World Distribution of Income for 1970-2010: Drastic Improvement in World Income Inequality during the 2000s

Soondong Hong* Heejoon Han† Chang Sik Kim‡

August 2017

Abstract

This paper nonparametrically estimates the distribution of world citizens' income and investigates the world income inequality for the period from 1970 to 2010. We consider 188 countries that account for 98.68% of the world population and almost 100% of the world GDP in the year 2010. Various income inequality indices such as the Gini coefficient show that the world income inequality drastically decreased during the 2000s while it slightly declined during the period from 1970 to 2000. This is because the inequality across countries substantially decreased during the 2000s even if the inequality within each country kept increasing during the 1990s and the 2000s. These findings still hold when we include top income tax data in the analysis. We also propose more sophisticated methods to impute missing top income shares and to combine them with income survey data.

JEL classification: H00, I30

Keywords: World distribution of income, income inequality, top income tax data, inequality decomposition

*Department of Economics, Sungkyunkwan University. E-mail: sunnyest10@skku.edu.
†Department of Economics, Sungkyunkwan University. E-mail: heejoonhan@skku.edu.
‡Department of Economics, Sungkyunkwan University. E-mail: skimcs@skku.edu.

1 Introduction

The global distribution of income and the world income inequality have been one of the major subjects of numerous research. Various approaches and assumptions have been made to more accurately estimate the world distribution of income (WDI) and the world income inequality in the literature. Recently, the related research has been quite active and gained much attention, which is partly because more data become available. Household income surveys and standardized databases of national accounts statistics have been updated and constantly extended to the global level. Moreover, researchers also employ new data such as top income data from tax records as in Atkinson et al. (2011) and Pikettty (2014) to study income inequality.

Recent studies on the world income inequality have focused on the concept of income differences of all individuals in the world.[1] These studies estimate the distribution of world citizens' income and also analyze the inequality of world citizens' income. See Anand and Segal (2014), Bourguignon and Morrisson (2002), Chotikapanich et al. (2012), Liberati (2015), Pinkovskiy and Sala-i-Martin (2009), Sala-i-Martin (2006), van Zanden et al. (2014) and Warner et al. (2014) among others. We also consider such a concept of world income inequality in this paper.

To estimate the WDI, some researchers adopted parametric methods, in which one should choose an appropriate form of income distribution. For example, Chotikapanich et al. (2012) and Warner et al. (2014) specified the income distribution of each country as the beta distribution and Liberati (2015), van Zanden et al. (2014) and Pinkovskiy and Sala-i-Martin (2009) used the log normal distribution. While the parametric method has its own merits, it inevitably involves a misspecification problem. The Monte Carlo simulation results in Krause (2014) show that a parametric specification of Lorenz Curve can lead to incorrectly-shaped income density function.[2] To overcome such a misspecification problem, nonparametric

[1] This corresponds to the concept referred to as "global inequality" in Milanovic (2005). See Anand and Segal (2008) and Milanovic (2005) for explanations on other concepts of the world income inequality.

[2] Krause (2014) shows that various parametric distributions can be derived from the shape of the same

estimation methods were instead adopted to estimate the WDI as in Sala-i-Martin (2002, 2006)

In this paper, we use a nonparametric method to estimate the WDI for recent four decades from 1970 to 2010[3] and examine how it changed over time. In addition, we analyze the world income inequality using various income inequality measures. In estimating the WDI and analyzing the world income inequality, there are some distinct features in our approaches compared to those in the existing literature. First, we consider more countries and populations than previous studies, which would reduce sample selection bias and, therefore, induce less biased results on the world income inequality. Total 188 countries are included in our analysis for the period from 1970 to 2010. In the year 2010, the total population of the 188 countries accounts for 98.68% of the world population and the total GDP of these 188 countries is close to 100% of the world GDP. Second, we adopt improved methods to impute missing income survey data. It is inevitable that many income survey data are missing particularly for low-income countries. We extend and modify the imputation methods by Sala-i-Martin (2006) and, specifically, use alternative interpolation methods.

Third, more importantly, we also analyze the case when top income tax data are combined with income survey data and adopt more sophisticated methods to impute missing top income shares and to combine them with income survey data. It is well known that household income surveys typically exclude the richest individuals or under-report their incomes, which can cause a substantial bias in estimating income inequality. To overcome such a problem, we additionally use top income shares based on income tax data that provide more precise information on top income shares (see Anand and Segal (2014), Atkinson et al. (2011) and Pikettty (2014) among others). To impute missing top income shares, we adopt a panel model allowing for both time and individual heterogeneity, which turns out to fit the data

Lorenz curve, and Lorenz curve estimation based on minimizing the MSE can lead to an Lorenz curve whose density has an incorrect modality.

[3]Readers are referred to Bourguignon and Morrisson (2002) and Van Zanden et al. (2014) for the global income inequality for the period before 1970. Both papers analyzed the long-term changes in the world income inequality beginning in 1820.

much better than the existing method by Anand and Segal (2014).

The main findings of this paper are following. First, the evolution of the estimated WDI shows that the mode of the distribution shifts rightward and the deviation from its center tends to be smaller over the last four decades. Moreover, it is shown that the WDI was bimodal in 1990 and changed to be unimodal in 2010. Second, most importantly, the Gini coefficient and other income inequality indices indicate that the world income inequality was drastically improved during the 2000s. The Gini coefficient decreased by 7.87% during the 2000s, which was a drastically rapid decrease considering that the Gini coefficient declined only by 0.91% for thirty years from 1970 to 2000.

Third, when we decompose the world income inequality into two components, across-country inequality and within-country inequality, each component provides more detailed explanations on the change of the world income inequality. The inequality within each country substantially increased during the 1990s and the 2000s, which is in accordance with the common perception that income inequality recently deteriorated. On the other hand, the inequality across countries decreased more substantially during the 2000s. Since the across-country inequality accounts for about 70% of the world income inequality, this is the main reason why the world income inequality was drastically improved during the 2000s.

Fourth, our results also show that China and India played the most important role in improving the world income inequality for the period from 1980 to 2010. However, the Sub-Saharan Africa region still played a role in deteriorating the world income inequality for the last four decades. Fifth, when we include top income tax data in our analysis, the main findings in this paper still hold; the world income inequality was drastically improved during the 2000s and it was mainly due to the rapid improvement in the across-country inequality during the period.

The rest of the paper is organized as follows. Section 2 explains the data and methodology to estimate the world distribution of income. Section 3 provides the results on the estimated world distribution of income and various income inequality indices. Section 4 presents the results for the case when top income shares are combined and Section 5 concludes the paper.

2 Method to Estimate World Distribution of Income

2.1 Data

While Sala-i-Martin (2006) used 138 countries to estimate the WDI for the period from 1970 to 2000, we use 188 countries for the period from 1970 to 2010. The added fifty countries[4] account for substantial portions of the world population and GDP: 5.63% of the world population and 4.29% of the world GDP in 2010. Inclusion of these fifty countries would reduce sample selection bias and, therefore, induce less biased results on world income inequality. In the year 2010, the total population of these 188 countries is 98.68% of the world population and the total GDP of these 188 countries is close to 100% of the world GDP. To estimate the WDI, we use GDP, population and quintile income shares of each country.

We obtain the real GDP and population of each country from the Penn World Table (version 7.1). Specifically, the GDP is based on 2005 USD and PPP-adjusted.[5] For countries with the missing GDP data, we impute them using the GDP growth rate[6]. However, there are some countries with the missing GDP data for a long period. For former Soviet Union Republics, we used the GDP growth rate of the Soviet Union until 1989 as in Sala-i-Martin (2006, Section II.D).

[4] Afghanistan, Albania, Bahamas, Bahrain, Bermuda, Bhutan, Bosnia and Herzegovina, Brunei, Bulgaria, Cambodia, Croatia, Cuba, Czech Republic, Djibouti, Eritrea, Iraq, Kiribati, Kuwait, Lao PDR, Lebanon, Liberia, Libya, Macao, Macedonia, Maldives, Malta, Marshall Islands, Micronesia Fed. Sts., Moldova, Mongolia, Montenegro, Oman, Palau, Puerto Rico, Qatar, Samoa, Saudi Arabia, Serbia, Slovak Republic, Slovenia, Solomon Islands, Somalia, Sudan, Suriname, Swaziland, Tonga, United Arab Emirates, Vanuatu, Vietnam, Yemen.

[5] While we do not address the detailed isssue of purchasing power parity (PPP) exchange rates in the paper, it is one of the important issues in estimating the WDI and measuring world income inequality. As noted in Milanovic (2012), the information about the level of prices in various countries plays a crucial role since the price information can gives us a criterion for the comparison of individual or average welfare in different countries.

[6] Available from the National Accounts Main Aggregate Database.

Considering data availability for more countries and a longer sample period, we use quintile income shares as in Sala-i-Martin (2006). We obtain quintile income shares of each country from the UNU-WIDER (United Nations University's World Institute for Development Research, version 3.0B). It provides information on the distribution of income based on microeconomic income surveys for each country.[7]

2.2 Methodology

To estimate the WDI, we basically follow the imputation method by Sala-i-Martin (2006). However, wider availability of survey data around world enables us to improve the estimation of the WDI and, therefore, we adopt a few modified approaches as explained below in details. We estimate an annual income distribution for each of 188 countries and integrate these country distributions for all levels of income to construct the WDI. We use the population-weighted income per capita as the mean of each country's income distribution (see Sala-i-Martin (2006) for more details). Next, we complement the mean of the distribution with within-country information on income distribution contained in income surveys. Specifically, we use the quintile income shares in the UNU-WIDER.

To estimate annual income distributions, we need to have the quintile income shares for every country and every year. However, since surveys are not available annually for every country, we need to impute the missing data. For example, Table 1 reports the quintile income shares in China. It shows that surveys are not conducted annually and therefore we need to impute the missing data to estimate the China's income distribution annually.

Sala-i-Martin (2006) divided the sample of countries into three groups based on data availability and applied different approaches for each group to impute the missing income share data.[8] Sala-i-Martin (2006) defined Group A as countries for which income surveys

[7] Deininger and Squire (1996) reported microeconomic income surverys, and extended and updated surveys are available in the UNU-WIDER.

[8] Bourguignon and Morrisson (2002), Pinkovskiy (2013), Pinkovskiy and Sala-i-Martin (2009) and Rougoor and van Marrewijk (2015) imputed the missing income share data according to economic and geographical characteristics without grouping of all the countries.

are reported for more than one year from 1970 to 2000, and used a simple linear time-trend forecast to estimate the missing values of quintile income shares. As explained in footnote 9 in Sala-i-Martin (2006), the regressions were estimated independently for each of the five quintiles.

However, when we apply his method on countries for which income surveys are available for more than one year, there are 26 countries whose linear time-trend extrapolations of quintile income shares severely violate the basic conditions for income shares; quintile shares are all positive and the last quintile income share is the largest. The linear time-trend extrapolations provide some negative quintile income shares for eighteen countries[9] and the fourth quintile income share becomes higher than the last quintile shares for a few years for eight country[10]. Compared to Sala-i-Martin (2006), we consider a longer sample period and a larger number of countries, and use more updated quintile income shares of individual countries. If basic conditions for income shares are violated, it will be more desirable to adopt another method to replace the linear time-trend interpolations method to estimate the missing values.

Consequently, we classify Group A and Group B differently from those in Sala-i-Martin (2006). Compared to Sala-i-Martin (2006), we impose an additional condition for Group A such that linear time-trend forecasts of missing quintile income shares satisfy basic conditions of income shares. If basic conditions for income shares are violated, we classify those countries as Group B while Sala-i-Martin (2006) defined Group B as countries for which only one income survey is reported between 1970 and 2000. We divide the sample of countries into the following three groups:

Group A – Countries for which linear time-trend forecasts of missing quintile income shares satisfy basic conditions for income shares (quintile shares are all positive and the last quintile share is the largest) and GDP per capita is available.

[9] Afghanistan, Angola, Bhutan, Cameroon, Central African Republic, Guinea, Guinea-Bissau, Iraq, Maldives, Micronesia, Namibia, Nicaragua, Singapore, Syria, Tanzania, Turkmenistan, Yemen and Zimbabwe.

[10] Albania, Belize, Bosnia and Herzegovina, Cuba, Gambia, Guyana, Iceland, Macedonia

Group B – Countries for which linear time-trend forecasts of missing quintile income shares violate basic conditions for income shares and GDP per capita is available.

Group C – Countries for which income surveys are not reported and GDP per capita is available.

The list of countries in each group is provided in Appendix.

Income Shares for Countries in Group A

There are 114 countries in Group A, which had 91.32% of the world population and 95.85% of the world GDP in the year 2010. For Group A, Sala-i-Martin (2006) used the linear time-trend forecast method to estimate the missing data. When we need to extrapolate the missing data, we also adopt the linear time-trend forecast method as in Sala-i-Martin (2006). However, when we need to interpolate the missing data, we use the linear interpolation method instead of the linear time-trend forecast method. We adopt such a modified approach because it can provide improved estimates of the missing data. Figures 1 and 2 present illustrations for China and India and they show that the linear interpolation provides better estimates, which is obvious particularly for the fourth and the last quintiles in China and for the first, the second and the last quintiles in India.

Since Group A covers the majority of the world population, one may want to consider only Group A to construct the WDI. However, as Sala-i-Martin (2006) pointed out, it would lead to sample selection bias because countries that do not belong to Group A tend to be poor and, therefore, their exclusion would induce biased results on world income inequality.

