GDAŃSK UNIVERSITY OF TECHNOLOGY
FACULTY OF MANAGEMENT AND ECONOMICS

ESTIMATES OF THE WORLD DISTRIBUTION OF PERSONAL INCOMES BASED ON COUNTRY SAMPLE CLONES

Stanislaw Maciej Kot *

GUT Faculty of Management and Economics

Working Paper Series A (Economics, Management, Statistics)

No 11/2016 (41)

November 2016

* Gdansk University of Technology, Faculty of Management and Economics,
skot@zie.pg.gda.pl (corresponding author)

ESTIMATES OF THE WORLD DISTRIBUTION OF PERSONAL INCOMES BASED ON COUNTRY SAMPLE CLONES

Stanislaw Maciej Kot*

Abstract

In this paper, the world distribution of personal incomes (WDPI) is estimated using a global sample comprising country sample clones. A clone is a random sample that reproduces – with predetermined high probability and precision – an unknown survey sample using information that is 'encoded' in the estimated parameters of a country's personal income distribution. The clone creation method is based on the sequential probability ratio test. The clones discussed in this paper are generated from the lognormal distribution using information encoded in countries' Gini indices, and are scaled to both per capita GDP and per capita household final consumption expenditures (HFCE). Statistical analysis of a global sample from the WDPI in the 1990-2010 period shows the following. The WDPI exhibited a twin-peaks shape in the initial years, but such bimodality disappeared in subsequent years. Inequality and poverty decreased in accordance with a three-phase pattern over the period. Whether clones are scaled to GDP or HFCE matters when evaluating inequality and poverty levels, but not when determining the general direction of their trends.

Keywords: global income distribution, global inequality, global poverty, estimation, sequential analysis

JEL classification: **D31, F01, O01**.

* Gdansk University of Technology, Department of Economic Sciences, Narutowicza 11/12 St., 80-233 Gdansk, Poland, e-mail: skot@zie.pg.gda.pl.

The author is grateful for the financial support of the Polish National Science Center, grant no. N N111 295638.

I. Introduction

This paper presents new estimates of the world distribution of personal incomes (WDPI) in the 1990-2010 period. Its main novelty is the use of a specially constructed global random sample comprising the 'clones' of unknown random samples drawn from national household surveys. The global sample of incomes is used to estimate all descriptive statistics of the WDPI and measures of inequality and poverty. Our biotechnological terminology is inspired by a parallel of DNA that encodes the complete structure of an organism. DNA from somatic cells is the basis of cloning, i.e. the process of creating a similar population of genetically identical individuals.

We define a clone as a random sample that reproduces – with predetermined high probability and precision – a survey sample using information that is 'encoded' in the estimated parameters of a country's personal income distribution. We propose a clone creation method that is based on the sequential probability ratio test (SPRT) (Wald, 1945, 1947; Wald and Wolfowitz, 1948; Fisz, 1967; Gosh and Sen, 1970). The global random sample from the WDPI is a collection of country clones in which each individual observation has a weight equal to the ratio of a country's population and the size of the clone.

In this paper, we create clones of unknown samples from national household surveys, assuming the lognormal distribution (Aitchison and Brown, 1997) of country-level personal incomes. The lognormal distribution is commonly used to assess global inequality and poverty (see, e.g., Chotikapanich et al., 1997, 1998; Milanovic, 2002; Lakner and Milanovic, 2013; López and Servén, 2006; Pinkovsky and Sala-i-Martin, 2009; Shorrocks and Wan, 2008; Liberati, 2015). Assuming the mean of unity,[1] *all information* on an unknown survey sample is *encoded* solely in estimates of the Gini index.

We use the Gini index estimates from the Standardised World Income Inequality Database (SWIID) developed by Solt (2014). We scale the unit-mean clones to both per capita GDP and per capita household final consumption expenditures (HFCE) (PPP adjusted, constant 2011 international $). Anand and Segal (2015) recommend HFCE rather than GDP for scaling. Our GDP and HFCE data come from the World Development Indicators (World Bank, 2015). The global sample from the WDPI comprises the rescaled clones of all countries. We apply our own elaborated FORTRAN90 programs to perform the sequential procedure, and use Stata12 software assisted by the DASP module (Araar and Duclos, 2009)

[1] This assumption plays a technical role. Every random variable Y with non-zero mean μ can be normalised to unit-mean random variable $X = Y/\mu$.

to estimate the inequality and poverty indices and their standard errors and the kernel density functions of the WDPI. Statistical analysis of the global sample for the 1990-2010 period corroborates the following hypotheses:

Hypothesis 1. The WDPI is bimodal in the initial years of the sample period (the twin-peaks hypothesis[2]), with bimodality gradually disappearing in subsequent years.

Hypothesis 2. Mean income and cosmopolitan social welfare are both increasing in the sample period.

Hypothesis 3. Inequality and poverty trends exhibit three phases separated by the years 1996 and 2002: an inverted U-shaped phase (or slowly decreasing phase), a plateau phase and a rapidly decreasing phase**.**

Hypothesis 4. Whether clones are scaled to per capita GDP or per capita HFCE matters when assessing global inequality and poverty levels, but not when determining the general direction of change over time.

The WDPI concept is a global analogue of a country's income distribution. However, its global scale renders analysis of the WDPI substantially more difficult than the typical analysis of income distributions at the country level (Lakner and Milanovic, 2013). In the absence of a global household survey, analysts must resort to combining national surveys. The collection of national household survey data from as many countries as possible (ideally all of them) would be ideal for this purpose, and has been accomplished by Milanovic (2002, 2005) and the World Bank (2005). However, such ideal data are available for only a limited number of countries and years. Moreover, even when such data exist, they are often rendered useless for spatial comparisons by differences in definitions of income, the recipient units used and the scope of coverage in different surveys (Fields, 1994; Deininger and Squire, 1996; Chotikapanich et al., 1997; Solt, 2014).

In recent years, two databases have been commonly used: the World Bank DS database compiled by Deininger and Squire (1996) and the World Income Inequality Database (WIID) compiled by the United Nations University's World Institute for Development Research (UNU-WIDER). These databases present national micro-data on incomes or expenditures in the form of summary statistics such as quantile income shares and Gini indices and/or means and/or medians.

Developing efficient methods of extracting underlying micro-data from existing summary statistics seems to be a remedy for the scarcity of available data from national

[2] See Quah (1993, 1997), Jones (1997), and Kremer, Onatsky and Stock (2001)

household surveys (Shorrocks and Wan, 2008). Several methods have been proposed. In the following paragraphs, we present those approaches most relevant to the problem in question.

The POVCAL software on the World Bank website[3] is commonly used to generate raw data from quantile data. However, Shorrocks and Wan (2008) show that this software can generate negative values even when the data are positive. Furthermore, the quantile shares obtained can differ significantly from the reported values with which the procedure begins (Minoiu and Reddy, 2006).

Sala-i-Martin (2006) uses quintile income shares from the DS and UNU-WIDER databases and approximates each country's annual income distribution by a kernel density function. He then decomposes the kernel density-based quintiles into kernels with 100 centiles, thereby obtaining a sample of 100 personal incomes for each country. Finally, he integrates the samples to estimate several measures of poverty and inequality.

Pinkovsky and Sala-i-Martin (2009) modify the aforementioned method by applying two-parameter lognormal distributions of personal income rather than the nonparametric kernel density function. For each year, the lognormal distributions of income for all countries are integrated to construct an estimate of the world distribution of income and poverty and inequality measures. However, the authors do not present details of this integration. It seems that they calculate the centiles of the fitted lognormal distributions.

Shorrocks and Wan (2008) make significant progress in ungrouping summary statistics to obtain a country's individual income observations. They use quantile (decile or quintile) income shares and propose a two-stage algorithm.[4] Stage I fits a parametric distribution (with the unit mean) to the grouped observations, and then generates a sample of quantiles of the order i/n, $i = 1,\ldots,n\text{-}1$ ($n = 1000$ or 2000) from the fitted distribution as an initial approximation to the synthetic observations. Stage II of the algorithm then takes the raw sample and adjusts the values until the sample statistics match the 'true' figures exactly.

Lakner and Milanovic (2013) recently used the Shorrocks-Wan algorithm to build a database of national household surveys across five benchmark years: 1988, 1993, 1998, 2003 and 2008. Their data consist of country-year average deciles of income/consumption.

The foregoing methods of ungrouping summary statistics suffer from three limitations. First, statistical inferences concerning the WDPI are limited when relaying the available data on quantile shares. The DS database and WIID contain data on only about 50% of countries,

[3] The POVCAL software can be accessed from www.worldbank.org/LSMS/tools/povcal/.

[4] This algorithm is available by calling up the 'ungroup' command in the DASP package (a submenu of the Stata software (Araar and Duclos, 2009).

and it is unclear whether such an incomplete sample is representative of the whole world. Almost all countries report Gini estimates, however, which means that 50% of potentially useful information is set aside when dealing with quantile income shares alone. There is thus a need to develop efficient methods of using Gini indices to extract raw data.

Second, the resulting samples of quantiles consist of *dependent* observations. Dependent data may give rise to serious statistical problems concerning, for example, estimation of the Gini index and its standard errors (e.g. Modarres and Gastwirth, 2006; Ogwang, 2006).

Third, the size of the generated samples is arbitrarily set, and the weights assigned to individual observations are population shares. Such arbitrary settings affect the properties of the estimators of the WDPI parameters, e.g. the standard errors of the estimators. In the stratifying sampling scheme, the weights are $w_i = pop_i/n_i$, i.e. the inverse of the probability of inclusion of an individual from the ith country in the global sample, where n_i and pop_i are the size of the generated sample and population of the ith country, respectively, and $i = 1,...,K$, K is the number of countries. If $n_i = n$ for all $i = 1,\ldots,K$, the weights are $w_i = pop_i/n$. Then, the size of the global sample is $\sum w_i = (1/n)\sum pop_i$. The cited authors apply wages $w_i^* = pop_i$ instead of pop_i/n, and therefore the size of the global sample is $\sum pop_1 = n\sum w_i$. In other words, it is n times greater than the properly calculated sample size.

Liberati (2015) recently generated random samples assuming the lognormal distribution of countries' personal incomes. His method uses Gini index estimates from the WIID and per capita GDP as the estimate of the mean, thereby overcoming the first two limitations discussed above. However, his proposal concerning the size of the generated samples is controversial. Liberati (2015) assumes sample size $n_{CHN} = 500{,}000$ for China, whilst other countries are assigned a number of observations n_i that is proportional to the ratio of their population to that of China, i.e. $n_j = 500{,}000 \cdot pop_i/pop_{CHN}$, where pop_{CHN} and pop_i are the populations of China and the ith country, respectively, and $i = 1,\ldots,K$. Unfortunately, the size of the Chinese sample far exceeds the number of persons in China's household surveys. For example, there were 102,860 surveyed households in 1995 (urban and rural), with an average household size of 3.23 persons (Fang et al., 1998). These figures give 332,238 surveyed persons. By assuming 500,000, Liberati is artificially 'improving' the data. He applies the weights n_i when forming the global sample, whereas properly calculated weights would be equal to pop_i/n_i. Therefore, Liberati's approach implies that weight w_i =

q_{CH}/500,000, i.e. the same for each of K countries, which means that he obtains something other than the total sample from the WDPI.

The cloning method proposed in this paper is free from all of the aforementioned limitations. It uses Gini indices alone to generate clones that are random samples of independent observations. The sequential procedure results in a sample size that is just sufficient to support the null hypothesis, with an assumed high degree of probability, that a clone comes from the same distribution as the cloned sample from a household survey.

The remainder of this paper is organised as follows. Section II presents our methodology for creating clones of micro-data from household surveys. Details of the data used are provided in Section III. The empirical results are given in Section IV, and Section V offers a summary and conclusions.

II. Methodology

2.1. The cloning of unknown samples

Let personal incomes in a given country and year be described by positive continuous random variable X with density function $f(x|\theta)$, $X \sim f(x|\theta)$ for short, where θ is a parameter or vector of parameters.[5] Let $\boldsymbol{x} = (x_1,\ldots,x_N)$ be a random sample from this distribution, and $\hat{\theta}_0$ be an estimate of θ based on that sample. Sample $\boldsymbol{x}$ may come from a national household survey, for example.

We apply the sequential statistical procedure (Wald, 1945; Fisz, 1967; Gosh and Sen, 1991).[6] We predetermine a small constant $\delta > 0$ and significance level α, as well as the probability β of committing an error of the second kind (i.e. accepting the null hypothesis when the alternative hypothesis is true). We generate successive random numbers x_i' from the distribution $X' \sim f(x|\hat{\theta}_0)$, $i = 1,2,\ldots,$ and control the procedure by using the SPRT as the stopping rule. Here, the SPRT is to verify the null hypothesis,

H_0: the sample comes from distribution $f(x|\hat{\theta}_0)$,

against the alternative hypothesis,

H_1: the sample comes from distribution $f(x|\hat{\theta}_1)$,

where $\theta_1 = \hat{\theta}_0 + \delta$.

The sequential procedure is halted at n when the null hypothesis is accepted.

[5] Hereafter, capital letters are reserved for random variables, and small letters for the values of random variables.

[6] The details of the sequential procedure are presented below.

Definition 1. Original sample $\boldsymbol{x} = (x_1,\ldots,x_N)$ is called the *cloned sample*.

Definition 2. Random sample $\boldsymbol{x}' = (x_1',\ldots,x_n')$ resulting from the foregoing sequential procedure is called the *δ-clone* of the original sample $\boldsymbol{x}$, $n \leq N$. We use the abbreviation 'the clone' when δ is known from the context.

Definition 3. Constant δ refers to the *precision of the cloning procedure*.

