IARIW 2021
Monday 23 – Friday 27 August
Statistisk sentralbyrå
Statistics Norway

Global Income Inequality: Concepts, Estimates and a First Ballpark for its Uncertainty

Michail Moatsos

(Utrecht University)

Paper prepared for the 36th IARIW Virtual General Conference
August 23-27, 2021
Theme 9: Comparing Income and Price Levels and International Income Inequality

Global Income Inequality: concepts, estimates and a first ballpark for its uncertainty*

Michail Moatsos†
Utrecht University, Department of History and Art History

VERY PRELIMINARY VERSION;
DO NOT CITE OR DISTRIBUTE FURTHER

New global income inequality concepts are introduced to complement those by Milanovic (2006) and Anand and Segal (2008) that can help painting a more complete picture for the evolution of global income inequality. Estimates of global income inequality since 1820 are provided along with their respective confidence intervals using Monte-Carlo experiments that account for a host of underlying uncertainties in data and definitions. Preliminary stylized estimations indicate an error term of 2.5 Gini points in 1820, dropping to 0.9 Gini points in 2016.

1 Introduction

Twenty years have passed since Atkinson (1997) reinvigorated the interest on the study of income inequality, when he brought "income distribution in from the cold". The more recent work of Piketty (2014) did its best to bring the field at the center of an ongoing global debate. In parallel, the work of Branko Milanovic has been a guiding compass to many in the field throughout the years (see for example Milanovic (2011)), along with other scholars such as James Galbraith[1] and many other contributors behind the very recent World Inequality Report Alvaredo et al. (2018). With respect to global income inequality on the long run, two important articles have been published by Bourguignon and Morrisson (2002) and Zanden van et al. (2013) that trace the evolution of income inequality on global scale for roughly the last 200 years. Both these articles employ Monte Carlo procedures to estimate a confidence interval, but only account the error implied to the global inequality estimates from the "sampling" errors in the country data or estimates on inequality. Also, it appears that

*I wish to thank Angus Deaton, Leandro Prados de la Escosura and Guido Alfani for kindly sharing their data, and James Galbraith for discussing his dataset with me. I also like to thank my promotor Jan Luiten van Zanden and my daily supervisor Auke Rijpma for their guidance and useful comments on my work. All analysis has been conducted with R open source statistical computing software (R Core Team, 2018). All remaining errors are my own.

†PhD candidate — Universiteit Utrecht — Social and Economic History Group — Drift 6, 3512 BS Utrecht — Room 2.07 — email: m.moatsos@uu.nl

[1]See for example Galbraith and Halbach (2016)

this has not been done conservatively at least with respect to (Bourguignon and Morrisson, 2002), as pointed out by (Anand and Segal, 2014, p.19).[2] In comparison, and for the more recent years, Milanovic (2012) estimates much higher uncertainty levels using a different set of assumptions.

The various calculation problems referred to in the literature[3] are rather complex in nature, and cover a wide range regarding both data and methodology (see for example Anand and Segal (2008) for an extensive overview). Prima facie, the occasional lack of data that most of the countries for recent years suffer from is certainly a headache, but perhaps it is not the most important among the problems. Rather, the limited coverage of certain territories for large time periods is perhaps a more fundamental concern. The importance of this aspect increases when the approach is a historical one, such as the one considered here, where for the periods further back in time we need to rely heavily on interpolations and synthetic estimates (Bourguignon and Morrisson, 2002; Zanden van et al., 2013). The uncertainty implied by all those indirect methods, as well of what uncertainty comes out of the direct methods, needs to be considered in an appropriate and extensive manner. Further, the purchasing power parity (PPP[4]) exchange rates used to convert all currencies into a common numaire (international PPP dollars) is an issue with perhaps wider implications in the details of the various calculations, as their estimates are far from being errorless (Deaton and Dupriez, 2011; Deaton, 2012; Rao et al., 2015); while historically there several options on how to incorporate the available PPP exchange rates from various years (Bolt et al., 2018).[5]

Finally, despite the wide spread use of the Gini coefficient to enumerate inequality, and in turn describe its trend, there is no well founded consensus about which income inequality indicator should be used to actually measure inequality of a given distribution. Thus, uncertainty in the proper mathematical specification with which to calculate inequality should also be incorporated in quantifying the total uncertainty of inequality trends and levels. Such an approach should be considering not only relative inequality indexes, but also indexes of absolute inequality (Goda and Torres Garcia, 2016; Niño-Zarazúa et al., 2016; Amiel and Cowell, 1992). As pointed out in (Ravallion, 2018b, p.12) "the stylized fact that overall inequality has been falling since around 1990 is not robust" depending on ones preferences in how we measure inequality.[6]

To contribute within the aforementioned gaps in the literature three distinct contributions are made here: (a) A set of new global income inequality concepts are introduced to complement those by Milanovic (2006) and Anand and Segal (2008) that help painting a more complete picture for global inequality evolution. (b) A Monte-Carlo pseudo-experiment, that accounts for a set of uncertainty (including PPP and survey sampling errors), is used to estimate the confidence intervals behind the various global inequality estimates. (c) an index of indices is compiled to reflect both uncertainties in data, as well as for uncertainty into which mathematical formulation is the most appropriate to express income inequalities.

[2] "In our view the other sources of error discussed above would imply much larger confidence intervals than these standard errors [in Bourguignon and Morrisson (2002)] suggest", Anand and Segal, (ibid).

[3] And here below in the methodological and data sections.

[4] See the methodological section for an explanation on the nature and purpose of PPP exchange rates.

[5] See also Aten and Heston (2010).

[6] Technically this is very possible since the Lorenz curve of the 1990 global income distribution is not entirely dominated by the Lorenz curve of the 2010 global income distribution. See Ravallion (2018b) for a discussion.