Income Shares for Countries in Group B

There are 49 countries in Group B, which had 6.50% of the world population and 2.13% of the world GDP in the year 2010. Sala-i-Martin (2006), for each country in Group B, imputed the shares for the missing years by averaging the trends for the "neighboring countries" in

Group A. "Neighboring countries" are those in "region" as defined by the World Bank[11] (see Sala-i-Martin (2006) for more details). In Sala-i-Martin (2006), quintile income shares for countries in Group B were available only for one year, and even if the available quintile income shares are quite different from the imputed estimates from the neighboring countries, such differences were ignored in Sala-i-Martin (2006). On the other hand, many countries in Group B have surveys reported for more than one year in our data set and it would be more desirable to account for the differences between the available quintile income shares and the imputed estimates from the neighboring countries. We employ the following modified approaches to obtain more reasonable imputed values for quintile income shares.

Suppose that surveys are available for the year 2001 and 2003 in an arbitrary country. We can also obtain the imputed estimates of quintile income shares from the neighboring countries as in Sala-i-Martin (2006). For example, available last quintile shares and corresponding neighboring averages are given as follows:

Year	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010
Actual data	40	(42)	45	(43)	(41)	(39)	(37)	(35)	(33)	(31)
Average	30	28	31	31	31	31	31	31	31	31
Deviation	10	(12)	14	(12)	(10)	(8)	(6)	(4)	(2)	(0)

Note: Average means the average of neighboring countries. Deviation implies the difference between an actual value and an average. The imputed values are in parentheses.

First, we consider the year 2002 for which we need to interpolate the missing data. If we simply use the neighboring average as the imputed value for the year 2002, it will be 28, which is too much different from actual income shares in the year 2001 and 2003. Instead, we adopt the following two-step procedure. First, we obtain the estimated deviation for the year 2002 by using the linear interpolation of deviations, which will be twelve. Second, we add it to the neighboring average and obtain the imputed value for the year 2002, which is

[11] The regions are East Asia and Pacific, Europe and Central Asia, Latin America and the Caribbean, Middle East and North Africa (MENA), South Asia, Sub-Saharan Africa, High-Income Non-OECE and High-Income OECD.

42. Obviously, it is a better estimate compared to the simple replacement of the neighboring average.

Next, we consider the period from 2004 to 2010 for which we need to extrapolate the missing data. We first set the deviation for the last year 2010 to be zero and obtain the estimated deviation by the linear interpolation. Here, we assume that the country's income shares will converge to the neighboring averages. Then, we combine the neighboring average and the estimated deviation for each missing year.

Income Shares for Countries in Group C

There are 25 countries in Group C, which had 0.85% of the world population and 2.03% of the world GDP in the year 2010. Compared to Group B, Group C has a much smaller population share but a higher GDP share. This is because there are many high income and non-OECD countries in Group C (for example, Bahrain, Kuwait, Macao, Qatar, Saudi Arabia and United Arab Emirates). For these countries, income surveys are not available. Following Sala-i-Martin (2006), we impute the neighboring countries' average quintile shares and the average time trend of each of the shares in groups A and B to construct the quintile income shares for each country in Group C.

Estimation of Country Distributions and Constructing World Distribution of Income

In the previous step, an income share is assigned to each quintile of each country for each year. The remaining steps are as follows; we approximate each country's annual income distribution and integrate the annual country distributions to estimate the annual WDI. For these remaining steps, we follow the method in Sala-i-Martin (2006). We use a nonparametric kernel method to approximate each country's annual income distribution and adopt the same bandwidth for all countries and periods (see Section II.F in Sala-i-Martin (2006) for details). We adopt the Gaussian kernel and select the bandwidth by following Silverman's (1986) rule of thumb: the bandwidth h is set to be $h = 1.06 \times \hat{\sigma} \times n^{-1/5}$, where $\hat{\sigma}$ is the standard deviation of the entire sample and n is the number of observations for each year. Once the kernel density function for a particular year and country is estimated, we anchor it so that its mean corresponds to the PPP-adjusted GDP per capita. Finally, we construct an annual WDI by integrating all the country distributions.

3 Results without Top Income Shares

3.1 Estimated World Distribution of Income

Figure 3(a) presents the estimates of annual WDI from 1970 to 2010 in a 3-D plot and Figure 3(b) provides those for 1970, 1980, 1990, 2000 and 2010. When we examine the evolutions of the WDI over time, the mode of the distribution shifts rightward and the deviation from its center tends to be smaller. This indicates that the mean of the WDI increases and the variance of the WDI decreases over time. When we examine the data, the mean increased by 12.00% and the variance decreased by 7.66% for the period from 1970 to 2010. The estimated WDI shows that the incomes of the world's citizens generally increased over time.

Another feature is that the WDI was bimodal in 1990 and changed to be unimodal in 2010. In Figures 3(a) and 3(b), the vertical line represents probability. Figure 4 also provides the WDI but the vertical line in Figure 4 represents population.[12] This figure corresponds to Figure IV in Sala-i-Martin (2006), where the WDI in 1990 was also slightly bimodal. Our result shows that the WDI was obviously bimodal from late 1980s to mid 1990s and "twin peaks" disappeared in 2010. The evolution of the WDI over time may indicate that, recently, global income inequality was substantially improved. In the next subsection, we adopt precise measures of income inequality to see the evolution of world income inequality over the last four decades.

[12] In estimating the WDI in Figure 4, the only different step is following. After the kernel density function for a particular year and country is estimated, we decompose it into 100 centiles and normalize it so that the area is equal to that year's total population of the country.

3.2 World Income Inequality

We first examine the Gini coefficient, which is the most typical measure of income inequality. We calculate the Gini coefficient using the income share data.[13] Figure 5 provides the plot of the world Gini coefficient from 1970 to 2010 and Table 2 reports the Gini coefficient with other measures of income inequality. The Gini coefficient exhibits a downward trend after it reaches its peak (0.692) in 1973. For the last four decades from 1970 to 2010, the Gini coefficient decreased by 8.71%, which implies that worldwide income inequality was substantially improved. One of the most important features is that global income inequality was markedly improved during the 2000s. For the period from 1970 to 2000, it fluctuated between 0.677 and 0.692 and slightly decreased; it decreased by 0.12%, 0.34% and 0.46% during the 1970s, the 1980s and the 1990s, respectively. However, the Gini coefficient exhibited a rapid decrease since 2000 and decreased by 7.87% during the 2000s. This is a drastically rapid decrease considering that it decreased only by 0.91% for thirty years from 1970 to 2000.

Sala-i-Martin (2006, Table III) reported that the Gini coefficient decreased by 2.4% for the period from 1970 to 2000. The difference between his result and ours is due to differences in data and methods to impute missing data. It should be noted that the difference between his result and ours is relatively minor. However, it should be noted that there can be a substantial difference if one adopts a different estimation method for the WDI. For example, Pinkovskiy and Sala-i-Martin (2009) adopted parametric estimations of the WDI and their result showed that the Gini coefficient decreased by 6.36% for the period from 1970 to 2000, which is quite different from the result in Sala-i-Martin (2006) or this paper.

It is worthwhile to briefly mention recent studies that covers the 2000s. Liberati (2015, Table 1) reported that the Gini coefficient decreased by 2.98% for the period from 1970 to 2000 and dropped by 4.97% for the period from 2000 to 2009. Pinkovskiy and Sala-i-Martin (2009, Table 3) reported that the Gini coefficient decreased by 6.36% for the period from

[13] We applied other methods to calculae the Gini coefficient and the results were similar. For example, we also calculated the Gini coefficient using the cumulative distribution function of the estimated WDI as in Cowell (2000) and the result was similar.

1970 to 2000 and decreased by 3.32% for the period from 2000 to 2006. Rougoor and van Marrewijk (2015, Figure 5) showed that the Gini coefficient slightly decreased during the 1990s and dropped substantially for the period from 2000 to 2009. Warner et al. (2014, Table 5) presented that the Gini coefficient decreased by 3.65% for the period from 2000 to 2005 while it decreased by 2.24% for the period from 1993 to 2000.

We also report seven other indices of income inequality in Table 2 and Figure 6: two Atkinson indices with coefficients 0.5 and 1, respectively, the variance of the logarithm of income, the ratio of the income of top 20 percent of the distribution to the bottom 20 percent and the ratio of the top 10 percent to the bottom 10 percent of the distribution, the Mean Logarithmic Deviation (MLD, which corresponds to the generalized entropy index with coefficient 0) and the Theil index (which corresponds to the generalized entropy index with coefficient 1). Figure 6 presents the plots of these seven indices. Except for the top-20-percent-to-bottom-20-percent ratio and the top-10-percent-to-bottom-10-percent ratio[14], most indices show similar results as the Gini coefficient; most importantly, global income inequality declined rapidly during the 2000s. Two Atkinson indices with coefficients 0.5 and 1, respectively, and MLD slightly decreased from 1970 to 2000 and rapidly decreased during the 2000s as the Gini coefficient did. The variance of log income and the Theil index slightly increased by 0.08% and 0.75%, respectively, from 1970 to 2000 but rapidly decreased by 8.46% and 18.97% during the 2000s.

Various income inequality indices indicate that world income inequality was drastically improved during the 2000s compared to the period from 1970 to 2000. One may think that this finding contradicts to the common perception that income inequality recently deteriorated. We address this issue in the next subsection by decomposing world income inequality into across-country inequality and within-country inequality.

[14]While the top-20-percent-to-bottom-20-percent ratio and the top-10-percent-to-bottom-10-percent ratio declined substantially by 18.61% and 20.37%, respectively, for the last four decades from 1970 to 2010, they exhibited different trends compared to the rest indices. For example, the top-20-percent-to-bottom-20-percent ratio rapidly decreased during the 1980s and the 1990s but increased during the 2000s.

3.3 Inequality Decomposition

We decompose global income inequality into two components: across-country inequality and within-country inequality. The "across-country" component is the amount of inequality that would exist in the world if all citizens within each country had the same level of income, but there were differences in per capita incomes across countries. For across-country inequality, every citizen's income is assumed to be his or her country's per capita income. The "within-country" is the difference between overall global inequality and across-country inequality. The within-country inequality is the amount of inequality that would exist in the world if all countries had the same income per capita but the actual within-country differences across individuals.

Among eight inequality indices we consider, the MLD and Theil index belong to the class of the generalized entropy index and are decomposable (see Cowell (1995) and Sala-i-Martin (2002) for details on the generalized entropy index and decomposition). Table 3 provides the decomposition of global income inequality using the MLD and Theil index.

In the year 1970, 79.5% of the global MLD was accounted for by the across-country inequality and only 20.5% of the global MLD was accounted for by the within-country inequality. This implies that the world income inequality was mainly due to the across-country inequality in the year 1970. While the global MLD declined by 19.0% during the last four decades, each component exhibited opposite trends; the across-country inequality decreased substantially by 30.6% but the within-country inequality increased by 26.1% from 1970 to 2010. The within-country inequality substantially increased during the 1990s and 2000s: 11.4% (during the 1990s) and 6.6% (during the 2000s). This is in accordance with the common perception that income inequality recently deteriorated. However, the across-country inequality decreased more substantially and, in particular, it rapidly declined by 23.9% during the 2000s. This is the main reason why the global income inequality was drastically improved during the 2000s considering that the across-country inequality takes up about 70% of the global MLD. For the whole period from 1970 to 2010, the decrease of the across-country inequality was larger than the increase of the within-country inequality, which consequently reduced overall global income inequality. In the year 2010, only 68.1% of global MLD came from the across-country component (down from 79.5% in 1970) and 31.9% was originated from the within-country component (up from 20.5% in 1970).

When we examine the decomposition of the Theil index, the results are generally similar to those for the MLD. In the year 1970, 76.1% of the global Theil index was accounted for by the across-country inequality and 23.9% was accounted for by the within-country inequality. For the last four decades, the across-country inequality declined by 27.5% while the within-country inequality increased by 10.6%. In particular, the across-country inequality rapidly declined by 27.2% during the 2000s. On the other hand, the within-country inequality increased during the 1990s and the 2000s: 12.5% (during the 1990s) and 6.0% (during the 2000s). It is interesting to note that the within-country inequality increased the most during the 1990s for both the MLD and the Theil index. In the year 2010, only 67.6% of global Theil index came from the across-country component (down from 76.1% in 1970) and 32.4% was originated from the within-country component (up from 23.9% in 1970).

For both the MLD and the Theil index, in general, the across-country inequality has been decreasing but the within-country inequality has been increasing over the last four decades. In particular, the across-country inequality rapidly declined during the 2000s while the within-country inequality substantially increased during the 1990s and the 2000s. We confirm that the common perception that income inequality recently deteriorated is indeed true when we consider the within-country income inequality. The within-country inequality shows that when we consider one country, income inequality recently deteriorated. Meanwhile, we also confirm that this does not contradict to our finding that the global income inequality was substantially improved during the 2000s. This is because the across-country inequality that accounts for about 70% of the global inequality became improved by a large amount during the 2000s. When we consider the whole population of the world, income inequality became substantially improved during the 2000s.

3.4 Main Convergers

Sala-i-Martin (2006) argued that the period from 1970 to 2000 was convergence period and that China played the most important role among "main convergers". It means that world income inequality declined for the period and one of the main reason was the rapid developments in countries like China. As shown in the previous subsections, world income inequality decreased drastically during the 2000s and it was mainly due to the improvement in the across-country inequality. In this subsection, we investigate which country or region played the role of main convergers for the period.