Our biotechnological terminology is inspired by a parallel of DNA, which encodes the whole structure of an organism. DNA from somatic cells is the basis of cloning, i.e. the process of creating similar populations of genetically identical individuals. Note that all information on original sample $\boldsymbol{x}$ is encoded in the $\hat{\theta}_0$ estimate. If the parametric form $f(x|\theta)$ of the income distribution and $\hat{\theta}_0$ are known, but $\boldsymbol{x}$ is unknown, the δ-clone retrieves the information contained in $\boldsymbol{x}$ with predetermined high probability $1-\beta$ and precision δ. The clone's size n is sufficient to distinguish between distributions $f(x|\hat{\theta}_0)$ and $f(x|\hat{\theta}_0+\delta)$. We can say that clone $\boldsymbol{x}'$ is genetically identical to unknown original cloned sample $\boldsymbol{x}$ with probability $1-\beta$ and precision δ.

2.2. Cloning samples from lognormal distributions

Let $X = Y/y$ be the normalised version of the distribution of personal incomes Y with mean $E[Y] = y > 0$. The mean of X equals 1, and the Gini index in the X distribution is exactly the same as that in the Y distribution. In this paper, we assume that X follows two-parameter lognormal distribution $\Lambda(\mu,\sigma)$ with density function

$$f(x) = \frac{1}{x\sigma\sqrt{2\pi}} \exp\left\{-\frac{(\log x - \mu)^2}{2\sigma^2}\right\}, \qquad x,\ \sigma > 0 \tag{1}$$

(Aitchison and Brown, 1957; Kleiber and Kotz, 2003 p. 107].

The lognormal distribution is commonly used as the theoretical distribution of country incomes when assessing global inequality and poverty (see, e.g. Chotikapanich et al., 1997, 1998; Milanovic, 2002; López and Servén, 2006; Pinkovsky and Sala-i-Martin, 2009; Shorrocks and Wan, 2009; Liberati, 2015]. Many forms other than lognormal are offered as approximations of empirical income distributions (see Kleiber and Kotz (2003) for a review), but to the best of our knowledge standard goodness-of-fit tests usually reject them when confronted with empirical data. Therefore, any discussion of the advantages of some distributions over others seems immaterial from a purely statistical point of view.

Furthermore, the practice provides ambiguous evidence concerning the poor performance of the lognormal distribution even in light of *ad hoc* chosen criteria. For

instance, Chotikapanich and Griffiths (2008) maintain that the gamma distribution outperforms the lognormal, Beta2, Dagum and Singh-Maddala distributions of income. However, their conjecture seems questionable when relying on the Kolmogorov-Smirnov test.[7] Also, the *ad hoc* choice of the root-mean-square error (RMSE) criterion does not seem decisive in favouring the gamma distribution over the lognormal and other distributions because it is unclear whether 'by-eye' observed differences between the RMSE values for the selected distributions are statistically significant. Shorrocks and Wan (2008), in contrast, observe the lognormal distribution to perform 'surprisingly well' in comparison with the general quadratic, beta, generalised beta and Singh-Maddala distributions. However, no goodness-of-fit test is referred to in their paper.

The lognormal distribution of incomes with the mean of unity, i.e. E[X] = 1, implies the following formulae for parameters σ and μ.

$$\sigma = \Phi^{-1}\left(\frac{G+1}{2}\right)\sqrt{2}\,; \tag{2}$$

$$\mu = -\frac{\sigma^2}{2} = -\left[\Phi^{-1}\left(\frac{G+1}{2}\right)\right]^2, \tag{3}$$

where G is the Gini index and $\Phi(\cdot)$ is the cumulative distribution function of standard normal distribution N(0,1).

It follows from (2) and (3) that the Gini index provides complete information on lognormal distribution $\Lambda(\mu,\sigma)$. In other words, *all information* contained in the random sample drawn from that distribution is *encoded* solely in the estimate of the Gini index. If the original sample of individual incomes obtained from national household surveys is unavailable, but the estimate of the Gini, say G_0, is, then we can use the information encoded in G_0 to generate a clone of the original sample. That clone contains the same 'genetic' information on the distribution of personal income in the general population as the original (cloned) sample.

We propose a sequential procedure (Wald, 1945; Fisz, 1967, Chapter 17) for creating the clone of the original (cloned) sample. Let G_0 be the Gini index estimate based on the original sample of incomes. Let $f_0(x)$ be the density function of lognormal distribution (1) $\Lambda(\mu_0,\sigma_0)$, where μ_0 and σ_0 are calculated from (2) and (3), respectively, given G_0. Let $G_1 = G_0$

[7] In fact, the authors use the Kolmogorov-Smirnov test erroneously. That test can be applied if the hypothesised distribution is fully specified, i.e. if *theoretical* distribution function $F(x)$ is known, but Chotikapanich and Griffiths (2008) use *estimated* distribution function $F(x|\mathbf{x})$ instead of $F(x)$. For such a case, there is no general asymptotic theory (D'Agostine and Stephens, 1986). Therefore, the authors' calculation of p values from the Kolmogorov λ distribution (or applying 5% critical value D_α= 1.358) is unjustified.

+ δ, where $\delta > 0$ is a small number. Let $f_1(x)$ be the density function of lognormal distribution (1) $\Lambda(\mu_1,\sigma_1)$, where μ_1 and σ_1 are calculated from (2) and (3), respectively, given G_1. We then generate a random sample from $\Lambda(\mu_0,\sigma_0)$ and apply the SPRT to verify the following hypotheses against each other.

H_0: The generated sample (clone) comes from lognormal distribution $\Lambda(\mu_0,\sigma_0)$.

H_1: The generated sample (clone) comes from lognormal distribution $\Lambda(\mu_1,\sigma_1)$.

The sequential procedure runs as follows. We first generate a random number, say u_1, from standard normal distribution U~N(0,1).[8] The corresponding value x_1 of the lognormal distribution under the null is[9]

$$x_1 = \exp\{\sigma_0 u_1 + \mu_0\}. \tag{4}$$

We then calculate the value of the SPRT:

$$Z_1 = \log \frac{f_1(x_1)}{f_0(x_1)}. \tag{5}$$

Let A and B satisfy inequality $0 < B < 1 < A$. We can write $a = \log A > 0$, $b = \log B < 0$. The decision regarding our hypotheses will depend on the position of Z_1 in relation to a and b, as follows.

C1. If $Z_1 \leq b$, we accept H_0.

C2. If $Z_1 \geq$ a, we reject H_0 and accept H_1.

C3. If $b < Z_1 < a$, we generate the next random number u_2 from the standard normal distribution, calculate x_2 using (4) and then calculate Z_2:

$$Z_2 = \log \frac{f_1(x_1)f_1(x_2)}{f_0(x_1)f_0(x_2)}. \tag{6}$$

Next, we compare Z_2 with a and b and make a decision according to conditions C1, C2 or C3.

In general, if a sample of size m-1 does not allow us to accept either H_0 or H_1, the decision-making in the next step depends on the following quantity.

$$Z_m = \log \frac{f_1(x_1)f_1(x_2)...f_1(x_m)}{f_0(x_1)f_0(x_2)...f_0(x_m)}, \; m=1, 2, \ldots, \tag{7}$$

which can be expressed as

$$Z_m = \sum_{i=1}^{m} \log \frac{f_1(x_i)}{f_0(x_i)}. \tag{8}$$

[8] In fact, generating random numbers from the uniform distribution (on the [0,1] interval) is sufficient to generate random samples from any probability distribution.

[9] For simplicity, we use the symbol x_i instead of the x_i' that was used in Subsection *II.1* for the clone values.

If $Z_m \leq b$, we accept H_0. If $Z_m \geq a$, we reject H_0 and accept H_1. If $a < Z_m < b$, we draw the $m+1$ element, calculate Z_{m+1} and compare it with a and b.

Wald (1945) proved the theorem that if the variance of the random variable[10]

$$Z = \log \frac{f_1(x)}{f_0(x)} \tag{9}$$

exists and differs from zero, then the sequential procedure is finite (almost surely), i.e. either H_0 or H_1 is accepted after a finite number of steps with probability 1 (see also Fisz [1967, p. 588]).

In the sequential test, as in the classical testing procedure, it is possible to commit two kinds of errors. Let α be the probability of committing an error of the first kind (rejecting H_0 when H_0 is true). Let β be the probability of committing an error of the second kind (accepting H_0 when H_1 is true). Wald (1945) showed that A and B can be approximated by

$$A \cong \frac{1-\beta}{\alpha}, \; B \cong \frac{\beta}{1-\alpha} \tag{10}$$

(see also Fisz [1967, p. 597]). Although the closed-form solutions of A and B are available for several probability models, it is standard practice to use Eq. (10) to approximate the stopping bounds in applications.

It is easy to show that random variable Z (9) has the following form for the lognormal distribution,

$$Z = \log \frac{\sigma_0}{\sigma_1} + \frac{1}{2}\left(\frac{\log x - \mu_0}{\sigma_0}\right)^2 - \frac{1}{2}\left(\frac{\log x - \mu_1}{\sigma_1}\right)^2, \tag{11}$$

and that its variance exists and differs from zero. From Wald's theorem, it follows that the sequential procedure of clone generation is finite (with probability 1). Given (11), Z_n (8) can be expressed as

$$Z_n = n \cdot \log \frac{\sigma_0}{\sigma_1} + \frac{1}{2}\sum_{i=1}^{n}\left(\frac{\log x_i - \mu_0}{\sigma_0}\right)^2 - \frac{1}{2}\sum_{i=1}^{n}\left(\frac{\log x_i - \mu_1}{\sigma_1}\right)^2. \tag{12}$$

To illustrate the stopping rule, we express (12) in an alternative form. From (4), we obtain $\log x_i = \sigma_0 \cdot u_i + \mu_0$ and substitute it in (12). We can also express μ_0 and μ_1 in (12) as $-\sigma_0^2/2$ and $-\sigma_1^2/2$, respectively. After simple algebra, we get

$$Z_n = d \cdot n + c\sum_{i=1}^{n}(u_i^2 - \sigma_0\sigma_1 u_i), \tag{13}$$

where

[10] In Wald's theorem, $f_0(x)$ and $f_1(x)$ denote two density functions of any continuous random variables.

(14) $$c = \frac{1}{2}\left(1 - \frac{\sigma_0^2}{\sigma_1^2}\right), \quad \text{and } d = \log\frac{\sigma_0}{\sigma_1} - \frac{\sigma_1^2 c^2}{2}.$$

Note that $c > 0$ (because $\sigma_0 < \sigma_1$) and $d < 0$.

Accordingly, the stopping rule is

(15) $$b < d \cdot n + c\sum_{i=1}^{n}(u_i^2 - \sigma_0\sigma_1 u_i) < a\,.$$

After rearranging, we finally find that

(16) $$\frac{b}{c} - \frac{d}{c}n < \sum_{i=1}^{n}(u_i^2 - \sigma_0\sigma_1 u_i) < \frac{a}{c} - \frac{d}{c}n\,.$$

Now, the thresholds are simply two parallel lines with positive slope *-d/c*. Sampling should stop when component $\sum_{i=1}^{n}(u_i^2 - \sigma_0\sigma_1 u_i)$ makes an excursion outside the continued-sampling region.

Fig. 1 illustrates the stopping rules for $G_0 = 0.3$ and $\delta = 0.03$. We assume that $\alpha = \beta = 0.05$. The u_i*s* are pseudo-random numbers from the standard normal distribution.

(Fig. 1 about here)

It can be seen from Fig.1 that a sample of size 90 is sufficient to accept H_0 with probability $1 - \beta = 0.95$ that it is true.[11] This sample can thus be recognised as the clone of the original sample. The clone can distinguish between G_0 and $G_0 + \delta$, i.e. between 0.3 and 0.33.

2.3. Expected clone size

An essential feature of the SPRT is that the number of observations required is not predetermined, but is a random variable. In general, the SPRT requires an expected number of observations that is considerably smaller than the fixed number of observations needed by the most powerful classical tests, which control errors of the first and second kind to exactly the same α and β. In this sense, the SPRT is an optimal test. Sequential analysis frequently results in approximately 50% savings in the number of observations compared with the most powerful classical tests (Wald, 1945).

The expected number of observations plays an important role in the sequential test. Let $E_0[Z_n]$ denote the conditional expected value of random variable Z_n, provided that $Z_n \leq \log B$, where the expectation is calculated with respect to $f_0(\cdot)$. We need not consider the case of

[11] This probability is outside the researcher's control in classical tests.

$E_1[Z_n]$, i.e. the mean with respect to $f_1(\cdot)$, because we are concerned with accepting the null hypothesis.

If we can neglect the difference $\log B$-Z_n, that is, if we can assume that $Z_n = \log B$ when the decision is made, then, for $E_0[Z]\neq 0$, the formula for the expected sample size is

$$E_0[N] \cong \frac{(1-\alpha)\log\frac{\beta}{1-\alpha} + \alpha\log\frac{1-\beta}{\alpha}}{E_0[Z]} \tag{17}$$

(see Fisz, 1967, p. 595).[12]

Note that the denominator of (17) can be expressed as

$$E_0[Z] = \int_{-\infty}^{\infty} f_0(x)\log\frac{f_1(x)}{f_0(x)}dx = -\int_{-\infty}^{\infty} f_0(x)\log\frac{f_0(x)}{f_1(x)}dx = -D(f_0 \,\|\, f_1), \tag{18}$$

where $D(f_0\|f_1)$ is the Kulback-Leiber (1951) divergence (KLD) between distributions $f_0(x)$ and $f_1(x)$. If we substitute (18) into formula (17), we obtain

$$E_0[N] \cong \frac{(1-\alpha)\log\frac{1-\alpha}{\beta} - \alpha\log\frac{1-\beta}{\alpha}}{D(f_0 \,\|\, f_1)}. \tag{19}$$

Note that decreasing α or β increases the expected sample size. The expected sample size also increases as $D(f_0||f_1)$ decreases, i.e. as the two densities become less distinguishable.