2 Methodology

2.1 Global Inequality Concepts in the literature

As Anand and Segal (2008, p.59) have long recognized "[t]he appropriate definition of 'global inequality' depends on the purpose at hand." Milanovic (2006) has distinguished three concepts of inequality: (1) between countries inequality where each country counts as one unit and the income used is that of per capita income; (2) as previously but each country is weighted by the number of its inhabitants; and (3) global inequality between individuals where each individual is being assigned an income value according to the income distribution in her own country. Concept 1 captures between countries inequality to the extend that we care about the world as if countries where the only entities that matter. This concept is also the one that economics mostly use in the "sigma convergence" literature.[7] However, as Anand and Segal (2014, p.9) show, a reduction of concept 1 does not imply a reduction or an increase for the other two inequality concepts. Based on this observation they conclude that "global inequality tout court is an under-specified concept, and estimates of different definitions of global inequality can move in different directions" (Anand and Segal, ibid).

Concept 2 global inequality re-introduces the individuals, but ignores the inequality component stemming from within country income distribution. Again it is only state entities that matter for inequality, with the role for the individual being reduced to a weight for the countries around the world.

Concept 3 corrects for this by assigning the pivotal role to the individual. In this concept countries do not matter at all, and it is only the distribution of individual income on a global scale that is measured.[8] Anand and Segal (2014, p.3) broadly point to the Wilkinson Hypothesis when arguing for such an approach in global inequality. As detailed in Wilkinson and Pickett (2009) there are important social consequences in living in a society that experiences high inequality levels, in terms for example population health or crime rates among many others. In so far the within country inequality component contributes to global inequality, this type of inequality (concept 3) would or should in turn concern us (Anand and Segal, ibid). Further, to the extend that the global individual income distribution is relevant for a global relative poverty deprivation, such as the type of poverty measured by Eurostat, then concept 3 income inequality becomes relevant from an additional perspective.

Very reasonably Anand and Segal (2008, p.60) add on this list of concepts a more fundamental concept 0, which is defined as between countries inequality this time using a per country total income. Here the individuals are completely ignored. Anand and Segal (2008) defend their suggestion by arguing that this better reflects the distribution of international power among countries, implying that the economic size of each country better reflects its negotiating position compared to its per

[7] Sigma convergence means that the dispersion of per capita national income between countries is reduced over time.

[8] Rather recently Ravallion (2017) has proposed global inequality approaches which relax this strict assumption and introduce the importance of higher national income to the individual. Such approaches are highly valuable, however they also fall outside the scope of this paper.

capita economic valuation (as captured by concept 1). Anand and Segal (2014, p.4) take a step further in arguing that:

> "[T]he institutional arrangements which exist in the global economy – the international rules and organizations that govern the flows of goods, capital and labour between countries – are sufficient to generate a normative concern for inequality among individuals in the world, even if they fall short of a global government. These international arrangements are largely determined by rich countries and tend to benefit citizens in rich countries at the expense of citizens in poor countries. Rich countries may therefore bear some responsibility for global inequality."

The concept of inequality implied by the first part of their argument is clearly concept 0, the later part of the argument though may be–to a greater or a lesser extend–of any of the concepts defined so far. In the Idea of Justice (Sen, 2009, p.409-10) makes a similar argument in the following words:

> "The distribution of the benefits of global relations depends not only on domestic policies, but also on a variety of international social arrangements, including trade agreements, patent laws, global health initiatives, international educational provisions, facilities for technological dissemination, ecological and environmental restraint, treatment of accumulated debts (often incurred by irresponsible military rulers of the past), and the restraining of conflicts and local wars."[9]

Again indicating that rich countries, and therefore concept 0 global inequality, has an important role to play for the wellbeing of individuals across the globe.

2.2 Relative and absolute inequality indices

Inequality indices can be categorized into three large families: relative, absolute and intermediate ones. Relative inequality indices have the characteristic that they remain unchanged under a common percentage change in all the incomes of the distribution. Absolute inequality indices remain unchanged when the same amount is added to all the incomes of the distribution. Intermediate or centrist indices also exist where only part of a relative increase would change the value of the index and likewise only a part of an absolute increase will again change the index' value. The following paragraphs concisely elaborate on these matters.

2.2.1 Relative indices

Typically the Gini index is being used in reporting inequality. A simple definition of this index is provided by (Bourguignon, 2015, p.18) as: "half the average absolute difference between two individuals chosen at random in the population, in relation to the average standard of living of

[9] As cited in Anand and Segal (2014, p.4).

the population as a whole." The usual way of explaining this index is via its visual relation to the Lorenz curve of absolute equality.[10] The Gini index can take values between 0 and 1, with 0 being the value of absolute equality, where every person gets the same income, and 1 of absolute inequality, where one person gets all the income. An important point with respect to the Gini index is that it considers the entire distribution, and it is therefore a summary inequality index. One of the caveats in using the Gini coefficient is that it is not decomposable.[11] Other members of the family of summary inequality indices that are used often are the Theil (T and L) indices and the Atkinson indices.[12] In particular, the Atkinson index allows the researcher to specify the desired level of inequality aversion which changes the sensitivity of the index to inequality at different levels of the distribution.[13] Similarly the Theil L index (also known as Mean Log Deviation or MLD) is more sensitive in changes in the low income groups while Theil T is more sensitive to changes in the upper part of the income distribution. The Gini index on the other hand is more sensitive in changes in the middle of the income distribution.

For these reasons different indices may evaluate a particular income distribution as more or less unequal compared to another distribution, depending on where the particular index is focusing the most. These considerations should be kept in mind when evaluating the evolution in inequality (global or local) as rightfully pointed out recently by Ravallion (2018b). It is only when the Lorenz curve of one distribution first-order dominates another that we know in advance that the former would be found less unequal to the other regardless of the (relative) inequality index we choose to use. Figure 1 offers an example of two income distributions where the one drawn above lies indeed above the other in all points except 0 and 1; this first-order dominance that represents the only clear cut case where *all* relative inequality indices agree on the valuation of which distribution is more unequal.