Table 4 provides the population ratios of eight regions of the world in the year 2010. China belongs to the East Asia and Pacific region and accounts for 19.4% of the world population. India belongs to the South Asia region and takes up 17.1% of the world population. The Sub-Saharan Africa region and the Latin America and Caribbean region account for 12.4% and 8.2%, respectively, of the world population. Figure 4 provides the plots of various Gini coefficients: 1) Gini coefficient including all countries, 2) Gini coefficient excluding China, 3) Gini coefficient excluding India, 4) Gini coefficient excluding Sub-Saharan Africa, 5) Gini coefficient excluding Latin America and Caribbean and 6) Gini coefficient excluding both China and India. Table 5 reports the change rates of these Gini coefficients for each decade.

When China was excluded, the Gini coefficient increased from 1978 to 2000 and increased by 5.06% during the period from 1970 to 2000. Considering that the Gini coefficient including China tended to slightly decrease by 0.91% until 2000 as shown in Figure 7, this implies that China was one of the main convergers until 2000. Even during the 2000s, China was still one of the main convergers; the Gini coefficient would decrease by 5.11% instead of 7.87% when China was excluded. While China still played the most important role as a main converger during the 2000s, India also played a role as a main converger for the period; the Gini coefficient would decrease by 7.08% instead of 7.87% during the 2000s when India was excluded. When both China and India were excluded, the Gini coefficient would increase by 4.06% and 4.83%, respectively, for the 1980s and the 1990s and would decrease only by 2.71% instead of 7.87% during the 2000s. Our result is in accordance with previous studies. Chotikapanich et al. (2012), Jones (1997), Liberati (2015) and Milanovic (2012) also emphasized the effect of China and India in the change of world income inequality.

However, when the Sub-Saharan Africa region was excluded, the Gini coefficient would decrease more substantially for the entire period including 2000s. The Gini coefficient would decrease by 9.58% instead of 7.87% during the 2000s when the Sub-Saharan Africa region was excluded. This indicates that the region was still one of "main divergers" during the 2000s. Meanwhile, the influence of the Latin America and Caribbean region was marginal; the Gini coefficient would decrease by 7.82% instead of 7.87% during the 2000s when the region was excluded.

4 Inclusion of Top Income Tax Data

The results in the previous section are based on the income survey data. However, it is well known that income survey data are biased particularly for top income shares due to under-reporting of the very rich. Therefore, one probably argues that the main findings in the previous section could be misleading due to the under-reporting in top income shares and it would be more desirable to include top income shares of each country in the analysis of world income inequality. We address this issue in this section and examine whether the main findings in the previous section still hold even when we include top income data in our analysis.

Recently, there have been a few attempts to combine top income shares with income survey data. Anand and Segal (2014) combined top 1% share of income tax data with income survey data. Their detailed method is explained in the next subsection. Meanwhile, Lakner and Milanovic (2013) imputed top 1% and 5% income shares by using the difference between survey incomes and the household final consumption expenditure from the national account and they assumed a specific distribution (Pareto distribution) for the WDI.

When we combine the data on top income shares with the income survey data, we use the income tax based top income shares provided by the World Top Income Database[15] that provide more precise information on top income shares. Recently, research based on top income tax data gained much attention (see Atkinson et al. (2011) and Pikettty (2014) among others) and Anand and Segal (2014) also used the same data set. It should be noted that the top income tax data considers taxable income while income survey data considers disposable income.[16] Due to such a difference in the definition of income, we should be cautious when we interpret the results in this section. It would be better if we use only one kind of income for the analysis. However, the data based on tax data are available only for small number of countries and are quite limited to estimate the distribution of world citizens' income. Hence, it is desirable to develop more rigorous methods to combine top income tax data with income survey data in estimating the distribution of world citizens' income. In the next subsection, we propose and investigate an alternative way to combine top income tax data with income survey data.

4.1 Data and Methodology

The top income shares are available only for 29 countries from the World Top Income Database. Table 6 reports countries for which top income shares are available. Among these 29 countries, 27 countries belong to Group A and two countries (Mauritius and Singapore) belong to Group B. Table 6 shows that top 10%, 5%, 1%, 0.05%, 0.1%, 0.05% and 0.01% income shares are available and that there are missing data for some countries. Considering data availability, we use top 1% and 5% income shares from the World Top Income Database.

We explain the method by Anand and Segal (2014) before we describe our approach to combine top income tax data with income survey data. Anand and Segal (2014) combined top 1% share of income tax data from the World Top Income Database with Milanovic's (2012)

[15] It is constructed by Facundo Alvaredo, Tony Atkinson, Thomas Piketty and Emmanuel Saez. Available at http://wid.world/.

[16] Lakner and Milanovic (2013) provide detailed explanations on the difference between tax data and survey data in defining income.

dataset[17] of household surveys for five years (1988, 1993, 1998, 2002 and 2005). Assuming that the survey data represent only the bottom 99% of the population in each country, they multiply the population in each income group in the surveys by 0.99 and append the top percentile with its income share from the tax data. In their work, the top 1% income share data were available for about twenty countries (ranging from 18 to 23 countries in each year). For those countries that do not have top income data, they impute top 1% shares using the following simple linear regression;

$$top1\%share_{it} = \beta_0 + \beta_1 top10\%share_{it} + \beta_2 mean\_income_{it} + \epsilon_{it} \quad (1)$$

where $top1\%share_{it}$ is top 1% income share from the tax data, $top10\%share_{it}$ is top 10% income share from survey data, and $mean\_income_{it}$ is the mean income from the surveys. They used 104 country-years observations and estimated eq. (1) using pooled OLS method. However, they did not consider either time or individual heterogeneity in eq. (1). Not controlling this heterogeneity runs the risk of yielding biased results, see e.g. Baltagi (2013), and using a panel model gives more degrees of freedom, more efficiency, less collinearity among the variables.

While Anand and Segal (2014) did not consider either time or individual heterogeneity, we adopt a model allowing for both time and individual heterogeneity;

$$y_{it} = \alpha + \mu_i + \lambda_t + x'_{it}\beta + \varepsilon_{it} \quad (2)$$

where y_{it} is top 1% or 5% income share for country i and year t and x_{it} =(top 20% income share$_{it}$, logged GDP per capita$_{it}$)$'$. In eq. (2), μ_i represents individual heterogeneity across countries and λ_t accounts for time heterogeneity. To reduce number of parameters to estimate, we let λ_t to have the same value for each decade. For 29 countries from the World Top

[17] The survey data in Milanovic (2012) consider between 92 and 119 countries and cover between 87% and 92% for the five years: 92 countries (87% of the world population) in 1998 and 119 countries (92% of the world population) in 2005.

Income Database, top 1% or 5% income share is missing for some years. Therefore, we have an unbalanced panel and we adopt the fixed effect estimation method. For comparison, we also estimate the pooled model

$$y_{it} = \alpha + x_{it}'\beta + \varepsilon_{it}, \tag{3}$$

which corresponds to the model adopted by Anand and Segal (2014).

Table 7 reports the estimation results of these two models. The model allowing for time and individual heterogeneity exhibits much higher adjusted R^2 than the pooled model; it is 0.83 in both top 1% and 5% income shares while it is 0.47 (top 1%) and 0.43 (top 5%) for the pooled model. This shows that the model allowing for time and individual heterogeneity fits the data much better than the pooled model. Hence, we use this model instead of the pooled model to impute missing top income shares. Both explanatory variables, top 20% income share and logged GDP per capita, are estimated to be significant in all cases and their coefficients are smaller for the model with time and individual heterogeneity than the pooled model.

Using the estimates of the model with time and individual effects, we impute missing top income shares. For countries without top income share data, we use the average of estimated individual effect μ_i in neighboring countries for each year. As explained in Section 2.2, neighboring countries are those that belong to the same region defined by the World Bank. However, there is no country for which top income share data are available in the region of 'Europe and Central Asia' and 'Middle East and North Africa'. For countries in these two regions, we use estimates of neighboring regions. For countries in 'Europe and Central Asia', we use estimates from 'South Asia'. For countries in 'Middle East and North Africa', we use estimates from 'Sub-Saharan Africa'.

Once the missing top income shares are imputed for every country and every year, we merge the top income shares with the quintile income shares obtained in Section 2. Particularly, we merge the fifth quintile income share with the top income shares for each country. Since top income shares nest other top income shares, i.e. top 5% income share includes top 1% income share, we can accordingly obtain income share and population share for each section such as income share between top 1% and top 5%.

Given any country and any year, if we calculate GDP per capita of each top income share, its magnitude should be in an order of top 1%, top 5%, top 10% and top 20% by construction; i.e., GDP per capita of top 1% income share is the largest and that of top 20% income share is the smallest. However, there are some countries for which such a natural condition does not hold for some years after we impute top 1% and 5% income shares. For example, for China and Rwanda, imputed GDP per capita of top 1% income share is smaller than that of top 5% income share for a few years. In case of Mauritius, imputed GDP per capita of top 5% income share is smaller than that of top 10% income share for a few years.[18] The numbers of countries for which such a natural condition is violated are given as follows:

Model	Top 1%	Top 5%	Top 10%
Model with time and individual effects	2	1	59
Pooled model	4	45	92

This table shows that when we use the pooled model to impute missing top income shares, there are more countries for which the condition is violated. For example, there is only one country where imputed GDP per capita of top 5% is smaller than that of top 10% if we adopt the model allowing for time and individual heterogeneity. On the other hand, there are 45 countries where such a condition is violated if we adopt the pooled model. Therefore, this result also supports our choice of the model allowing for time and individual heterogeneity to impute missing top income shares. Meanwhile, there are 59 or 92 countries for which imputed GDP per capita of top 10% is smaller than that of top 20%. This is why we decide to exclude the top 10% income share data in estimating the WDI.

Once we merge the top 1% and 5% income shares with the quintile income shares, we estimate the WDI by following the method described in Section 2. Using this estimated

[18] For these years in three countries, we use change rate of top 20% income share to impute top 1% or 5% income shares.

WDI, we adopt precise measures of income inequality to see the evolution of world income inequality over the last four decades.[19]

4.2 Results with Top Income Shares

When top income shares are combined, the results are generally similar to those in Section 3 where top income share are not considered. While there exist some differences, the main findings in Section 3 still hold even when we combine top income tax data. We can still observe that world income inequality was rapidly improved during the 2000s and it was mainly due to the rapid improvement in the across-country inequality during the period.

Figure 8(a) provides the estimates of annual WDI for 1970, 1980, 1990, 2000 and 2010 for the cases with top income shares. Compared to the estimates of WDI in Figure 3(b) where top income shares are not included, the shape of WDI is generally similar but right tail of the WDI is longer due to the inclusion of top income shares. This is obvious when we look at Figure 8(b).

Figure 9 provides the world Gini coefficients with/without top income shares and Table 8 reports various measures of income inequality for the cases with top income shares. Not surprisingly, the value of the Gini coefficient is higher for all years when top income shares are included. Nevertheless, the trend of the Gini coefficient is similar in both cases. During the period from 1970 to 2000, it slightly decreased by 0.44% while it did by 0.91% when top income shares are excluded. During the 2000s, it rapidly decreased by 7.17% while it did by 7.87% when top income shares are excluded. During the last four decades from 1970 to 2010, it decreased by 7.58% while it did by 8.71% when top income shares are excluded. When top income shares were included, the improvement of world income inequality slightly slowed down. This implies that top income shares increased more than the rest shares over the last four decades. However, it should be noted that even if there exists such a difference,

[19] We also estimated a model allowing for only individual heterogeneity and imputed missing top income shares by using its estimates. The estimated WDI and income inequality indices were still similar in this case.

the trend of the Gini coefficient is generally similar regardless of top income shares.

We also report seven other indices of income inequality in Figure 10 and Table 8. In general, these other indices of income inequality also exhibit similar trends as those in Figure 6 and Table 2 where top income shares are excluded. Two Atkinson indices with coefficients 0.5 and 1, respectively, the variance of log income, the top-10-percent-to-bottom-10-percent ratio, MLD and Theil index also rapidly declined during the 2000s as the Gini coefficient did. Meanwhile, the top-20-percent-to-bottom-20-percent ratio still shows a different trend; it rapidly decreased during the 1980s and the 1990s but increased during the 2000s.

Finally, we decompose global income inequality into two components: across-country inequality and within-country inequality. Table 9 provides the decomposition of global income inequality using the MLD and Theil index. Compared to the case where top income shares are excluded, the within-country component accounts for a larger portion of overall global inequality for both the MLD and the Theil index. In the year 1970, the within-country component takes up 22.3% of the MLD and 30.3% of the Theil index while it accounts for 20.5% of the MLD and 23.9% of the Theil index when top income shares are excluded. This difference is not surprising because the within-country inequality deteriorates as top income shares are included.

We can still observe the same features for both the MLD and the Theil index as those in the case where top income shares are excluded. First, the across-country inequality decreased for the last four decades; -30.3% for the MLD and -27.5% for the Theil index. Second, the within-country inequality increased for the last four decades; 28.3% for the MLD and 32.6% for the Theil index. Third, the across-country component rapidly declined during the 2000s (-23.9% for the MLD and -27.2% for the Theil index). Fourth, the within-country component rapidly increased during the 1990s (14.3% for the MLD and 22.0% for the Theil index). In case of the MLD, the inclusion of top income shares does not make any significant difference; the aggregate MLD and each component exhibit similar trends as those in the case where top income shares are excluded. However, the Theil index turns out to be substantially affected. In particular, the within-country component increased by 32.6% for the last four decades while it did by only 10.6% when top income share were excluded. Consequently, the within-country component accounts for 44.3% of the global Theil index in the year 2010, which is larger that 32.4% in the case where top income shares were excluded.

