h}_1\mathbf{h}_2' \\ \frac{1}{T}\sum_{t=1}^{T}\mathbf{h}_2\mathbf{h}_1' & \frac{1}{T}\sum_{t=1}^{T}\mathbf{h}_2\mathbf{h}_2' \end{pmatrix}$$

$$= \text{plim}\begin{pmatrix} \mathbf{P}_{(N-1)\times(N-1)} & \mathbf{Q}_{(N-1)\times N} \\ \mathbf{Q'}_{N\times(N-1)} & \mathbf{M}_{N\times N} \end{pmatrix}$$

where the elements of $\mathbf{P}$, $\mathbf{Q}$ and $\mathbf{M}$ are

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

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

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

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

$$m_{ii} = \frac{1}{T}\sum_{t=1}^{T}\left[y_t g_i(y_t) - \mu_i\right]^2 = \frac{1}{T}\sum_{t=1}^{T} y_t^2 g_i(y_t) + \mu_i^2 - 2\tilde{y}_i\mu_i$$

$$m_{ij} = \frac{1}{T}\sum_{t=1}^{T}\left[y_t g_i(y_t) - \mu_i\right]\left[y_t g_j(y_t) - \mu_j\right] = -\tilde{y}_i\mu_j - \mu_i\tilde{y}_j + \mu_i\mu_j \quad (i \neq j)$$

Using $\text{plim}\, c_i = k_i$, $\text{plim}\, \tilde{y}_i = \mu_i$, and $\text{plim}\, T^{-1}\sum_{t=1}^{T} y_t^2 g_i(y_t) = \mu_i^{(2)}$, the elements of $\mathbf{W}^{-1}$ are

$$\text{plim}\, p_{ii} = k_i(1 - k_i) \qquad \text{plim}\, p_{ij} = -k_i k_j \qquad \text{plim}\, q_{ii} = \mu_i(1 - k_i)$$

$$\text{plim}\, q_{ij} = -k_i\mu_j \qquad \text{plim}\, m_{ii} = \mu_i^{(2)} - \mu_i^2 \qquad \text{plim}\, m_{ij} = -\mu_i\mu_j$$

Using these results and the notation defined in the body of the paper, we can write

$$\mathbf{W}^{-1} = \left[\begin{array}{c:c} \mathbf{D}(\mathbf{k}_{-N}) - \mathbf{k}_{-N}\mathbf{k}'_{-N} & \left[\mathbf{D}(\boldsymbol{\mu}_{-N}) \quad \mathbf{0}_{N-1}\right] - \mathbf{k}_{-N}\boldsymbol{\mu}' \\ \hdashline \begin{bmatrix} \mathbf{D}(\boldsymbol{\mu}_{-N}) \\ \mathbf{0}'_{N-1} \end{bmatrix} - \boldsymbol{\mu}\mathbf{k}'_{-N} & \mathbf{D}(\boldsymbol{\mu}^{(2)}) - \boldsymbol{\mu}\boldsymbol{\mu}' \end{array}\right]$$

A.2 Deriving $\mathbf{W}$ from $\mathbf{W}^{-1}$

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

$$\mathbf{W}^{-1} = \begin{bmatrix} \mathbf{D}(\mathbf{k}_{-N}) & \left[\mathbf{D}(\boldsymbol{\mu}_{-N}) \quad \mathbf{0}_{N-1}\right] \\ \begin{bmatrix} \mathbf{D}(\boldsymbol{\mu}_{-N}) \\ \mathbf{0}'_{N-1} \end{bmatrix} & \mathbf{D}(\boldsymbol{\mu}^{(2)}) \end{bmatrix} - \begin{bmatrix} \mathbf{k}_{-N} \\ \boldsymbol{\mu} \end{bmatrix}\left[\mathbf{k}'_{-N} \quad \boldsymbol{\mu}'\right] = \mathbf{A} - \mathbf{c}\mathbf{c}'$$

To invert this matrix, we use the result

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

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

$$\mathbf{A}^{11}=\left[\mathbf{D}\left(\mathbf{k}_{-N}\right)-\left[\mathbf{D}\left(\boldsymbol{\mu}_{-N}\right)\quad \mathbf{0}_{N-1}\right]\mathbf{D}\left(\boldsymbol{\mu}^{(2)}\right)^{-1}\begin{bmatrix}\mathbf{D}\left(\boldsymbol{\mu}_{-N}\right)\\ \mathbf{0}'_{N-1}\end{bmatrix}\right]^{-1}$$

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

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

where $v_i=k_i\mu_i^{(2)}-\mu_i^2$. Using notation $\mathbf{D}(\mathbf{x}/\mathbf{s})$ 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 $\mathbf{s}$, we can write $\mathbf{A}^{11}=\mathbf{D}\left(\boldsymbol{\mu}_{-N}^{(2)}/\mathbf{v}_{-N}\right)$.

