THE UNIVERSITY OF MELBOURNE

Department of Economics

Working Paper Series

Pareto-Lognormal Income Distributions: Inequality and Poverty Measures, Estimation and Performance

William E. Griffiths and Gholamreza Hajargasht

June 2012

Research Paper Number 1149

ISSN: 0819 2642
ISBN: 978 0 7340 4499 0

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

Pareto-Lognormal Income Distributions:
Inequality and Poverty Measures, Estimation and Performance

Gholamreza Hajargasht and William E. Griffiths

Department of Economics
University of Melbourne, Australia

June 26, 2012

ABSTRACT

The (double) Pareto-lognormal is an emerging parametric distribution for income that has a sound underlying generating process, good theoretical properties, and favourable evidence of its fit to data. We extend existing results for this distribution in 3 directions. We derive closed form formula for its moment distribution functions, and for various inequality and poverty measures. We show how it can be estimated from grouped data using the GMM method developed in Hajargasht et al. (2012). Using grouped data from ten countries, we compare its performance with that of another leading 4-parameter income distribution, the generalized beta-2 distribution. The results confirm that both distributions provide a good fit, with the double Pareto-lognormal distribution outperforming the beta distribution in some but not all cases.

Keywords: GB2 distribution, GMM, moment distributions, double-Pareto.

JEL Classification: C13, C16, D31

I. Introduction

For more than a century considerable effort has been directed towards providing new functional forms for income distributions. A major motivation of such studies has been to suggest functions with theoretical properties and statistical fits that are better than those of functions already appearing in the literature. The large number of functional forms that has been suggested includes, but is not limited to, the lognormal, gamma, Pareto, Weibull, Dagum, Singh-Maddala, beta-2, and generalized beta-2 distributions, and the quest continues. See Kleiber and Kotz (2003) for a comprehensive review. An emerging 4-parameter distribution for income and other phenomena is the double Pareto-lognormal distribution (hereafter dPLN) developed by Reed (2003) and studied further by Reed and Jorgensen (2004)[1]. It has good properties with a sound theoretical justification (Reed 2003), and, according to Reed and Wu (2008), provides a very good fit to income data.

The purpose of this paper is to provide some further results for this distribution. We derive closed form solutions for various inequality and poverty measures in terms of the parameters of the dPLN distribution and show how they relate to the corresponding inequality and poverty measures for the lognormal and Pareto distributions. We also provide the moment distribution functions which are required for estimating the dPLN distribution from grouped data within the GMM framework developed in Hajargasht el. (2012). Results from estimating this model are compared with those from estimating another leading 4-parameter distribution known as the generalized beta of the second kind (hereafter GB2), developed by McDonald (1984) and McDonald and Xu (1995)[2]. Using grouped data from ten regions (China rural, China urban, India rural, India urban, Pakistan, Russia, Poland, Brazil, Nigeria and Iran), we estimate the GB2, a 3-parameter Pareto-lognormal, and the dPLN distributions, and compare their performance in terms of goodness of fit. The results

[1] - According to “GoogleScholar”, these papers have 107 and 114 citations, respectively.
[2] - According to “GoogleScholar”, these papers have 420 and 215 citations, respectively.

suggest that all three distributions provide good fits, and there is not one particular distribution that dominates the others over all data sets.

The paper is organized as follows. In Section II we briefly review the dPLN distribution. Expressions are provided for its moment distribution functions needed for GMM estimation and for various inequality and poverty measures. The GB2 distribution that we later compare with the dPLN is reviewed in Section III. In Section IV we summarize the GMM methodology developed in Hajargasht et al. (2012) for estimating the parameters of a general income distribution, and show how it can be applied to the dPLN distribution. Section V contains a description of the data used to illustrate the theoretical framework, and the results. The results include parameter estimates and their standard errors, test results for excess moment conditions, mean-square-error comparisons for goodness-of-fit, and Gini and Theil coefficients. Concluding remarks are provided in Section VI.

II. Pareto-lognormal income distributions

The probability density function (pdf) of the dPLN distribution with parameters (m,σ,α,β) is

$$f^{dPLN}(y;m,\sigma,\alpha,\beta)=\frac{\alpha\beta}{(\alpha+\beta)y}\phi\left(\frac{\ln y-m}{\sigma}\right)\left\{R\left(x_1\right)+R\left(x_2\right)\right\} \tag{1}$$

where $R(t)=\left[1-\Phi(t)\right]/\phi(t)$ is a Mill's ratio, $\phi(.)$ and $\Phi(.)$ are, respectively, the pdf and cumulative distribution function (cdf) for a standard normal random variable,

$$x_1=\alpha\sigma-\frac{\ln y-m}{\sigma} \quad \text{and} \quad x_2=\beta\sigma+\frac{\ln y-m}{\sigma}$$

The cdf of the dPLN distribution has been derived as (Reed and Jorgensen 2004)

$$F^{dPLN}(y;m,\sigma,\alpha,\beta)=\int_0^y f(t)dt=\Phi\left(\frac{\ln y-m}{\sigma}\right)-\phi\left(\frac{\ln y-m}{\sigma}\right)\left(\frac{\beta R(x_1)-\alpha R(x_2)}{\alpha+\beta}\right) \tag{2}$$

Two attractive features of the dPLN distribution are that it can be derived from reasonable assumptions about the generation of incomes, and that it accords with empirical observation. It is derived by relaxing some of the assumptions of earlier models of income generation. Gibrat (1931) showed that if the income of each individual at time *t* grows randomly but proportionately with respect to their income in the previous time, the emerging distribution is lognormal. Variations on this theme were shown to generate features of the Pareto distribution (Champernowne 1953, Mandelbrot 1960). Reed (2001, 2003) and Reed and Jorgensen (2004) developed and extended these earlier models, using the assumption that income generation follows a geometric Brownian motion (GBM) process, the continuous time version of assumptions made by Gibrat and Champernowne. Within this framework, if income growth is modeled by GBM with a fixed initial state, then after a fixed time *T*, incomes follow a lognormal distribution. If we relax the assumption that *T* is fixed, recognizing that income earning ceases at different times for different individuals, assume that *T* follows an exponential distribution, and assume that the all income earners start with the same income, normalized to unity, then the resulting income distribution is a double-Pareto distribution with pdf

$$f^{dP}(y;\alpha,\beta)=\begin{cases}\dfrac{\alpha\beta}{(\alpha+\beta)}y^{\beta-1} & \text{for } 0<y\le 1\\[2ex] \dfrac{\alpha\beta}{(\alpha+\beta)}y^{-\alpha-1} & \text{for } y>1\end{cases}\qquad \alpha,\beta>0 \qquad (3)$$

This pdf exhibits Pareto behaviour in both tails. It is proportional to *y* raised to a negative power in the right tail; in the left tail it is proportional to *y* raised to a positive or negative power depending on whether or not $\beta>1$. It has also been derived by Benhabib and Zhu (2008) as an emerging distribution for wealth in an elaborate overlapping generation model with utility maximizing agents.

The dPLN distribution in (1) is a generalization of the double-Pareto distribution obtained by relaxing the assumption that all income earners have the same initial income, and assuming instead that starting incomes are lognormally distributed, and then evolve according to GBM. It can be derived as the distribution of the product of a double-Pareto random variable with parameters (α,β) and a lognormal random variable with parameters (m,σ). It has Pareto behaviour in both the upper and lower tails. That is,

$$f(y) \to c_1 y^{-\alpha-1} \ (y \to \infty) \quad \text{and} \quad f(y) \to c_2 y^{\beta-1} \ (y \to 0)$$

where c_1 and c_2 are constants. Apart from the tail behavior, it is shaped somewhat like a lognormal distribution, providing $\beta > 1$; it is monotonically decreasing for $0 < \beta < 1$. The desirability of Pareto behavior in the upper tail has been well documented; there have been claims that similar behavior can be observed in the left tail (Champernowne 1953). If $(\alpha,\beta) \to \infty$, the dPLN tends to a lognormal distribution and for large values of (α,β), it is close to lognormal. If we have only $\beta \to \infty$, we obtain the 3-parameter Pareto-lognormal (PLN) distribution that shows Pareto behavior only in the upper tail. This distribution has studied by Colombi (1990). We include it as one of the distributions we estimate in our empirical work. Its pdf and cdf are given respectively by

$$f^{PLN}(y;m,\sigma,\alpha) = \frac{\alpha}{y}\phi\left(\frac{\ln y - m}{\sigma}\right)R\left(x_1\right) \tag{4}$$

$$F^{PLN}(y;m,\sigma,\alpha) = \Phi\left(\frac{\ln y - m}{\sigma}\right) - \phi\left(\frac{\ln y - m}{\sigma}\right)R(x_1) \tag{5}$$

For estimating the dPLN and PLN distributions using the GMM method of Hajargasht et al. (2012), and for finding closed form formulas for some poverty measures, we need the dPLN and PLN moments and moment distribution functions of orders one and two. From Reed and Jorgensen (2004), the *k*-th moments (which exist for $\alpha > k$) are given by

$$\mu_k^{dPLN} = \frac{\alpha\beta}{(\alpha-k)(\beta+k)}\exp\left\{km + \frac{k^2\sigma^2}{2}\right\} \tag{6}$$

$$\mu_k^{PLN} = \frac{\alpha}{\alpha-k}\exp\left\{km + \frac{k^2\sigma^2}{2}\right\} \tag{7}$$

We show in Appendix A that the moment distribution functions are given by

$$F_k^{dPLN}(y;m,\sigma,\alpha,\beta) = \frac{1}{\mu_k^{dPLN}} \int_0^y t^k f^{dPLN}(t)dt$$
$$= \Phi\left(\frac{\ln y - m}{\sigma} - k\sigma\right) - \phi\left(\frac{\ln y - m}{\sigma} - k\sigma\right)\left(\frac{(\beta+k)R(x_1) - (\alpha-k)R(x_2)}{\alpha+\beta}\right) \tag{8}$$

$$F_k^{PLN}(y;m,\sigma,\alpha) = \frac{1}{\mu_k^{PLN}} \int_0^y t^k f^{PLN}(t)dt$$
$$= \Phi\left(\frac{\ln y - m}{\sigma} - k\sigma\right) - \phi\left(\frac{\ln y - m}{\sigma} - k\sigma\right)R(x_1) \tag{9}$$

Because income distributions are often used to assess inequality and poverty, it is desirable to express inequality and poverty measures in terms of the parameters of those distributions. Such expressions show explicitly the dependence of inequality or poverty on each of the parameters, and they facilitate calculation of standard errors for inequality or poverty indices. In Appendix B we derive expressions for the Gini, generalized entropy, Theil, and Atkinson inequality measures in terms of the parameters of the dPLN and PLN distributions. Details of these measures can be found in a variety of sources. See, for example, Cowell (2007), and Kleiber and Kotz (2003). Here we report the results for the Gini and Theil measures; results for the others are given in Appendix B.