During the last decade or so researchers behind the World Inequality Report (Alvaredo et al., 2018) show strong preference in the use of income shares instead of summary indices. Income shares denote the percentage of income that goes to a particular population group (say the 10% of the population with the highest incomes or the other 10% with the lowest incomes, etc). A derivative of this family of inequality indices are the income ratios which are the ratios of income shares of particular income groups (say 10th income decile over the income of the poorest 1st income decile). Such indices are not summary indices as they focus only on specific parts of the distribution. A strong case can be made in favor of these indices along the lines of their transparency and clarity in what they measure and how. At the same time they typically ignore most of the information

[10]Explaining the details of Gini or any other inequality index used here goes beyond the purpose of this paper, and the interested reader is deferred to the specialized literature. An example of Lorenz curve is shown in figure 1, and the figure is discussed in the following paragraph.

[11]Decomposability is a highly desired property of inequaity measures that allow one to express total inequality as a function of inequality within sub-groups and inequality between those sub-groups within a society. The Gini index fails to fulfill this requirement.

[12]Ravallion (2016) offers a very informative introduction to these concepts, and so does Cowell (2009).

[13]For example, Ravallion (2018a) finds that when one uses an Atkinson parameter that represents strong aversion to inequality (e.g. values of 5 or 6), then global inequality increased over the period 1988-2008. For the typical values of the Atkinson's parameter (e.g. less than 2) global inequality has declined in the same period according to the same calculations.

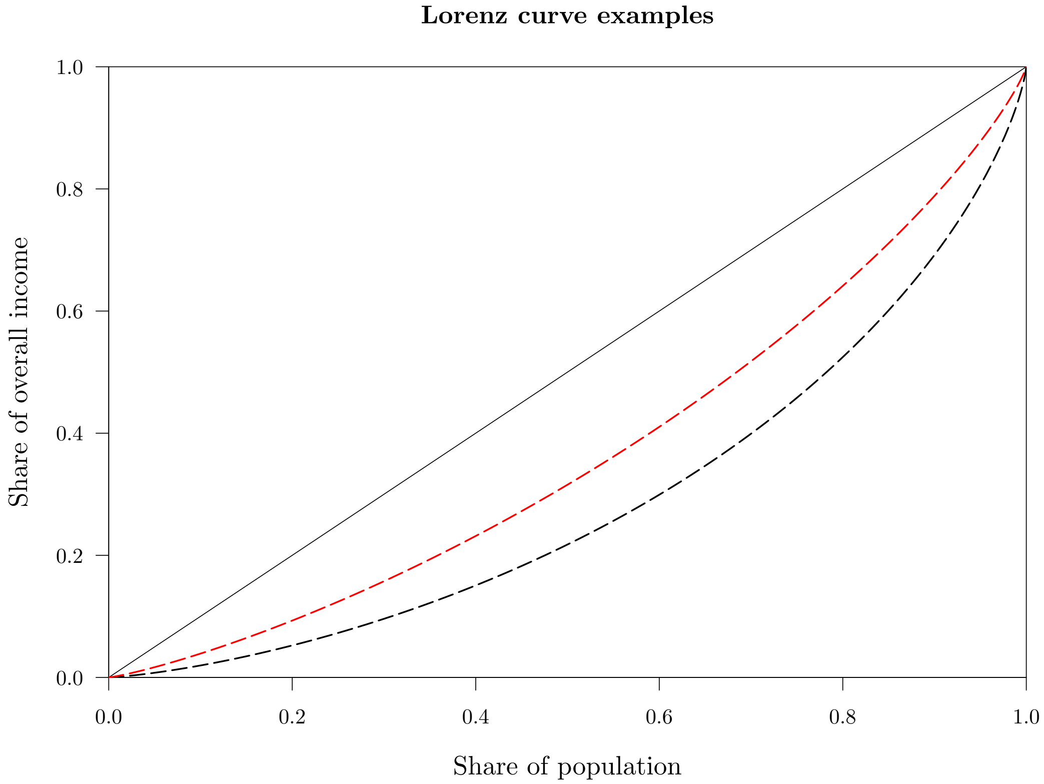

Figure 1: One of the two Lorenz curves shown (red) first-order dominates the other (black), this is a case where all relative inequality indices will agree that the red line represents an income distribution that is less unequal than the black one. The characteristic is that that two lines never intersect between the numbers 0 and 1.

available in the form of the income distribution.

2.2.2 Absolute indices

As Niño-Zarazúa et al. (2016, p.5) put it: "[o]ne especially important normative judgement regards the manner in which inequality is deemed to change as economies grow and the size of the pie to be divided increases". And as noted also by (Ravallion, 2018b, p.17), the empirical literature has "almost solely" focused on relative inequality as captured by indices discussed above. The main selling point for considering inequality as a relative phenomenon is the so called "scale independence axiom". This axiom exactly describes the main property of relative inequality indices, that is their fixity under conditions of equiproportional growth. However, this is simply an axiom and there is no strict law in adhering to this axiom as the only way in looking into inequality changes and valuations. Which class of inequality indices to use is a key normative judgment. As a matter of fact there exists an alternative axiom, called the "translation invariance" axiom, according to which an inequality index should remain unchanged when one adds a constant. Inequality indices that fulfill this axiom are called absolute.

Two important members of this class of inequality indices are the Standard Deviation (SD) and the Absolute Gini. The Standard Deviation is a common concept in statistics representing the dispersion of a set of data points.[14] While the Absolute Gini is the simple product of average income multiplied with the Gini index.