Also,

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

and

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

Then,

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

and

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

Thus,

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

Finally,

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

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

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

By using matrix multiplication and then simplifying, we can show that, in line with equations (16) and (17), the GMM objective function can be written as

$$\begin{bmatrix} \mathbf{c}_{-N} - \mathbf{k}_{-N} \\ \tilde{\mathbf{y}} - \boldsymbol{\mu} \end{bmatrix}' \mathbf{W} \begin{bmatrix} \mathbf{c}_{-N} - \mathbf{k}_{-N} \\ \tilde{\mathbf{y}} - \boldsymbol{\mu} \end{bmatrix} = \begin{bmatrix} \mathbf{c} - \mathbf{k} \\ \tilde{\mathbf{y}} - \boldsymbol{\mu} \end{bmatrix}' \begin{bmatrix} D\left(\boldsymbol{\mu}^{(2)}/\mathbf{v}\right) & -D\left(\boldsymbol{\mu}/\mathbf{v}\right) \\ -D\left(\boldsymbol{\mu}/\mathbf{v}\right) & D\left(\mathbf{k}/\mathbf{v}\right) \end{bmatrix} \begin{bmatrix} \mathbf{c} - \mathbf{k} \\ \tilde{\mathbf{y}} - \boldsymbol{\mu} \end{bmatrix}$$

APPENDIX B: 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 $\boldsymbol{\theta}' = \left(\mathbf{z}', \boldsymbol{\phi}'\right)$ where $\mathbf{z}' = \left(z_1, z_2, \ldots, z_{N-1}\right)$ 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, $k_i(\boldsymbol{\theta}) = F(z_i;\phi) - F(z_{i-1};\phi)$, and $\mu_i(\boldsymbol{\theta}) = \mu\left(F_1(z_i;\phi) - F_1(z_{i-1};\phi)\right)$. We focus on the derivatives of $F(z_i;\phi)$ and $F_1(z_i;\phi)$ with respect to z_i and the elements in ϕ. Finding the derivatives of μ and using them along with $F(z_i;\phi)$ and $F_1(z_i;\phi)$ to find the required derivatives of $\mu_i(\boldsymbol{\theta})$ is straightforward, although tedious. The basic tool used to find the derivatives of $F(z_i;\phi)$ and $F_1(z_i;\phi)$ 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 $u_i = (z_i/b)^a \Big/ \left[1 + (z_i/b)^a\right] = z_i^a \Big/ \left(b^a + z_i^a\right)$. Derivatives for $B_{u_i}\left(p + k/a, q - k/a\right)$ are provided. Setting $k = 0$ gives the required expressions for $F(z_i;\phi)$; setting $k = 1$ will give the required expressions for $F_1(z_i;\phi)$. Derivatives for the beta-2, Singh-Maddala and Dagum distributions, can be obtained by setting $a = 1$, $p = 1$ and $q = 1$, respectively.

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

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

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

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

$$\frac{\partial\left(B_{u_i}\left(p+k/a,q-k/a\right)\right)}{\partial p}=\left[\psi(p+q)-\psi\left(p+\frac{k}{a}\right)\right]B_{u_i}\left(p+\frac{k}{a},q-\frac{k}{a}\right)$$
$$+\int_0^{u_i}\frac{x^{p+k/a-1}(1-x)^{q-k/a-1}}{B(p+k/a,q-k/a)}\ln(x)dx$$
$$=\left[\psi(p+q)-\psi(p+k/a)+\ln\left(u_i\right)\right]B_{u_i}\left(p+k/a,q-k/a\right)$$
$$-\frac{x^{p+k/a}}{\left(p+k/a\right)^2B\left(p+k/a,q-k/a\right)}$$
$$\times\,{}_3F_2\left(p+k/a,p+k/a,1+k/a-q;p+k/a+1,p+k/a+1;u_i\right)$$

$$\frac{\partial\left(B_{u_i}\left(p+k/a,q-k/a\right)\right)}{\partial q}=\left[\psi(p+q)-\psi\left(q-\frac{k}{a}\right)\right]B_{u_i}\left(p+\frac{k}{a},q-\frac{k}{a}\right)$$
$$+\int_0^{u_i}\frac{x^{p+k/a-1}(1-x)^{q-k/a-1}}{B(p+k/a,q-k/a)}\ln(1-x)dx$$
$$=-\left[\psi(p+q)-\psi(q-k/a)+\ln\left(1-u_i\right)\right]B_{u_i}\left(p+k/a,q-k/a\right)$$
$$+\frac{(1-x)^{q-k/a}}{\left(q-k/a\right)^2B\left(p+k/a,q-k/a\right)}$$
$$\times\,{}_3F_2\left(q-k/a,q-k/a,1-k/a-p;q-k/a+1,q-k/a+1;1-u_i\right)$$

$$\frac{\partial\left(B_{u_i}\left(p+k/a,q-k/a\right)\right)}{\partial a}=\frac{k}{a^2}\left(-\frac{\partial}{\partial p}\left(B_{u_i}\left(p+\frac{k}{a},q-\frac{k}{a}\right)\right)+\frac{\partial}{\partial q}\left(B_{u_i}\left(p+\frac{k}{a},q-\frac{k}{a}\right)\right)\right)$$
$$+\frac{a_i^{ap+k}b^{aq-k}(\ln a_i-\ln b)}{\left(b^a+a_i^a\right)^{p+q}B\left(p+k/a,q-k/a\right)}$$