It is easy to show that the KLD between two considered lognormal distributions is equal to

$$D(f_0 \,\|\, f_1) = \log\frac{\sigma_1}{\sigma_0} + \frac{\sigma_0^2}{2\sigma_1^2} + \frac{1}{8\sigma_1^2}[\sigma_0^2 - \sigma_1^2]^2 - \frac{1}{2}, \tag{20}$$

and differs from zero because $\sigma_1 \neq \sigma_0$. After substituting (20) into (19), we obtain the expected number of observations that allows us to accept the null hypothesis:

$$E_0[N] \cong \frac{(1-\alpha)\log\frac{1-\alpha}{\beta} - \alpha\log\frac{1-\beta}{\alpha}}{\log\frac{\sigma_1}{\sigma_0} + \frac{\sigma_0^2}{2\sigma_1^2} + \frac{1}{8\sigma_1^2}[\sigma_0^2 - \sigma_1^2]^2 - \frac{1}{2}}. \tag{21}$$

According to (2) and the aforesaid hypotheses, $\sigma_0 = \Phi^{-1}[(G_0+1)/2]\sqrt{2}$ and $\sigma_1 = \Phi^{-1}[(G_0+\delta+1)/2]\sqrt{2}$. It can be seen that the smaller δ is, the smaller $D(f_0||f_1)$ (the denominator of (21)) is, which means that decreasing δ increases the expected size of the generated clone when accepting the null hypothesis.

[12] We adopt the original Fisz formula (17.6.5) for our symbols.

An analyst may apply either an absolute δ, i.e. the same for all countries' estimates of G_0, or a relative δ, e.g. a fraction of G_0. Fig. 2 illustrates the expected sample sizes for absolute $\delta = 0.006$ and relative $\delta = 0.015$f G_0 for G_0 ranges from [0.15 to 0.8]. We assume that $\alpha = \beta = 0.05$.

(Fig. 2 about here)

It can be seen from Fig. 2 that the choice of δ (absolute or relative precision) influences the expected sample size necessary to accept the null hypothesis. However, the empirical results presented in Section III suggest that the choice between absolute and relative δ does not matter when analysing inequality in the WPDI.

2.4. Forming a global sample from the WDPI

The created clone $(x_1,...,x_n)$ has a unit mean by assumption. We can scale the values of the clone to any desired constant $h > 0$ by multiplying each x_i by h. If $h = \bar{y}$, the scaled clone has the same mean as the original sample $(y_1,...,y_N)$. However, neither the DS database nor WIID provides mean income estimates for all collected countries and years. For this reason, economists commonly use per capita GDP data from national accounts to scale results derived from quantile income shares[13] (see, among others, Korzeniewicz and Moran, 1997; Chotikapanich et al., 1997; Bourguignon and Morrisson, 2002; Dikhanov and Ward, 2002; Dowrick and Akmal, 2005; Sala-i-Martin, 2006; Bourguignon, 2011). Anand and Segal (2015) argue for scaling samples to per capita HFCE rather than per capita GDP, and Lakner and Milanovic (2013) were the first to scale survey incomes to HFCE.

The proposed method for generating a random sample from the WDPI can be treated as a stratifying sampling scheme. We define the general population as the set of incomes of persons (statistical units) living in the selected K countries (strata). If we have scaled clones of size $n_1,...,n_K$ for each K country, then we can form a global sample of size $n_1 + n_2 + ... + n_K$. The probability of including individual personal income x_{ij} to the jth stratum is equal to n_j/q_j, where q_j is the population of the jth country, $i = 1,...,n_j$, $j = 1,...,K$. Hence, the weight w_{ij} assigned to each individual observation in the total sample is equal to q_j/n_j, i.e. to the inverse of the probability of inclusion.

[13] Quantile income shares enable the income distribution function to be calculated up to a scale parameter (see, e.g. Iritani and Kuga, 1983).

III. Statistical data

In this paper, we use the Gini index estimates from the SWIID developed by Solt (2014). Assuming LIS standards, Solt employs a custom missing-data algorithm. He uses data drawn from regional collections, national statistical offices and academic studies. As a result, SWIID provides comparable estimates of the Gini index of net- and market-income inequality for 174 countries from 1960 to 2013. As SWIID's coverage and comparability far exceed those of alternate datasets, it is better suited to broad cross-national research on income inequality than other sources. Version 5.0 of the SWIID (SWIID5) dataset is available on Solt's website: http://myweb.uiowa.edu/fsolt.

Our analysis covers the 1990-2010 period. We generate K clones of household samples of personal income in K countries that provide statistical data. We assume relative precision $\delta = 0.015 \cdot G_0$ and $\alpha = \beta = 0.05$ when generating the clones.

A problem with outliers can arise when pseudo-random numbers are generated[14] to create clones. It is well known that welfare indicators estimated from micro-data can be highly sensitive to the presence of a few extreme incomes (Cowell and Victoria-Feser, 1996a, 1996b, 2002). Several approaches have been proposed to make estimates robust to outlying observations (see Van Kerm (2007) for a review). However, we decided not to follow those approaches. Lakner and Milanovic (2013) show that the inclusion of top incomes makes the resulting income distributions more reliable than those without top incomes. Allowing our procedure to generate unrestricted incomes renders the appearance of top incomes more probable than a procedure with restrictions. Obviously, the thin upper tail of the lognormal distribution does not guarantee that top incomes are generated as frequently as they are in the Pareto distribution, for example.

We scale the clones to both per capita GDP and per capita HFCE (both PPP adjusted, constant 2011 international $). The GDP and HFCE data come from the World Development Indicators (World Bank, 2015). We form the global sample of personal incomes by pulling all of the clones together. We calculate the weights $w_{ij} = q_j/n_j$ assigned to each individual income in the global sample, where q_j is the population of the jth country and n_j is the size of the clone for the jth country, $j = 1,...K,\ i = 1,...,n_j$. The country population data also come from the World Development Indicators (World Bank, 2015).

We use the World Bank's estimates of GDP rather than those in the Penn World Table (PWT). The World Bank and PWT use different methods to calculate PPP. The World Bank

[14] We use the subroutine RNNOA from the FORTRAN90 MSIMSL library, assuming the same seed of 23457 for every country when generating random numbers.

uses the Eletö-Köves-Szulc (EKS) method, whereas the PWT uses the Geary-Khamis (GK) method. Anand and Segal (2015) argue that the EKS method has an advantage over the GK when the concern is with international comparisons of living standards and inequality.

The procedure for scaling the clones to per capita GDP or per capita HFCE reduces the initial number of countries in SWIID5 because of gaps in the national account data on GDP and HFCE in the World Development Indicators (World Bank, 2015). Table 1 presents the number of countries and their share of the world population for SWIID5 and its subsets according to the version used to scale the clones (i.e. GDP or HFCE).

(Table 1 about here)

It can be seen from Table 1 that the SWIID5 database covers a large fraction of the world population.[15] The extent to which the scaling version used reduces the initial number of countries in that database can also be seen from the table.

IV. Results

4.1. The shape of the WDPI

To aid visualisation of the WDPI's shape, we estimate the density function by the kernel method using the 'Distributions' module of the DAD4.5 package (Araar and Duclos, 2006). To make that visualisation feasible, we plot the estimated density functions for only three years: 1990, 2000 and 2010. The natural logarithms of income are on the horizontal axis.

Fig. 3 displays the kernel estimates of the WDPI density functions when the clones are scaled to per capita HFCE.

(Fig. 3 about here)

We can see in Fig. 3 that the WDPI is bimodal in 1990. That bimodality then vanishes gradually over the next few years, eventually disappearing entirely in 2010.

Fig. 4 displays the kernel estimates of the WDPI density functions when the clones are scaled to per capita GDP.

(Fig. 4 about here)

The bimodality of the WDPI in 1990 is more visible in Fig. 4 than in Fig. 3. We can also observe the gradual disappearance of that bimodality, with a unimodal WDPI appearing in the final year.

[15] The number of countries is smaller than that in SWIID5 because of the lack of population data for a small group of countries.

It follows from Figs. 3 and 4 that the version of scaling used, i.e. whether the clones are scaled to per capita GDP or HFCE, is unimportant for general changes in the WDPI's shape, although the position of its probability mass does depend on the scaling version.

The basic descriptive statistics of the WDPI are presented in Tables 2 and 3, which show scaling to per capita HFCE and GDP, respectively. For the sake of simplicity, the terms 'GDP' and 'HFCE' following the given descriptive statistics denote the version used in the country's sample scaling.

(Tables 2 and 3 about here)

Analysis of the results in Tables 2 and 3 reveals an increase in the mean and median of the WDPI over time. The WDPI in the general population is positively skewed, which means that the percentage of people with incomes below the mean is greater than that of those with incomes above the mean. It can also be seen that the HFCE version of the WDPI exhibits greater dispersion, skewness and concentration than the GDP version in all years.

Fig. 5 displays the means of personal incomes scaled to per capita HFCE and per capita GDP.

(Fig. 5 about here)

4.2. Global inequality

Tables 4 and 5 show global inequality and global (or cosmopolitan) social welfare.

(Tables 4 and 5 about here)

It can be seen from the two tables that the level of global inequality, measured by the Gini index and Theil's (1967) entropy index (with $\theta = 1$), is high. The Gini index exceeds 0.6 in all years considered. Fig. 6 shows the trends in the index for both HFCE and GDP scaling.

(Fig. 6 about here)

The graphs in Fig. 6 reveal three phases of global inequality changes separated by the years 1996 and 2002. In the first phase, inequality follows an inverted U-shaped trend. In the second phase, inequality remains stable (HFCE scaling) or decreases slowly (GDP scaling). Finally, in the third phase, a rapid decrease in inequality can be observed.

Global inequality trends measured by the Theil entropy index are plotted in Fig. 7.

(Fig. 7 about here)

It can be seen from Fig. 7 that the trends of the Theil index are similar to those of the Gini index in Fig. 6. Analysis confirms the qualitative hypothesis of the three-phase pattern of global inequality changes, although the quantitative differences are remarkable. Analysis also shows that the sample scaling version used, i.e. scaling to either HFCE or GDP, is important

in assessments of inequality levels, but has no influence on the general pattern of global inequality trends.

Our evaluation of global inequality and its trends in the 1990-2010 period contradicts the findings of Lakner and Milanovic (2013). They make Pareto-type adjustments to address the possible under-reporting of top incomes in household surveys and obtain a Gini index estimate of about 0.76 without a visible downward trend.

The cosmopolitan welfare function (CWF) is the global counterpart of the social welfare function (SWF). Here, we use the cosmopolitan SWF (CSWF) from Sen (1973), where *CSWF* = mean(1-Gini). Fig. 8 displays the CSWF trends.

(Fig. 8 about here)

Fig. 8 reveals three phases of change in global welfare, separated by the years 1994 and 2002. In the first phase, global welfare is constant. In the second phase, it is increasing, and, in the third phase, accelerated welfare growth can be observed.

4.3. Global poverty

We apply two poverty line measures in analysing global poverty. The first measure, an absolute poverty line of $1.25/person/day is used in the World Development Indicators (World Bank, 2015). The second, the relative poverty line, is equal to half the per capita mean income (for scaling to HFCE or GDP, see Tables 2 and 3, respectively). We use the FGT_0 poverty measure from Foster, Greer and Thornbecke (1984), i.e. a head-count ratio that measures the *incidence* of poverty. Table 6 presents the estimates of this poverty index for the absolute and relative poverty lines, with both versions of clone scaling considered.

(Table 6 about here)

Fig. 9 displays the global poverty trends for the absolute poverty line.

(Fig. 9 about here)

In Fig. 9, the head-count ratio based on HFCE-scaled data shows that the world's percentage of poor declined from 14% in 1990 to 8% in 1998, stabilised at the 8% level in the 1998-2006 period and then declined further to 6% in 2010. The GDP-scaled time-series of the FGT_0 also indicates a general decrease in the incidence of poverty from about 5% in the initial years of the period analysed to about 2% in 2010, although the level and pattern of changes in this measure of poverty differ from those in the HFCE-scaled time-series. The GDP-FGT_0 is more or less constant from 1990 to 1994, then decreases in the 1995-1999 period, stabilises in the 2000-2005 period and decreases again in the following years.

Fig. 10 displays the trends of the FGT_0 indices when the relative poverty line (equal to half the mean income) is used instead of the absolute.

(Fig. 10 about here)

The HFCE-scaled graph of FGT_0 shows that the incidence of poverty was initially stable at a very high level of about 61% in the years from 1990-1999, and then declined gradually to reach 55% in 2010. The GDP-scaled graph, in contrast, suggests a slow decline in the incidence of poverty in the 1990-1999 period and then a more rapid decrease in subsequent years.

4.4. Absolute or relative δ?

As noted in Section II, an analyst may apply either absolute δ, i.e. the same for all countries' estimates of G_0, or relative δ, i.e. a fraction of G_0, when generating clones. Fig. 2 shows that the version of δ used influences the expected clone size. However, a more important question is the extent to which that version influences the estimates of inequality and poverty measures in the WDPI.

To answer that question, we present six figures (Figs. 11-16) plotting the time-series of the Gini index and FGT_0. In each figure, the estimates of inequality and poverty measures and the 95% confidence interval limits are displayed for relative $\delta = 0.015 \cdot G_0$. For comparison, the estimates of these measures for absolute $\delta = 0.01$ are plotted without confidence interval limits.