The most widely used inequality measure is the Gini coefficient. For the dPLN distribution, it is given by

$$G^{dPLN} = -1 + 2\left[\Phi\left(\frac{\sigma}{\sqrt{2}}\right) + \frac{\beta(1+\beta)\exp\{\alpha(\alpha-1)\sigma^2\}}{(\alpha+\beta)(2\alpha-1)(1-\alpha+\beta)}\Phi\left(\frac{(1-2\alpha)\sigma}{\sqrt{2}}\right) - \frac{\alpha(\alpha-1)\exp\{\beta(\beta+1)\sigma^2\}}{(\alpha+\beta)(1+2\beta)(1-\alpha+\beta)}\Phi\left(\frac{(-1-2\beta)\sigma}{\sqrt{2}}\right)\right] \tag{10}$$

and for the PLN model it simplifies to

$$G^{PLN} = \frac{2\exp\{\alpha(\alpha-1)\sigma^2\}}{(2\alpha-1)}\Phi\left(\frac{(1-2\alpha)\sigma}{\sqrt{2}}\right) + 2\Phi\left(\frac{\sigma}{\sqrt{2}}\right) - 1 \tag{11}$$

The Gini coefficient for lognormal distribution with parameter σ is $2\Phi\left(\sigma/\sqrt{2}\right)-1$, that for a left tail Pareto is $(1+2\beta)^{-1}$, and that for a right tail Pareto is $(2\alpha-1)^{-1}$. Thus, G^{dPLN} and G^{PLN} are both equal to the Gini coefficient for a lognormal distribution plus weighted adjustment(s) for the Pareto tails.

A second commonly used inequality measure is the Theil index defined as

$$T=\int_0^\infty \frac{y}{\mu}\ln\left(\frac{y}{\mu}\right) f(y)dy$$

In Appendix B we show that, for the dPLN distribution,

$$T^{dPLN}=\ln\left(\frac{\alpha-1}{\alpha}\right)+\frac{1}{(\alpha-1)}+\ln\left(\frac{1+\beta}{\beta}\right)-\frac{1}{(\beta+1)}+\frac{\sigma^2}{2} \tag{12}$$

and for the PLN distribution it becomes

$$T^{PLN}=\ln\left(\frac{\alpha-1}{\alpha}\right)+\frac{1}{\alpha-1}+\frac{\sigma^2}{2} \tag{13}$$

In this case we have the interesting result that the Theil measure is equal to the sum of the Theil measure for a lognormal distribution $\left(\sigma^2/2\right)$ and the relevant measure(s) for the Pareto distributions, $\ln\left((\alpha-1)/\alpha\right)+\left(\alpha-1\right)^{-1}$ and $\ln\left((\beta+1)/\beta\right)-\left(\beta+1\right)^{-1}$.

We provide details for five poverty measures – the headcount ratio, poverty gap, Foster-Greer-Thorbecke $FGT(2)$, Atkinson and Watts measures. With the exception of the Watts measure, all these quantities can be written in terms of the moments and moment distribution functions provided in equations (2) and (5)-(9). Thus, for computing in these cases, it is sufficient to give general expressions that hold for all distributions and to rely on the earlier moment and moment-distribution equations for the dPLN and PLN distributions. The Watts measure requires more work. Derivation of expressions for this measure for the dPLN and PLN distributions are given in Appendix C. General details of the various measures and many others are conveniently summarized in Kakwani (1999).

The head-count ratio is given by the percentage of population below a target level of income y_P, called the poverty line. That is, $F(y_P)$. The poverty gap is a modification that takes into account how far the poor are below the poverty line. It is given by

$$PG = \int_0^{y_P} \frac{y_P - y}{y_P} f(y)dy = F(y_P) - \frac{\mu}{y_P} F_1(y_P) \tag{14}$$

The measure $FGT(2)$ is similar, but weights the gap between income and the poverty line more heavily. It is given by

$$FGT(2) = \int_0^{y_P} \left(\frac{y_P - y}{y_P} \right)^2 f(y)dy = F(y_P) - \frac{2\mu}{y_P} F_1(y_P) + \frac{\mu_2}{y_P^2} F_2(y_P) \tag{15}$$

Another generalization of the poverty gap is the Atkinson index given by

$$A(e) = \frac{1}{e} \int_0^{y_P} \left[1 - \left(\frac{y}{y_P} \right)^e \right] f(y)dy = \frac{1}{e} \left[F(y_P) - \frac{\mu_e}{y_P^e} F_e(y_P) \right] \tag{16}$$

where e is a parameter for the degree of inequality aversion.

The Watts index, shown by Zheng (1993) to satisfy all axiomatic conditions, is given by

$$W = \int_0^{y_P} \ln\left(y_P / y \right) f(y)dy$$

In Appendix C we show that, for the dPLN distribution,

$$W^{dPLN} = \frac{\alpha}{\beta(\alpha+\beta)} F^{PLN(\beta)}(y_P) - \frac{\beta}{\alpha(\alpha+\beta)} F^{PLN(\alpha)}(y_P) + \sigma \left\{ \frac{\ln y_P - m}{\sigma} \Phi\left(y_P \right) + \phi\left(\frac{\ln y_P - m}{\sigma} \right) \right\} \tag{17}$$

For the DPL distribution it becomes

$$W^{PLN} = -\frac{1}{\alpha} F^{PLN(\alpha)}(y_P) + \sigma \left\{ \frac{\ln y_P - m}{\sigma} \Phi\left(y_P \right) + \phi\left(\frac{\ln y_P - m}{\sigma} \right) \right\} \tag{18}$$

where $F^{PLN(\alpha)}(y)$ is the cdf of the PLN distribution with Pareto behaviour in the right tail, $F^{PLN(\beta)}(y)$ is cdf of the PLN distribution with Pareto behaviour in the left tail, and the last term in both equations is the Watts index for a lognormal distribution. See Muller (2001).

Once the parameters of the distributions have been estimated, these expressions are convenient ones for computing the various poverty measures.

III. The generalized beta distribution of the second kind

For a distribution with which to compare the performance of the PLN and dPLN distributions, we chose the GB2 distribution whose pdf with positive parameters (a,b,p,q) is

$$f(y;a,b,p,q)=\frac{ay^{ap-1}}{b^{ap}B(p,q)\left(1+\left(\frac{y}{b}\right)^{a}\right)^{p+q}} \qquad y>0 \tag{19}$$

where $B(\cdot,\cdot)$ is the beta function. Like the dPLN, the GB2 income distribution is derived from a reasonable economic model. Parker (1999) shows how it arises from a neoclassical model with optimizing firm behaviour under uncertainty where the shape parameters p and q become functions of the output-labor elasticity and the elasticity of income returns with respect to human capital. Another very useful feature of the GB2 is that it nests many popular income distributions as special or limiting cases, including the beta-2, Singh-Maddala, Dagum, generalized gamma and lognormal distributions. Its detailed properties are given by McDonald (1984), McDonald and Xu (1995), and Kleiber and Kotz (2003), with further information on inequality measures provided by McDonald and Ransom (2008) and Jenkins (2009). Hajargasht et al. (2012) show how to find GMM estimates of its parameters from grouped data. The quantities needed for GMM estimation, and also for computing poverty measures, are the moments and moment distribution functions given respectively by

$$\mu_k^{GB2}=\frac{b^{k}B\left(p+k/a\,,\,q-k/a\right)}{B\left(p,q\right)} \tag{20}$$

$$F_k^{GB2}(y;a,b,p,q) = F^{GB2}\left(y;a,b,p+\frac{k}{a},q-\frac{k}{a}\right) = B_u\left(p+\frac{k}{a},q-\frac{k}{a}\right) \tag{21}$$

where $u = (y/b)^a \Big/ \left[1+(y/b)^a\right]$ and $B_u(p,q) = \int_0^u t^{p-1}(1-t)^{q-1}dt \Big/ B(p,q)$ is the cdf for a standard beta random variable defined on the $(0,1)$ interval. Note that the cdf for the GB2 is given by $F^{GB2}(y;a,b,p,q) = F_0^{GB2}(y;a,b,p,q)$. Also, the moments of order k exist only if $-ap < k < aq$.