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

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

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

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

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

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

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

B.3 Derivatives for the Lognormal Distribution

Derivatives of the lognormal cdf $\Phi\left(\left[\ln(z_i)-\mu-k\sigma^2\right]/\sigma\right)$ for $k=0$ and $k=1$ 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.

ACKNOWLEDGEMENTS

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

REFERENCES

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 Mantrala, A. (1996), "Something New, Something Old: Parametric Models for the Size Distribution of Income", *Journal of Income Distribution*, 6, 91-103.

Bourguignon, F., and Morrisson, C. (2002), "Inequality Among World Citizens: 1820-1992", *American Economic Review*, 92, 727-744.

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

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

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

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.E., Rao, D.S.P., and Valencia, V. (2012), "Global Income Distributions and Inequality, 1993 and 2000: Incorporating Country-Level Inequality Modelled with Beta Distributions", *The Review of Economics and Statistics* forthcoming.

Cox, C. (1984), "An Elementary Introduction to Maximum Likelihood Estimation for Multinomial Models: Birch's Theorem and the Delta Method", *The American Statistician*, 38, 283-287.

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

Hall, A.R. (2005), *Generalized Method of Moments*, Oxford: Oxford University Press.

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

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*, 47**,** 1513 – 1525.

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

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. Moments, distribution functions and Gini coefficients for income distributions

	Density Function	Moments	Distribution Function	Moment Distribution Functions	Gini Coefficient
GB2	$f(y;\phi)=\frac{ay^{ap-1}}{b^{ap}B(p,q)\left(1+(y/b)^a\right)^{p+q}}$	$\mu^{(k)}=\frac{b^kB(p+k/a,\,q-k/a)}{B(p,q)}$	$F(y;\phi)=B_u(p,q)$ with $u=(y/b)^a\big/\left[1+(y/b)^a\right]$	$F_k(y;\phi)=B_u(p+k/a,q-k/a)$ with $u=(y/b)^a\big/\left[1+(y/b)^a\right]$	Integral evaluated numerically
Beta-2 ($a=1$)	$f(y;\phi)=\frac{y^{p-1}}{b^pB(p,q)\left(1+(y/b)\right)^{p+q}}$	$\mu=bp/(q-1)$ $\mu^{(2)}=bp(p+1)/(q-1)(q-2)$	$F(y;\phi)=B_u(p,q)$ with $u=(y/b)\big/\left[1+(y/b)\right]$	$F_k(y;\phi)=B_u(p+k,q-k)$ with $u=(y/b)\big/\left[1+(y/b)\right]$	$G=\frac{2B(2p,2q-1)}{pB^2(p,q)}$
Singh-Maddala ($p=1$)	$f(y;\phi)=\frac{aqy^{a-1}}{b^a\left(1+(y/b)^a\right)^{1+q}}$	$\mu^{(k)}=\frac{b^k\Gamma(1+k/a,\,q-k/a)}{\Gamma(q)}$	$F(y;\phi)=1-\left[1+\left(\frac{y}{b}\right)^a\right]^{-q}$	$F_k(y;\phi)=B_u(1+k/a,q-k/a)$ with $u=(y/b)^a\big/\left[1+(y/b)^a\right]$	$G=1-\frac{\Gamma(q)\Gamma(2q-1/a)}{\Gamma(q-1/a)\Gamma(2q)}$
Dagum ($q=1$)	$f(y;\phi)=\frac{apy^{ap-1}}{b^{ap}\left(1+(y/b)^a\right)^{p+1}}$	$\mu^{(k)}=\frac{b^k\Gamma(p+k/a,\,1-k/a)}{\Gamma(p)}$	$F(y;\phi)=\left[1+\left(\frac{y}{b}\right)^{-a}\right]^{-p}$	$F_k(y;\phi)=B_u(p+k/a,1-k/a)$ with $u=(y/b)^a\big/\left[1+(y/b)^a\right]$	$G=\frac{\Gamma(p)\Gamma(2p+1/a)}{\Gamma(p+1/a)\Gamma(2p)}-1$
Generalized Gamma	$f(y;\phi)=\frac{a}{\beta^{ap}\Gamma(p)}y^{ap-1}\exp\left(-\left(\frac{y}{\beta}\right)^a\right)$	$\mu^{(k)}=\frac{\beta^k\Gamma(p+k/a)}{\Gamma(p)}$	$F(y;\phi)=G_u(p,\beta^a)$ with $u=y^a$	$F_k(y;\phi)=G_u(p+k/a,\beta^a)$	Integral evaluated numerically
Lognormal	$f(y;\phi)=\frac{1}{y\sqrt{2\pi}\sigma}\exp\left(-\frac{(\ln y-\mu)^2}{2\sigma^2}\right)$	$\mu^{(k)}=\exp\left(k\mu+\frac{k^2\sigma^2}{2}\right)$	$F(y;\phi)=\Phi\left(\frac{\ln(y)-\mu}{\sigma}\right)$	$F_k(y;\phi)=\Phi\left(\frac{\ln(y)-\mu-k\sigma^2}{\sigma}\right)$	$G=2\Phi\left(\frac{\sigma}{\sqrt{2}}\right)-1$