It is remarkable how recent the interest in absolute inequality indices has been, especially in global inequality research. According to Anand and Segal (2014, p.35) “the first thorough investigation of measures of absolute global inequality is by Atkinson and Brandolini (2010).” Absolute inequality indices are very clear in the direction of global inequality, as it is found to be rising at different speeds according to the exact absolute index one is using. To a certain point the increase in global inequality according to absolute inequality criteria is almost mechanical, and to prevent an increase in absolute inequality, then according to (Anand and Segal, 2014, p.36) “relative inequality has to decrease at a faster rate than the rise in mean incomes”; arguably a very unlikely development.

There is experimental evidence summarized in Ravallion (2016), that supports the view that the opinions of university students about whether inequality should be evaluated using relative or absolute inequality indices is more or less split evenly between the two options. This can also be seen as a support for more intermediate approaches (see next sub-section). As a final note on absolute inequality considerations, Ravallion (2018b, p.2) warns us about an implicit trade-off that exists between absolute inequality and absolute poverty, where a reduction in the former would typically bring about an increase in the later.[15]

2.2.3 Intermediate measures

A very important property among inequality measures is unit consistency. This property requires that the ranking of income distributions remains the same regardless of the denomination unit used for income. Think for instance if one would use 2017 fixed dollars to get different results in terms of which of two income distributions is more unequal, compared to using say 2010 fixed dollars instead. Such inconsistency would be highly problematic. Several well-know absolute or intermediate inequality measures fail to meet this property (Niño-Zarazúa et al., 2016, p.6).

The Krtscha measure stands almost in the middle between the absolute and the relative inequality measures, and can be simply expressed as the product of the coefficient of variation and standard deviation.[16] By its nature it represents a split compromise between relative and absolute indices. Very desirably the Krtscha measure satisfies both unit-consistency and subgroup decomposability

[14] The exact formula is: $SD = \sqrt{\frac{\sum_{i=1}^{N}(x_i - \bar{x})^2}{N-1}}$, with obvious notation.

[15] This would hold for absolute poverty using methods similar to the dollar-a-day method promoted by the World Bank, and several scholars. It is not necassarily the case when absolute poverty is measured using other methods such as the cost of basic needs approach.

[16] The coefficient of variation is another famous relative inequality index, and it is defined as the square root of the income variance divided by mean income. It is therefore a standardized form of variance, and variance is given by the formula: $V = \frac{1}{n}\sum_{i=1}^{N}(x_i - \bar{x})^2$. The standard deviation is the absolute inequality index defined in the previous section.

"unlike other known intermediate indices" (Subramanian and Jayaraj, 2015, p.7); see also Zheng (2007).

2.3 Placing new Global Inequality Concepts among the existing ones

In estimating income inequality around the world three building blocks are essential: (a) the concept whose unequal distribution we seek to estimate (typically a form of income, although consumption or wealth are also used), (b) the treatment of observation units (typically countries, which can be treated as of equal populations or weighted by their actual population size), and (c) disparities of the concept selected in (a) between, within, or between and within the observation units as selected in block (b).

Using these building blocks the four fundamental concepts of global inequality as discussed above are numbered 0 through 3 in the list below. Now these concepts alternate between total and per capita income, accounting for population size in each country or not, and accounting for differences in income within each country or not. Along the building blocks (a)-(c), as discussed in the previous paragraph, there are a number of key possibilities left outside the conceptual framework of global income inequality concepts 0 through 3. For example, total or per capita income can also be discounted for within country inequality as advocated by Atkinson (1970) or Sen (1976)[17]. The use of such an equivalized income would still result in a per capita valuation, however the within country inequality component is not totally ignored as in concepts 0, 1, and 2 do. In the list below the variations on concepts 0, 1 and 2 that use this equivalized income are marked by ".e".

Five new concept variations that fall along the ranks of building blocks (a)-(c) are introduced below and provide for some missing links among the four already available concepts:

- Concept 0: total income per country, with each country counting as one unit (Anand and Segal, 2008, p.59).
- Concept 0.e: total income discounted for inequality (using Atkinson's or Sen's formulations) per country, with each country counting as one unit. Such a concept represents an intermediate consideration of global inequality between economic and political power (concept 0) and other inequality concepts that account for inequality within countries.
- Concept 1: per capita income of each country, with each country counting as one unit (therefore differences in population size are ignored) (Milanovic, 2006, p.2), also known as "sigma convergence".
- Concept 1.e: per capita income discounted for inequality per country, with each country counting as one unit, and can be seen as the sigma convergence of a welfare equivalent income. Such a concept represents the dispersion on a global scale of the per country required income to achieve the same welfare level among individuals from the same country.

[17]Referring to the contributions of Sen and Atkinson in equivalent or real incomes, which is the equivalent income distributed equally among individuals that would accrue the same amount of individual welfare as the present unequally distributed income (see the "Methodological Details" sub-section for more details).

- Concept 2: per capita income of each country, with each country weighted by the number of inhabitants (assuming inequality within countries being zero; also known as between countries inequality) (Milanovic, 2006, p.2)[18]

- Concept 2.max: per capita income of each country, with each country weighted by the number of inhabitants assuming inequality within countries being one, which results in the attainable concept 2 under conditions of maximum inequality within countries given the prevailing between countries income differences. It can be thought of as the complementary of concept 2, since that concept assumes zero inequality within countries.

- Concept 2.e: per capita income discounted for inequality per country with each country weighted with the number of inhabitants. Similar to 1.e, but this time the differences in population size between countries are not ignored.

- Concept 2.5: per capita income of each country, with each country weighted by the number of inhabitants, assuming inequality within countries being equal to a fixed "resting" number. Arguably, since concept 2 ignores within country inequality or equivalently assumes a Gini of the within country income distribution equal to 0, using a value of 0.5 Gini is half way between the current and the max concept 2 definitions. This concept is capable to produce an upward biased concept 3 estimator.[19]

- Concept 3: income inequality of the distribution of individuals according to their personal income ignoring their country of residence. This concept fully accounts for the within country inequality, and it is mostly known as global interpersonal inequality and typically referred to simply as global inequality.