5 Conclusion

In the paper, we used GDP, population and quintile income shares of each country to estimate the world distribution of income. In particular, the quintile income share data are from the UNU-WIDER income survey data. We nonparametrically estimated the world distribution of income and calculated the world income inequality indices for the period from 1970 to 2010. We adopted improved methods to impute the missing income share data of individual countries. Specifically, we applied alternative interpolation methods in group specific imputation procedures. Moreover, we also investigated the case when top income tax data are included. It is well known that income survey data have the problem of non-response and under-reporting of top income groups. To take this into account, we used the most recent available income share data based on the World Top Income Database. We proposed an alternative way to impute missing top income shares by using a panel model allowing for time and individual heterogeneity.

Regardless that top income tax data are included or not, our results clearly show that the world income inequality was drastically improved during the 2000s compared to the period from 1970 to 2000. One may think that this finding contradicts to the common perception that income inequality recently deteriorated. However, the analysis of inequality decomposition provides proper explanations on such a seemingly contradictory result. The inequality within each country indeed increased during the 1990s and 2000s, which supports the common perception. However, the income inequality across countries, accounting for about 70% of the whole inequality, substantially decreased during the 2000s, which leaded to the drastic improvement of the world income inequality during the 2000s. It is shown that China and India played the most important role in improving the world income inequality for the period from 1980 to 2010. On the other hand, Sub-Saharan Africa region still played a role in deteriorating the world income inequality for the last four decades from 1970 to 2010.

Our analysis included total 188 countries, which accounted for 98.68% of the world population and almost 100% of the world GDP in the year 2010. To the best of our knowledge, this is the largest number of countries that are considered in the related literature. To analyze the distribution and inequality of world citizens' income, it is desirable to include as many countries as possible. Meanwhile, it is inevitable that many income survey data or top income tax data are missing particularly for low-income countries. Therefore, it is important to apply appropriate methods to impute missing income data to precisely estimate the world distribution of income and the world income inequality. In the paper, we adopted several improved methods to impute missing data and proposed an alternative method to combine top income tax data with income survey data. There are still rooms to improve imputation methods as more data will become available. Household income survey data, top income tax data and national accounts statistics are constantly being updated and extended to the global level. Moreover, it would be of interest if one extend/improve the method to combine top income tax data with income survey data.

A Tables and Figures

Table 1. Available quintile income shares in China

Year	1st quintile	2nd quintile	3rd quintile	4th quintile	5th quintile
1970	8.40	13.30	17.10	22.50	38.70
1972	8.60	13.20	17.10	22.50	38.60
1975	8.90	13.70	17.20	22.30	37.90
1980	7.93	12.27	18.42	24.72	36.66
1982	8.47	13.73	17.94	22.26	37.60
1983	8.65	14.57	17.01	24.26	35.51
1984	10.08	13.61	19.08	23.18	34.05
1985	8.71	12.91	16.25	23.38	38.75
1986	7.56	11.94	15.96	25.94	38.60
1987	6.92	11.12	15.95	28.44	37.57
1988	6.60	10.92	16.12	28.84	37.52
1989	6.46	11.58	15.87	24.06	42.03
1990	7.01	11.89	16.14	23.98	40.98
1991	6.44	11.40	14.85	31.25	36.06
1992	6.02	10.70	15.81	25.82	41.65
1993	7.35	11.32	15.80	22.30	43.23
1995	5.00	8.80	13.60	22.10	50.60
1998	5.86	10.20	15.10	22.20	46.64
2001	4.66	9.00	14.22	22.13	49.99
2002	4.55	8.45	13.65	23.41	49.96
2004	4.25	8.48	13.68	21.73	51.86
2005	4.99	9.85	14.99	22.24	47.93

Note: Available from the UNU-WIDER (United Nations University's World Institute for Development Research, version 3.0B).

Table 2. Various measures of World income inequality

Year	Gini	A(0.5)	A(1)	Variance	20/20	10/10	MLD	Theil
1970	0.687	0.389	0.630	1.826	12.43	34.55	0.993	0.898
1971	0.687	0.390	0.630	1.828	13.23	34.19	0.995	0.899
1972	0.691	0.395	0.638	1.877	13.95	37.51	1.016	0.908
1973	0.692	0.397	0.642	1.904	13.19	38.15	1.025	0.912
1974	0.689	0.393	0.638	1.898	13.26	39.72	1.016	0.898
1975	0.686	0.389	0.633	1.871	13.57	38.40	1.001	0.887
1976	0.689	0.394	0.641	1.923	13.71	42.93	1.024	0.897
1977	0.688	0.393	0.639	1.922	13.57	42.74	1.019	0.893
1978	0.685	0.389	0.634	1.896	13.92	39.85	1.004	0.884
1979	0.688	0.392	0.638	1.916	14.28	40.09	1.016	0.894
1980	0.686	0.390	0.637	1.933	14.79	42.17	1.014	0.884
1981	0.684	0.387	0.632	1.895	13.79	39.35	0.998	0.878
1982	0.680	0.381	0.623	1.848	13.24	36.18	0.976	0.867
1983	0.679	0.380	0.621	1.822	13.30	34.66	0.969	0.869
1984	0.679	0.380	0.621	1.824	13.14	33.76	0.969	0.872
1985	0.679	0.379	0.617	1.789	12.41	32.45	0.959	0.874
1986	0.679	0.379	0.616	1.783	12.19	34.66	0.958	0.876
1987	0.678	0.378	0.614	1.766	11.06	36.31	0.951	0.877
1988	0.678	0.378	0.615	1.777	11.62	34.89	0.954	0.875
1989	0.683	0.384	0.622	1.810	11.19	34.25	0.973	0.892
1990	0.684	0.385	0.623	1.824	10.10	35.24	0.976	0.895
1991	0.685	0.386	0.626	1.848	10.71	36.18	0.983	0.898
1992	0.685	0.386	0.624	1.832	9.70	33.67	0.979	0.902
1993	0.685	0.386	0.621	1.796	10.15	32.72	0.970	0.908
1994	0.683	0.384	0.620	1.804	10.13	32.05	0.967	0.905
1995	0.677	0.377	0.612	1.784	9.21	30.92	0.947	0.888
1996	0.679	0.379	0.614	1.794	9.00	31.24	0.952	0.893
1997	0.678	0.379	0.615	1.808	8.58	31.74	0.953	0.894
1998	0.682	0.383	0.619	1.826	8.70	30.60	0.965	0.906
1999	0.679	0.380	0.615	1.813	8.56	28.77	0.955	0.901
2000	0.680	0.382	0.617	1.828	8.43	30.07	0.961	0.905
2001	0.677	0.377	0.611	1.790	8.04	28.77	0.943	0.894
2002	0.676	0.376	0.613	1.816	9.20	27.32	0.948	0.887
2003	0.671	0.369	0.604	1.786	8.95	27.40	0.926	0.869
2004	0.664	0.362	0.595	1.756	7.96	28.16	0.903	0.850
2005	0.662	0.360	0.596	1.779	10.73	27.19	0.905	0.842
2006	0.655	0.352	0.586	1.747	9.32	27.13	0.881	0.820
2007	0.648	0.344	0.577	1.731	9.72	29.10	0.861	0.795
2008	0.642	0.338	0.571	1.717	10.13	30.60	0.845	0.776
2009	0.628	0.324	0.553	1.662	10.27	27.74	0.806	0.739
2010	0.627	0.323	0.553	1.673	10.12	27.51	0.805	0.733

Table 2. Continued

Year	Gini	A(0.5)	A(1)	Variance	20/20	10/10	MLD	Theil
% Change 1970-1980	-0.12%	0.12%	1.20%	5.83%	18.99%	22.07%	2.08%	-1.58%
% Change 1980-1990	-0.34%	-1.33%	-2.19%	-5.63%	-31.71%	-16.43%	-3.72%	1.21%
% Change 1990-2000	-0.46%	-0.85%	-0.94%	0.20%	-16.60%	-14.67%	-1.58%	1.14%
% Change 2000-2010	-7.87%	-15.41%	-10.46%	-8.46%	20.08%	-8.52%	-16.24%	-18.97%
% Change 1970-2000	-0.91%	-2.06%	-1.95%	0.08%	-32.22%	-12.95%	-3.27%	0.75%
% Change 1970-2010	-8.71%	-17.16%	-12.20%	-8.39%	-18.61%	-20.37%	-18.98%	-18.37%

Note: Gini is the Gini coefficient. A(0.5) is the Atkinson index with coefficient 0.5. A(1) is the Atkinson index with coefficient 1. Variance is the variance of log income. 20/20 is the ratio of the income of top 20 centile to botoom 20 centile. 10/10 is the ratio of the income of top 10 centile to bottom 10 centile. MLD is the Mean Logarithmic Deviation. Theil is the Theil index of income inequality.

Table 3. Decomposition of world income inequality

	Mean log deviation					Theil index				
Year	Global	Across	% Across	Within	%Within	Global	Across	% Across	Within	%Within
1970	0.993	0.789	79.5%	0.204	20.5%	0.898	0.684	76.1%	0.215	23.9%
1971	0.995	0.790	79.4%	0.205	20.6%	0.899	0.682	75.8%	0.217	24.2%
1972	1.016	0.811	79.9%	0.205	20.1%	0.908	0.694	76.4%	0.214	23.6%
1973	1.025	0.821	80.1%	0.204	19.9%	0.912	0.698	76.6%	0.214	23.4%
1974	1.016	0.820	80.7%	0.197	19.3%	0.898	0.693	77.1%	0.205	22.9%
1975	1.001	0.800	79.9%	0.201	20.1%	0.887	0.679	76.5%	0.208	23.5%
1976	1.024	0.821	80.2%	0.202	19.8%	0.897	0.690	76.9%	0.207	23.1%
1977	1.019	0.816	80.1%	0.203	19.9%	0.893	0.688	77.0%	0.205	23.0%
1978	1.004	0.802	79.8%	0.202	20.2%	0.884	0.683	77.2%	0.201	22.8%
1979	1.016	0.811	79.8%	0.205	20.2%	0.894	0.689	77.1%	0.205	22.9%
1980	1.014	0.800	78.9%	0.214	21.1%	0.884	0.681	77.0%	0.203	23.0%
1981	0.998	0.788	78.9%	0.210	21.1%	0.878	0.678	77.2%	0.200	22.8%
1982	0.976	0.769	78.8%	0.207	21.2%	0.867	0.667	77.0%	0.200	23.0%
1983	0.969	0.766	79.0%	0.203	21.0%	0.869	0.670	77.1%	0.199	22.9%
1984	0.969	0.762	78.7%	0.207	21.3%	0.872	0.673	77.2%	0.199	22.8%
1985	0.959	0.757	78.9%	0.202	21.1%	0.874	0.676	77.3%	0.198	22.7%
1986	0.958	0.756	79.0%	0.201	21.0%	0.876	0.677	77.3%	0.199	22.7%
1987	0.951	0.753	79.2%	0.198	20.8%	0.877	0.677	77.2%	0.200	22.8%
1988	0.954	0.752	78.8%	0.202	21.2%	0.875	0.682	77.9%	0.194	22.1%
1989	0.973	0.760	78.2%	0.212	21.8%	0.892	0.693	77.7%	0.199	22.3%
1990	0.976	0.760	77.8%	0.216	22.2%	0.895	0.696	77.7%	0.199	22.3%
1991	0.983	0.757	77.0%	0.226	23.0%	0.898	0.695	77.4%	0.203	22.6%
1992	0.979	0.748	76.4%	0.231	23.6%	0.902	0.693	76.8%	0.210	23.2%
1993	0.970	0.741	76.4%	0.229	23.6%	0.908	0.690	76.0%	0.218	24.0%
1994	0.967	0.735	76.0%	0.232	24.0%	0.905	0.686	75.8%	0.219	24.2%
1995	0.947	0.713	75.2%	0.234	24.8%	0.888	0.671	75.5%	0.217	24.5%
1996	0.952	0.712	74.9%	0.239	25.1%	0.893	0.672	75.2%	0.221	24.8%
1997	0.953	0.708	74.3%	0.245	25.7%	0.894	0.668	74.8%	0.225	25.2%
1998	0.965	0.716	74.2%	0.249	25.8%	0.906	0.680	75.0%	0.226	25.0%

Table 3. Continued

	Mean log deviation					Theil index				
Year	Global	Across	% Across	Within	%Within	Global	Across	% Across	Within	%Within
1999	0.955	0.713	74.6%	0.243	25.4%	0.901	0.679	75.4%	0.222	24.6%
2000	0.961	0.720	74.9%	0.241	25.1%	0.905	0.681	75.2%	0.224	24.8%
2001	0.943	0.706	74.8%	0.238	25.2%	0.894	0.669	74.9%	0.225	25.1%
2002	0.948	0.688	72.5%	0.261	27.5%	0.887	0.652	73.5%	0.235	26.5%
2003	0.926	0.676	72.9%	0.251	27.1%	0.869	0.639	73.5%	0.231	26.5%
2004	0.903	0.659	73.0%	0.244	27.0%	0.850	0.621	73.1%	0.229	26.9%
2005	0.905	0.646	71.3%	0.260	28.7%	0.842	0.606	72.0%	0.235	28.0%
2006	0.881	0.630	71.6%	0.251	28.4%	0.820	0.588	71.7%	0.232	28.3%
2007	0.861	0.611	70.9%	0.250	29.1%	0.795	0.566	71.1%	0.230	28.9%
2008	0.845	0.595	70.4%	0.250	29.6%	0.776	0.545	70.2%	0.231	29.8%
2009	0.806	0.557	69.1%	0.249	30.9%	0.739	0.507	68.7%	0.231	31.3%
2010	0.802	0.548	68.1%	0.257	31.9%	0.733	0.496	67.6%	0.238	32.4%
% Change 1970-1980	2.1%	1.3%		4.9%		-1.6%	-0.4%		-5.4%	
% Change 1980-1990	-3.7%	-5.0%		1.2%		1.2%	2.2%		-2.0%	
% Change 1990-2000	-1.6%	-5.3%		11.4%		1.1%	-2.1%		12.5%	
% Change 2000-2010	-16.2%	-23.9%		6.6%		-19.0%	-27.2%		6.0%	
% Change 1970-2000	-3.3%	-8.9%		18.3%		0.7%	-0.4%		4.3%	
% Change 1970-2010	-19.0%	-30.6%		26.1%		-18.4%	-27.5%		10.6%	

Note: Gobal measures indicate the overall index of inequality for the Mean log deviation and the Theil index, respectively. Across represents the across-country inequality. The column %Across provides the percentage of the global index that can be attributed to across-country inequality. Within refers to the within-country inequality. The column %Within shows the percentage of the global index that can be attributed to the within-country inequality.