Table 2: Estimated coefficients from GB2 distributions

	China Rural					China Urban				
	CGR	Two-Step	SE	One-Step	SE	CGR	Two-Step	SE	One-Step	SE
z_1	22.334	22.437	0.108	22.438	0.108	52.402	52.557	0.262	52.562	0.262
z_2	27.972	27.882	0.079	27.882	0.079	65.688	65.016	0.209	65.014	0.210
z_3	33.607	33.476	0.061	33.476	0.061	85.274	84.522	0.202	84.522	0.202
z_4	36.411	36.263	0.060	36.263	0.060	102.100	101.420	0.197	101.420	0.197
z_5	41.893	41.828	0.072	41.828	0.072	118.740	118.600	0.210	118.600	0.210
z_6	47.344	47.403	0.082	47.403	0.082	136.760	137.190	0.239	137.190	0.239
z_7	55.631	55.743	0.110	55.743	0.110	157.800	158.460	0.292	158.460	0.292
z_8	69.582	69.710	0.153	69.710	0.153	184.610	185.380	0.393	185.380	0.393
z_9	83.573	83.600	0.213	83.600	0.213	223.940	223.490	0.614	223.480	0.614
z_{10}	112.010	111.320	0.387	111.320	0.387	298.540	293.870	1.180	293.880	1.181
z_{11}	140.940	139.820	0.756	139.820	0.757	386.530	381.070	2.657	380.990	2.657
b	32.887	25.098	4.888	25.377	4.808	122.260	107.220	4.981	107.340	4.849
p	3.774	6.094	1.954	5.970	1.884	1.371	2.437	0.409	2.366	0.389
q	1.761	2.209	0.407	2.180	0.398	1.217	1.810	0.259	1.755	0.247
a	1.666	1.410	0.176	1.422	0.175	2.424	1.817	0.168	1.849	0.168
	India Rural					**India Urban**				
	CGR	**Two-Step**	**SE**	**One-Step**	**SE**	**CGR**	**Two-Step**	**SE**	**One-Step**	**SE**
z_1	20.748	20.897	0.075	20.899	0.075	19.668	19.871	0.079	19.870	0.078
z_2	24.284	24.183	0.054	24.183	0.054	23.479	23.503	0.063	23.503	0.063
z_3	28.858	28.550	0.049	28.550	0.049	29.153	28.857	0.065	28.857	0.064
z_4	32.827	32.603	0.046	32.603	0.046	34.692	34.404	0.066	34.404	0.066
z_5	36.611	36.623	0.048	36.623	0.048	40.188	40.028	0.073	40.028	0.073
z_6	40.400	40.613	0.053	40.613	0.053	46.578	46.811	0.088	46.811	0.088
z_7	45.117	45.490	0.066	45.490	0.066	54.746	55.319	0.111	55.319	0.111
z_8	51.502	51.836	0.091	51.836	0.092	64.993	65.511	0.151	65.511	0.151
z_9	61.535	61.518	0.149	61.518	0.149	81.310	81.477	0.249	81.477	0.249
z_{10}	80.837	79.464	0.297	79.463	0.297	112.840	111.860	0.499	111.860	0.499
z_{11}	105.730	103.040	0.708	103.050	0.709	153.210	150.740	1.157	150.740	1.157
b	32.369	28.454	1.094	28.576	1.061	9.019	2.259	5.589	2.265	5.587
p	1.086	2.046	0.308	1.982	0.294	15.470	60.168	135.207	60.169	135.360
q	0.547	0.814	0.083	0.789	0.080	2.253	2.906	0.800	2.882	0.788
a	4.705	3.404	0.262	3.470	0.265	1.252	1.038	0.204	1.042	0.204

Table 2. Estimated coefficients from GB2 distributions (cont.)