(Figures. 11-16 about here)

It can be seen from Figs. 11-16 that all time-series of the considered measures for absolute $\delta = 0.01$ lie within the confidence intervals of those of the measures for relative δ, which means that the choice of absolute or relative δ does not influence the estimates of global inequality and poverty to any significant degree.

The practical consequence of the foregoing observation is that a radical reduction in clone size can be achieved in assessing global inequality and poverty. Tables 8 and 9 present the global sizes of the samples and average sizes of the clones in the 1990-2010 period according to absolute and relative δ, respectively.

(Tables 8 and 9 about here)

We can see from the tables that the application of 1.5% relative precision to countries' Gini indices requires the creation of clones five times larger in size than the application of absolute precision of 0.01. Hence, $\delta = 0.01$ seems to be reasonable because most countries present Gini index estimates with two decimal places. However, this degree of precision

requires more sizable samples than the 2000 recommended by Shorrocks and Wan (2008). Relative precision of 1.5%, in contrast, provides samples of a size close to that recommended by EUROSTAT as the minimum efficient sample size for European countries (Wolff et al., 2010).

V. Conclusions

The main contribution of this paper is that it presents a new method of creating a global sample from the WDPI. That sample consists of clones, i.e. sets of independent observations that reconstruct unknown country samples with predetermined precision and high probability. The size of a given clone is not predetermined, but is a random variable. The bases of the proposed cloning method are estimates of summary statistics. In this paper, Gini index estimates are applied to provide the lognormal distribution of personal incomes, although other forms of theoretical distributions and summary statistics could be applied. The personal income databases created contain approximately 28 million and 34 million observations when the clones are scaled to per capita HFCE and per capita GDP, respectively, for the 1990-2010 period.

Our approach returns some empirical evidence on the evolution of the global income distribution during the period under study. In the initial years of the period, that distribution displayed a twin-peaks shape that subsequently disappeared. This result confirms earlier findings. In addition, the mean and cosmopolitan social welfare were increasing in the period.

Our other results concern inequality and poverty. Global inequality, measured by the Gini and Theil indices, was generally high, but decreased over the 1990-2010 period. The inequality trends exhibited three phases separated by the years 1996 and 2002: an inverted U-shaped phase (or slowly decreasing phase), a plateau phase and a rapidly decreasing phase. These results contradict Milanovic's (20013) hypothesis of constant inequality in the 1988-2008 period. In our work, the incidence of poverty, according to the absolute $1.25 standard, also displayed a three-phase pattern separated by the years 1999-2002: an inverted U-shaped phase, a plateau phase and a rapidly decreasing phase. When a relative standard, i.e. half the mean income, was applied instead, the level of global poverty was very high (61-62%) and remained stable over the 1990-2000 period before declining to the 53-55% level in 2002.

This paper also sheds light on the problem of scaling income data. It shows that the scaling method selected, i.e. scaling to per capita GDP or to per capita HFCE, is important when evaluating the level of inequality and poverty but not when determining the general direction of their trends.

Despite its many attractive properties, the proposed method of cloning unknown samples from national household survey data has several limitations. For example, the method works quite well when the parametric form of the income distribution is known and reliable estimates of summary statistics are available. The normalised version of the lognormal distribution requires knowledge of the Gini index alone, and therefore only a simple statistical hypothesis is tested herein. More elaborate theoretical forms of countries' income distributions may require the testing of complex statistical hypotheses. In such cases, the generalised SPRT can be applied (e.g. Pavlov, 1987).

BIBLIOGRAPHY

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

Anand, S. and P. Segal. 2008. "The Global Distribution of Income." In: Atkinson, A.B., and F. Bourguignon (Eds) *Handbook of Income Distribution.* Amsterdam: North-Holland.

Araar, A., and J.Y. Duclos. 2009. "DASP: Distributive Analysis Stata Package." Université Laval, PEP, CIRPÉE and World Bank.

Bourguignon, F. 2011. "A Turning Point in Global Inequality… and Beyond." In: Krull, W. (Ed.) *Research and Responsibility: Reflection on Our Common Future.* Leipzig: CEP Europäiche Verlanganstalt Leipzig Gmbh.

Bourguignon, F. and C. Morrisson. 2002. "The Size Distribution of Income Among World Citizens, 1820-1990." *American Economic Review*: 727-744.

Chotikapanich, D., and W.E. Griffiths. 2008. "Estimating Income Distributions Using a Mixture of Gamma Densities." University of Melbourne, Department of Economics Research Paper No. 1034.

Chotikapanich, D., Valencia, V., and D.S.P. Rao. 1997. "Global and Regional Inequality in the Distribution of Income: Estimation with Limited and Incomplete Data." *Empirical Economics* 22:533-546.

Chotikapanich, D., and D. S. P. Rao. 1998. "Inequality in Asia 1975–1990: A Decomposition Analysis." *The Asia Pacific Journal of Economics and Business* 2(1): 63–78.

Cowell, F. A. and M.P. Victoria-Feser. 1996a. "Poverty Measurement with Contaminated Data: A Robust Approach." *European Economic Review* 40(9):1761–1771.

Cowell, F. A. and M.P. Victoria-Feser. 1996b. Robustness Properties of Inequality Measures." *Econometrica* 64(1):77–101.

Cowell, F. A. and M.P. Victoria-Feser 2002. “Welfare Rankings in the Presence of Contaminated Data.” *Econometrica* 70(3):1221–33.

D’Agostino, R.B., and M.A. Stephens. 1986. *Goodness-of-Fit Techniques.* New York and Base: Marcel Dekker, Inc.

Deininger, K., and L. Squire. 1996. “A New Data Set Measuring Income Inequality.” *The World Bank Economic Review*. 10 (3): 565-591.

Dikhanov, Y., and M. Ward. 2002. ““Evolution of the Global Distribution of Income, 1970–99.” Paper prepared for the 53rd Session of the International Statistical Institute held in Seoul, Republic of Korea, August 22-29, 2001 and updated for the 5th Conference on Globalization, Growth and (In)Equality held in Warwick, England, March 15-17, 2002.

Dowrick, S., and M. Akmal. 2005. “Contradictory Trends in Global Income Inequality: A Tale of Two Biases.” *Review of Income and Wealth* 51(2): 201-229.

Fang, C., Wailes, E., and G. Cramer. 1998. “China’s Rural and Urban Household Survey Data: Collection, Availability, and Problems.” CARD Working Paper 98-WP 202, Center for Agricultural and Rural Development. Iowa State University.

Fields, G.S. 1994. “Data for Measuring Poverty and Inequality Changes in the Developing Countries.” *Journal of Development Economics* 44(1): 87-102.

Fisz, M. 1967. *Probability Theory and Mathematical Statistics.* Huntington, N.Y.: Krieger Publ. Co.

Foster, J.E., Greer, J., and E. Thorbecke. 1984. “A Class of Decomposable Poverty Indices." *Econometrica* 52: 761-766.

Ghosh, B.K. and B.K. Sen. 1991. *Handbook of Sequential Analysis*, New York, Basel, Hong Kong: Marcel Dekker Inc.

Iritani, J., and K.Kuga. 1983. “Duality Between the Lorenz Curves and the Income Distribution Functions.” *Economic Studies Quarterly* 34: 9–21.

Jones, C. 1997. “On the Evolution of the World Income Distribution.” *Journal of Economic Perspectives* XI : 19–36.

Kleiber, C., and S. Kotz. 2003. *Statistical Size Distributions in Economics and Actuarial Sciences.* New Jersey: John Wiley & Sons, Inc.

Korzeniewicz, R. and T. Moran. 1997. „World-Economic Trends in the Distribution of Income, 1965-1992.” *American Journal of sociology* 102(4): 1000-1039.

Kremer, M., Onatski, A. and J. Stock. 2001. “Searching for Prosperity.” *Carnegie-Rochester Conference Series on Public Policy* LV: 275–303.

Kullback, S., and R. A. Leibler. 1951. "On Information and Sufficiency." *The Annals of Mathematical Statistics* 22: 79–86.

Lakner, C., and B. Milanovic. 2013. "Global Income Distribution From the Fall of the Berlin Wall to the Great Recession." World Bank Policy Research Working Paper No. 6719.

Liberati, P. 2015. "The World Distribution of Income and Its Inequality, 1970-2009. *Review of Income and Wealth* 61(2): 248-273.

López, J. H. and L. Servén. 2006. "A Normal Relationship? Poverty, Growth and Inequality." World Bank Policy Research Working Paper 3814.

Milanovic, B. 2002. "True World Income Distribution, 1988 and 1993: First Calculations Based on Household Surveys Alone." *Economic Journal* 112(476): 51-92.

Milanovic, B. 2005. *Worlds Apart: Global and International Inequality 1950-2000*, Princeton: Princeton University Press.

Minoiu, C., and , S.G. Reddy. 2006. "The Estimation of Poverty and Inequality from Grouped Data Using Parametric Curve Fitting: an Evaluation of POVCAL. mimeo.

Modarres, R. and J.L. Gastwirth 2006. "Practitioner's Corner: A Cautionary Note on Estimating the Standard Error of the Gini Index of Inequality." *Oxford Bulletin of Economics and Statistics* 68 (3): 385-390.

Ogwang, T. 2006. "A Cautionary Note on Estimating the Standard Error of the Gini Index of Inequality: Comment." *Oxford Bulletin of Economics and Statistics* 68 (3): 391-393.

Pavlov, I.V. 1987. "A Sequential Procedure for Testing Many Composite Hypotheses." *Theory of Probability and its Applications* 21(1):138-142.

Pinkovsky, M. and X. Sala-i-Martin. 2009. "Parametric Estimation of the World Distribution of Income." Working Paper 15433 http://www.nber.org/papers/w15433.

Quah, D. 1996). "Twin Peaks: Growth and Convergence in Models of Distribution Dynamics." *Economic Journa,* CVI: 1045–1055.

Sala-i-Martin X. 2006. "The World Distribution of Income: Falling Poverty and… Convergence Period." *Quarterly Journal of Economics* 121(2): 351–97.

Sen, A. 1973. *On Economic Inequality.* Oxford: Clarendon Press.

Shorrocks, A., and G. Wan 2008. "Ungrouping Income Distributions. Synthesizing Samples for Inequality and Poverty Analysis." UNU-WIDER Research Paper No. 2008/16.

Solt, F. 2014 (October). "The Standardized World Income Inequality Database.", Working Paper. SWIID Version 5.0.

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

UNU-WIDER. 2015 (September). "World Income Inequality Database (WIID3c)." http://www.wider.unu.edu/research/Database/en_GB/wiid/

Van Kerm, P. 2007. "Extreme Incomes and the Estimation of Poverty and Inequality Indicators from EU-SILC." IRISS Working Paper 2007-01. Luxembourg: CEPS/INSTEAD, Differdange.

Wald, A. 1945. Sequential Tests of Statistical Hypotheses." *Annals of Mathematical Statistics* 16 (2): 117-186.

Wald, A. 1947. *Sequential Analysis.* New York: John Wiley.

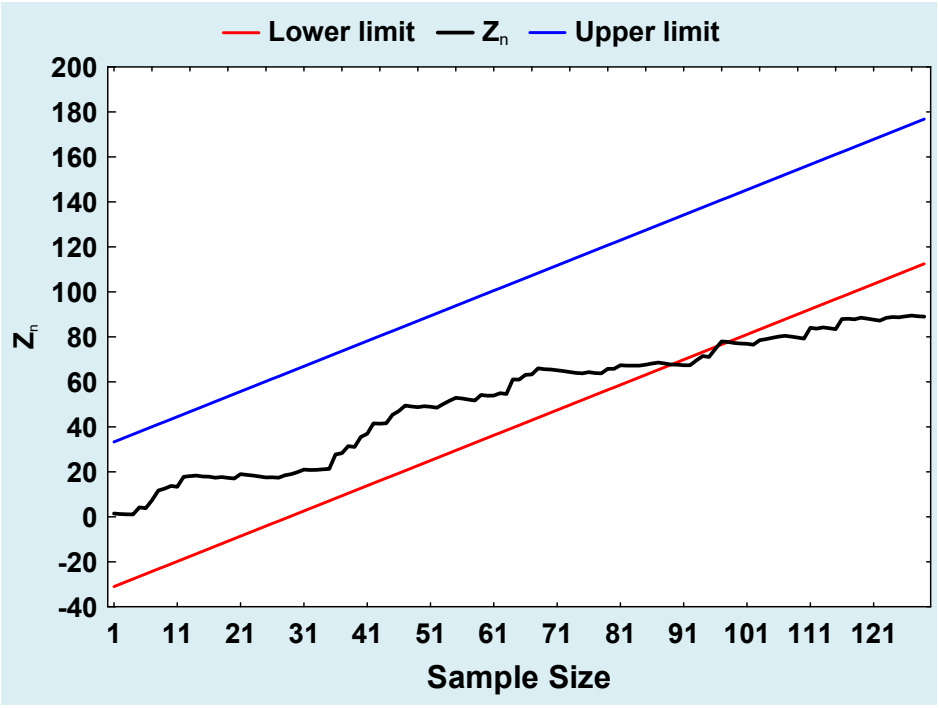
Wald, A. and J. Wolfowitz. 1948. "Optimum Character of the Sequential Probability Ratio Test." *The Annals of Mathematical Statistics* 19(3):326-339.

Wolff, P., Montaigne, F. and G. R. González. 2010. "Investing in Statistics: EU-SILC." In: Atkinson, A. B and E.Malier, (Eds.) *Income and Living Conditions in Europe.* Luxembourg: Publications Office of the European Union.

World Bank. 2005. *World Development Report 2006: Equity and Development.* Washington, D.C.: World Bank and Oxford University Press.