IV. The GMM estimator

To describe the general form of the GMM estimator for estimating an income distribution from grouped data, we begin with a sample of T observations $(y_1, y_2, \ldots, y_T)$, assumed to be randomly drawn from an income distribution $f(y;\phi)$, where ϕ is a vector of unknown parameters, 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 $z_N = \infty$. The available data are the mean class incomes $\bar{y}_1, \bar{y}_2, \ldots, \bar{y}_N$ and the proportions of observations in each class $c_1, c_2, \ldots, c_N$. The estimation problem is to estimate ϕ, along with unknown class limits $z_1, z_2, \ldots, z_{N-1}$. To tackle this problem, Hajargasht et al. (2012) proposed a GMM estimator given by

$$\hat{\boldsymbol{\theta}} = \arg\min_{\boldsymbol{\theta}} \mathbf{H}(\boldsymbol{\theta})'\boldsymbol{\Omega}\,\mathbf{H}(\boldsymbol{\theta}) \tag{22}$$

where $\boldsymbol{\theta} = (z_1, z_2, \ldots, z_{N-1}, \phi')'$,

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

is a set of moments set up for the c_i and $\bar{y}_i$ such that $E\left[\mathbf{H}(\boldsymbol{\theta})\right] = \mathbf{0}$, and $\boldsymbol{\Omega}$ is the weight matrix

$$\boldsymbol{\Omega} = \left[\operatorname{plim}\frac{1}{T}\sum_{t=1}^{T}\mathbf{h}(y_t,\boldsymbol{\theta})\mathbf{h}(y_t,\boldsymbol{\theta})'\right]^{-1} \tag{23}$$

The first N elements in the $(2N \times 1)$ vector $\mathbf{h}(y_t, \boldsymbol{\theta})$ are

$$g_i(y_t) - k_i(\boldsymbol{\theta}) \qquad i = 1, 2, \ldots, N \tag{24}$$

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

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

and

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

The second set of N elements in the $(2N \times 1)$ vector $\mathbf{h}(y_t, \boldsymbol{\theta})$ are

$$y_t g_i(y_t) - \lambda_i(\boldsymbol{\theta}) \qquad i = 1, 2, \ldots, N \tag{25}$$

where

$$\lambda_i(\boldsymbol{\theta}) = \int_{z_{i-1}}^{z_i} y f(y;\boldsymbol{\phi})dy = \int_0^\infty y g_i(y) f(y;\boldsymbol{\phi}) dy = E\left[y g_i(y)\right] \tag{26}$$

With these definitions, we can write

$$\mathbf{H}(\boldsymbol{\theta}) = \frac{1}{T}\sum_{t=1}^{T}\mathbf{h}(y_t;\boldsymbol{\theta}) == \begin{bmatrix} \mathbf{c} - \mathbf{k} \\ \tilde{\mathbf{y}} - \boldsymbol{\lambda} \end{bmatrix} \tag{27}$$

where $\mathbf{c}$, $\mathbf{k}$, and $\boldsymbol{\lambda}$ are N-dimensional vectors containing the elements c_i, k_i and λ_i, respectively, and, using $T_i = \sum_{t=1}^{T} g_i(y_t)$, the i-th element of $\tilde{\mathbf{y}}$ is given by

$$\tilde{y}_i = c_i \bar{y}_i = \frac{T_i}{T}\frac{1}{T_i}\sum_{t=1}^{T} y_t g_i(y_t) = \frac{1}{T}\sum_{t=1}^{T} y_t g_i(y_t) \tag{28}$$

Hajargasht et al. (2012) show that the weight matrix can be written as

$$\boldsymbol{\Omega} = \begin{bmatrix} \mathbf{D}(\omega_1) & -\mathbf{D}(\omega_3) \\ -\mathbf{D}(\omega_3) & \mathbf{D}(\omega_2) \end{bmatrix} \tag{29}$$

where $\mathbf{D}(\boldsymbol{\omega})$ denotes a diagonal matrix with elements of the vector $\boldsymbol{\omega}$ on the diagonal. The elements in the vectors $\boldsymbol{\omega}_1$, $\boldsymbol{\omega}_2$, and $\boldsymbol{\omega}_3$ are

$\omega_{1i} = \lambda_i^{(2)} / v_i$, $\omega_{2i} = k_i / v_i$, and $\omega_{3i} = \lambda_i / v_i$, where

$$\lambda_i^{(2)}(\boldsymbol{\theta}) = \int_{z_{i-1}}^{z_i} y^2 f(y;\phi) dy = \int_0^{\infty} y^2 g_i(y) f(y;\phi) dy = E\left[y^2 g_i(y) \right] \tag{30}$$

and $v_i = k_i \lambda_i^{(2)} - \lambda_i^2$. Collecting all these various terms, the GMM objective function in (22) can be written as

$$\begin{aligned} GMM &= \mathbf{H}'\boldsymbol{\Omega}\mathbf{H} = \begin{bmatrix} \mathbf{c}-\mathbf{k} \\ \tilde{\mathbf{y}}-\boldsymbol{\lambda} \end{bmatrix}' \begin{bmatrix} \mathbf{D}(\boldsymbol{\omega}_1) & -\mathbf{D}(\boldsymbol{\omega}_3) \\ -\mathbf{D}(\boldsymbol{\omega}_3) & \mathbf{D}(\boldsymbol{\omega}_2) \end{bmatrix} \begin{bmatrix} \mathbf{c}-\mathbf{k} \\ \tilde{\mathbf{y}}-\boldsymbol{\lambda} \end{bmatrix} \\ &= \sum_{i=1}^{N} \omega_{1i} (c_i - k_i)^2 + \sum_{i=1}^{N} \omega_{2i} (\tilde{y}_i - \lambda_i)^2 - 2\sum_{i=1}^{N} \omega_{3i} (c_i - k_i)(\tilde{y}_i - \lambda_i) \end{aligned} \tag{31}$$

Equations (22) to (31) are a useful summary of the results in Hajargasht et al. (2012), and equation (31) is a computationally convenient expression for finding the GMM estimator for θ. However, the above results mask much of the development that led to the final result in (31). Because $\sum_{i=1}^{N} k_i(\boldsymbol{\theta}) = \sum_{i=1}^{N} c_i = 1$, one of the moment conditions in $\mathbf{H}(\boldsymbol{\theta})$ (see equation (27)) is redundant, and, unless we consider a generalized inverse, the inverse defined in (23) does not exist. Hajargasht et al. set up $(2N-1)$ non-redundant moment conditions, derived the corresponding matrix $\boldsymbol{\Omega}^{-1}$, and found its inverse $\boldsymbol{\Omega}$. They then showed that the relatively complicated objective function, expressed in terms of a $[(2N-1)\times 1]$ vector $\mathbf{H}$, and a $[(2N-1)\times(2N-1)]$ matrix $\boldsymbol{\Omega}$, can be written much more simply in terms of the $2N$-dimensional versions of $\mathbf{H}$ and $\boldsymbol{\Omega}$ given in equation (31). If K is the dimension of ϕ (the number of unknown parameters in the income density), there are a total of $(N-1+K)$ unknown parameters. Given there are $(2N-1)$ non-redundant moment conditions, the number of excess moment conditions is $(N-K)$.

The quantities k_i, λ_i and $\lambda_i^{(2)}$ are all functions of the unknown parameters θ, and will depend on the assumed form of the income distribution. For most distributions it is convenient to compute these quantities by expressing them in terms of the distribution function and the first and second moment distribution functions of the assumed distribution. Specifically,

$$k_i(\boldsymbol{\theta}) = F(z_i;\phi) - F(z_{i-1};\phi) \tag{32}$$

$$\lambda_i(\boldsymbol{\theta}) = \mu\left(F_1(z_i;\phi) - F_1(z_{i-1};\phi)\right) \tag{33}$$

and

$$\lambda_i^{(2)}(\theta) = \mu_2\left(F_2(z_i;\phi) - F_2(z_{i-1};\phi)\right) \tag{34}$$

The right-hand-side quantities in these equations can be obtained from equations (2), (6) and (8) for the dPLN distribution, (5), (7) and (9) for the PLN distribution and (20) and (21) for the GB2 distribution.

In our empirical work we employed an iterative two-step GMM estimator. In the first stage we find $\hat{\boldsymbol{\theta}}_1 = \arg\min_{\boldsymbol{\theta}} \mathbf{H}(\boldsymbol{\theta})'\boldsymbol{\Omega}_0\, \mathbf{H}(\boldsymbol{\theta})$ where $\boldsymbol{\Omega}_0 = \mathbf{D}(c_1^{-2}, c_2^{-2}, \ldots, c_N^{-2}, \tilde{y}_1^{-2}, \tilde{y}_2^{-2}, \ldots, \tilde{y}_N^{-2})$. In the second stage we find $\hat{\boldsymbol{\theta}}_2 = \arg\min_{\boldsymbol{\theta}} \mathbf{H}(\boldsymbol{\theta})'\boldsymbol{\Omega}\left(\hat{\boldsymbol{\theta}}_1\right)\mathbf{H}(\boldsymbol{\theta})$, and then we iterate until convergence. The rationale behind using $\boldsymbol{\Omega}_0$ in the first step is that it leads to an estimator that minimizes the sum of squares of the percentage errors in the moment conditions.

The asymptotic covariance for $\hat{\theta}$ can be shown to be

$$\operatorname{var}\left(\hat{\boldsymbol{\theta}}\right) = \frac{1}{T}\left(\begin{bmatrix} \dfrac{\partial \mathbf{k}'}{\partial \boldsymbol{\theta}} & \dfrac{\partial \mathbf{m}'}{\partial \boldsymbol{\theta}} \end{bmatrix} \begin{bmatrix} \mathbf{D}(\omega_1) & -\mathbf{D}(\omega_3) \\ -\mathbf{D}(\omega_3) & \mathbf{D}(\omega_2) \end{bmatrix} \begin{bmatrix} \partial \mathbf{k}/\partial \boldsymbol{\theta}' \\ \partial \mathbf{m}/\partial \boldsymbol{\theta}' \end{bmatrix}\right)^{-1} \tag{35}$$

This expression can be used to compute standard errors for the elements in $\hat{\theta}$ and functions of them. In our empirical work we successfully used numerical derivatives to compute (35).

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

$$J = T\,\mathbf{H}\left(\hat{\boldsymbol{\theta}}\right)'\mathbf{W}\left(\hat{\boldsymbol{\theta}}\right)\mathbf{H}\left(\hat{\boldsymbol{\theta}}\right) \xrightarrow{d} \chi^2_{N-K} \tag{36}$$

where $N-K$ is the number of excess moment conditions. 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.