	Pakistan					Russia				
	CGR	**Two-Step**	**SE**	**One-Step**	**SE**	**CGR**	**Two-Step**	**SE**	**One-Step**	**SE**
z_1	26.890	26.788	0.100	26.788	0.100	81.818	81.272	0.475	81.276	0.475
z_2	31.053	31.151	0.073	31.151	0.073	103.310	103.990	0.388	103.990	0.388
z_3	36.996	37.204	0.065	37.204	0.065	137.030	137.790	0.381	137.790	0.381
z_4	42.141	42.335	0.061	42.335	0.061	168.190	168.520	0.379	168.520	0.379
z_5	47.307	47.315	0.063	47.315	0.063	200.690	200.560	0.410	200.560	0.410
z_6	52.992	52.796	0.072	52.796	0.072	237.150	237.050	0.476	237.050	0.476
z_7	59.760	59.415	0.091	59.415	0.091	280.880	281.450	0.590	281.450	0.590
z_8	68.605	68.580	0.129	68.580	0.129	337.590	337.910	0.797	337.910	0.797
z_9	81.967	82.670	0.216	82.671	0.216	420.190	419.610	1.257	419.610	1.258
z_{10}	108.540	109.080	0.440	109.080	0.440	573.990	577.570	2.442	577.570	2.444
z_{11}	141.770	141.720	1.028	141.730	1.028	749.470	756.180	5.208	756.190	5.213
b	33.786	39.114	1.140	39.140	1.135	70.013	174.940	22.264	174.940	21.735
p	2.647	1.562	0.206	1.555	0.204	17.791	5.418	1.717	5.293	1.642
q	0.908	0.691	0.067	0.688	0.066	8.294	4.036	1.074	3.930	1.022
a	2.979	3.722	0.277	3.732	0.277	0.643	1.025	0.161	1.039	0.160
	Poland					**Brazil**				
	CGR	**Two-Step**	**SE**	**One-Step**	**SE**	**CGR**	**Two-Step**	**SE**	**One-Step**	**SE**
z_1	93.757	94.504	0.459	94.503	0.459	30.844	30.667	0.298	30.671	0.297
z_2	115.550	115.460	0.367	115.460	0.367	45.493	46.067	0.273	46.066	0.273
z_3	148.320	146.970	0.356	146.970	0.356	70.465	70.652	0.299	70.652	0.299
z_4	177.540	175.990	0.353	175.990	0.353	95.646	96.128	0.323	96.128	0.323
z_5	207.340	206.650	0.380	206.650	0.380	124.040	125.040	0.376	125.040	0.377
z_6	240.250	240.580	0.437	240.580	0.437	157.780	161.170	0.460	161.170	0.460
z_7	279.230	281.510	0.540	281.510	0.540	201.330	202.490	0.584	202.490	0.584
z_8	329.470	333.120	0.720	333.120	0.720	261.620	256.870	0.847	256.870	0.847
z_9	402.520	403.370	1.098	403.370	1.098	358.960	355.000	1.561	355.000	1.560
z_{10}	539.020	531.820	2.044	531.820	2.044	566.870	577.410	3.804	577.450	3.804
z_{11}	696.790	686.110	4.400	686.110	4.400	854.590	892.510	10.975	892.450	10.972
b	170.600	133.470	23.178	133.470	22.926	123.720	157.830	8.564	158.240	8.589
p	3.588	5.990	1.951	5.930	1.914	2.754	1.539	0.240	1.533	0.239
q	2.260	3.081	0.695	3.039	0.678	2.220	1.525	0.260	1.523	0.259
a	1.519	1.227	0.176	1.235	0.176	1.001	1.311	0.134	1.313	0.134

Table 3. *p*-values from *J*-statistics

	GB2	B2	SM	Dagum	GGamma	LogN
China Rural	0.8463	0.3142	0.0000	0.0001	0.0000	0.0000
China Urban	0.0297	0.0000	0.0000	0.0000	0.0000	0.0000
India Rural	0.0944	0.0000	0.0000	0.0121	0.0000	0.0000
India Urban	0.0110	0.0141	0.0000	0.0000	0.0000	0.0000
Pakistan	0.1404	0.0000	0.0042	0.0041	0.0000	0.0000
Russia	0.4973	0.5933	0.0000	0.0000	0.0000	0.0000
Poland	0.0555	0.0424	0.0000	0.0000	0.0000	0.0000
Brazil	0.0000	0.0000	0.0000	0.0000	0.0000	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	1.548	1.553	1.717	1.709	1.332	1.332
0.052	0.054	2.083	2.052	2.325	2.310	1.774	1.788
0.171	0.174	2.442	2.416	5.283	5.240	4.143	4.115
1.007	1.002	2.678	2.665	6.480	6.416	5.270	5.225
2.105	2.091	2.938	2.925	7.095	7.057	5.822	5.776
3.334	3.317	3.134	3.124	7.269	7.246	6.888	6.877
2.060	2.051	3.338	3.332	8.520	8.506	8.364	8.407
4.604	4.591	3.580	3.587	9.926	9.888	9.301	9.360
5.055	5.059	3.791	3.801	11.800	11.699	11.897	11.901
7.985	8.001	4.032	4.052	13.785	13.552	15.105	15.017
12.726	12.741	4.296	4.316	9.147	8.921	10.674	10.477
10.788	10.781	4.545	4.561	16.654	17.456	19.431	19.726
8.873	8.836	4.881	4.895				
6.902	6.852	5.187	5.188				
5.683	5.624	5.668	5.663				
4.314	4.258	6.167	6.138				
24.328	24.554	6.820	6.727				
		7.777	7.632				
		9.292	9.091				
		15.802	16.280				