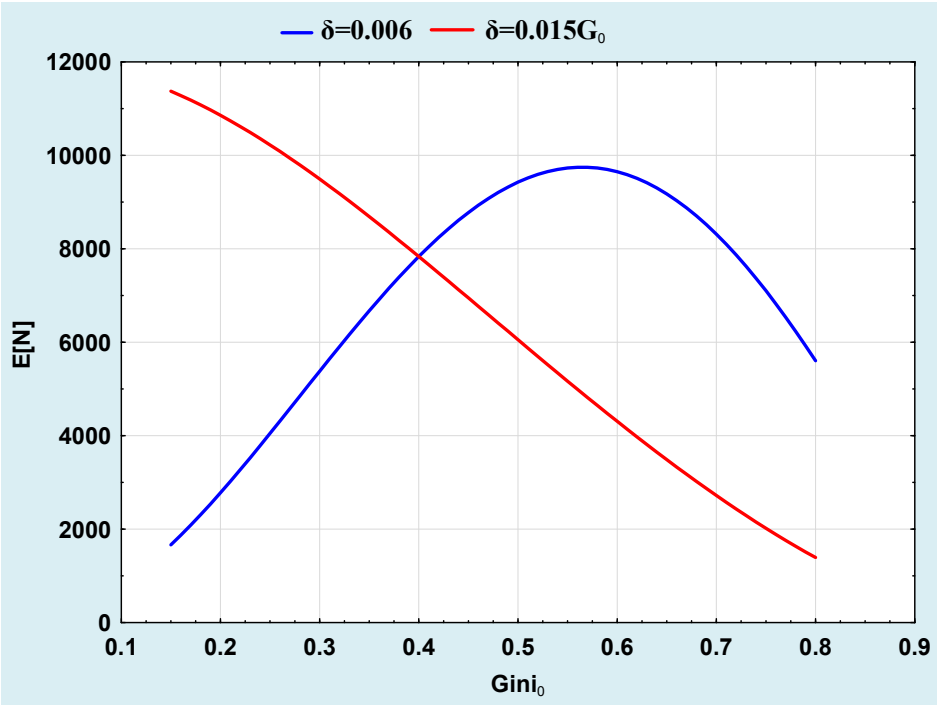
World Bank. 2015. *World Development Indicators*. Washington D.C.: World Bank and Oxford University Press.

FIGURE1. The Size of the Clone Provided by Sequential Procedure (G_0=0.3, δ=0.03)



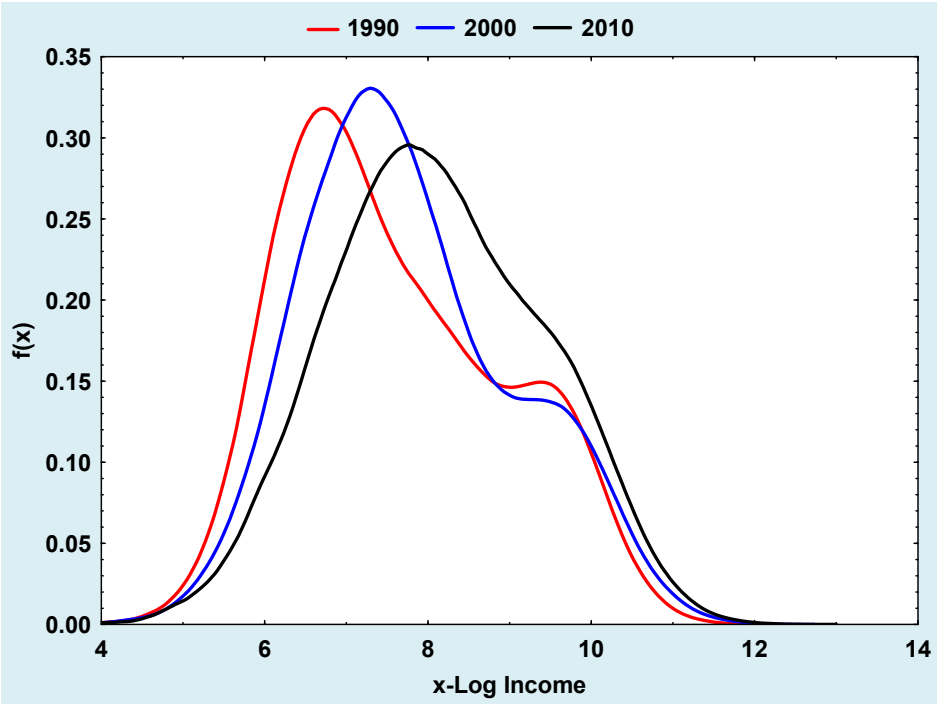
Source: Own elaboration.

FIGURE 2. The Mean Sample Size for Absolute and Relative δ



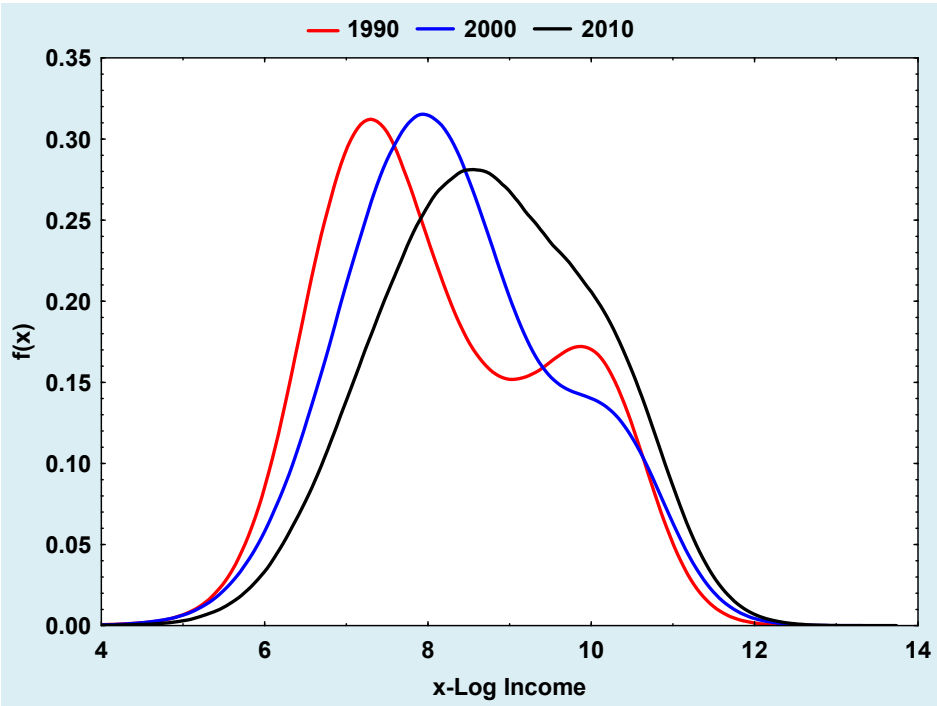
Source: Own elaboration.

FIGURE 3. Kernel Estimate of the Density Function of the WPDI (Clones Scaled to HFCE)



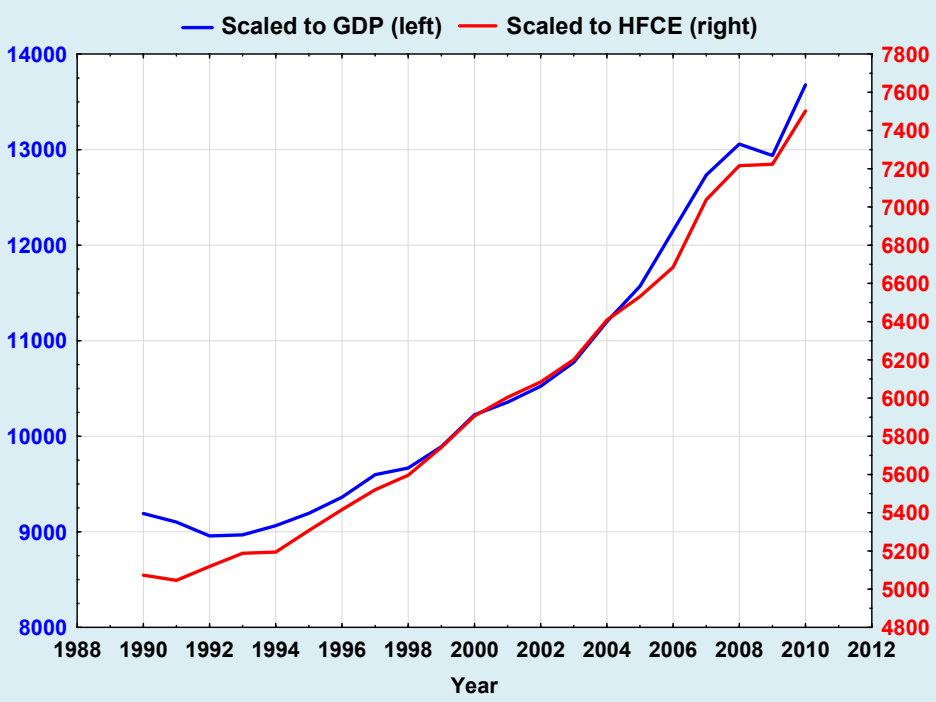
Source: Own elaboration.

FIGURE 4. Kernel Estimate of the Density Function of the WPDI (Clones Scaled to GDP)



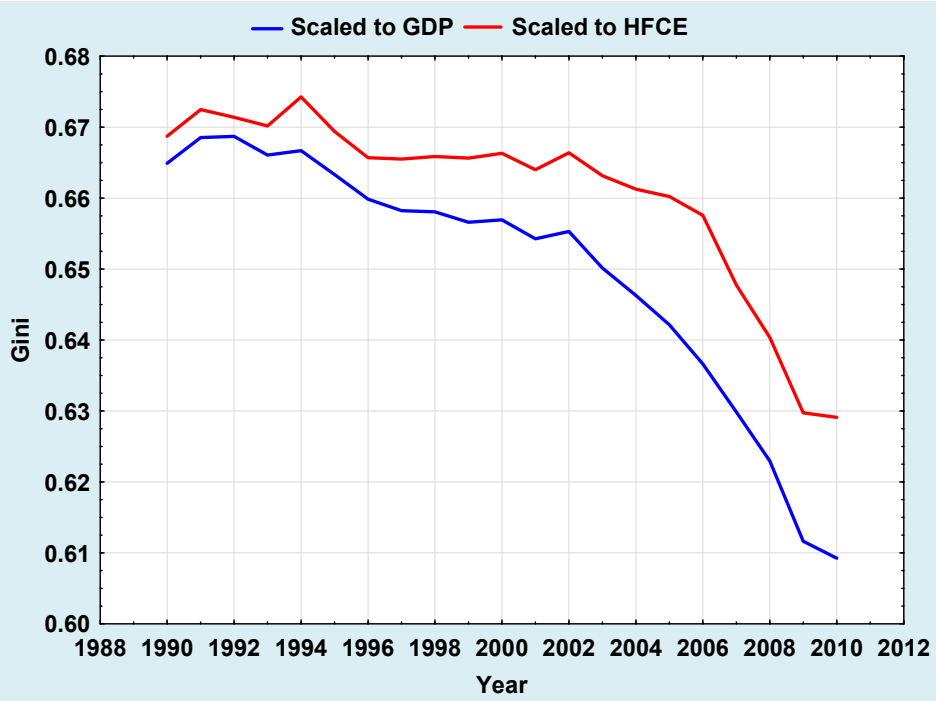
Source: Own elaboration.

FIGURE 5. The Mean of the WDPI



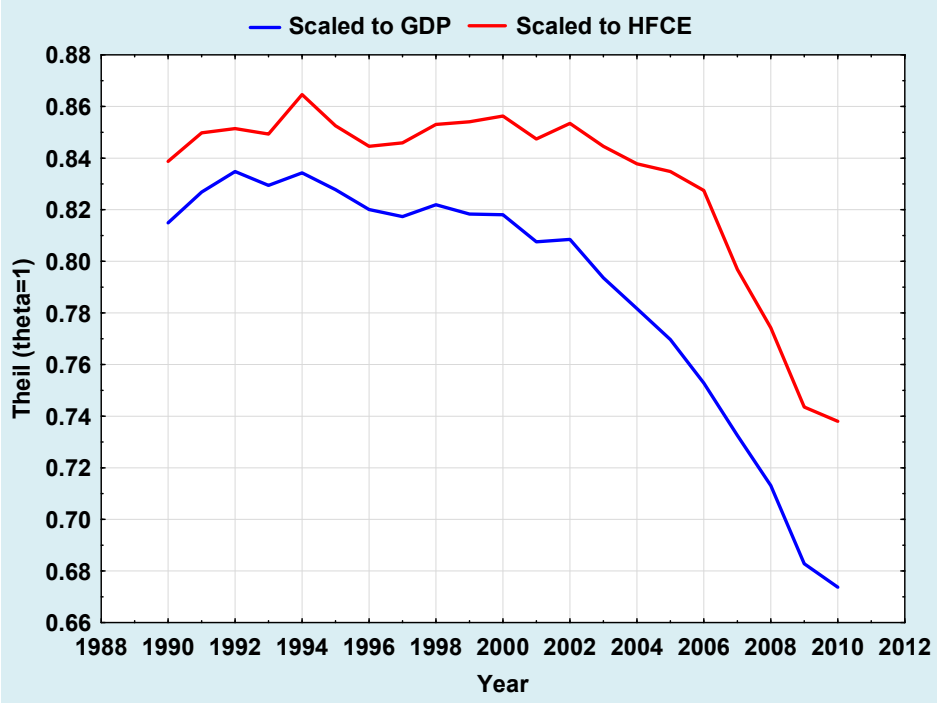
Source: Own elaboration.