V. Description of data and empirical analysis

To compare GB2, PLN and dPLN distributions, we used grouped data from the ten "regions" China urban, China rural, India urban, India rural, Pakistan, Russia, Poland, Brazil, Nigeria and Iran. These countries represent a mix of different regions and different inequality levels. The data were downloaded from the World Bank web site http://go.worldbank.org/WE8P1I8250. Population shares c_i and the corresponding income shares s_i were available for 20 groups for all regions except for China rural (17 groups) and India rural and urban (12 groups). For consistency, we converted the data from all regions into 12 groups. Also available from the World Bank website is each region's mean monthly income $\bar{y}$, found from surveys and then converted using a 2005 purchasing-power-parity exchange rate. To implement the methodology described in Section 4 we compute $\tilde{y}_i = c_i\bar{y}_i = s_i\bar{y}$ for each of the classes, where the class mean incomes are given by $\bar{y}_i = s_i\,\bar{y}/c_i$.

Sample sizes T for each of the surveys from which the grouped data were computed are not available. For calculating standard errors, we used $T = 20{,}000$. This is a conservative value since most of surveys have sample sizes which are larger. If standard errors for other sample sizes are of interest, they can be obtained from our results by multiplying by the appropriate scaling factor.

In what follows we first examine the GMM estimates and standard errors for the parameters and class limits estimated using the PLN, dPLN and GB2 distributions. Second, we consider the results of J tests for the validity of the excess moment conditions, comparing the GB2 distribution with the PLN and dPLN distributions. Then, a goodness-of-fit comparison of these distributions is made on the basis of their ability to predict the observed income shares for each group. Finally, we present the Gini and Theil coefficient estimates and their standard errors obtained under the alternative distributional assumptions.

Estimates of the parameters for the three distributions, and the class limits for the grouped data, are reported in Table 1. We can make the following observations. Estimates of the class limits are almost identical for each of the models, and their standard errors are small. Estimates of the parameter α for the PLN and dPLN distributions are relatively small, and significantly different from zero, indicating significant Pareto behaviour in the upper tail of the distribution. However, estimates for β are large in some regions (China rural, India rural and urban, Pakistan, Poland and Iran), indicating that Pareto behaviour is not significant in the lower tail. In these instances there was very little difference between corresponding parameters for the PLN and dPLN models. All standard errors are calculated numerically. They are relatively small providing reassurance that we are estimating the models accurately, at least for our assumed sample size. A potential problem that may occur where there is a high level of inequality and which we encountered when estimating the GB2 and dPLN distributions for Brazil (but not the PLN distribution), is the non-existence of the second moment. Hajargasht et al. (2012) found the same problem. For the existence of the *k*-th moment, the GB2 distribution requires $aq > k$, and the DPLN and PLN distributions require $\alpha > k$. If, in the first step of estimation using the weighting matrix $\mathbf{\Omega}_0$, the estimates are such that $\hat{a}\hat{q} < 2$ or $\hat{\alpha} < 2$, the optimal weighting matrix in the second step, which requires the existence of second order moments, cannot be computed. We are unable to proceed with iterative GMM estimation without modifying the algorithm. In the results reported in Table 1 we overcame the problem for Brazil by minimizing the objective function subject to the constraint $aq > 2$ for the GB2 distribution, and subject to the constraint $\alpha > 2$ for the dPLN distribution. This solution may not be entirely satisfactory. The underlying income distribution may indeed not have second moments, and the standard errors for the boundary solutions that result may not be valid.

When sample sizes are very large, the probability of rejecting a false null hypothesis can be close to one, even when the difference between an actual parameter value and the value hypothesized under the null is so small as to be meaningless. In the context of testing excess moment conditions to assess the validity for a particular income distribution, this means that a particular income distribution can be rejected even when it fits the data well, both visually and in terms of predicting income shares. Given these circumstances, and given that sample sizes are not available for the regions for which we have data, we use the magnitudes of the *J* statistics as a basis for comparing goodness of fit of the three distributions, rather than as a significance test for the validity of excess moment conditions. Using a sample size of $T = 20,000$, Table 2 provides *J*-statistic values for all three distributions and for all regions. The smaller the test value, the better the fit of the distribution. In this sense, the results indicate better performance of dPLN relative to GB2 for China rural and urban, India rural and Poland, and better performance of GB2 for the other regions. Also, there are a few regions (China rural and urban, India urban and Poland) where PLN is marginally better than dPLN. For Brazil all distributions appear to be a relatively poor fit, with GB2 being the best of the three.

Goodness-of-fit in terms of predicting income shares was also carried out by comparing the observed income shares s_i with the predicted income shares derived from the estimated distributions, defined as $\hat{s}_i = F_1\left(\hat{z}_i;\hat{\phi}\right) - F_1\left(\hat{z}_{i-1};\hat{\phi}\right)$ where $\hat{z}_i$ are the estimates of the class limits. Table 2 contains the percentage 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. The first impressive thing to report about these values is that they are very small. For example, the RMSE for China rural, using the 3-parameter Pareto-lognormal is approximately 0.08%. That means that the average error (in the RMSE sense) from predicting the income shares for this region is 0.0008, a very accurate prediction indeed. One of the worst performers using this measure is for India rural, using PLN, but even here the average error is only 0.0083. A comparison of the RMSEs for the three distributions yields similar conclusions to those reached by examining the *J*-statistics. The dPLN has better performance relative to GB2 for 5 regions (China rural and urban, India rural, Poland and Iran) but worse performance for the other 5 regions. The dPLN has a substantially smaller RMSE than PLN for those regions where the estimate for β was relatively small, and a similar RMSE for those regions where the estimate for β was relatively large.

Estimates of the Gini and Theil inequality measures are given in Table 3. Estimates of the Gini coefficient are generally comparable across the 3 distributions, with the largest discrepancy being 0.014 for India urban. A similar remark can be made for the Theil coefficient, but in this case there are discrepancies for both India urban and Brazil, and the discrepancies are larger. In the case of India urban, the GB2 leads to a smaller estimate of the degree of inequality. The restricted parameter estimates for Brazil's GB2 and dPLN distributions led to a higher degree of inequality than that estimated from PLN. The regions with the greatest inequality are Brazil, Nigeria and India urban. For Iran, the Gini coefficient is comparable to that of India urban, but the Theil coefficient suggests a lower degree of inequality.

VI. Concluding remarks

The Pareto-lognormal and double Pareto-lognormal distributions have been advocated as good choices for modelling income distributions because of their sound theoretical base and their superior empirical goodness-of-fit. We have added to the existing literature in four ways: (1) Expressions for inequality measures in terms of the parameters of the distributions have been derived. (2) We show how the distributions can be estimated from grouped data using the generalized method of moments, and we derive the moment distribution functions required for GMM estimation. (3) We indicate how the moments and moment distribution functions can be used to compute a number of poverty measures. An alternative expression is derived for the Watts poverty index that cannot be expressed in terms of moment distribution functions. (4) For ten example countries, the performance of the PLN and dPLN distributions is compared with another popular income distribution, the generalized beta distribution of the second kind. We find that all three distributions fit the data relatively well, but there is no clear evidence to suggest that the PLN or the dPLN is superior to the GB2 distribution.

Appendix A Moment Distribution Functions

First consider the following Pareto-lognormal density function

$$f_1^{PLN}(y) = \frac{\alpha}{y}\phi\left(\frac{\ln y - m}{\sigma}\right)\left(\frac{1-\Phi\left(\alpha\sigma - (\ln y - m)/\sigma\right)}{\phi\left(\alpha\sigma - (\ln y - m)/\sigma\right)}\right)$$

which can be written equivalently as

$$f_1^{PLN}(y) = \frac{\alpha}{y}\exp\left\{\frac{\alpha^2\sigma^2}{2} - \frac{\alpha\sigma(\ln y - m)}{\sigma}\right\}\left(1-\Phi\left(\alpha\sigma - \frac{\ln y - m}{\sigma}\right)\right)$$

To obtain the moment distribution function of order *k* we must compute

$$\int_0^y t^k f_1^{PLN}(t)dt = \int_0^y \alpha t^{k-1}\exp\left\{\frac{\alpha^2\sigma^2}{2} - \frac{\alpha\sigma(\ln t - m)}{\sigma}\right\}\left(1-\Phi\left(\alpha\sigma - \frac{\ln t - m}{\sigma}\right)\right)dt$$

Let $u = (\ln t - m)/\sigma$ so that $t = \exp\{m + \sigma u\}$ and $dt = \sigma \exp\{m + \sigma u\} du$. Then,

$$\int_0^y t^k f_1^{PLN}(t)dt = \int_{-\infty}^{(\ln y - m)/\sigma} \alpha\sigma \exp\left\{km + \frac{\alpha^2\sigma^2}{2}\right\} \exp\{(k-\alpha)\sigma u\} \left[1 - \Phi(\alpha\sigma - u)\right] du$$

Integration by parts leads to

$$\int_0^y t^k f_1^{PLN}(t)dt = \frac{\alpha}{k-\alpha} \exp\left\{km + \frac{\alpha^2\sigma^2}{2}\right\}$$
$$\times \left[\exp\{(k-\alpha)(\ln y - m)\} \left(1 - \Phi\left(\alpha\sigma - \frac{\ln y - m}{\sigma}\right)\right) - \int_{-\infty}^{(\ln y - m)/\sigma} \exp\{(k-\alpha)\sigma u\} \phi(\alpha\sigma - u) du \right]$$

Now,

$$\int_{-\infty}^{(\ln y - m)/\sigma} \exp\{(k-\alpha)\sigma u\} \phi(\alpha\sigma - u) du = \exp\left\{\frac{(k^2 - \alpha^2)\sigma^2}{2}\right\} \int_{-\infty}^{(\ln y - m)/\sigma} \frac{1}{\sqrt{2\pi}} \exp\left\{-\frac{1}{2}(u - k\sigma)^2\right\} du$$
$$= \exp\left\{\frac{(k^2 - \alpha^2)\sigma^2}{2}\right\} \Phi\left(\frac{\ln y - m}{\sigma} - k\sigma\right)$$

Using this result, we obtain, after some algebra,

$$\int_0^y t^k f_1^{PLN}(t)dt = \frac{\alpha}{\alpha - k} \exp\left\{km + \frac{k^2\sigma^2}{2}\right\}$$
$$\times \left[\Phi\left(\frac{\ln y - m}{\sigma} - k\sigma\right) - \left(1 - \Phi\left(\alpha\sigma - \frac{\ln y - m}{\sigma}\right)\right) \exp\left\{k\sigma\frac{\ln y - m}{\sigma} - \alpha\sigma\frac{\ln y - m}{\sigma} - \frac{k^2\sigma^2}{2} + \frac{\alpha^2\sigma^2}{2}\right\} \right]$$
$$= \frac{\alpha}{\alpha - k} \exp\left\{km + \frac{k^2\sigma^2}{2}\right\} \left[\Phi\left(\frac{\ln y - m}{\sigma} - k\sigma\right) - \phi\left(\frac{\ln y - m}{\sigma} - k\sigma\right) R(x_1) \right]$$

From this result we obtain equation (9), which gives the *k*-th moment distribution function for the PLN, with Pareto behavior in the right tail.