Pakistan		Russia		Poland		Brazil	
Estimated	Observed	Estimated	Observed	Estimated	Observed	Estimated	Observed
1.694	1.700	1.021	1.017	1.261	1.264	0.343	0.343
2.192	2.212	1.510	1.528	1.753	1.755	0.664	0.698
2.483	2.495	1.831	1.851	2.065	2.057	0.911	0.936
2.700	2.720	2.111	2.125	2.332	2.320	1.139	1.157
2.910	2.929	2.373	2.378	2.581	2.565	1.364	1.372
3.110	3.134	2.630	2.635	2.823	2.800	1.596	1.621
3.295	3.312	2.890	2.887	3.065	3.044	1.882	1.894
3.488	3.501	3.157	3.146	3.313	3.306	2.068	2.089
3.687	3.699	3.436	3.424	3.572	3.570	2.492	2.529
3.899	3.905	3.734	3.719	3.847	3.850	2.496	2.547
4.127	4.133	4.056	4.051	4.143	4.154	2.979	3.064
4.378	4.378	4.411	4.408	4.469	4.490	3.363	3.443
4.653	4.659	4.809	4.818	4.834	4.878	4.105	4.114
4.991	5.017	5.266	5.277	5.252	5.305	4.045	4.037
5.387	5.443	5.807	5.818	5.746	5.792	5.098	5.080
5.897	5.991	6.470	6.470	6.351	6.361	5.799	5.792
6.537	6.664	7.330	7.377	7.137	7.108	6.960	7.000
7.535	7.626	8.548	8.572	8.249	8.143	8.701	8.890
9.313	9.332	10.593	10.732	10.125	9.915	12.121	12.755
17.727	17.150	18.018	17.767	17.085	17.322	31.875	30.639

Table 5. Root-Mean-Square errors

	China Rural	China Urban	India Rural	India Urban	Pakistan	Russia	Poland	Brazil
GB2	0.0607	0.1230	0.2531	0.1098	0.1367	0.0661	0.0779	0.3156
B2	0.1609	0.2549	0.8231	0.0907	0.4354	0.0515	0.1150	0.1827
SM	1.0281	0.1481	0.0862	1.3835	0.3723	1.2971	0.5378	0.2122
Dagum	0.5643	0.2045	0.3987	0.9539	0.1184	1.0691	0.5789	0.4505
GGamma	1.2267	0.7425	1.9744	1.8218	1.2565	0.3649	0.6740	1.3612
LogN	0.8973	0.5883	1.1341	1.0306	0.6445	0.3463	0.5108	0.9255

Table 6. Gini and Theil coefficients and their standard errors

	B2		SM		Dagum		GB2	
	Gini	SE	Gini	SE	Gini	SE	Gini	SE
China Rural	0.355	0.003	0.386	0.006	0.375	0.004	0.357	0.004
China Urban	0.344	0.003	0.352	0.004	0.355	0.004	0.345	0.005
India Rural	0.287	0.002	0.306	0.004	0.295	0.003	0.298	0.004
India Urban	0.376	0.004	0.408	0.007	0.399	0.005	0.374	0.004
Pakistan	0.299	0.002	0.323	0.004	0.307	0.003	0.315	0.005
Russia	0.376	0.003	0.401	0.005	0.404	0.004	0.376	0.004
Poland	0.347	0.003	0.364	0.005	0.367	0.004	0.348	0.003
Brazil	0.546	0.006	0.540	0.008	0.523	0.006	0.555	0.008
	Theil	SE	Theil	SE	Theil	SE	Theil	SE
China Rural	0.230	0.006	0.326	0.015	0.289	0.009	0.238	0.008
China Urban	0.206	0.005	0.240	0.008	0.243	0.007	0.216	0.007
India Rural	0.145	0.003	0.198	0.007	0.169	0.004	0.178	0.007
India Urban	0.271	0.008	0.393	0.019	0.349	0.012	0.268	0.010
Pakistan	0.159	0.003	0.225	0.009	0.185	0.005	0.205	0.009
Russia	0.249	0.006	0.328	0.009	0.332	0.010	0.250	0.006
Poland	0.213	0.005	0.266	0.010	0.268	0.008	0.216	0.006
Brazil	0.614	0.023	0.633	0.033	0.593	0.021	0.679	0.057