FIGURE 6. Trends of the Global Inequality Measured by the Gini Index



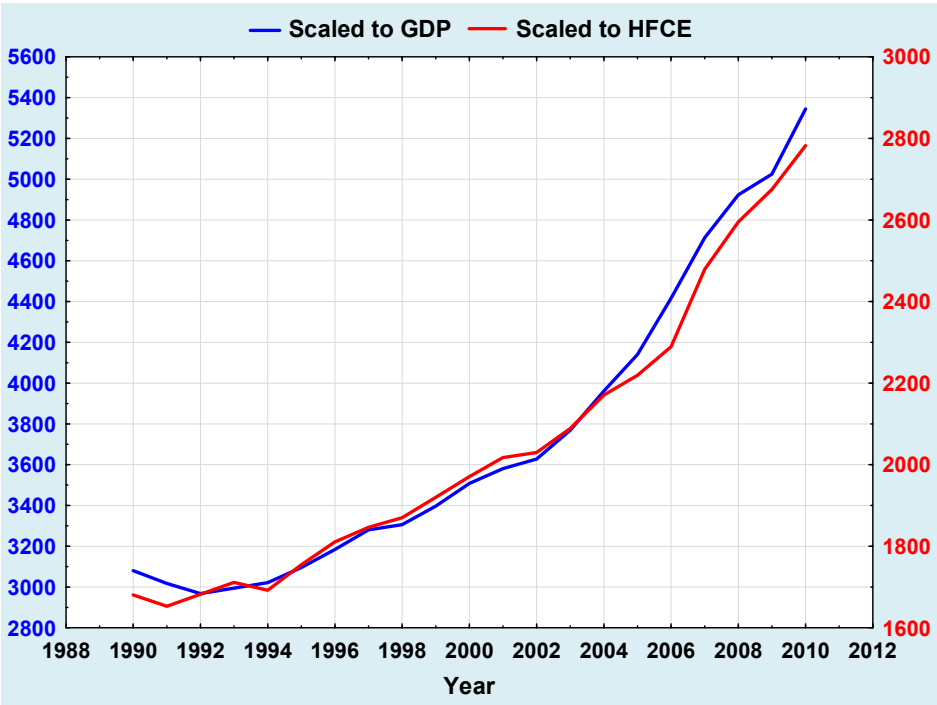
Source: Own elaboration.

FIGURE 7. Trends of the Global Inequality Measured by the Theil Index



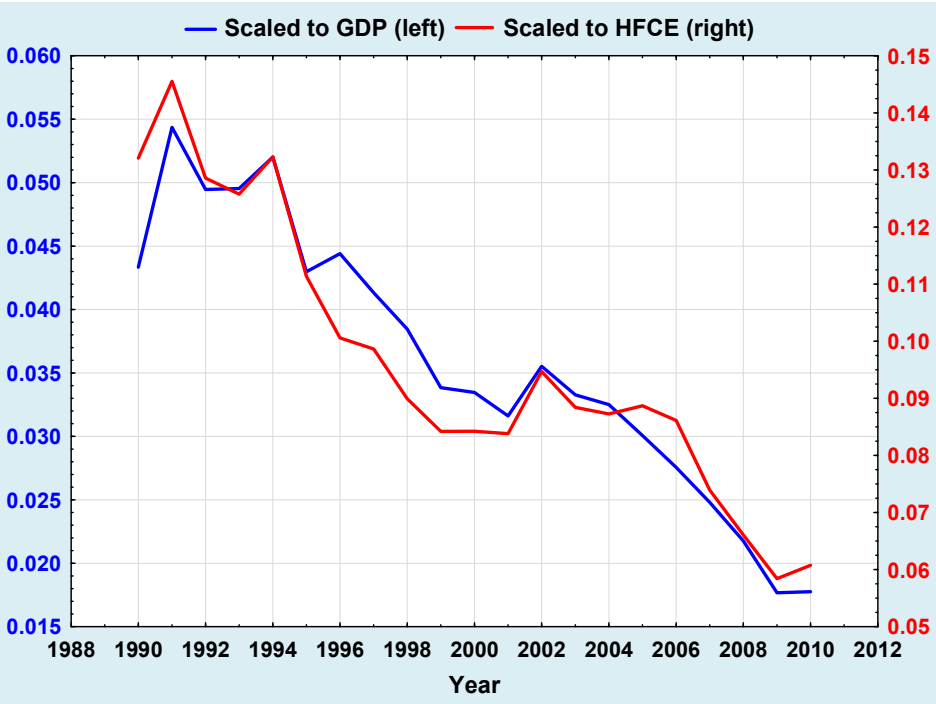
Source: Own elaboration.

FIGURE 8. Cosmopolitan Social Welfare Function: CSWF=Mean·(1-Gini)



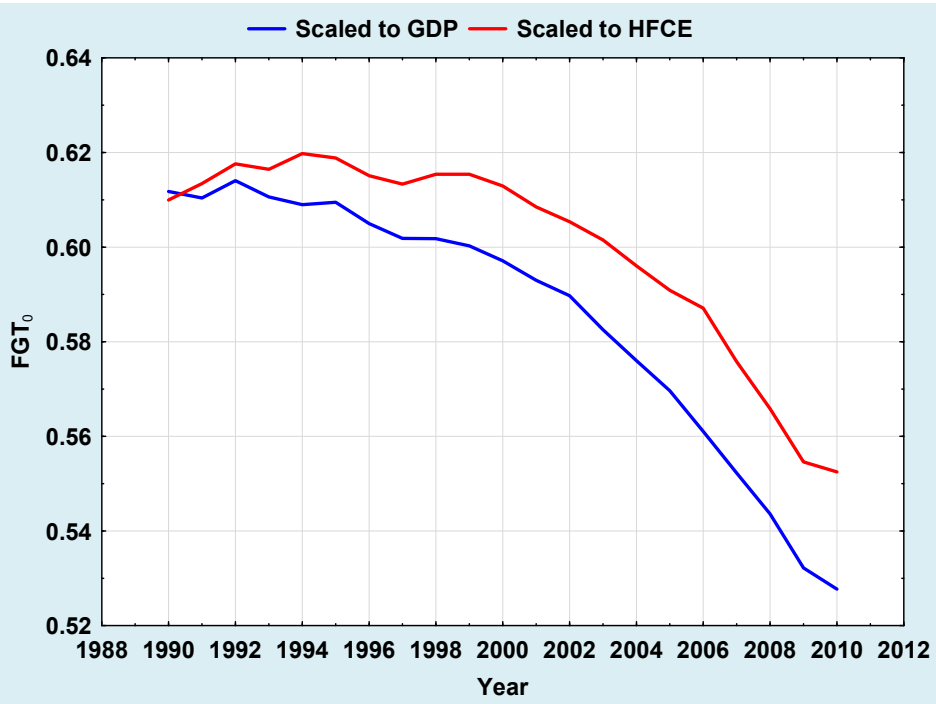
Source: Own elaboration.

FIGURE 9. The Incidence of Global Poverty (FGT_0). Absolute Poverty Line = $1.25/person/ day



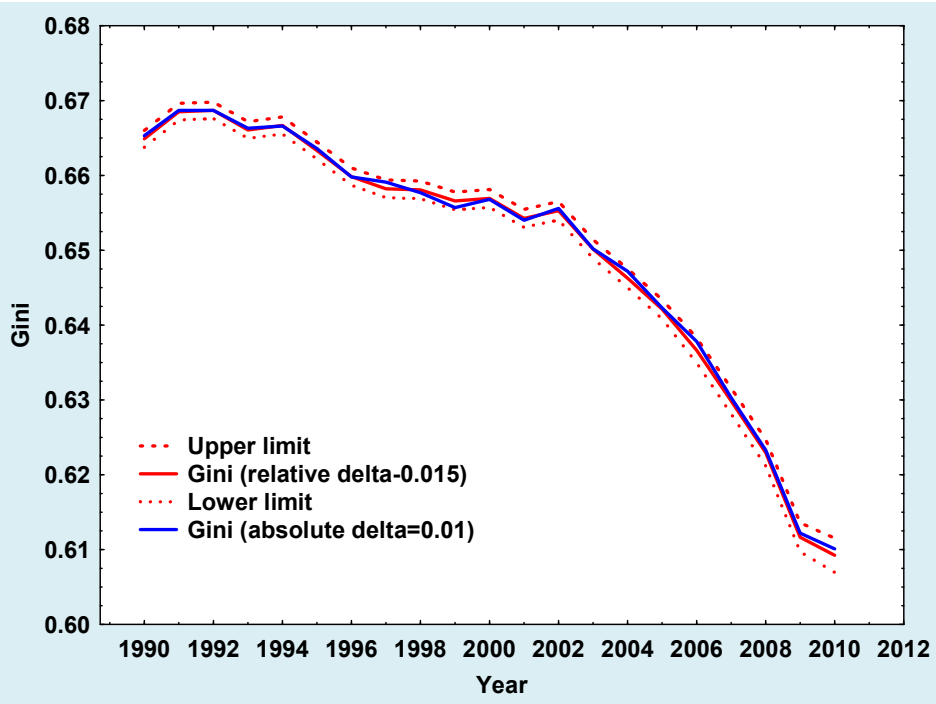
Source: Own elaboration.

FIGURE 10. The Incidence of Global Poverty (FGT_0). Relative Poverty Line = Mean/ 2



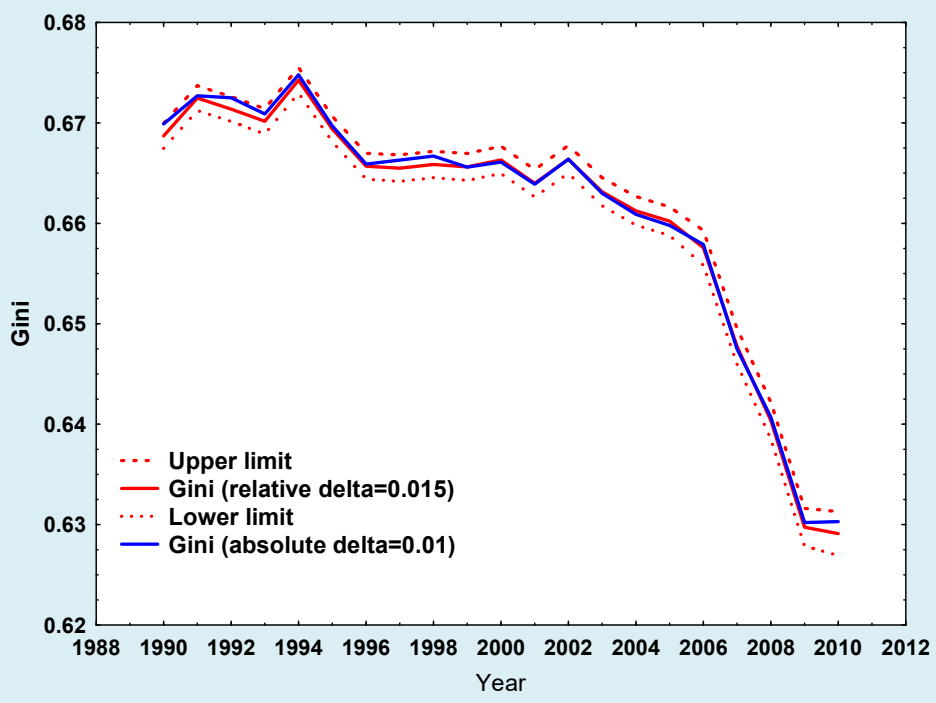
Source: Own elaboration.

FIGURE 11. The Estimates of Gini for Relative and Absolute Precision *δ*. (Countries' Clones Scaled to Per Capita GDP)



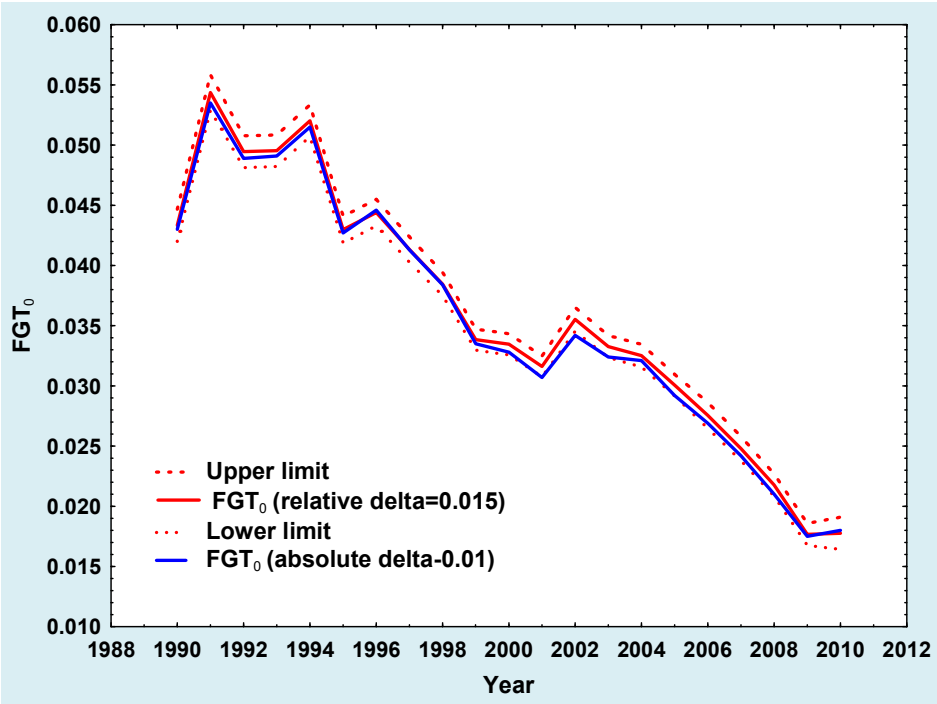
Note: The limits of 95% confidence interval for Gini, relative $\delta=0.015 \cdot G_0$
Source: Own elaboration based on data from Table 5.