Table 4. Population share of each region

Region	Population Shares
East Asia and Pacific	27.6%
South Asia	23.2%
High income: OECD	15.1%
Sub-Saharan Africa	12.4%
Latin America and Caribbean	8.2%
Middle East and North Africa	4.8%
Europe and Central Asia	4.0%
High income: non OECD	3.1%

Note: The table reports the population share of each region in the year 2010. The region is defined by the World Bank.

Table 5. Change rate of Gini coefficient for each decade

	All countries	No China	No India	No Africa	No Latin	No China and Inida
% Change 1970-1980	-0.12%	-0.15%	-0.65%	-0.65%	0.32%	-0.97%
% Change 1980-1990	-0.34%	2.10%	0.61%	-0.99%	-0.51%	4.06%
% Change 1990-2000	-0.46%	3.06%	-0.01%	-1.17%	-0.49%	4.84%
% Change 2000-2010	-7.87%	-5.11%	-7.08%	-9.58%	-7.82%	-2.71%
% Change 1970-2000	-0.91%	5.06%	-0.06%	-2.78%	-0.68%	8.04%
% Change 1970-2010	-8.71%	-0.30%	-7.14%	-12.09%	-8.44%	5.11%

Note: Table reports change rate of Gini coefficient for each decade and each case. For example, ‘No China’ means the case where the population in China is excluded.

Table 6. Countries for which top income shares are available

	10%	5%	1%	0.50%	0.10%	0.05%	0.01%
Argentina		o	o	o	o		o
Australia	o	o	o	o	o	o	o
Canada	o	o	o	o	o		o
China	o	o	o	o	o		
Colombia			o	o	o	o	o
Denmark	o	o	o	o	o	o	o
Finland	o	o	o				
France	o	o	o	o	o		o
Germany	o	o	o	o	o		o
India			o	o	o		o
Indonesia			o		o	o	o
Ireland	o		o	o	o		
Italy	o	o	o	o	o		o
Japan	o	o	o	o	o	o	o
Korea, Rep.	o	o	o	o	o	o	o
Malaysia	o	o	o	o	o	o	o
Mauritius	o	o	o	o	o	o	
Netherlands	o	o	o	o	o	o	o
New Zealand	o	o	o	o	o		
Norway	o	o	o	o	o	o	
Portugal	o	o	o	o	o		o
Singapore	o	o	o	o	o	o	o
South Africa	o	o	o	o	o	o	o
Spain	o	o	o	o	o		o
Sweden	o	o	o	o	o	o	o
Switzerland	o	o	o	o	o		o
Taiwan	o	o	o		o		o
United Kingdom	o	o	o	o	o	o	o
United States	o	o	o	o	o		o

Note: The table reports countries and percentages for which top income shares are available. For example, top 5%, 1%, 0.5%, 0.1% and 0.01% income shares are available for Argentian. Source: the World Top Incomes Database.

Table 7. Estimation results of top income shares

	β_1	β_2	adjusted R^2
y_{it} =top 1% income share			
Model with time and individual heterogeneity	0.25***	0.21***	0.83
	(0.02)	(0.08)	
Pooled model	0.36***	1.20***	0.47
	(0.01)	(0.10)	
y_{it} =top 5% income share			
Model with time and individual heterogeneity	0.35***	1.16***	0.83
	(0.03)	(0.12)	
Pooled model	0.55***	3.88***	0.43
	(0.02)	(0.23)	

Note: The table reports the estimation results of the following two models. The model with time and individual heterogeneity is given as

$$y_{it} = \alpha + \mu_i + \lambda_t + x'_{it}\beta + \varepsilon_{it},$$

where $y_{it}$is top 1% or 5% income share for country $i$and year $t$and x_{it}=(top 20% income share$_{it}$, logged GDP per capita$_{it}$)$'$and $\beta = (\beta_1, \beta_2)'$ .We let $\lambda_t$to have the same value for each decade. The pooled model is given as

$$y_{it} = \alpha + x'_{it}\beta + \varepsilon_{it}.$$

Standard errors are given in parentheses. ***denotes significance at the 1% level.

Table 8. Various measures of World income inequality when top income shares are combined

Year	Gini	A(0.5)	A(1)	Variance	20/20	10/10	MLD	Theil
1970	0.696	0.404	0.638	1.809	12.58	32.69	1.016	0.981
1971	0.696	0.403	0.638	1.802	13.17	32.50	1.014	0.979
1972	0.699	0.408	0.644	1.845	13.84	34.61	1.033	0.989
1973	0.701	0.410	0.648	1.872	12.87	35.22	1.043	0.993
1974	0.697	0.406	0.645	1.867	12.79	36.89	1.034	0.978
1975	0.694	0.402	0.639	1.842	13.07	34.50	1.019	0.967
1976	0.698	0.407	0.647	1.893	13.38	39.31	1.041	0.976
1977	0.696	0.405	0.645	1.891	13.51	37.47	1.035	0.969
1978	0.693	0.401	0.640	1.864	13.54	35.88	1.020	0.961
1979	0.696	0.404	0.644	1.884	14.16	36.36	1.032	0.970
1980	0.694	0.402	0.643	1.900	14.73	39.53	1.029	0.960
1981	0.691	0.398	0.637	1.864	13.21	36.71	1.012	0.951
1982	0.687	0.392	0.629	1.817	13.24	33.91	0.990	0.939
1983	0.687	0.392	0.627	1.795	13.25	33.37	0.985	0.942
1984	0.687	0.392	0.627	1.796	12.94	32.21	0.984	0.947
1985	0.687	0.392	0.624	1.766	12.64	32.06	0.977	0.951
1986	0.687	0.392	0.624	1.760	12.25	33.47	0.977	0.955
1987	0.686	0.392	0.621	1.738	11.17	33.66	0.969	0.964
1988	0.687	0.393	0.623	1.747	11.62	31.87	0.975	0.976
1989	0.692	0.399	0.630	1.779	10.90	30.71	0.994	0.993
1990	0.693	0.401	0.631	1.784	10.14	31.69	0.997	1.005
1991	0.694	0.402	0.634	1.808	10.71	33.03	1.004	1.004
1992	0.694	0.402	0.632	1.790	9.86	32.08	0.999	1.013
1993	0.694	0.402	0.630	1.759	10.15	32.48	0.992	1.016
1994	0.693	0.401	0.628	1.767	10.20	31.67	0.989	1.014
1995	0.688	0.395	0.621	1.748	9.41	30.08	0.970	0.999
1996	0.689	0.397	0.623	1.755	9.72	30.53	0.975	1.011
1997	0.689	0.398	0.625	1.771	9.53	30.67	0.980	1.016
1998	0.692	0.402	0.629	1.783	9.18	31.22	0.990	1.034
1999	0.690	0.399	0.625	1.772	9.32	29.11	0.981	1.034
2000	0.693	0.404	0.629	1.777	8.47	28.73	0.991	1.058
2001	0.689	0.399	0.623	1.745	8.47	27.27	0.974	1.038
2002	0.689	0.397	0.624	1.768	10.30	26.72	0.976	1.025
2003	0.683	0.391	0.616	1.739	9.96	27.12	0.955	1.009
2004	0.677	0.385	0.607	1.706	8.70	26.82	0.934	1.000
2005	0.677	0.386	0.609	1.723	10.07	26.02	0.938	1.007
2006	0.670	0.379	0.600	1.689	8.67	26.70	0.915	0.989
2007	0.664	0.371	0.592	1.673	8.47	28.33	0.896	0.968
2008	0.658	0.365	0.585	1.661	8.22	27.10	0.879	0.942
2009	0.645	0.350	0.568	1.611	8.86	25.64	0.840	0.894
2010	0.644	0.349	0.568	1.620	9.49	24.23	0.839	0.890

Table 8. Continued

Year	Gini	A(0.5)	A(1)	Variance	20/20	10/10	MLD	THEIL
% Change 1970-1980	-0.35%	-0.48%	0.69%	5.05%	17.09%	20.93%	1.22%	-2.15%
% Change 1980-1990	-0.08%	-0.17%	-1.78%	-6.12%	-31.16%	-19.82%	-3.05%	4.65%
% Change 1990-2000	-0.01%	0.70%	-0.37%	-0.39%	-16.50%	-9.35%	-0.61%	5.28%
% Change 2000-2010	-7.17%	-13.65%	-9.71%	-8.82%	12.09%	-15.67%	-15.38%	-15.84%
% Change 1970-2000	-0.44%	0.06%	-1.47%	-1.76%	-32.69%	-12.11%	-2.48%	7.80%
% Change 1970-2010	-7.58%	-13.60%	-11.03%	-10.43%	-24.56%	-25.88%	-17.47%	-9.28%

Note: Same as Table 2.

Table 9. Decomposition of world income inequality with top income share

	Mean log deviation					Theil index				
Year	Global	Across	% Across	Within	%Within	Global	Across	% Across	Within	%Within
1970	1.016	0.790	77.7%	0.226	22.3%	0.981	0.684	69.7%	0.297	30.3%
1971	1.014	0.790	77.9%	0.224	22.1%	0.979	0.682	69.7%	0.297	30.3%
1972	1.033	0.811	78.6%	0.221	21.4%	0.989	0.694	70.2%	0.295	29.8%
1973	1.043	0.822	78.8%	0.221	21.2%	0.993	0.699	70.3%	0.295	29.7%
1974	1.034	0.820	79.3%	0.214	20.7%	0.978	0.693	70.8%	0.285	29.2%
1975	1.019	0.800	78.5%	0.219	21.5%	0.967	0.679	70.2%	0.288	29.8%
1976	1.041	0.821	78.9%	0.220	21.1%	0.976	0.690	70.7%	0.286	29.3%
1977	1.035	0.816	78.8%	0.219	21.2%	0.969	0.688	71.0%	0.280	29.0%
1978	1.020	0.802	78.6%	0.219	21.4%	0.961	0.683	71.1%	0.278	28.9%
1979	1.032	0.811	78.6%	0.221	21.4%	0.970	0.689	71.0%	0.281	29.0%
1980	1.029	0.800	77.8%	0.228	22.2%	0.960	0.681	70.9%	0.279	29.1%
1981	1.012	0.788	77.8%	0.224	22.2%	0.951	0.678	71.3%	0.273	28.7%
1982	0.990	0.769	77.7%	0.221	22.3%	0.939	0.667	71.0%	0.272	29.0%
1983	0.985	0.766	77.8%	0.219	22.2%	0.942	0.670	71.2%	0.272	28.8%
1984	0.984	0.762	77.4%	0.222	22.6%	0.947	0.673	71.1%	0.274	28.9%
1985	0.977	0.757	77.5%	0.220	22.5%	0.951	0.676	71.1%	0.275	28.9%
1986	0.977	0.757	77.5%	0.220	22.5%	0.955	0.677	70.9%	0.278	29.1%
1987	0.969	0.753	77.7%	0.216	22.3%	0.964	0.677	70.2%	0.287	29.8%
1988	0.975	0.752	77.1%	0.223	22.9%	0.976	0.682	69.9%	0.294	30.1%
1989	0.994	0.761	76.5%	0.233	23.5%	0.993	0.693	69.8%	0.300	30.2%
1990	0.997	0.760	76.2%	0.237	23.8%	1.005	0.696	69.3%	0.309	30.7%
1991	1.004	0.757	75.5%	0.246	24.5%	1.004	0.695	69.2%	0.309	30.8%
1992	0.999	0.748	74.9%	0.251	25.1%	1.013	0.693	68.4%	0.320	31.6%
1993	0.992	0.741	74.7%	0.251	25.3%	1.016	0.690	67.9%	0.326	32.1%
1994	0.989	0.735	74.3%	0.254	25.7%	1.014	0.687	67.7%	0.328	32.3%
1995	0.970	0.713	73.4%	0.258	26.6%	0.999	0.671	67.2%	0.328	32.8%
1996	0.975	0.713	73.1%	0.263	26.9%	1.011	0.672	66.5%	0.339	33.5%
1997	0.980	0.709	72.3%	0.271	27.7%	1.016	0.668	65.7%	0.348	34.3%
1998	0.990	0.716	72.4%	0.274	27.6%	1.034	0.680	65.7%	0.354	34.3%

Table 9. Continued

	Mean log deviation					Theil index				
Year	Global	Across	% Across	Within	%Within	Global	Across	% Across	Within	%Within
1999	0.981	0.713	72.7%	0.268	27.3%	1.034	0.679	65.7%	0.355	34.3%
2000	0.991	0.720	72.6%	0.271	27.4%	1.058	0.681	64.4%	0.377	35.6%
2001	0.974	0.706	72.5%	0.268	27.5%	1.038	0.669	64.5%	0.368	35.5%
2002	0.976	0.688	70.5%	0.288	29.5%	1.025	0.652	63.6%	0.373	36.4%
2003	0.955	0.676	70.8%	0.279	29.2%	1.009	0.639	63.3%	0.371	36.7%
2004	0.934	0.659	70.6%	0.275	29.4%	1.000	0.621	62.1%	0.379	37.9%
2005	0.938	0.646	68.8%	0.293	31.2%	1.007	0.606	60.2%	0.400	39.8%
2006	0.915	0.631	68.9%	0.285	31.1%	0.989	0.588	59.5%	0.401	40.5%
2007	0.896	0.611	68.2%	0.285	31.8%	0.968	0.566	58.4%	0.403	41.6%
2008	0.879	0.595	67.6%	0.285	32.4%	0.942	0.545	57.8%	0.397	42.2%
2009	0.840	0.557	66.3%	0.283	33.7%	0.894	0.507	56.8%	0.386	43.2%
2010	0.839	0.548	65.3%	0.291	34.7%	0.890	0.496	55.7%	0.394	44.3%
% Change 1970-1980	1.2%	1.3%		0.8%		-2.2%	-0.4%		-6.2%	
% Change 1980-1990	-3.1%	-5.0%		3.9%		4.6%	2.2%		10.7%	
% Change 1990-2000	-0.6%	-5.3%		14.3%		5.3%	-2.1%		22.0%	
% Change 2000-2010	-15.4%	-23.9%		7.2%		-15.8%	-27.2%		4.7%	
% Change 1970-2000	-2.5%	-8.8%		19.7%		7.8%	-0.4%		26.7%	
% Change 1970-2010	-17.5%	-30.6%		28.3%		-9.3%	-27.5%		32.6%	

Note: Same as Table 5.