To obtain the *k*-th moment distribution function for the dPLN distribution it is convenient to first consider the PLN distribution with Pareto behavior in the left tail. Its pdf is

$$f_2^{PLN}(y) = \frac{\beta}{y} \phi\left(\frac{\ln y - \mu}{\sigma}\right) \left(\frac{1 - \Phi(\beta\sigma + (\ln y - m)/\sigma)}{\phi(\beta\sigma + (\ln y - m)/\sigma)}\right)$$

Then, using similar steps to those used to derive $\int_0^y t^k f_1^{PLN}(t)dt$, we can show that

$$\int_0^y t^k f_2^{PLN}(t)dt = \frac{\beta}{\beta+k}\exp\left\{km+\frac{k^2\sigma^2}{2}\right\}\left[\Phi\left(\frac{\ln y-m}{\sigma}-k\sigma\right)+\phi\left(\frac{\ln y-m}{\sigma}-k\sigma\right)R(x_2)\right]$$

From Reed and Jorgensen (2004), the density of the dPLN can be expressed as the mixture

$$f^{dPLN}(y) = \frac{\beta}{\alpha+\beta} f_1^{PLN}(y) + \frac{\alpha}{\alpha+\beta} f_2^{PLN}(y)$$

Thus,

$$\begin{aligned}\int_0^y t^k f^{dPLN}(t)dt &= \frac{\beta}{\alpha+\beta}\int_0^y t^k f_1^{PLN}(t)dt + \frac{\alpha}{\alpha+\beta}\int_0^y t^k f_2^{PLN}(t)dt \\ &= \frac{\alpha\beta}{(\alpha-k)(\beta+k)}\exp\left\{km+\frac{k^2\sigma^2}{2}\right\} \\ &\quad\times\left[\Phi\left(\frac{\ln y-m}{\sigma}-k\sigma\right)-\phi\left(\frac{\ln y-m}{\sigma}-k\sigma\right)\left(\frac{(\beta+k)R(x_1)-(\alpha-k)R(x_2)}{\alpha+\beta}\right)\right]\end{aligned}$$

From this result, we obtain *k*-th moment distribution function $F_k^{dPLN}(y;m,\sigma,\alpha,\beta)$ given in (8).

Appendix B Inequality Measures

1. The Gini Coefficient

To derive the Gini coefficient for the dPLN distribution we use the formula

$$G = \frac{2\int_0^\infty yF(y)f(y)dy}{\mu} - 1$$

where the mean is given by

$$\mu = \frac{\alpha\beta}{(\alpha-1)(\beta+1)}\exp\left\{m+\frac{\sigma^2}{2}\right\}$$

The required integral can be written in terms of six components. That is,

$$I = \int_0^\infty yF(y)f(y)dy = I_1+I_2+I_3+I_4+I_5+I_6$$

where the six components arise from the product of two components in the pdf $f^{dPLN}(y)$ and three components in the cdf $F^{dPLN}(y)$. See equations (1) and (2). Each of the six components and their solutions follow. We use the transformation $u = (\ln y - m)/\sigma$, integration by parts, and the following result from Gupta and Pillai (1965),

$$\int_{-\infty}^{\infty} \frac{\exp\{-x^2/2\}}{\sqrt{2\pi}} \Phi(ax+b)\,dx = \Phi\left(\frac{b}{\sqrt{1+a^2}}\right)$$

$$I_1 = \frac{\alpha\beta}{\alpha+\beta}\exp\left\{\frac{\alpha^2\sigma^2}{2}\right\}\int_{-\infty}^{\infty}\exp\{(1-\alpha)\sigma u\}\,\Phi(u)\big(\Phi(u-\alpha\sigma)\big)\,du$$

$$= \frac{\alpha\beta}{(\alpha+\beta)(\alpha-1)\sigma}\left[\exp\left\{\frac{\left(\alpha^2+(\alpha-1)^2\right)\sigma^2}{2}\right\}\Phi\left(\frac{(1-2\alpha)\sigma}{\sqrt{2}}\right)+\exp\left\{\frac{\sigma^2}{2}\right\}\Phi\left(\frac{\sigma}{\sqrt{2}}\right)\right]$$

$$I_2 = \frac{\alpha\beta}{\alpha+\beta}\exp\left\{\frac{\beta^2\sigma^2}{2}\right\}\int_{-\infty}^{\infty}\exp\{(1+\beta)\sigma u\}\,\Phi(u)\big(\Phi(-u-\beta\sigma)\big)\,du$$

$$= \frac{\alpha\beta}{(\alpha+\beta)(\beta+1)\sigma}\left[-\exp\left\{\frac{(\beta^2+(\beta+1)^2)\sigma^2}{2}\right\}\Phi\left(\frac{(-1-2\beta)\sigma}{\sqrt{2}}\right)+\exp\left\{\frac{\sigma^2}{2}\right\}\Phi\left(\frac{\sigma}{\sqrt{2}}\right)\right]$$

$$I_3 = -\frac{\alpha\beta^2}{(\alpha+\beta)^2}\exp\{\alpha^2\sigma^2\}\int_{-\infty}^{\infty}\exp\{(1-2\alpha)\sigma u\}\big[\Phi(u-\alpha\sigma)\big]^2\,du$$

$$= -\frac{2\alpha\beta^2}{(\alpha+\beta)^2(2\alpha-1)\sigma}\exp\left\{\frac{\left(\alpha^2+(\alpha-1)^2\right)\sigma^2}{2}\right\}\Phi\left(\frac{(1-2\alpha)\sigma}{\sqrt{2}}\right)$$

$$I_4 = -\frac{\alpha\beta^2}{(\alpha+\beta)^2}\exp\left\{\left(\alpha^2+\beta^2\right)\frac{\sigma^2}{2}\right\}\int_{-\infty}^{\infty}\exp\{(1-\alpha+\beta)\sigma u\}\,\Phi(u-\alpha\sigma)\,\Phi(-u-\beta\sigma)\,du$$

$$= -\frac{\alpha\beta^2}{(\alpha+\beta)^2(1-\alpha+\beta)\sigma}\Bigg[-\exp\left\{\left(2\beta(1+\beta)+1\right)\frac{\sigma^2}{2}\right\}\Phi\left(\frac{(-1-2\beta)\sigma}{\sqrt{2}}\right)$$

$$+\exp\left\{\left(2\alpha(\alpha-1)+1\right)\frac{\sigma^2}{2}\right\}\Phi\left(\frac{(1-2\alpha)\sigma}{\sqrt{2}}\right)\Bigg]$$

$$I_5 = \frac{\alpha^2\beta}{(\alpha+\beta)^2}\exp\left\{(\alpha^2+\beta^2)\frac{\sigma^2}{2}\right\}\int_{-\infty}^{\infty}\exp\{(1-\alpha+\beta)\sigma u\}\Phi(u-\alpha\sigma)\Phi(-u-\beta\sigma)\,du$$

$$= \frac{\alpha^2\beta}{(\alpha+\beta)^2(1-\alpha+\beta)\sigma}\left[-\exp\left\{(2\beta(1+\beta)+1)\frac{\sigma^2}{2}\right\}\Phi\left(\frac{(-1-2\beta)\sigma}{\sqrt{2}}\right)\right.$$
$$\left. + \exp\left\{(2\alpha(\alpha-1)+1)\frac{\sigma^2}{2}\right\}\Phi\left(\frac{(1-2\alpha)\sigma}{\sqrt{2}}\right)\right]$$

$$I_6 = \frac{\alpha^2\beta}{(\alpha+\beta)^2}\exp\{\beta^2\sigma^2\}\int_{-\infty}^{\infty}\exp\{(1+2\beta)\sigma u\}\left(\Phi(-u-\beta\sigma)\right)^2 du$$

$$= \frac{2\alpha^2\beta}{(\alpha+\beta)^2(1+2\beta)\sigma}\exp\left\{\left(\beta^2+(\beta+1)^2\right)\frac{\sigma^2}{2}\right\}\Phi\left(\frac{(-1-2\beta)\sigma}{\sqrt{2}}\right)$$

Collecting all the terms yields

$$G = 2\left[\Phi\left(\frac{\sigma}{\sqrt{2}}\right) + \left(\frac{1+\beta}{(\alpha+\beta)}\left(1 - \frac{2\beta(\alpha-1)}{(\alpha+\beta)(2\alpha-1)} + \frac{(\alpha-\beta)(\alpha-1)}{(\alpha+\beta)(1-\alpha+\beta)}\right)\right)\exp\{\alpha(\alpha-1)\sigma^2\}\Phi\left(\frac{(1-2\alpha)\sigma}{\sqrt{2}}\right)\right.$$
$$\left. + \left(\frac{\alpha-1}{(\alpha+\beta)}\left(-1 + \frac{2\alpha(1+\beta)}{(\alpha+\beta)(1+2\beta)} - \frac{(\alpha-\beta)(1+\beta)}{(\alpha+\beta)(1-\alpha+\beta)}\right)\right)\exp\{\beta(\beta+1)\sigma^2\}\Phi\left(\frac{(-1-2\beta)\sigma}{\sqrt{2}}\right)\right] - 1$$

which can be further simplified to

$$G^{dPLN} = 2\left[\Phi\left(\frac{\sigma}{\sqrt{2}}\right) + \frac{\beta(1+\beta)\exp\{\alpha(\alpha-1)\sigma^2\}}{(\alpha+\beta)(2\alpha-1)(1-\alpha+\beta)}\Phi\left(\frac{(1-2\alpha)\sigma}{\sqrt{2}}\right)\right.$$
$$\left. - \frac{\alpha(\alpha-1)\exp\{\beta(\beta+1)\sigma^2\}}{(\alpha+\beta)(1+2\beta)(1-\alpha+\beta)}\Phi\left(\frac{(-1-2\beta)\sigma}{\sqrt{2}}\right)\right] - 1$$