FIGURE 12. The Estimates of Gini for Relative and Absolute Precision *δ*. (Countries' Clones Scaled to Per Capita HFCE)



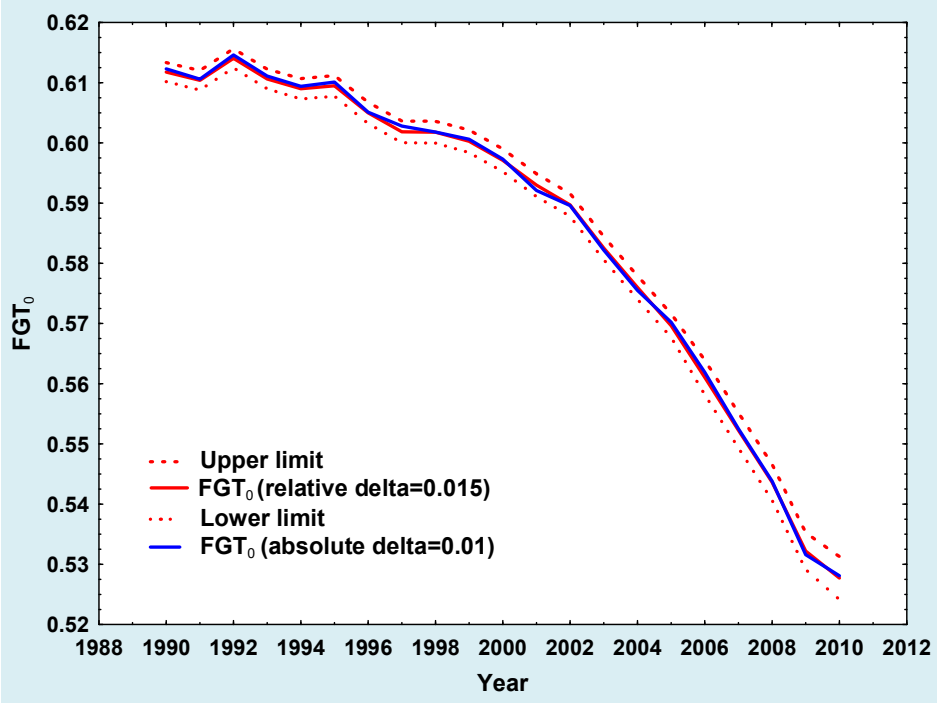
Note: The limits of 95% confidence interval for Gini, relative $\delta=0.015 \cdot G_0$
Source: Own elaboration based on data from Table 4.

FIGURE 13. The Estimates of FGT_0 for Relative and Absolute Precision *δ*. (Absolute poverty line, Countries' Clones Scaled to Per Capita GDP)



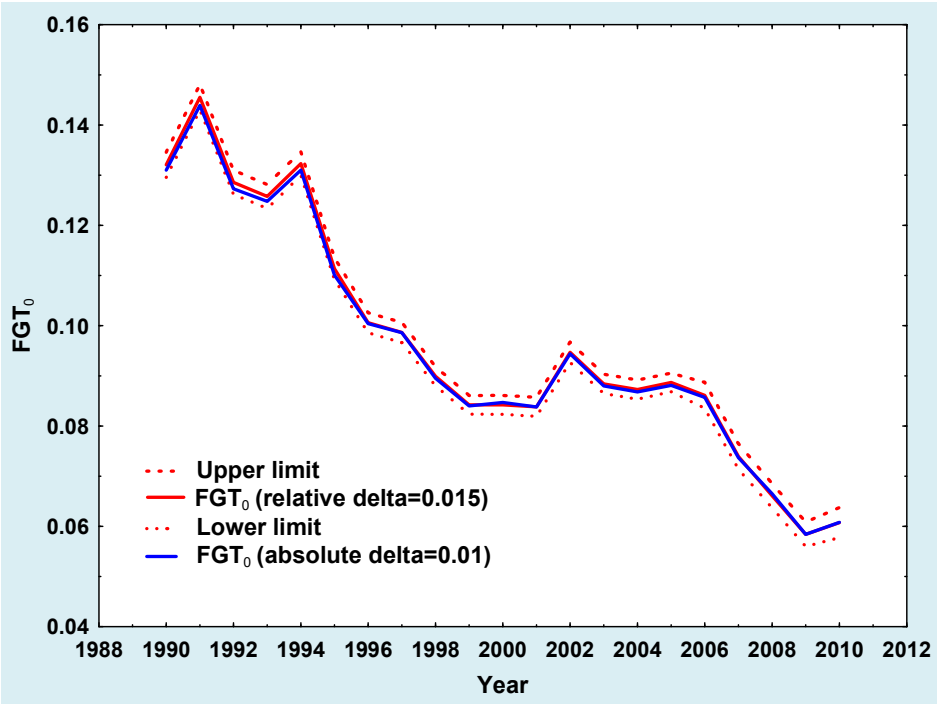
Note: The limits of 95% confidence interval for FGT_0 relative $\delta=0.015 \cdot G_0$
Source: Own elaboration based on data from Table 6.

FIGURE 14. The Estimates of FGT_0 for Relative and Absolute Precision *δ*. (Relative poverty line, Countries' Clones Scaled to Per Capita GDP)



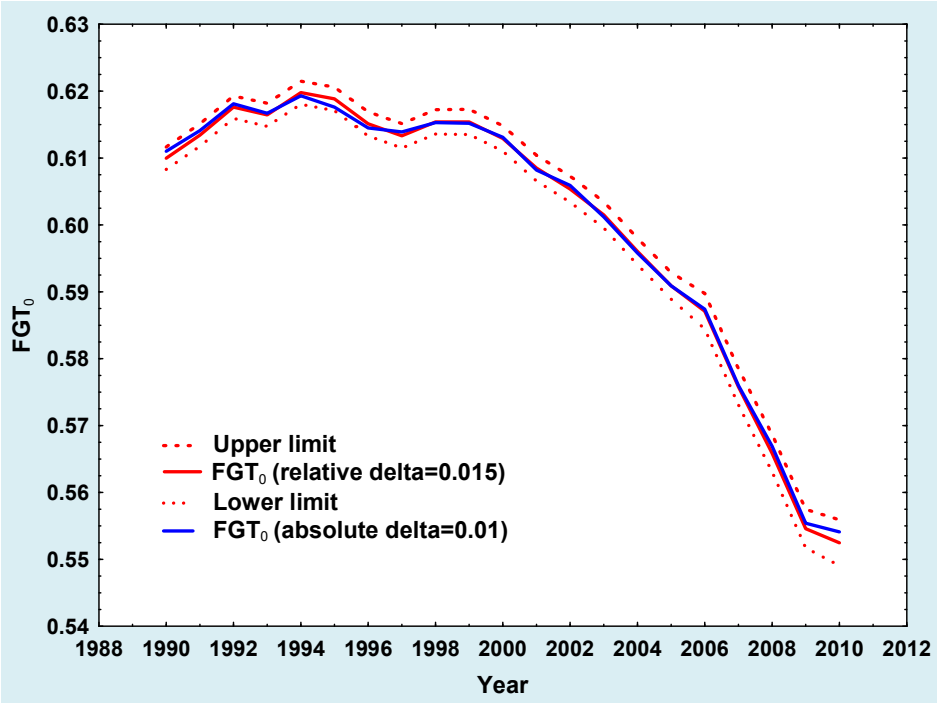
Note: The limits of 95% confidence interval for FGT_0 relative $\delta=0.015 \cdot G_0$
Source: Own elaboration based on data from Table 6.

FIGURE 15. The Estimates of FGT_0 for Relative and Absolute Precision δ. (Absolute poverty line, Countries' Clones Scaled to Per Capita HFCE)



Note: The limits of 95% confidence interval for FGT_0 relative $\delta=0.015 \cdot G_0$
Source: Own elaboration based on data from Table 6.

FIGURE 16. The Estimates of FGT_0 for Relative and Absolute Precision δ. (Relative poverty line, Countries' Clones Scaled to Per Capita HFCE)



Note: The limits of 95% confidence interval for FGT_0 relative $\delta=0.015 \cdot G_0$
Source: Own elaboration based on data from Table 6.

TABLE 1. Countries and Populations in the Years 1990-2010

Year	*World Population [mln]*	*SWIID5*		*Scaling to HFCE*		*Scaling to GDP*	
		No. of countries	% World population	No. of countries	% World population	No. of countries	% World population
1990	5278.91	106	90.38	79	84.19	96	88.67
1991	5365.43	114	91.35	84	85.11	104	89.76
1992	5448.30	123	93.21	88	85.25	114	91.74
1993	5531.85	127	93.47	93	85.51	118	92.01
1994	5614.43	130	93.76	99	87.21	122	92.39
1995	5698.21	136	94.35	102	87.30	132	93.16
1996	5780.01	139	94.89	104	87.74	135	93.70
1997	5861.97	141	95.16	106	87.82	136	93.72
1998	5942.98	142	95.14	106	87.74	137	93.83
1999	6023.02	145	95.11	106	87.67	139	93.81
2000	6101.96	146	95.09	108	87.73	141	93.79
2001	6179.98	149	95.18	111	88.05	143	93.76
2002	6257.40	150	95.24	113	88.31	143	93.79
2003	6334.78	151	95.32	115	88.30	144	93.87
2004	6412.47	148	95.21	112	88.01	142	93.76
2005	6490.29	150	95.27	120	89.38	146	93.92
2006	6568.34	136	93.47	119	89.86	132	92.11
2007	6646.37	131	92.65	115	88.39	127	91.42
2008	6725.58	121	91.05	108	86.72	119	90.11
2009	6804.92	114	90.24	103	86.21	112	89.31
2010	6884.35	105	87.54	95	83.50	103	86.62

Source: Own elaboration using data from The World Development Indicators (2015) and the SWIID5 database.

TABLE 2. Descriptive Statistics of the WDPI (Country Samples are Scaled to Per Capita HFCE)

Year	*Mean*	*Median*	*Std.dev*	*Skewness*	*Kurtosis*	*N*
1990	5073	1574	8438.37	3.73	22.89	1014452
1991	5046	1560	8458.75	3.74	23.09	1069710
1992	5119	1599	8671.47	3.88	25.04	1129251
1993	5188	1641	8817.47	3.99	26.83	1169207
1994	5194	1632	8984.86	4.13	29.00	1258594
1995	5307	1701	9137.77	4.18	29.75	1280002
1996	5416	1779	9343.89	4.36	33.27	1294550
1997	5520	1835	9582.89	4.50	36.05	1323441
1998	5596	1879	9860.48	4.62	37.53	1323839
1999	5742	1935	10162.79	4.71	39.26	1346946
2000	5906	2005	10480.83	4.73	39.40	1367859
2001	6004	2070	10551.36	4.65	38.29	1388836
2002	6084	2102	10751.82	4.75	40.12	1416106
2003	6201	2180	10919.12	4.84	41.99	1469341
2004	6408	2284	11226.99	4.89	43.37	1472569
2005	6531	2360	11431.91	4.94	44.11	1558333
2006	6685	2458	11656.70	4.99	45.28	1548972
2007	7038	2699	11945.05	4.95	45.03	1506093
2008	7216	2851	11995.50	4.92	45.18	1414276
2009	7224	2989	11648.73	4.74	41.33	1364408
2010	7502	3124	11941.84	4.59	38.71	1258774

Source: Own elaboration.

TABLE 3. Descriptive Statistics of the WDPI (Country Samples are Scaled to Per Capita GDP)

Year	*Mean*	*Median*	*Std.dev*	*Skewness*	*Kurtosis*	*N*
1990	9191	2714	14545.55	3.29	17.73	1229596
1991	9102	2735	14604.98	3.34	18.17	1298921
1992	8957	2764	14702.35	3.54	20.26	1437433
1993	8968	2874	14755.99	3.65	21.88	1453008
1994	9065	2974	15078.49	3.77	23.34	1532151
1995	9194	3072	15304.21	3.82	24.18	1638261
1996	9362	3218	15593.78	3.97	26.85	1674719
1997	9598	3361	16036.20	4.10	29.06	1689050
1998	9668	3432	16350.99	4.19	30.10	1693150
1999	9891	3535	16717.52	4.25	31.36	1741773
2000	10225	3673	17253.62	4.23	30.82	1754019
2001	10357	3756	17245.68	4.13	29.54	1788872
2002	10523	3809	17512.71	4.18	30.62	1792656
2003	10774	4018	17738.09	4.23	31.95	1822924
2004	11199	4269	18258.18	4.29	33.64	1848816
2005	11569	4515	18673.60	4.27	33.04	1892887
2006	12152	4877	19286.90	4.24	32.90	1717677
2007	12734	5265	19790.45	4.18	32.37	1659867
2008	13058	5584	19907.97	4.15	32.18	1556737
2009	12939	5798	19143.24	4.04	30.37	1479822
2010	13678	6234	19911.08	3.92	28.48	1360289

Source: Own elaboration.