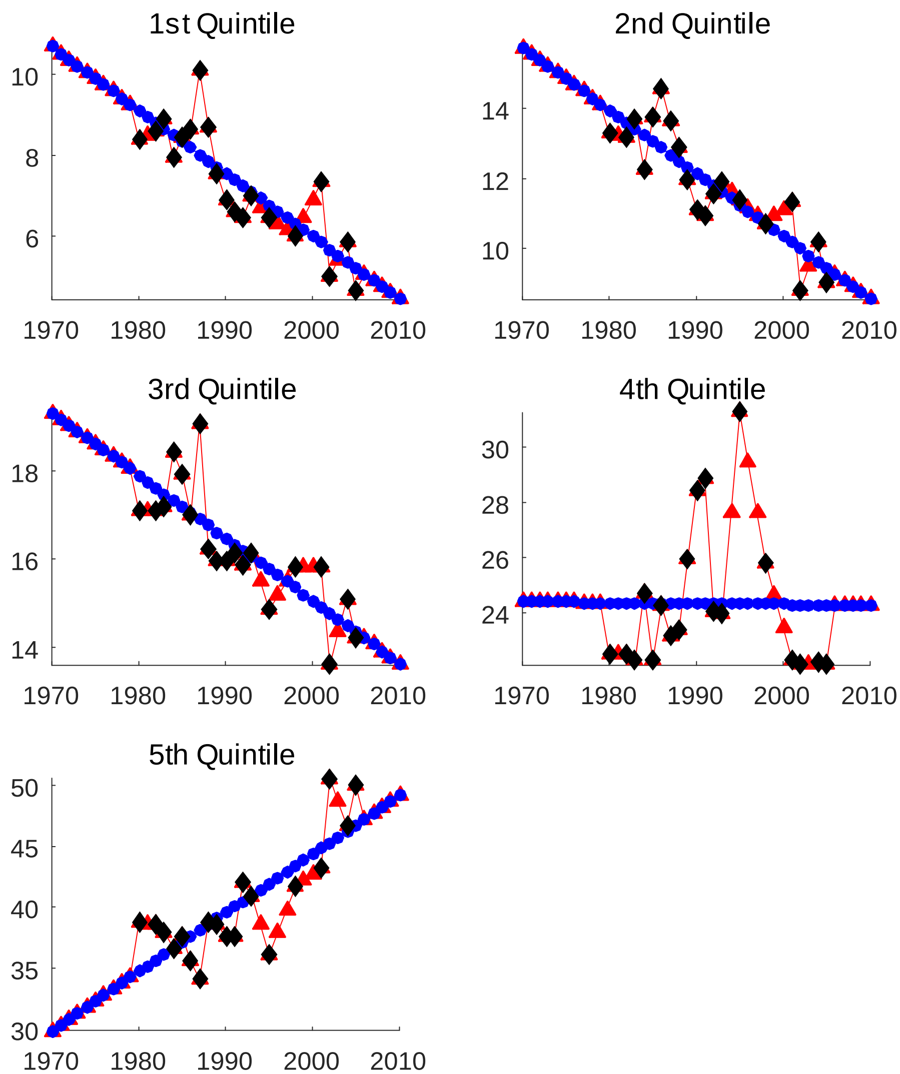


Figure 1. Imputation of income shares for China. Diamonds indicate actual available data. Circles refer to the imputed values by the linear time-trend forecast method. Triangles represent the imputed values by both linear interpolation and extrapolation using linear time-trend forecast.

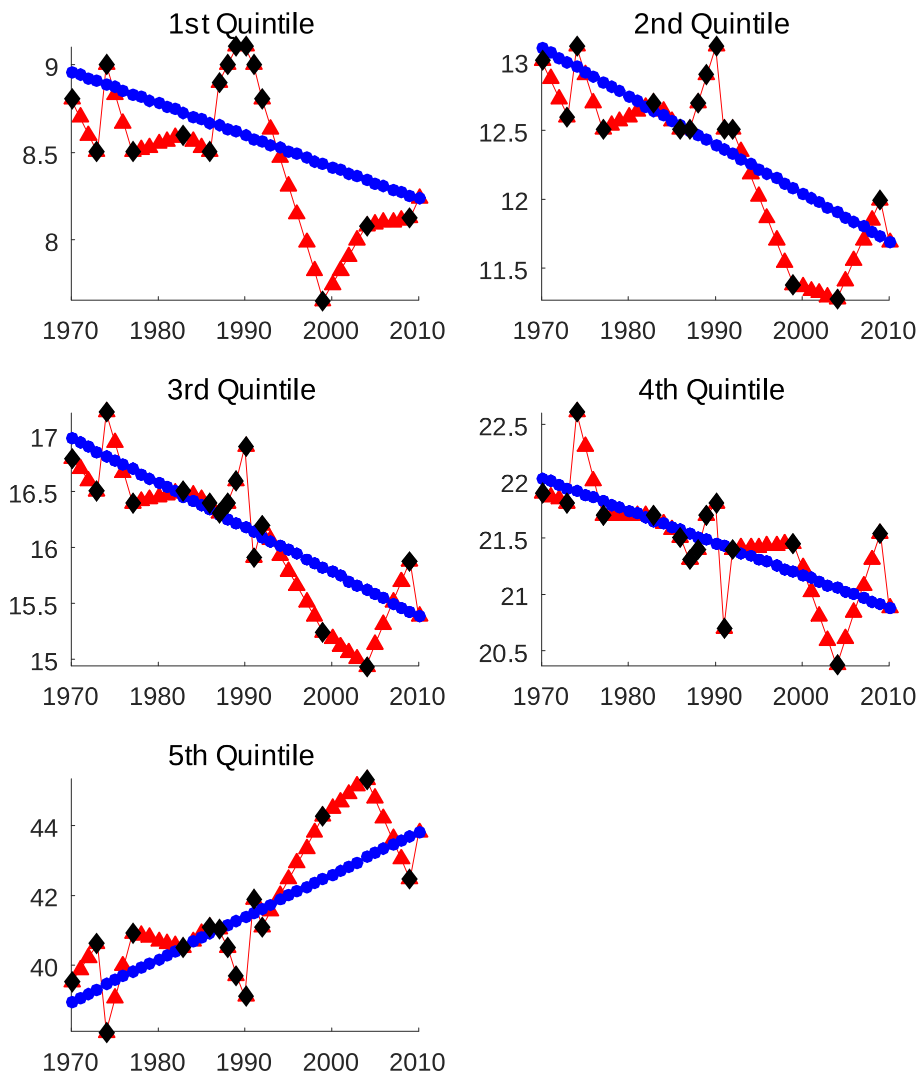


Figure 2. Imputation of income shares for India. Same as Figure 1.

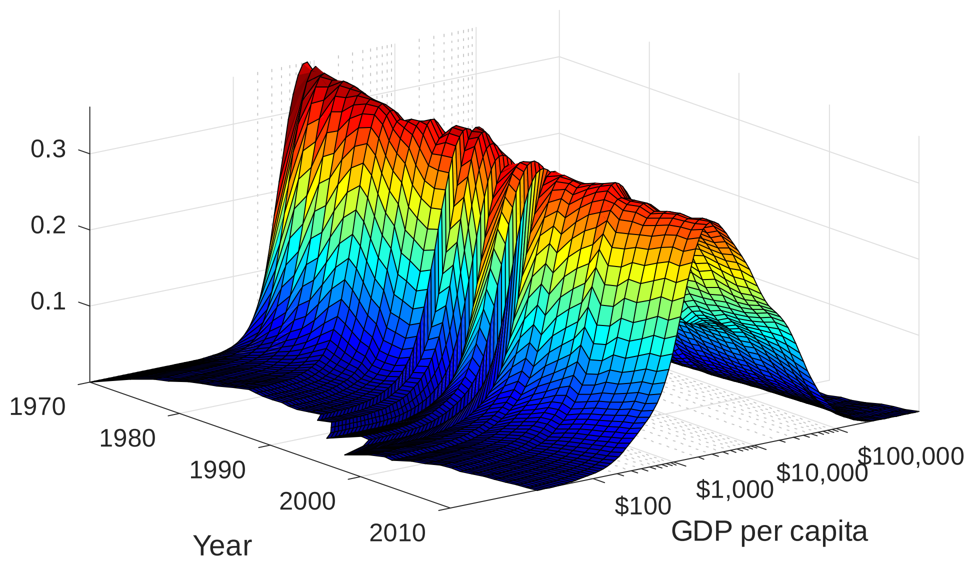


Figure 3(a). 3-D plot of the estimated world distribution of income

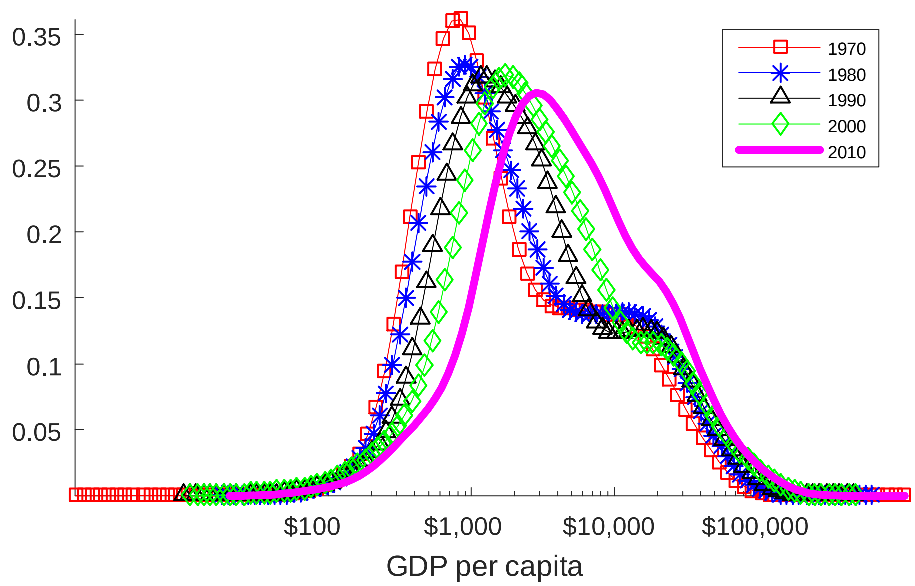


Figure 3(b). Estimated world distribution of income in various years

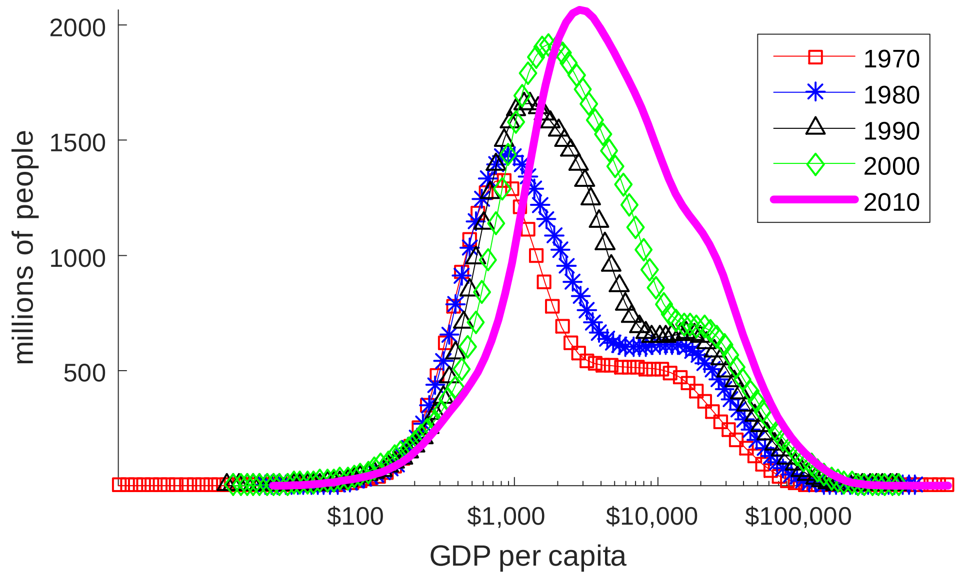


Figure 4. Estimated world distribution of income in various years (population-normalized version)