Table 7. Poverty measures (%) and their standard errors

	B2		SM		Dagum		GB2	
	estimates	se	estimates	se	estimates	se	estimates	se
China R								
HCR	24.667	0.363	23.608	2.177	23.944	0.370	24.363	0.381
FGT2	2.172	0.062	2.409	0.679	2.181	0.067	2.169	0.063
China U								
HCR	1.559	0.097	2.395	0.502	2.037	0.113	1.806	0.112
FGT2	0.076	0.008	0.242	0.073	0.176	0.016	0.124	0.015
India R								
HCR	41.490	0.393	42.905	1.357	42.735	0.434	42.833	0.440
FGT2	3.704	0.073	3.498	0.658	3.454	0.070	3.432	0.071
India U								
HCR	35.553	0.416	34.968	1.828	35.403	0.422	35.412	0.418
FGT2	3.655	0.079	3.9326	0.840	3.605	0.081	3.662	0.080
Pakistan								
HCR	22.194	0.338	20.262	1.743	21.038	0.350	20.634	0.365
FGT2	1.306	0.037	1.263	0.505	1.220	0.042	1.218	0.044
Russia								
HCR	0.299	0.033	0.637	0.329	0.471	0.045	0.302	0.044
FGT2	0.017	0.003	0.070	0.061	0.043	0.006	0.017	0.004
Poland								
HCR	0.046	0.009	0.347	0.187	0.165	0.022	0.062	0.016
FGT2	0.001	0.000	0.033	0.029	0.011	0.002	0.002	0.001
Brazil								
HCR	7.530	0.100	7.226	1.645	7.691	0.224	7.434	0.219
FGT2	1.325	0.053	1.488	N/A	1.763	0.085	1.388	0.076