Letting $\beta \to \infty$, we get, for the PLN distribution,

$$G^{PLN} = \frac{2\exp\{\alpha(\alpha-1)\sigma^2\}}{(2\alpha-1)}\Phi\left(\frac{(1-2\alpha)\sigma}{\sqrt{2}}\right) + 2\Phi\left(\frac{\sigma}{\sqrt{2}}\right) - 1$$

2. Generalized Entropy

The generalized entropy inequality measure for the dPLN distribution is

$$GE^{dPLN} = \frac{1}{\upsilon^2 - \upsilon}\left[\int_0^{+\infty}\left(\frac{y}{\mu}\right)^{\upsilon} f(y)dy - 1\right] = \frac{1}{\upsilon^2 - \upsilon}\left(\frac{\mu_{\upsilon}}{\mu^{\upsilon}} - 1\right)$$

$$= \frac{1}{\upsilon^2 - \upsilon}\left(\frac{\alpha\beta\exp\left\{\upsilon\mu + \frac{\upsilon^2\sigma^2}{2}\right\}}{(\alpha-\upsilon)(\beta+\upsilon)} \Bigg/ \frac{(\alpha\beta)^{\upsilon}\exp\left\{\upsilon\mu + \frac{\upsilon\sigma^2}{2}\right\}}{(\alpha-1)^{\upsilon}(\beta+1)^{\upsilon}} - 1\right)$$

$$= \frac{1}{\upsilon^2 - \upsilon}\left(\frac{(\alpha-1)^{\upsilon}(\beta+1)^{\upsilon}\exp\left\{(\upsilon-1)\upsilon\sigma^2/2\right\}}{(\alpha\beta)^{\upsilon-1}(\alpha-\upsilon)(\beta+\upsilon)} - 1\right)$$

For the PLN distribution it becomes

$$GE^{PLN} = \frac{1}{\upsilon^2 - \upsilon}\left(\frac{(\alpha-1)^{\upsilon}\exp\left\{(\upsilon-1)\upsilon\sigma^2/2\right\}}{\alpha^{\upsilon-1}(\alpha-\upsilon)} - 1\right)$$

3. Theil Index

Theil (1967) proposed two inequality measures. The most common form and the one we consider is a special case of the generalized entropy measure obtained by letting $\upsilon \to 1$. It is given by

$$T = \int_0^{\infty}\left(\frac{y}{\mu}\right)\ln\left(\frac{y}{\mu}\right) f(y)dy$$

To obtain

$$T^{dPLN} = \lim_{\upsilon\to 1} GE^{dPLN}$$

$$= \lim_{\upsilon\to 1}\frac{(\alpha-1)^{\upsilon}(\beta+1)^{\upsilon}\exp\left\{(\upsilon-1)\upsilon\sigma^2/2\right\} - (\alpha\beta)^{\upsilon-1}(\alpha-\upsilon)(\beta+\upsilon)}{(\alpha\beta)^{\upsilon-1}(\alpha-\upsilon)(\beta+\upsilon)(\upsilon^2-\upsilon)}$$

we use L'Hôpital's rule, for which we need the following derivatives

$$\frac{d}{d\upsilon}(\alpha-1)^{\upsilon}(\beta+1)^{\upsilon}\exp\left\{(\upsilon-1)\upsilon\sigma^2/2\right\}=[\ln(\alpha-1)](\alpha-1)^{\upsilon}(\beta+1)^{\upsilon}\exp\left\{(\upsilon-1)\upsilon\sigma^2/2\right\}$$
$$+(\alpha-1)^{\upsilon}[\ln(\beta+1)](\beta+1)^{\upsilon}\exp\left\{(\upsilon-1)\upsilon\sigma^2/2\right\}$$
$$+(\alpha-1)^{\upsilon}(\beta+1)^{\upsilon}(2\upsilon-1)\sigma^2/2\exp\left\{(\upsilon-1)\upsilon\sigma^2/2\right\}$$

$$\frac{d}{d\upsilon}(\alpha\beta)^{\upsilon-1}(\alpha-\upsilon)(\beta+\upsilon)=[\ln(\alpha\beta)](\alpha\beta)^{\upsilon-1}(\alpha-\upsilon)(\beta+\upsilon)-(\alpha\beta)^{\upsilon-1}(\beta+\upsilon)+(\alpha\beta)^{\upsilon-1}(\alpha-\upsilon)$$

$$\frac{d}{d\upsilon}(\alpha\beta)^{\upsilon-1}(\alpha-\upsilon)(\beta+\upsilon)(\upsilon^2-\upsilon)=\left(\frac{d}{d\upsilon}(\alpha\beta)^{\upsilon-1}(\alpha-\upsilon)(\beta+\upsilon)\right)(\upsilon^2-\upsilon)$$
$$+(\alpha\beta)^{\upsilon-1}(\alpha-\upsilon)(\beta+\upsilon)(2\upsilon-1)$$

Evaluating these derivatives at $\upsilon=1$, forming the ratio and simplifying leads to

$$T^{dPLN}=\ln\left(\frac{\alpha-1}{\alpha}\right)+\ln\left(\frac{\beta+1}{\beta}\right)+\frac{\sigma^2}{2}+\frac{1}{\alpha-1}-\frac{1}{\beta+1}$$

The Theil index for PLN, obtained by letting $\beta\to\infty$ is

$$T^{PLN}=\ln\left(\frac{\alpha-1}{\alpha}\right)+\frac{1}{\alpha-1}+\frac{\sigma^2}{2}$$

4. Atkinson Index

$$A^{dPLN}=1-\left[\int_0^{+\infty}\left(\frac{y}{\mu}\right)^{1-\varepsilon}f(y)dy\right]^{\frac{1}{1-\varepsilon}}=1-\left(\frac{\mu_{1-\varepsilon}}{\mu^{1-\varepsilon}}\right)^{\frac{1}{1-\varepsilon}}$$
$$=1-\left(\frac{(\alpha\beta)^{1/(1-\varepsilon)}\exp\left\{m+(1-\varepsilon)\sigma^2/2\right\}\Big/\left[(\alpha+\varepsilon-1)(\beta-\varepsilon+1)\right]^{1/(1-\varepsilon)}}{\alpha\beta\exp\left\{m+\sigma^2/2\right\}\Big/(\alpha-1)(\beta+1)}\right)$$
$$=1-\frac{(\alpha\beta)^{\varepsilon/(1-\varepsilon)}(\alpha-1)(\beta+1)\exp\{-\varepsilon\sigma^2/2\}}{\left[(\alpha+\varepsilon-1)(\beta-\varepsilon+1)\right]^{1/(1-\varepsilon)}}$$

Then for the PLN distribution, we have

$$A^{PLN}=1-\frac{\alpha^{\varepsilon/(1-\varepsilon)}(\alpha-1)\exp\{-\varepsilon\sigma^2/2\}}{(\alpha+\varepsilon-1)^{1/(1-\varepsilon)}}$$

Appendix C Watts Poverty Measure

The Watts poverty measure is defined by

$$W=\int_0^{y_P}\left(\ln y_P-\ln t\right)f(t)dt=\left(\ln y_P\right)F(y_P)-\int_0^{y_P}\left(\ln t\right)f(t)dt$$

We set $u=\left(\ln t-m\right)/\sigma$, and $u_P=\left(\ln y_P-m\right)/\sigma$, so that $\ln t=\sigma u+m$, $t=\exp\{\sigma u+m\}$ and $dt/t=\sigma du$ Then, using the definition for $f^{dPLN}(t)$ in equation (1), and collecting all terms under the exponent, we can write

$$\int_0^{y_P}\left(\ln t\right)f(t)dt=\frac{\alpha\beta\sigma}{\alpha+\beta}\left[\exp\left\{\frac{\alpha^2\sigma^2}{2}\right\}I_1+\exp\left\{\frac{\beta^2\sigma^2}{2}\right\}I_2\right]$$

where, using integration by parts,

$$\begin{aligned}I_1&=\int_{-\infty}^{u_P}(\sigma u+m)\exp\{-\alpha\sigma u\}\Phi\left(u-\alpha\sigma\right)du\\&=\frac{1}{\alpha^2\sigma^2}\left[-\exp\{-\alpha\sigma u_P\}\left(\alpha\sigma^2u_P+\alpha\sigma m+\sigma\right)\Phi\left(u_P-\alpha\sigma\right)\right]+I_3\end{aligned}$$

$$\begin{aligned}I_3&=\frac{1}{\alpha^2\sigma^2}\int_{-\infty}^{u_P}\exp\{-\alpha\sigma u\}\left(\alpha\sigma^2u+\alpha\sigma m+\sigma\right)\phi\left(u-\alpha\sigma\right)du\\&=\frac{\exp\left\{-\alpha^2\sigma^2/2\right\}}{\alpha^2\sigma^2}\left[\sigma(\alpha m+1)\Phi(u_P)-\alpha\sigma^2\phi(u_P)\right]\end{aligned}$$