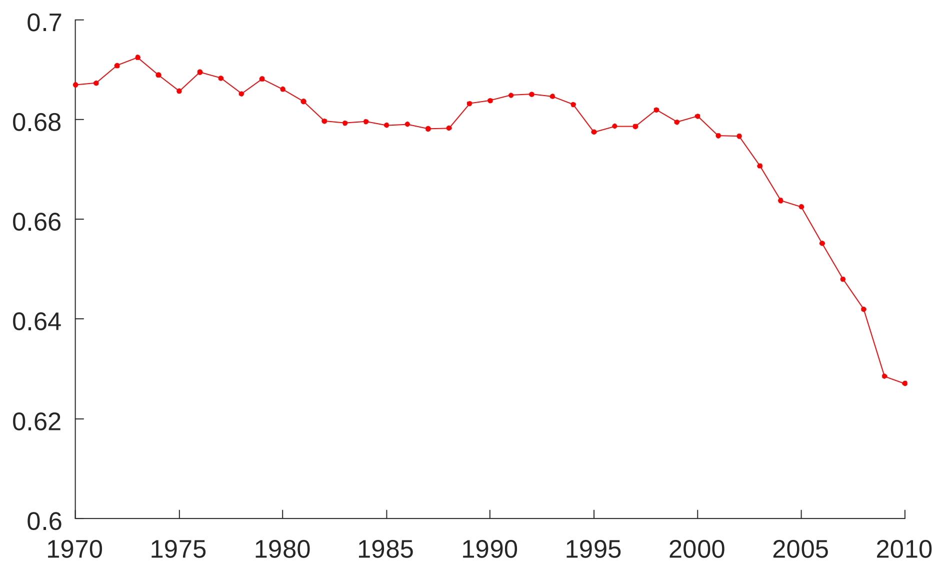


Figure 5. Global Gini coefficient

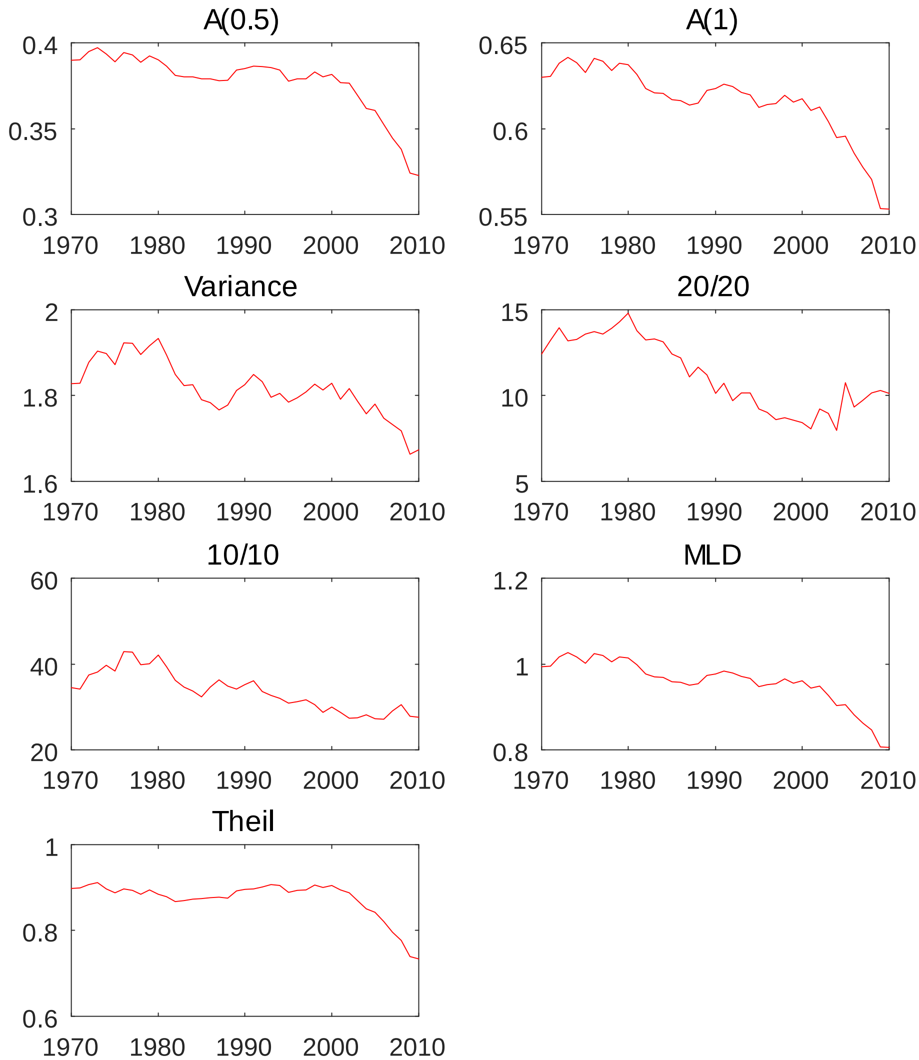


Figure 6. Seven other global inequality indices. Same as the note of Table

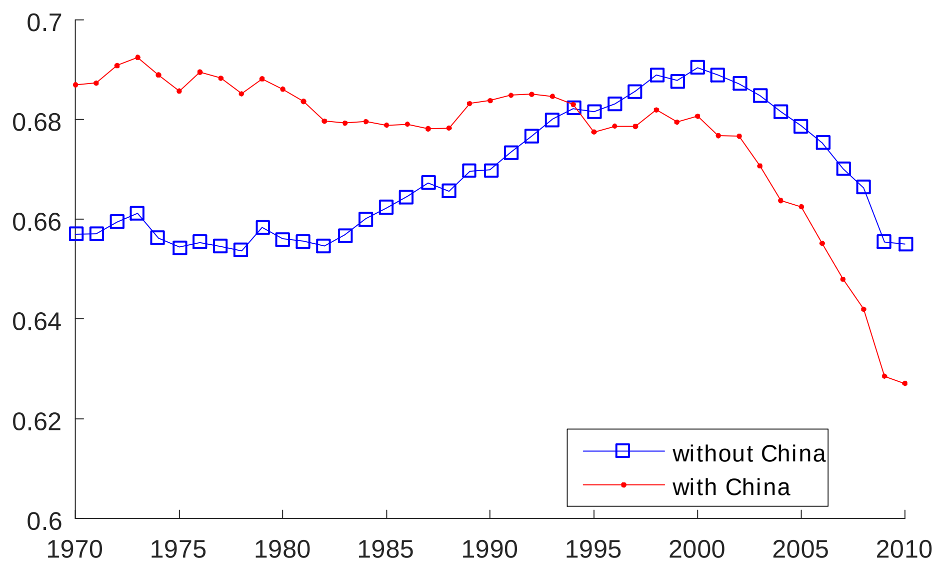


Figure 7. Global Gini coefficients for the cases with and without China

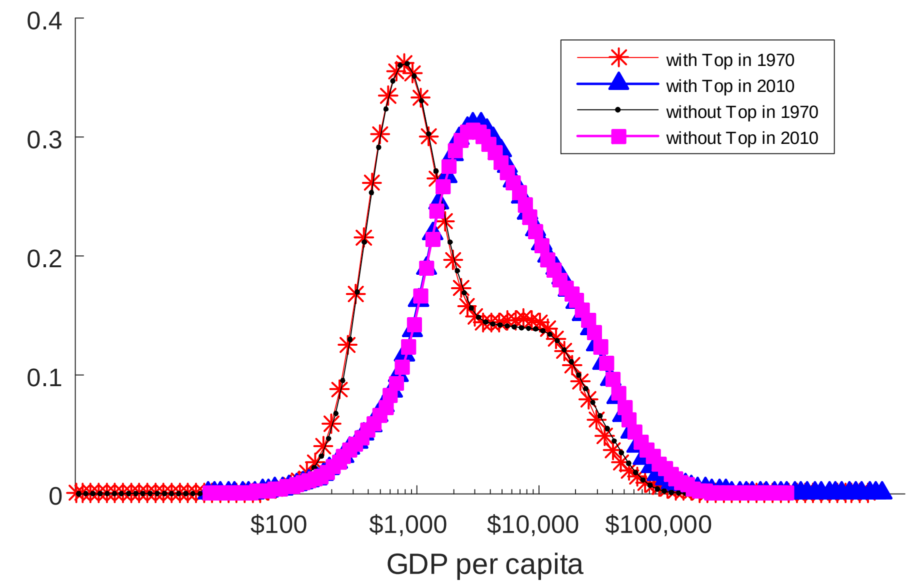


Figure 8. Estimated world distribution of income when top income shares are combined (**change the plots later**)

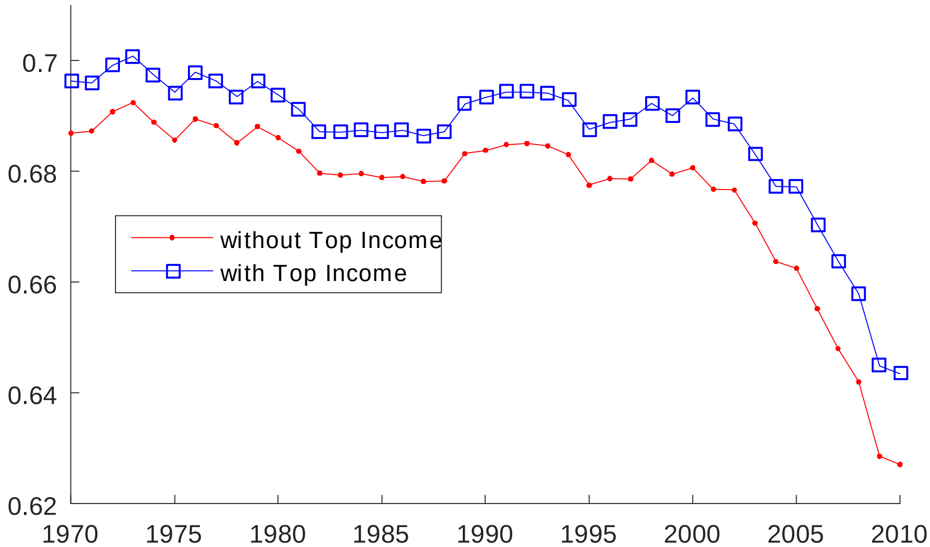


Figure 9. Global Gini coefficients for the cases with and without top income shares

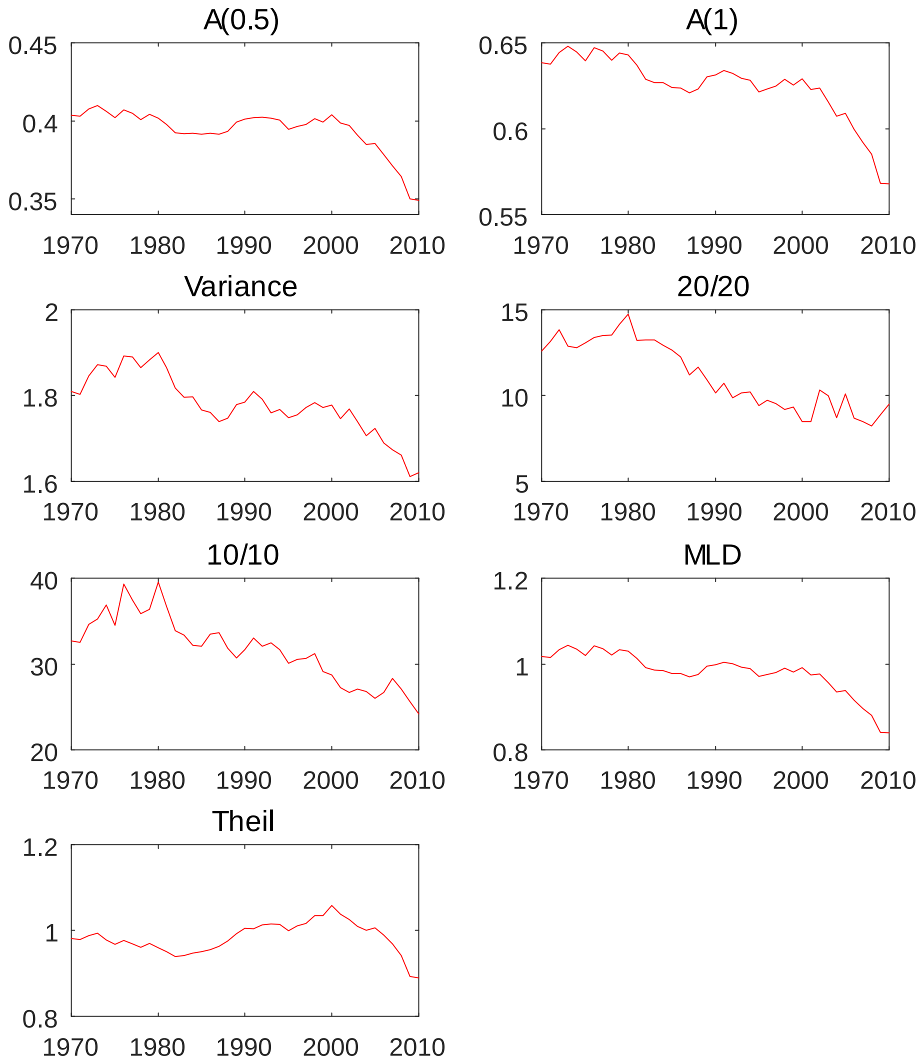
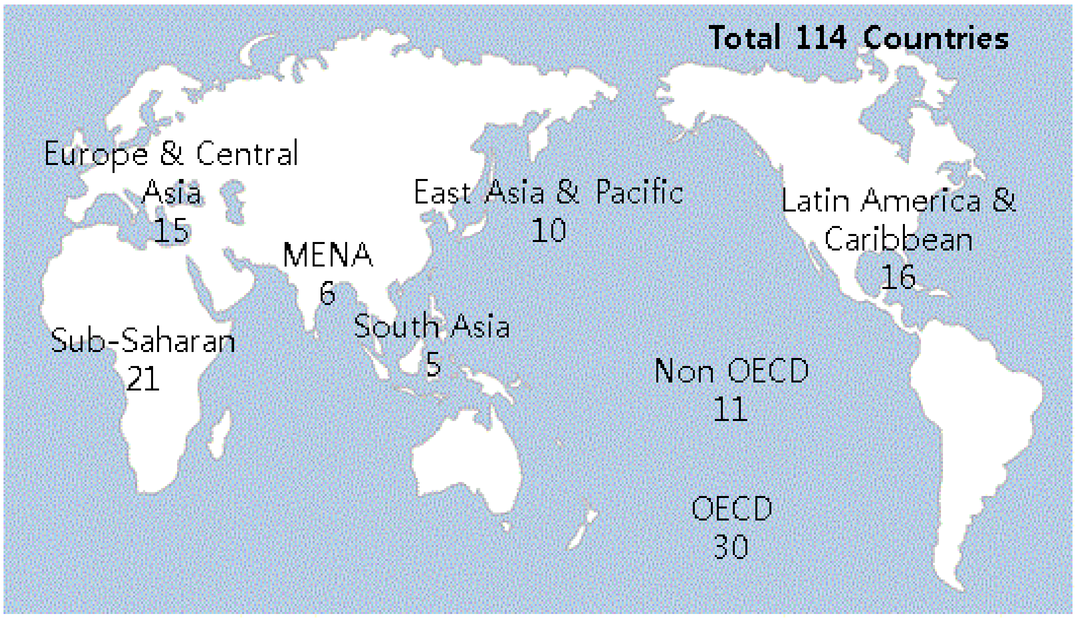