TABLE 4. Global inequality and Cosmopolitan Welfare in the years 1990-2010 (Country Samples are Scaled to Per Capita HFCE)

Year	*Gini*				*Theil*				*CSWF*
	Estimate	*SE*	*LB*	*UB*	*Estimate*	*SE*	*LB*	*UB*	
1990	0.6687	0.0006	0.6675	0.6700	0.8387	0.0023	0.8342	0.8433	1681
1991	0.6725	0.0006	0.6712	0.6737	0.8498	0.0023	0.8452	0.8544	1653
1992	0.6714	0.0006	0.6701	0.6726	0.8514	0.0024	0.8468	0.8561	1682
1993	0.6702	0.0006	0.6689	0.6714	0.8493	0.0024	0.8446	0.8541	1711
1994	0.6743	0.0006	0.6730	0.6755	0.8646	0.0025	0.8598	0.8695	1692
1995	0.6694	0.0006	0.6681	0.6707	0.8525	0.0025	0.8477	0.8574	1754
1996	0.6657	0.0007	0.6644	0.6670	0.8446	0.0025	0.8396	0.8496	1811
1997	0.6655	0.0007	0.6642	0.6668	0.8459	0.0026	0.8408	0.8511	1846
1998	0.6659	0.0007	0.6645	0.6672	0.8531	0.0026	0.8479	0.8582	1870
1999	0.6656	0.0007	0.6643	0.6670	0.8541	0.0027	0.8488	0.8594	1920
2000	0.6663	0.0007	0.6650	0.6677	0.8563	0.0027	0.8510	0.8617	1971
2001	0.6640	0.0007	0.6626	0.6653	0.8474	0.0027	0.8422	0.8526	2017
2002	0.6664	0.0007	0.6650	0.6678	0.8535	0.0028	0.8481	0.8589	2030
2003	0.6631	0.0007	0.6617	0.6646	0.8446	0.0028	0.8391	0.8500	2089
2004	0.6612	0.0007	0.6598	0.6627	0.8378	0.0028	0.8323	0.8433	2171
2005	0.6602	0.0007	0.6588	0.6616	0.8348	0.0028	0.8293	0.8403	2219
2006	0.6576	0.0009	0.6558	0.6593	0.8274	0.0032	0.8212	0.8337	2289
2007	0.6478	0.0009	0.6460	0.6495	0.7969	0.0031	0.7907	0.8030	2479
2008	0.6404	0.0009	0.6385	0.6422	0.7743	0.0031	0.7681	0.7804	2595
2009	0.6297	0.0010	0.6279	0.6316	0.7435	0.0031	0.7375	0.7495	2675
2010	0.6291	0.0011	0.6269	0.6313	0.7380	0.0034	0.7314	0.7447	2782

Note:
CSWF – cosmopolitan social welfare function. *SE*-standard error, *LB, LU*- lower and upper boundary of 95% confidence interval.
Source: Author's elaboration

TABLE 5. Global Inequality and Cosmopolitan Social Welfare in the Years 1990-2010 (Country Samples are Scaled to Per Capita GDP)

Year	*Gini*				*Theil*				*CSWF*
	Estimate	*SE*	*LB*	*UB*	*Estimate*	*SE*	*LB*	*UB*	
1990	0.6649	0.0006	0.6638	0.6660	0.8149	0.0020	0.8109	0.8188	3080
1991	0.6685	0.0006	0.6674	0.6696	0.8269	0.0020	0.8229	0.8308	3017
1992	0.6687	0.0006	0.6676	0.6698	0.8348	0.0020	0.8308	0.8388	2967
1993	0.6661	0.0006	0.6649	0.6672	0.8294	0.0021	0.8254	0.8334	2995
1994	0.6667	0.0006	0.6655	0.6678	0.8343	0.0021	0.8302	0.8385	3022
1995	0.6633	0.0006	0.6622	0.6645	0.8278	0.0021	0.8237	0.8319	3095
1996	0.6598	0.0006	0.6587	0.6610	0.8201	0.0021	0.8159	0.8243	3184
1997	0.6582	0.0006	0.6570	0.6594	0.8173	0.0022	0.8130	0.8217	3280
1998	0.6581	0.0006	0.6569	0.6592	0.8219	0.0022	0.8176	0.8263	3306
1999	0.6566	0.0006	0.6554	0.6578	0.8183	0.0022	0.8139	0.8227	3397
2000	0.6569	0.0006	0.6557	0.6581	0.8181	0.0022	0.8137	0.8224	3508
2001	0.6543	0.0006	0.6531	0.6555	0.8076	0.0022	0.8033	0.8118	3581
2002	0.6553	0.0006	0.6541	0.6565	0.8085	0.0022	0.8041	0.8129	3627
2003	0.6502	0.0006	0.6489	0.6514	0.7936	0.0022	0.7892	0.7980	3769
2004	0.6463	0.0006	0.6450	0.6475	0.7817	0.0023	0.7773	0.7861	3961
2005	0.6421	0.0007	0.6409	0.6434	0.7697	0.0022	0.7653	0.7741	4140
2006	0.6366	0.0009	0.6350	0.6383	0.7528	0.0027	0.7476	0.7581	4416
2007	0.6299	0.0009	0.6281	0.6316	0.7326	0.0027	0.7273	0.7378	4714
2008	0.6230	0.0009	0.6212	0.6247	0.7131	0.0027	0.7078	0.7184	4923
2009	0.6116	0.0010	0.6097	0.6136	0.6828	0.0029	0.6772	0.6884	5025
2010	0.6092	0.0012	0.6070	0.6115	0.6737	0.0033	0.6673	0.6802	5345

Note: see Table 4.
Source: Author's elaboration

TABLE 6. Global Poverty Incidences (FGT_0) 1990-2010 (Absolute Poverty Line=$1.25/person/day)

	Scaling to HFCE				*Scaling to GDP*			
Year	*Estimate*	*SE*	*LB*	*UB*	*Estimate*	*SE*	*LB*	*UB*
1990	0.1321	0.0013	0.1296	0.1346	0.0433	0.0007	0.0420	0.0447
1991	0.1455	0.0013	0.1430	0.1481	0.0544	0.0008	0.0529	0.0558
1992	0.1286	0.0012	0.1262	0.1310	0.0495	0.0007	0.0481	0.0508
1993	0.1258	0.0012	0.1234	0.1281	0.0495	0.0007	0.0482	0.0509
1994	0.1323	0.0012	0.1299	0.1347	0.0520	0.0007	0.0507	0.0533
1995	0.1114	0.0011	0.1092	0.1135	0.0430	0.0006	0.0419	0.0441
1996	0.1006	0.0010	0.0986	0.1026	0.0444	0.0006	0.0433	0.0455
1997	0.0987	0.0010	0.0966	0.1007	0.0413	0.0005	0.0403	0.0424
1998	0.0899	0.0010	0.0880	0.0918	0.0385	0.0005	0.0375	0.0394
1999	0.0842	0.0010	0.0823	0.0861	0.0338	0.0004	0.0330	0.0347
2000	0.0842	0.0010	0.0823	0.0861	0.0335	0.0005	0.0326	0.0343
2001	0.0838	0.0010	0.0819	0.0857	0.0316	0.0004	0.0308	0.0325
2002	0.0947	0.0010	0.0927	0.0967	0.0355	0.0005	0.0345	0.0365
2003	0.0884	0.0010	0.0865	0.0903	0.0333	0.0005	0.0324	0.0342
2004	0.0873	0.0010	0.0853	0.0892	0.0325	0.0005	0.0316	0.0335
2005	0.0887	0.0009	0.0868	0.0905	0.0301	0.0005	0.0292	0.0310
2006	0.0861	0.0013	0.0835	0.0887	0.0276	0.0005	0.0265	0.0286
2007	0.0739	0.0013	0.0714	0.0765	0.0248	0.0005	0.0238	0.0258
2008	0.0661	0.0012	0.0637	0.0685	0.0218	0.0005	0.0208	0.0227
2009	0.0584	0.0012	0.0560	0.0608	0.0177	0.0005	0.0168	0.0186
2010	0.0608	0.0015	0.0577	0.0638	0.0178	0.0007	0.0164	0.0191

Note:
SE-standard error, *LB, LU*- lower and upper boundary of 95% confidence interval
Source: Author's elaboration.

TABLE 7. Global Poverty Incidences (FGT_0) 1990-2010 (Relative Poverty Line=Mean/2)

	Scaling to HFCE				*Scaling to GDP*			
Year	*Estimate*	*SE*	*LB*	*UB*	*Estimate*	*SE*	*LB*	*UB*
1990	0.6100	0.0009	0.6083	0.6117	0.6118	0.0008	0.6102	0.6134
1991	0.6134	0.0009	0.6118	0.6151	0.6104	0.0008	0.6088	0.6120
1992	0.6176	0.0009	0.6159	0.6193	0.6141	0.0008	0.6125	0.6157
1993	0.6165	0.0009	0.6147	0.6182	0.6106	0.0008	0.6090	0.6123
1994	0.6198	0.0009	0.6180	0.6215	0.6090	0.0009	0.6073	0.6107
1995	0.6188	0.0009	0.6171	0.6206	0.6095	0.0009	0.6078	0.6112
1996	0.6151	0.0009	0.6133	0.6169	0.6050	0.0009	0.6032	0.6068
1997	0.6133	0.0009	0.6115	0.6151	0.6018	0.0009	0.6000	0.6036
1998	0.6154	0.0009	0.6136	0.6172	0.6018	0.0009	0.6000	0.6036
1999	0.6154	0.0010	0.6135	0.6173	0.6003	0.0010	0.5984	0.6021
2000	0.6129	0.0010	0.6110	0.6148	0.5971	0.0010	0.5952	0.5990
2001	0.6085	0.0010	0.6066	0.6104	0.5930	0.0010	0.5911	0.5949
2002	0.6054	0.0010	0.6034	0.6073	0.5897	0.0010	0.5878	0.5916
2003	0.6015	0.0010	0.5995	0.6035	0.5825	0.0010	0.5806	0.5845
2004	0.5960	0.0010	0.5941	0.5980	0.5760	0.0010	0.5740	0.5779
2005	0.5909	0.0010	0.5889	0.5929	0.5696	0.0010	0.5677	0.5716
2006	0.5871	0.0014	0.5844	0.5897	0.5611	0.0015	0.5582	0.5639
2007	0.5758	0.0014	0.5731	0.5785	0.5523	0.0015	0.5494	0.5552
2008	0.5658	0.0014	0.5630	0.5687	0.5436	0.0015	0.5406	0.5466
2009	0.5546	0.0015	0.5517	0.5575	0.5322	0.0016	0.5291	0.5353
2010	0.5525	0.0017	0.5491	0.5559	0.5277	0.0018	0.5241	0.5313

Note:
SE-standard error, *LB, LU*- lower and upper boundary of 95% confidence interval
Source: Author's elaboration.

TABLE 8. Global Sample Size (N) and Average Size of Clones for Absolute δ=0.01 and Relative δ=0.015·G_0. (Country Samples are Scaled to Per Capita HFCE)

Year	*Absolute δ*		*Relative δ*		*No.of*
	Total N	Average N	Total N	Average N	*Countries*
1990	217511	2540	1014452	12842	79
1991	238098	2617	1069710	12735	84
1992	250721	2632	1129251	12833	88
1993	266039	2645	1169207	12573	93
1994	287352	2676	1258594	12714	99
1995	297618	2691	1280002	12550	102
1996	304667	2704	1294550	12448	104
1997	309598	2703	1323441	12486	106
1998	309872	2701	1323839	12490	106
1999	310777	2701	1346946	12708	106
2000	316375	2706	1367859	12666	108
2001	326252	2710	1388836	12513	111
2002	331720	2706	1416106	12532	113
2003	338924	2711	1469341	12777	115
2004	329961	2711	1472569	13148	112
2005	355048	2730	1558333	12987	120
2006	350764	2720	1548972	13017	119
2007	339631	2723	1506093	13097	115
2008	317916	2710	1414276	13096	108
2009	301995	2704	1364408	13247	103
2010	277050	2688	1258774	13251	95
All years	5882127	2691	27975559	12798	2186

Source: Author's elaboration.

TABLE 9. Global Sample Size (N) and Average Size of Clones for Absolute δ=0.01 and Relative δ=0.015·G_0 (Country Samples are Scaled to Per Capita GDP)

Year	*Absolute δ*		*Relative δ*		*No.*
	Total N	Average N	Total N	Average N	*Countries*
1990	246945	2573	1229596	12809	96
1991	275113	2646	1298921	12490	104
1992	303904	2666	1437433	12610	114
1993	318210	2697	1453008	12314	118
1994	331654	2719	1532151	12559	122
1995	360920	2735	1638261	12412	132
1996	370247	2743	1674719	12406	135
1997	372759	2741	1689050	12420	136
1998	375580	2742	1693150	12359	137
1999	381169	2743	1741773	12531	139
2000	386888	2744	1754019	12440	141
2001	392103	2742	1788872	12510	143
2002	390244	2729	1792656	12537	143
2003	393442	2733	1822924	12660	144
2004	388240	2735	1848816	13020	142
2005	400806	2746	1892887	12965	146
2006	359690	2725	1717677	13013	132
2007	346340	2728	1659867	13070	127
2008	322693	2712	1556737	13082	119
2009	303291	2708	1479822	13213	112
2010	277263	2692	1360289	13207	103
All years	7297501	2718	34062631	12687	2685

Source: Author's elaboration.

Original citation:
Kot, S.M. (2016). Estimates of the World Distribution of Personal Incomes Based on Country Sample Clones. GUT FME Working Paper Series A, No 11/2016(41). Gdansk (Poland): Gdansk University of Technology, Faculty of Management and Economics.

All GUT Working Papers are downloadable at: http://zie.pg.edu.pl/working-papers

GUT Working Papers are listed in Repec/Ideas
http://ideas.repec.org/s/gdk/wpaper.html

Gdańsk University of Technology, Faculty of Management and Economics

Narutowicza 11/12, (premises at Traugutta 79)

80-233 Gdańsk, Poland phone: 58 347-18-99

www.zie.pg.edu.pl