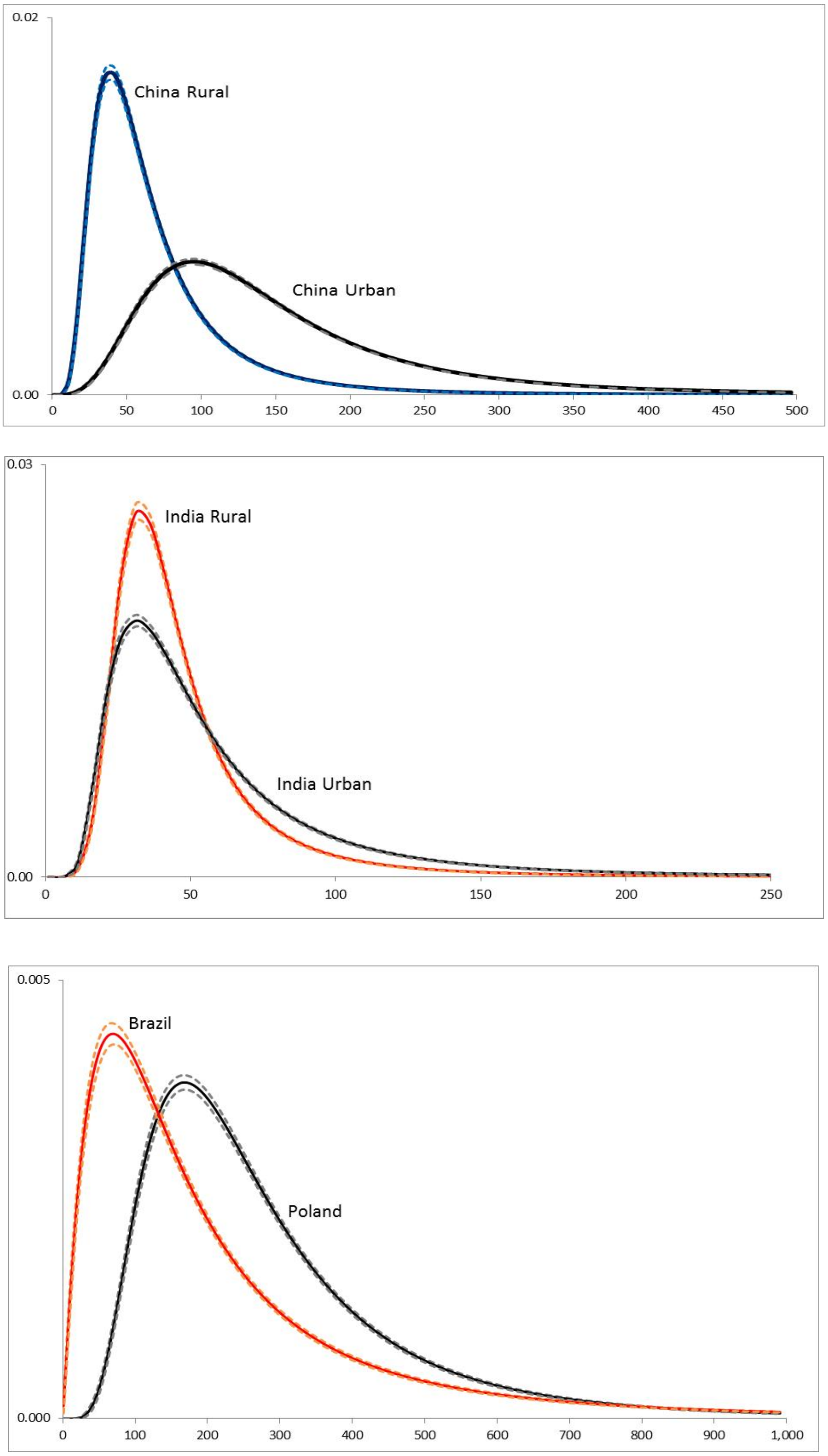

Figure 1. GB2-estimated pdfs for selected regions

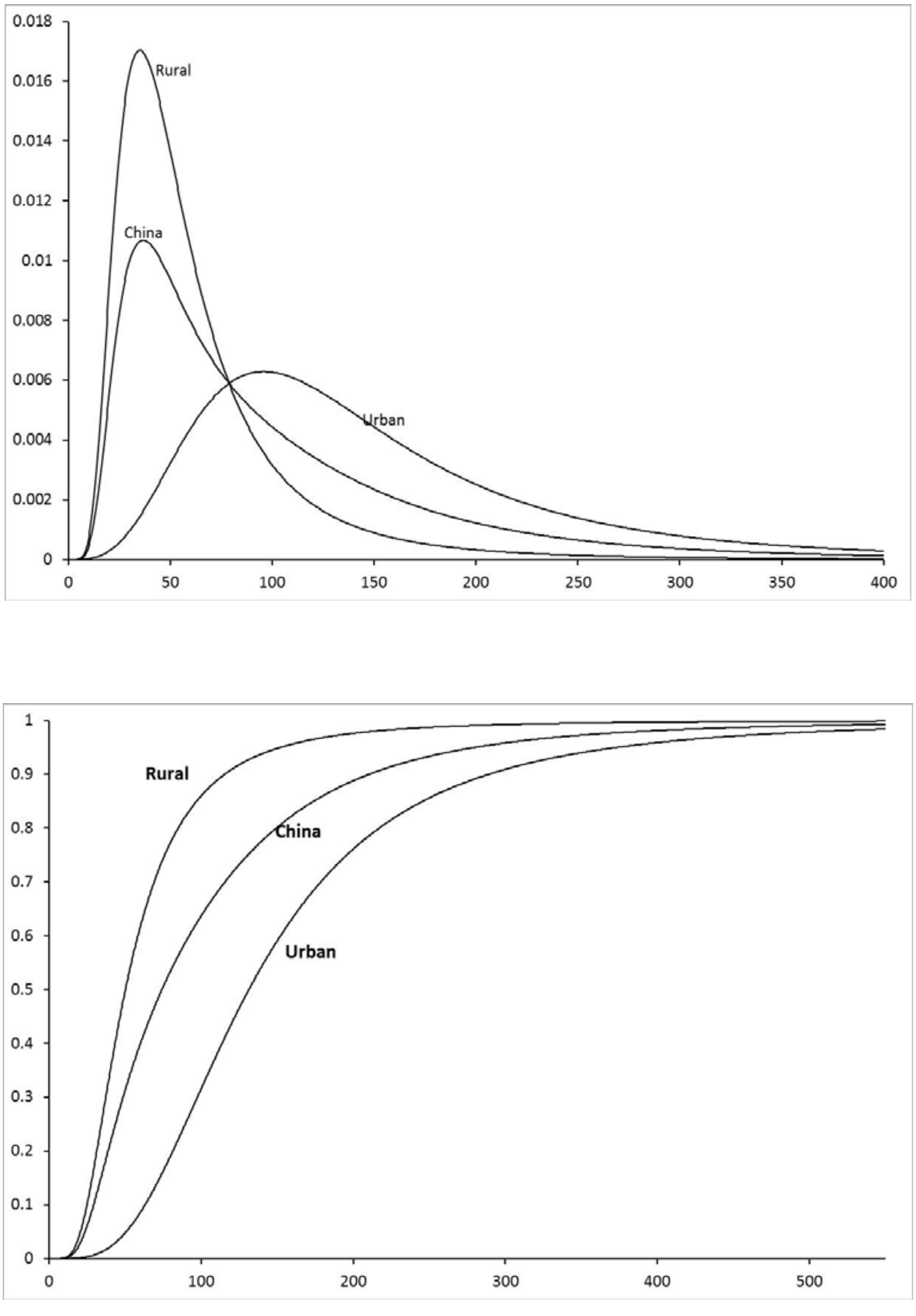

Figure 2. GB2-estimated pdf and cdf for all of China