The diagram in figure 2 places these concepts along the lines of the building blocks introduced above. From left to right they move from the global inter-personal inequality of concept 3, down to the global inequality according to the concept that is the most distant to the individual (concept 0). The specificity axis at the bottom of the diagram follows this logic and it traces those inequality concepts from the most specific (or narrow) concepts to the individual (on the left side) down to the most distant ones (or wider in terms of specificity to the individual) concepts. Do note that the consideration of within inequality in concepts marked with ".e" is only partial, since it takes place through the equivalent, or real, income approach as explained above. The only concept that (in theory) fully accounts for within countries inequality is concept 3.[20]

Although there are clear links among the various concepts of inequality, there is no mechanical way of how one concept can be used to pin-point the level of other concepts. However, there are some

[18]As pointed out in Anand and Segal (2008, p.60) "some studies (e.g., Firebaugh (1999, 2003)) have used it as a downward-biased estimator for concept 3 inequality."

[19]One can think of scenarios where this is not the case, and concept 3 is actually more unequal than 2.5, but for the vast majority of actual cases this concept will indeed work as an upward biased estimator for concept 3.

[20]The necessity for the qualification here draws upon the empirical problems with missing information, incomplete surveys, and various other problems that does not allow one to fully account for inequality within a country, let alone to account for inequality within each country around the world and for each year. Data section provides more information on these issues.

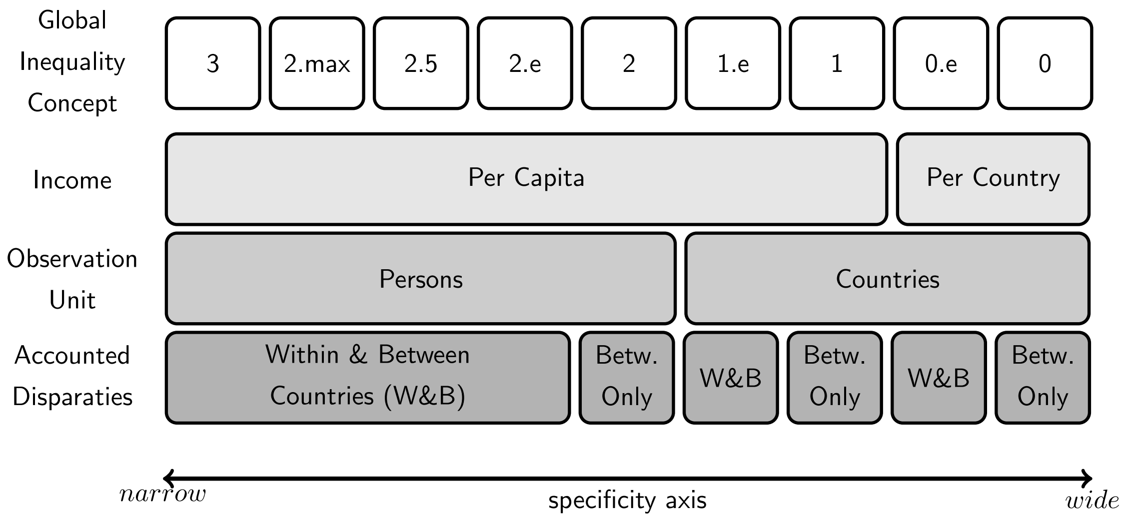

Figure 2: Table showing the overlaps among various global inequality concepts in terms of the essential building blocks.

broad empirical relations among the concepts. For example, absolute inequality in terms of concept 2.e would typically be lower than absolute inequality in terms of concept 2, and the same applies among concepts 1.e and 1, and 0.e and 0. This is expected as the real income that is discounted for inequality would result on average in lower absolute differences among observation units. Another example is that concepts 2 and 2.5 would provide respectively a downward and an upward biased estimator for concept 3.[21] Similarly between concepts 1 and 0, and 1 and 2 most of the variation can be expected to come from the relative population sizes of each country, therefore again no mechanical relationship can be identified.

Anand and Segal (2014, p.8) discuss China as an instructive example of how a particular country can push an inequality concept in different directions. On the one hand, China has a total income which is above average it will introduce a disequalizing force in terms of concept 0 inequality. On the other hand though, in terms of GDP per capita China is below average and growing faster than the world average, which implies that in terms of both concepts 1 and 2 China is an equalizing force. In turn this also implies that China is an equalizing force according to concept 3 as well.

With respect to the observation units specified in figure 2, it needs to be noted that when we talk about a "person" as an observation unit the actual meaning of the word may differ depending on the treatment this variable gets during the household survey. Typically it would mean an adult equivalent, although in other cases it can also represent an actual person. An adult equivalent is used when we wish to account for the fact that since data are gathered on a household level we cannot simply rank all households according to the size of their income and ignore the fact that households differ in size of members as well. If we do not account for this variation we would say that a household of 2 people with 20000 euro income is poorer than a household with 5 members and

[21] Concept 2.max is an extreme case and it is not clear in advance what the interplay between the between and the extreme within component will bring about to the level of inequality compared to concept 3.

21000 euro income. Very few would argue that this is proper (since our ultimate point of reference is the person and not the house; hold). Instead the method used to account for the household size are the equivalence scales, and there exist a number of methods that can be applied. Two typically used are the square root of the household size or simply to assign a weight to each member: 1 for the head, 0.8 for the next adult, 0.5 for each older child and 0.3 for each young child for example. The exact method used depends on the national statistical office or the international organization that homogenizes the data (such as Luxemburg Income Study (LIS), see data section for more details on available sources).

What the above discussion (and almost the entire literature on global and country inequality) ignores is the allocation of resources within households. As mentioned above the household survey data are collected on the level of the household and we are not able to discern the actual allocation of income or resources among the members within the household. The use of equivalence scales only solves the most obvious of the problems and in an ad-hoc manner. It does not actually inform us about how income is dispersed among the household members in reality. Therefore an argument can be made that the estimate of inequality among person from household data would only provide a downward biased estimation. The exact size of this bias is unknown and it would vary depending on the time and place and the equivalence scale used (theoretically it could also be an upward bias as well).[22] This is by its nature an empirical problem, but it must be kept in mind when interpreting the levels and trends of the various inequality concepts.