Similarly,

$$\begin{aligned}I_2&=\int_{-\infty}^{u_P}(\sigma u+m)\exp\{\beta\sigma u\}\Phi\left(-u-\beta\sigma\right)du\\&=\frac{1}{\beta^2\sigma^2}\left[\exp\{\beta\sigma u_P\}\left(\beta\sigma^2u_P+\beta\sigma m-\sigma\right)\Phi\left(-u_P-\beta\sigma\right)\right]+I_4\end{aligned}$$

$$\begin{aligned}I_4&=\frac{1}{\beta^2\sigma^2}\int_{-\infty}^{u_P}\exp\{\beta\sigma u\}\left(\beta\sigma^2u+\beta\sigma m-\sigma\right)\phi\left(-u-\beta\sigma\right)du\\&=\frac{\exp\left\{-\beta^2\sigma^2/2\right\}}{\beta^2\sigma^2}\left[\sigma(\beta m-1)\Phi(u_P)-\beta\sigma^2\phi(u_P)\right]\end{aligned}$$

Collecting all these terms and carrying out some convenient algebraic manipulations, we have

$$\int_0^{y_P} (\ln t)\, f(t)dt = (m+\sigma u_P)\left[\Phi(u_P) - \frac{\beta}{\alpha+\beta}\exp\left\{\frac{\alpha^2\sigma^2}{2} - \alpha\sigma u_P\right\}\Phi(u_P - \alpha\sigma)\right.$$
$$\left. + \frac{\alpha}{\alpha+\beta}\exp\left\{\frac{\beta^2\sigma^2}{2} + \beta\sigma u_P\right\}\Phi(-u_P - \beta\sigma)\right] - \sigma[u_P\Phi(u_P) + \phi(u_P)]$$
$$+ \frac{\beta}{\alpha(\alpha+\beta)}\left[\Phi(u_P) - \exp\left\{\frac{\alpha^2\sigma^2}{2} - \alpha\sigma u_P\right\}\Phi(u_P - \alpha\sigma)\right]$$
$$- \frac{\alpha}{\beta(\alpha+\beta)}\left[\Phi(u_P) + \exp\left\{\frac{\beta^2\sigma^2}{2} + \beta\sigma u_P\right\}\Phi(-u_P - \beta\sigma)\right]$$
$$= (m+\sigma u_P)F^{dPLN}(y_P) - \sigma[u_P\Phi(u_P) + \phi(u_P)]$$
$$+ \frac{\beta}{\alpha(\alpha+\beta)}F^{PLN(\alpha)}(y_P) - \frac{\alpha}{\beta(\alpha+\beta)}F^{PLN(\beta)}(y_P)$$

where $F^{dPLN}(y)$ is the cdf of the dPLN distribution, $F^{PLN(\alpha)}(y)$ is the cdf of the PLN distribution with Pareto behaviour in the right tail, and $F^{PLN(\beta)}(y)$ is cdf of the PLN distribution with Pareto behaviour in the left tail. Noting that $\ln y_p F^{dPLN}(y_p) = (m+\sigma u_P)F^{dPLN}(y_P)$, the Watts index for the dPLN distribution can then be written as

$$W^{dPLN} = (\ln y_P)F^{dPLN}(y_P) - \int_0^{y_p} (\ln t)\, f^{dPLN}(t)dt$$
$$= -\frac{\beta}{\alpha(\alpha+\beta)}F^{PLN(\alpha)}(y_p) + \frac{\alpha}{\beta(\alpha+\beta)}F^{PLN(\beta)}(y_p) + \sigma[u_p\Phi(u_p) + \phi(u_p)]$$

Letting $\beta \to \infty$, we get, for the PLN distribution,

$$W^{PLN} = -\frac{1}{\alpha}F_{PLN}(y_p) + \sigma[u_p\Phi(u_p) + \phi(u_p)]$$

References

Benhabib, J., and Zhu, S. (2008). ‘Age, luck and inheritance’, NBER Working Paper 14128.

Champernowne, D. A. (1953). ‘A model of income distribution’, *Economic Journal*, Vol. 63, pp. 318-351.

Colombi, R. (1990). ‘A new model of income distribution: The Pareto-lognormal distribution’, in C. Dagum and M. Zenga (ed.), *Income and Wealth Distribution, Inequality and Poverty*, Springer, Berlin, pp.18-32.

Cowell, F. A. (2007). *Measuring Inequality* (3rd ed.), Oxford UniversityPress, Oxford.

Gibrat R. (1931). *Les Inégalités Économiques*, Librairie du Recueil Sirey, Paris.

Gupta, S. and Pillai, S. (1965). 'On linear functions of ordered correlated normal random variables, *Biometrika*, Vol. 52, pp.367-379.

Hajargasht, G., Griffiths, W., Brice, J., Rao, D. S. P. and Chotikapanich, D. (2012). 'Inference for income distributions using grouped data', *Journal of Business of Economic Statistics*, in press.

Jenkins, S. (2009). 'Distributionally-sensitive inequality indices and the GB2 income distribution', *Review of Income and Wealth*, Vol. 55, pp. 392-398.

Kakwani, N. (1999), 'Inequality, welfare and poverty: three interrelated phenomena', in J. Silber (ed), *Handbook on Income Inequality Measurement*, Kluwer, Dordrecht, pp. 599-628.

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

Mandelbrot, B. (1960). 'The Pareto-Lévy law and the distribution of income', *International Economic Review*, Vol. 1, pp. 79-106.

McDonald, J.B. (1984). 'Some generalized functions of the size distribution of income', *Econometrica*, Vol. 52, pp. 647-663.

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 D. Chotikapanich (ed.), *Modeling Income Distributions and Lorenz Curves*, Springer, New York, pp. 147-166.

McDonald, J.B. and Xu, Y.J. (1995). 'A generalization of the beta distribution with applications', *Journal of Econometrics*, Vol. 66, pp. 133-152. Errata (1995), Vol. 69, pp. 427-428.

Muller, C. (2001). 'The properties of the Watts poverty index under lognormality', *Economics Bulletin*, Vol. 9, pp. 1-9.

Parker, S.C. (1999). 'The generalized beta as a model for the distribution of earnings', *Economics Letters*, Vol. 62, pp. 197-200.

Reed, W. J. (2001). ‘The Pareto, Zipf and other power laws’, *Economics Letters*, Vol. 74, pp.15-19.

Reed, W.J. (2003). ‘The Pareto law of incomes – an explanation and extension’, *Physica A*, Vol. 319, pp. 469-486.

Reed, W.J. and Jorgensen M. (2004). ‘The double Pareto-lognormal distribution – a new parametric model for size distributions’, *Communications in Statistics Theory and Methods*, Vo. 33, pp.1733-1753.

Reed, W.J. and Wu, F. (2008). ‘New four- and five-parameter models for income distributions’, in D. Chotikapanich (ed.), *Modeling Income Distributions and Lorenz Curves*, Springer, New York, pp. 211-224.

Theil, H. (1967). *Economics and Information Theory*, North-Holland, Amsterdam.

Zheng, B. (1993). ‘An axiomatic characterization of the Watts index’, *Economics Letters*. Vol. 42, pp. 81-86.

Table 1. Estimated Coefficients and Standard Errors from all Distributions

	China Rural						China U					
	PL		DPL		GB2		PL		DPL		GB2	
z1	22.429	0.075	22.432	0.075	22.437	0.076	52.291	0.174	52.537	0.185	52.557	0.185
z2	27.884	0.056	27.884	0.056	27.882	0.056	65.090	0.144	65.021	0.147	65.016	0.148
z3	33.476	0.043	33.476	0.044	33.476	0.043	84.504	0.142	84.520	0.143	84.522	0.143
z4	36.263	0.043	36.263	0.043	36.263	0.042	101.420	0.141	101.420	0.140	101.420	0.139
z5	41.828	0.051	41.828	0.051	41.828	0.051	118.600	0.151	118.600	0.149	118.600	0.149
z6	47.403	0.058	47.403	0.058	47.403	0.058	137.190	0.172	137.190	0.169	137.190	0.169
z7	55.744	0.077	55.743	0.083	55.743	0.078	158.460	0.209	158.460	0.206	158.460	0.206
z8	69.710	0.107	69.709	0.219	69.710	0.108	185.380	0.280	185.380	0.277	185.380	0.278
z9	83.596	0.149	83.594	0.221	83.600	0.150	223.490	0.432	223.480	0.431	223.490	0.434
z10	111.360	0.274	111.374	1.084	111.320	0.274	293.890	0.820	293.970	0.833	293.870	0.834
z11	139.560	0.547	139.767	0.807	139.820	0.535	380.830	1.835	380.300	1.898	381.070	1.879
m/b	3.676	0.009	3.754	0.001	25.098	3.456	4.577	0.010	4.786	0.013	107.220	3.522
σ/p	0.483	0.005	0.476	0.004	6.094	1.381	0.528	0.005	0.454	0.011	2.437	0.289
α/q	2.638	0.062	2.629	0.032	2.209	0.288	3.264	0.105	2.994	0.090	1.810	0.183
β/*a*			12.632	0.160	1.410	0.124			4.222	0.249	1.817	0.119

	India R						India U					
	PL		DPL		GB2		PL		DPL		GB2	
z1	20.827	0.050	20.876	0.053	20.897	0.053	19.926	0.057	19.921	0.057	19.871	0.056
z2	24.197	0.038	24.187	0.038	24.183	0.038	23.491	0.046	23.493	0.046	23.503	0.045
z3	28.546	0.035	28.549	0.035	28.550	0.034	28.860	0.046	28.860	0.047	28.857	0.046
z4	32.605	0.033	32.604	0.033	32.603	0.033	34.404	0.047	34.404	0.047	34.404	0.047
z5	36.623	0.034	36.623	0.034	36.623	0.034	40.028	0.051	40.021	0.052	40.028	0.052
z6	40.613	0.038	40.613	0.038	40.613	0.038	46.811	0.061	46.809	0.073	46.811	0.062
z7	45.491	0.047	45.490	0.046	45.490	0.047	55.319	0.077	55.322	0.137	55.319	0.078
z8	51.837	0.064	51.837	0.064	51.836	0.065	65.509	0.105	65.518	0.187	65.511	0.107
z9	61.517	0.104	61.516	0.105	61.518	0.105	81.482	0.177	81.418	0.524	81.477	0.176
z10	79.465	0.210	79.462	0.212	79.464	0.210	111.780	0.367	111.892	0.411	111.860	0.353
z11	102.990	0.506	103.050	0.513	103.040	0.501	151.290	0.907	151.282	0.911	150.740	0.818
m/b	3.388	0.006	3.512	0.012	28.454	0.773	3.454	0.008	3.530	0.006	2.259	3.952
σ/p	0.334	0.004	0.297	0.009	2.046	0.218	0.453	0.005	0.449	0.000	60.168	95.606
α/q	2.720	0.044	2.652	0.045	0.814	0.059	2.224	0.042	2.243	0.031	2.906	0.566
β/*a*			7.507	0.723	3.404	0.185			13.855	0.088	1.038	0.144