Figure 10. Seven other global inequality indices when top income shares are combined.

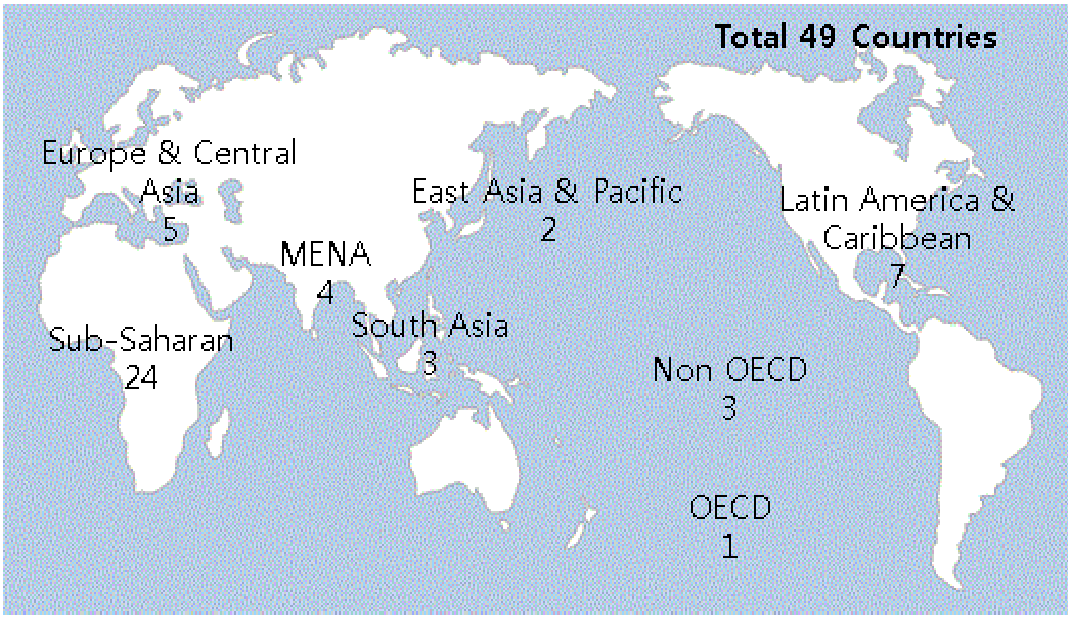
B Countries in Each Group

Countries in Group A:



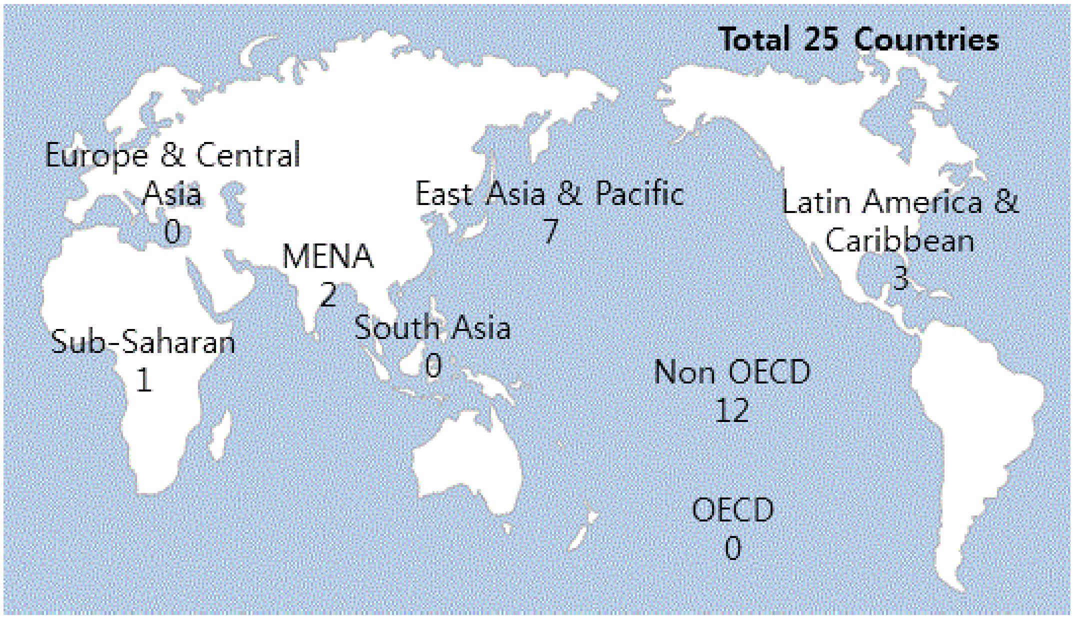
Algeria, Argentina, Armenia, Australia, Austria, Azerbaijan, Bangladesh, Barbados, Belarus, Belgium, Bolivia, Botswana, Brazil, Bulgaria, Burkina Faso, Burundi, Cambodia, Canada, Chile, China, Colombia, Costa Rica, C?e d'Ivoire, Croatia, Cyprus, Czech Republic, Denmark, Dominican Republic, Ecuador, Egypt, El Salvador, Estonia, Ethiopia, Fiji, Finland, France, Georgia, Germany, Ghana, Greece, Guatemala, Honduras, Hong Kong, Hungary, India, Indonesia, Iran, Ireland, Israel, Italy, Jamaica, Japan, Jordan, Kazakhstan, Kenya, Korea, Rep. , Kyrgyz Republic, Lao PDR, Latvia, Lesotho, Lithuania, Luxembourg, Madagascar, Malawi, Malaysia, Mali, Malta, Mauritania, Mexico, Moldova, Mongolia, Montenegro, Morocco, Mozambique, Nepal, Netherlands, New Zealand, Niger, Nigeria, Norway, Pakistan, Panama, Paraguay, Peru, Philippines, Poland, Portugal, Romania, Russia, Rwanda, Senegal, Slovak Republic, Slovenia, South Africa, Spain, Sri Lanka, Swaziland, Sweden, Switzerland, Taiwan, Tajikistan, Thailand, Trinidad and Tobago, Tunisia, Turkey, Uganda, Ukraine, United Kingdom, United States, Uruguay, Uzbekistan, Venezuela, Vietnam, Zambia

Countries in Group B:



Afghanistan, Albania, Angola, Bahamas, Belize, Benin, Bhutan, Bosnia and Herzegovina, Cabo Verde, Cameroon, Central African Republic, Chad, Comoros, Congo, Dem. Rep., Congo, Rep. ,Cuba, Djibouti, Gabon, Gambia, Guinea, Guinea-Bissau, Guyana, Haiti, Iceland, Iraq, Liberia, Macedonia, Maldives, Mauritius, Micronesia, Fed. Sts., Namibia, Nicaragua, Papua New Guinea, Puerto Rico, Sao Tome and Principe, Serbia, Seychelles, Sierra Leone, Singapore, Somalia, St. Lucia, Sudan, Suriname, Syria, Tanzania, Togo, Turkmenistan, Yemen, Zimbabwe

Countries in Group C:



Antigua and Barbuda, Bahrain, Bermuda, Brunei, Dominica, Equatorial Guinea, Eritrea, Grenada, Kiribati, Kuwait, Lebanon, Libya, Macao, Marshall Islands, Oman, Palau, Qatar, Samoa, Saudi Arabia, Solomon Islands, St. Kitts & Nevis, St.Vincent & Grenadines, Tonga, United Arab Emirates, Vanuatu

References

Anand, S. and P. Segal (2008) What do we know about global income inequality?, *Journal of Economic Literature,* 46, 57-94.

Anand, S. and P. Segal (2014) The global distribution of income, *International Development Institute Working Paper*, 2014-01.

Atkinson, A., T. Piketty and E. Saez (2011) Top incomes in the long run of history, *Journal of Economic Literature*, 49, 3-71.

Baltagi, B. (2013) *Econometric Analysis of Panel Data*, 5th ed. Chichester, UK: Wiley.

Bourguignon, F. and C. Morrisson (2002) Inequality among world citizens: 1820-1992, *American Economic Review,* 92, 727-744.

Chotikapanich, D., W.E. Griffiths, D.S.P. Rao and V. Valencia (2012) Global income distributions and inequaltity, 1993 and 2000: Incorporating country–level inequaltity modeled with beta distributions, *Review of Economics and Statistics,* 94, 52-73.

Cowell, F.A. (1995) *Measuring Inequality*, 2nd ed. Harvester Wheatsheaf, Hemel Hempstead.

Cowell, F.A. (2000) Measurement of inequality, *Handbook of income distribution,* 1, 87-166.

Davies, J.B. and A.F. Shorrocks (1989) Optimal grouping of income and wealth data, *Journal of econometrics,* 42, 97-108.

Deininger, K. and L. Squire (1996) A new data set measuring income inequality, *The World Bank Economic Review,* 10, 565-591.

Jones, C. I. (1997) On the evolution of the world income distribution, *The Journal of Economic Perspectives,* 11(No.3), 19-36.

Krause, M. (2014) Parametric Lorenz curves and the modality of the income density function, *Review of Income and Wealth,* 60, 905-929.

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,* 6719.

Liberati, P. (2015) The world distribution of income and its inequality, 1970-2009, *Review of Income and Wealth,* 61, 248-273.

Milanovic, B. (2005). *Worlds apart: Global and international inequality 1950-2000*, Princeton University Press Princeton, NJ.

Milanovic, B. (2012) Global inequality recalculated and updated: The effect of new PPP estimates on global inequality and 2005 estimates, *Journal of Economic Inequality,* 10, 1-18.

Milanovic, B. (2015) Global inequaltity of opportunity: How much of our income is determined by where we live?, *Review of Economics and Statistics,* 97, 452-460.

Minoiu, C. and S.G. Reddy (2014) Kernel density estimation on grouped data: The case of poverty assessment, *Journal of Economic Inequality,* 12, 163-189.

Piketty, T. (2014) *Capital in the 21st Century*, Harvard University Press, Cambridge.

Pinkovskiy, M. and X. Sala-i-Martin (2009) Parametric estimations of the world distribution of income, *NBER Working Paper*, 15433, National Bureau of Economic Research.

Pinkovskiy, M.L. (2013) World welfare is rising: Estimation using nonparametric bounds on welfare measures, *Journal of Public Economics,* 97, 176-195.

Quah, D.T. (1996) Twin peaks: Growth and convergence in models of distribution dynamics, *Economic Journal, 106,* 1045-1055.

Rougoor, W. and C. van Marrewijk (2015) Demography, growth, and global income iInequality, *World Development,* 74, 220-232.

Sala-i-Martin, X. (2002) The disturbing "rise" of global income inequality, *NBER Working Paper*, 8904, National Bureau of Economic Research.

Sala-I-Martin, X. (2006) The world distribution of income: Falling poverty and ... convergence, period, *Quarterly Journal of Economics,* 121, 351-397.

van Zanden, J.L., J. Baten and P. Foldvari (2014) The changing shape of global inequality 1820-2000; exploring a new dataset, *Review of Income and Wealth,* 60, 279-297.

Warner, D., D.S.P. Rao, W.E. Griffiths and D. Chotikapanich (2014) Global inequality; levels and trends, 1993-2005: How sensitive are these to the choice of PPPs and real income measures?, *Review of Income and Wealth,* 60, S281-S304.