A few additional remarks are in order with respect to concept 2.5, which is of interest in particular for the following reasons:

1. It can be used inversely after calculating concept 3 inequality to come up with the effective mean within country inequality that produces the same global Gini as concept 3 has. This function in the domain of inequality is similar to the function of equivalent income in the domain of national income discussed above.

2. It can be used as a substitute of concept 3 inequality for those cases that we do not have inequality data for a particular country. This would result in a hybrid concept, where countries with available data are treated within concept 3, while the remaining countries get a fixed inequality value–around the average within country inequality index–a la concept 2.5. This hybrid would reflect concept 3 global inequality better than the concept 3 inequality estimating with missing countries when inequality information does not exist for those countries.

3. It is also a sort of a global inequality possibility frontier[23], as it could take the value of concept 2 inequality (where within country inequality has a Gini of 0), and also take any other fixed

[22] This important issue remains unresolved to this day. See for example the World Bank's cover note on the Report of the Commission on Global Poverty (also known as the Atkinson Commission), particularly last paragraph on page 3.

[23] Inequality possibility frontier (IPF) introduced by Milanovic et al. (2007) represents the maximum attainable level on inequality at different stages of economic development. The term here is used for a slightly different context as the one in the original paper by Milanovic et al.. Here in particular it only reflects the possible global inequality levels for different fixed within country inequality values.

Gini value up to 1 (when it becomes 2.max), marking a trajectory with all possible inequality values for the given income per capita differences among countries as a function of different fixed within country inequality levels.

Figure 3 demonstrates these points. The black line is concept 2.5 for various levels of within country inequality (the particular values used are indicated by the x-axis, and at the most left the red dot marks the value of concept 2 global inequality at a Gini of 0.53 on the y-axis). The dotted red line in the middle of the figure is concept 3 inequality where the countries without within inequality data are treated with the concept 2.5, which means they take a range of within country inequality values.[24] Overall the figure shows how when the various within country inequality levels get fixed by known data the dotted red line occurs that intersects with the Concept 2.5 curvilinear trajectory. This concept visually gives an understanding of the underlying uncertainties introduced by not having inequality data for all countries. For the inequality data in 2000 the uncertainty band is within quite reasonable limits 0.66-0.69, as the available data account for most population.[25]

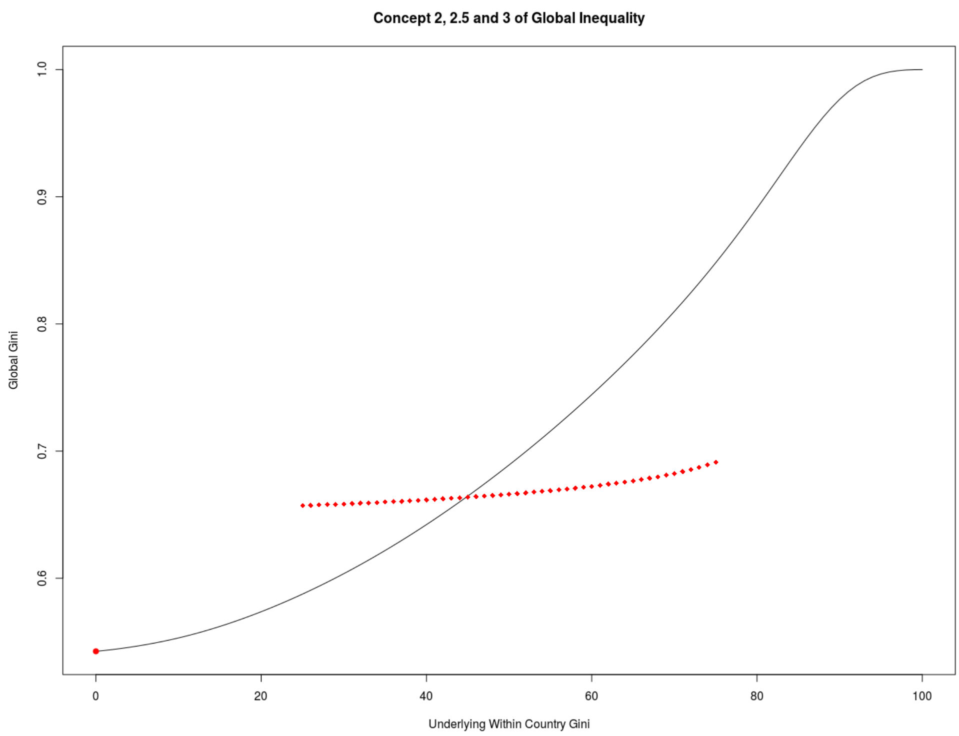

Figure 3: Concepts 2, 2.5 and 3 of Global Inequality for the year 2000. For income GDP data from Maddison (2010) are used, and the within country inequality data are drawn from Zanden van et al. (2013). See data section for details on these sources.

[24]The minimum and maximum available Gini values in the dataset have been used in this case.

[25]Do note that even with this calculation one does not account for 100% of world population but around 95%. Another limitation is that when the within country inequality is not know then according to concept 2.5 for all these countries it will take the same value. The Monte Carlo and the results sections below provide a more complete method to properly estimate confidence intervals.

The intersection point of Concept 3 and Concept 2.5 gives the effective within country inequality level which is around a Gini of 0.45. When this Gini value is used with the 2.5 concept then the resulting global inequality is the same as with concept 3 (which in this case has a Gini of 0.66).[26] Overall concept 2.5 transcends other definitions and helps in providing a basic understanding of global inequality and the empirical limitations at a high level (the section on Monte-Carlo simulations below provides more details in how the empirical limitations can better be captured at a higher level of detail).