Table 1 (continued). Estimated Coefficients and Standard Errors from all Distributions

	Pakistan						Russia					
	PL		DPL		GB2		PL		DPL		GB2	
z1	26.635	0.066	26.786	0.071	26.788	0.071	81.023	0.324	81.261	0.337	81.272	0.336
z2	31.179	0.051	31.151	0.052	31.151	0.052	104.050	0.271	104.000	0.274	103.990	0.274
z3	37.199	0.046	37.204	0.046	37.204	0.046	137.780	0.269	137.790	0.270	137.790	0.269
z4	42.336	0.044	42.335	0.043	42.335	0.043	168.520	0.269	168.520	0.269	168.520	0.268
z5	47.315	0.046	47.315	0.045	47.315	0.045	200.560	0.292	200.560	0.291	200.560	0.290
z6	52.796	0.052	52.796	0.051	52.796	0.051	237.050	0.339	237.050	0.337	237.050	0.336
z7	59.416	0.064	59.415	0.064	59.415	0.064	281.450	0.420	281.450	0.417	281.450	0.417
z8	68.580	0.091	68.580	0.091	68.580	0.091	337.910	0.565	337.910	0.562	337.910	0.564
z9	82.671	0.150	82.669	0.153	82.670	0.153	419.610	0.886	419.600	0.884	419.610	0.889
z10	109.090	0.307	109.050	0.316	109.080	0.311	577.590	1.709	577.690	1.721	577.570	1.727
z11	141.550	0.722	142.000	0.746	141.720	0.727	755.930	3.633	755.470	3.710	756.180	3.683
m/b	3.644	0.006	3.796	0.009	39.114	0.806	5.219	0.015	5.400	0.022	174.940	15.743
σ/p	0.352	0.004	0.280	0.008	1.562	0.145	0.623	0.006	0.571	0.014	5.418	1.214
α/q	2.617	0.043	2.449	0.038	0.691	0.047	3.752	0.200	3.353	0.162	4.036	0.759
β/*a*			5.638	0.273	3.722	0.196			4.684	0.549	1.025	0.114

	Poland						Brazil					
	PL		DPL		GB2		PL		DPL		GB2	
z1	94.401	0.317	94.462	0.326	94.504	0.325	29.810	0.191	30.710	0.210	30.667	0.210
z2	115.480	0.259	115.470	0.260	115.460	0.259	46.220	0.184	46.057	0.192	46.067	0.193
z3	146.960	0.253	146.970	0.253	146.970	0.252	70.609	0.206	70.653	0.211	70.652	0.211
z4	176.000	0.251	175.990	0.251	175.990	0.250	96.137	0.230	96.126	0.230	96.128	0.229
z5	206.650	0.269	206.650	0.269	206.650	0.269	125.034	0.273	125.040	0.268	125.040	0.266
z6	240.580	0.309	240.580	0.309	240.580	0.309	161.171	0.337	161.170	0.327	161.170	0.326
z7	281.510	0.381	281.510	0.380	281.510	0.382	202.500	0.427	202.500	0.414	202.490	0.413
z8	333.130	0.507	333.130	0.505	333.120	0.509	256.863	0.614	256.870	0.597	256.870	0.599
z9	403.350	0.771	403.340	0.770	403.370	0.776	355.020	1.107	354.960	1.092	355.000	1.104
z10	532.040	1.441	532.050	1.444	531.820	1.446	577.928	2.564	578.210	2.661	577.410	2.690
z11	684.480	3.135	684.380	3.161	686.110	3.112	888.152	6.590	887.120	7.784	892.510	7.761
m/b	5.190	0.010	5.317	0.038	133.470	16.389	4.685	0.024	5.097	0.023	157.830	6.056
σ/p	0.524	0.005	0.504	0.016	5.990	1.380	0.943	0.010	0.720	0.022	1.539	0.170
α/q	3.172	0.098	3.128	0.112	3.081	0.492	2.578	0.156	2.000	0.106	1.525	0.184
β/*a*			7.595	2.486	1.227	0.125			1.891	0.064	1.311	0.094

Table 1 (continued). Estimated Coefficients and Standard Errors from all Distributions

	Nigeria						Iran					
	PL		DPL		GB2		PL		DPL		GB2	
z1	7.894	0.039	7.960	0.042	7.955	0.019	52.228	0.204	51.925	1.057	52.492	0.096
z2	10.763	0.035	10.745	0.036	10.747	0.016	66.533	0.171	67.392	0.294	66.473	0.078
z3	15.372	0.036	15.377	0.037	15.376	0.016	88.035	0.170	89.219	0.222	88.049	0.077
z4	19.722	0.037	19.721	0.037	19.721	0.017	107.646	0.171	108.972	0.252	107.644	0.076
z5	24.263	0.041	24.263	0.041	24.263	0.018	128.359	0.186	129.798	0.274	128.359	0.082
z6	29.400	0.048	29.400	0.048	29.400	0.021	152.032	0.216	153.520	0.311	152.032	0.095
z7	35.529	0.061	35.528	0.060	35.528	0.027	179.861	0.267	181.479	0.362	179.861	0.118
z8	43.728	0.084	43.729	0.083	43.729	0.037	215.887	0.361	217.552	0.450	215.886	0.160
z9	56.104	0.135	56.095	0.134	56.096	0.060	266.966	0.558	268.671	0.633	266.962	0.250
z10	79.576	0.261	79.635	0.264	79.616	0.119	359.753	1.044	361.502	1.096	359.704	0.473
z11	108.210	0.571	107.825	0.593	107.957	0.265	471.554	2.277	473.196	2.321	471.948	1.037
m/b	3.126	0.023	3.382	0.018	29.375	0.623	4.716	0.012	4.780	0.013	116.849	2.602
σ/p	0.751	0.008	0.631	0.014	3.051	0.197	0.597	0.006	0.595	0.006	3.201	0.206
α/q	3.934	0.349	2.938	0.140	3.053	0.193	3.111	0.113	3.114	0.114	2.323	0.126
β/*a*			2.918	0.147	1.107	0.040			15.769	1.060	1.397	0.047

Table 2. *J*-Statistics and Root-Mean-Square Errors

	PL		DPL		GB2	
	J-Stat	MSE	J-Stat	MSE	J-Stat	MSE
China Rural	4.8492	0.0855	5.7774	0.0855	8.2369	0.1133
China Urban	59.0320	0.3602	20.3260	0.2072	34.0740	0.2447
India Rural	29.7400	0.2521	16.0010	0.1675	27.0900	0.2629
India Urban	104.5200	0.8320	107.5038	0.7710	39.6500	0.2310
Pakistan	91.1590	0.3607	41.2590	0.5864	24.5010	0.3647
Russia	28.0730	0.2297	20.9920	0.2054	14.7390	0.1524
Poland	20.3750	0.2498	21.1840	0.2496	30.3820	0.3144
Brazil	314.8090	0.9306	116.4900	0.7407	98.24700	0.3599
Nigeria	54.9190	0.2773	5.9641	0.0893	4.4500	0.0660
Iran	42.7804	0.2281	35.4120	0.2243	28.6437	0.3091

Table 3. Gini and Theil Indexes and their Standard Errors

Gini and Theil Measures						
	PL		DPL		GB2	
	Gini	**SE**	**Gini**	**SE**	**Gini**	**SE**
China Rural	0.35980	0.00330	0.36000	0.00280	0.35740	0.00268
China Urban	0.34630	0.00250	0.34580	0.00270	0.34516	0.00322
India Rural	0.29840	0.00280	0.29980	0.00290	0.29825	0.00297
India Urban	0.38780	0.00420	0.38650	0.00420	0.37429	0.00316
Pakistan	0.31300	0.00300	0.31930	0.00330	0.31522	0.00381
Russia	0.37570	0.00240	0.37610	0.00250	0.37606	0.00250
Poland	0.34810	0.00260	0.34810	0.00260	0.34798	0.00234
Brazil	0.54460	0.00430	0.54665	0.00825	0.55523	0.00546
Nigeria	0.43046	0.00261	0.42980	0.00310	0.42948	0.00145
Iran	0.38180	0.00280	0.38200	0.00280	0.3813	0.0013
	Theil	**SE**	**Theil**	**SE**	**Theil**	**SE**
China Rural	0.25080	0.00720	0.25140	0.00570	0.23829	0.00590
China Urban	0.21510	0.00460	0.21940	0.00530	0.21592	0.00469
India Rural	0.17880	0.00480	0.18370	0.00520	0.17800	0.00505
India Urban	0.32250	0.01130	0.31880	0.01080	0.26816	0.00736
Pakistan	0.19890	0.00550	0.21720	0.00660	0.20538	0.00622
Russia	0.24740	0.00450	0.25140	0.00510	0.24968	0.00451
Poland	0.21890	0.00480	0.21910	0.00490	0.21605	0.00427
Brazil	0.58470	0.01630	0.64498	0.04208	0.67901	0.04027
Nigeria	0.32943	0.00593	0.33860	0.00810	0.33589	0.00313
Iran	0.26410	0.00590	0.26440	0.00580	0.2654	0.0026