Finally it should be noted that the list of concepts presented above includes no 3.e concept. Such a concept would only attenuate the consideration of each individual's income across the globe, and effectively produce a concept 2.e-like instead. However, there is room for rather an inverse concept, or e3 if you will. In this case concept 3 inequality can be used to estimate the equivalized or real income which would have the same welfare equivalence across the world if it was evenly distributed.[27]

2.4 Fundamental concerns of methodological consistency

The focus in the empirical part of this paper is on the estimation of confidence intervals for all the concepts and inequality measures presented above. This subsection discusses a few of the many methodological details that play a role in the final estimates, and provides anchors to literature that discusses other details more extensively. Foremost are discussed: the type of income to be used; the exchange rates to be used to convert all incomes to one common currency; and finally how one can use available inequality information to re-construct the entire distribution. A couple of other issues are also discussed at the end of this section.

2.4.1 Income variable

Unfortunately there is a discrepancy between the total household income as captured by the Household Surveys (hereafter also HHS[28]) and its most directly corresponding income variable in the National Account Statistics which is Household Final Consumption Expenditure (HFCE). What is perhaps more of a problem is that there is also a divergence in the growth rates of incomes calculated based on the one or the other source. Ferreira et al. (2015) recognizing this problem introduce a simple method to account for this discrepancy. Deaton (2005, table 2) with data from 272 surveys of household income internationally, on average the income as captured by HHS is 57% of GDP, but in terms of HFCE it stands at 90% (and 101% when weighted by population). Still, as Anand and Segal (2014, p.2) note "[u]sing HFCE rather than the mean incomes from household surveys changes both the level and trend of estimated global inequality."

[26] This however might be slightly different from the effective within country inequality level of the average Concept 3, which is at a Gini of 0.46. This is the average for concept 3 given that it takes various values depending on which value is assigned to countries without inequality data.

[27] Since this income concept is a direct product of the inequality concepts presented above, in the results section I provide estimates for its level and its evolution at time.

[28] These surveys provide us with the insightful data about within countries distribution of income. Usually conducted by National Statistical Offices (NSO) or other agencies in collaboration with international donors or the World Bank.

In practice, some authors prefer to use the income variable as captured by Household Surveys (Milanovic, 2002b). While others choose to use the shape of the distribution from the Household Surveys and the average income from the National Account Statistics (Dikhanov and Ward, 2001). In addition other authors choose to use the overall economic activity in a country (as it is captured by its gross domestic product, or GDP) in place of the income variable (Zanden van et al., 2013; Sala-i Martin, 2006; Bourguignon and Morrisson, 2002). My approach here is to follow the advice by Anand and Segal (2014) and link the exact choice of income variable to the underlying concept of inequality one tries to estimate.

Concept 0 focuses on the discrepancies in terms of economic, and by implication political, volume among countries. Therefore the GDP variable would be the corresponding variable, since this is exactly what this variable tries to do in economic terms. As concept 0.e is an intermediate concept between the individual and a country, the answer is less obvious. For the purpose of the present investigation I will use again the GDP variable to allow for the difference between concepts 0 and 0.e to be more easily attributable to the effect of the within country inequality component, and not perplexed by the contribution of a change in the underlying income variable.

For concept 1 the center of focus moves further in the direction of the individual, but still it is the country that counts as a unit, and therefore it seems more plausible to work again with income data from the National Account Statistics instead of HHS. Nevertheless, since we have moved one step away from a strict country-level perspective toward a concept more related to the individual, the use of the HFCE variable instead of GDP appears to be more appropriate, since HFCE is also one step away from the country aggregate to the household aggregate. Therefore HFCE is more in line with the spirit of this inequality concept than using the overarching GDP data. The same applies to concept 1.e, also for reasons of comparability as in the case of concepts 0 and 0.e above.

Finally, concept 2 and its variants have moved much closer to the individual, and they make use of the distributional components which typically are sourced from household surveys.[29] Thereafter, it is a matter of simple consistency to use income and distributional data from the same source. Logically, the same applies–and even more so–to concept 3, which due to its sharp focus on the individual disregards national borders as much as the data would allow. For all those income inequality concepts the household surveys should be the preferred data source for estimating global inequality.[30]

2.4.2 Exchange rates

The same argument used above about the specificity of the type of income type used per concept is similarly applicable in the case of the exchange rate one has to use (Anand and Segal, 2014, p.5). For Anand and Segal (ibid), because concept 0 "is most relevant for questions of geopolitics and market access" should be estimated using the usual market exchange rates, as it is the "country's

[29]As discussed in the data section that follows for the historical periods where household sources are unavailable, other type of data like social tables are used instead.

[30]Again additional restrains due to data limitations and availability are applied. See section 3 for details.

weight in international markets – its command over internationally–traded goods and services, or financial assets–that matters". Therefore, for concept 0 estimates, the obvious choice is the market exchange rate.

For the remaining concepts one additional candidate, known as the Purchasing Power Parity (PPP) exchange rates provides a key advantage. Most of the literature on global inequality is using these exchange rates to denominate all currencies in a single one (the so called numeraire currency or country, typically the USA[31]). The advantage of the PPP rates is that they are attempting to account for the fact that the market exchange rates reflect the relative value of currency in the domain of tradeable goods only. The PPP rates allow one to convert all currencies to a single denomination that captures the purchasing power of each currency within each country, and does this by accounting for the non-tradeable sector as well (such as rents and services). Typically accounting for the non-tradeable sector increases the purchasing power of the poorer countries.[32] PPPs are the exchange rate choice made here for all inequality concepts except 0 and 0.e.

The PPP exchange rates are calculated using two main approaches, Gearis-Karmis (GK) and Gini-Eltetö-Köves-Szulc (GEKS). GK estimates impose a price structure to all countries that resemble the price structure of the richest ones, and by implication the incomes of poorer countries become upward biased relative to the richer ones (Anand and Segal (2008, p.72);Ackland et al. (2013)). GEKS on the other hand may be introducing a similar bias due to the rising share of services in consumption, while services are relatively cheaper in low income countries. This will mechanically introduce an increasing bias over time that would inflate the incomes of poorer countries relative to the richer ones (Reddy and Pogge, 2010). However, in practice GEKS PPPs inflate incomes of poor countries less so than GK PPPs.

Deaton and Heston (2010) estimate that the Gini for GDP-based concept 2 global inequality is rather marginally higher when GEKS PPP rates are applied comparet to GK (0.533 vs 0.527). To contextualize those differences do note that the choice of the type of exchange rates can have a considerably higher impact. As Milanovic (2007) shows, the concept 3 global Gini shifts from 0.65 using PPP rates to 0.81 using the official market exchange rates. Thus having a much higher impact than the choice between GK or GEKS type of PPP exchange rates.

A source of concern is that the estimation of the PPP exchange rates works better when the price trends are similar across countries, and worryingly (Dowrick and Akmal, 2005, p.213, fig.5) find that the "price structures became less similar over the period 1980 to 1991" as noted also by Anand and Segal (2008, p.72).

Another problem with the PPP exchange rates is that they are not estimated on a yearly basis.

[31] Deaton and Dupriez (2009) use India as the numeraire country in constructing PPPs for the poor.

[32] According to Ghosh (2018, p.30), this takes place exactly because the labor force in poorer countries receives very low wages or even remain effectively unpaid. Which therefore means that "it is clear that the greater purchasing power of that currency reflects conditions of indigence and low or no remuneration for what could even be the majority of workers". And by implication is one is using PPP rates adds insult to injury by using a highly problematic situation as an "advantage". This is a serious point that needs to be considered in more detail both empirically and methodologically, however such an investigation is beyond the scope of this paper.

Instead they are being produced in quasi-regular intervals, with 2011 being the latest benchmark year for which the have become available and 2005 the previous one. Every time a new round of PPP exchange rates is available all PPP-based statistical data are re-based with the new PPP rates. The advice that Aten and Heston (2001) offer though runs against this widely held practice. They propose a "reconciliation process" where they average out the PPP rates in a process that assigns more wait to the more recent estimates. The Maddison project historical GDP estimates lately use such approaches similar to this one in their estimates Bolt et al. (2018); on the contrary the World Bank data do not do so.

2.4.3 From the Gini index back to the distribution

Unfortunately the actual distributions from the various household surveys are not readily available, and in many cases they are simply lost, and only some summary or partial inequality indices are all that one can use. The problem then is how one can convert an inequality index back to a full distribution. When a Gini index is available and the underlying welfare measure is income (either gross or net), the log-normal hypothesis cannot be rejected, as (Lopez and Servén, 2006, tab.3) have shown.[33] Therefore one cannot show that I would be wrong in assuming that the underlying distribution has a log-normal form, and apply the conversion formulas described in Aitchison and Brown (1957):

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

where σ is the standard deviation of income, Φ is the cumulative normal distribution (with the -1 exponent denoting its inverse form), and G is the Gini index. This formula combined with the formula of the Lorenz curve:

$$L(p) = \Phi(\Phi^{-1}(p) - \sigma) \tag{2}$$

we obtain the entire distribution. One issue with this approach, on which I return in the next section, has to do with the fact of under-reporting by the rich. Research on tax data from the perspective of econophysics (Silva and Yakovenko, 2005) has shown that the higher 1-3% of the richest households are "super-thermal" which means that this section of the distribution follows a Pareto law distribution and not a log-normal. The difference being that one gets higher possible incomes with the former than with the latter. This observation coupled with the information on top incomes by Alvaredo et al. (2013) and Atkinson et al. (2011) indicate that simply using the log-normal assumption on Gini indices from HHS may under-estimate the breadth of possible incomes in the distribution and local and global inequality by implication. Put differently, the incomes in the high end of the HHS data are much lower than the incomes captured by the tax records.

There is one other concern that one needs to address in the process of estimating the complete distribution. This problem stems from the observation that the log-normal distribution derived

[33] On expenditure data though this assumption is rejected at a 5% level.

from "inverting" the Gini as described above, may not be in agreement with additional information about the same distribution (for instance with separate information about income shares or deciles). Shorrocks and Wan (2008) propose that consists of two stages. First stage uses the log-normal assumption to derive a rough distribution. And second stage uses a synthetic sample of observations drawn with various techniques that result in a final distribution which matches as much as possible to the available information on quantile shares from the data they use. For the "vast majority" of final distribution they estimate the difference between the "'true' Gini value and the 'synthetic' estimate was less than 0.003 (approximately 1 per cent)" (Shorrocks and Wan, 2008, p.10).[34] Given the appeal of this procedure, I follow its spirit with a small difference: in the second stage I do not use their algorithm, but I will simple draw random samples from the log-normal distribution until the distribution that minimizes the error below a threshold is found (see the Monte Carlo procedure section for more details).

2.4.4 Two last concerns

There are two final issues to discuss before introducing the data. The first deals with the problem of using income distributions outside of their year of origin. This is common practice in this line of research and it is required because not all countries have distributions for all the years. Otherwise, one will end up having considerably low country coverage per year, and a sample of countries with substantial variability on a year-by-year basis. The second deals with how the literature accounts for persons from different countries in the final distribution.

It is only in the last 30 years that one is able to estimate a form of global income distribution (following a concept like concept 3), since before that period the extensive lack of data over the within country distributions did not allow such an estimate. The scene has changed a lot with major statistical undertakings that will be discussed in more detail in the data section. Regardless, still we are missing more income distribution estimates in terms of country-years than we actually have. Milanovic (2019) estimates the ratio of available distributions (derived from actual household surveys) over the require distributions to cover all countries for all years in the 1948-2017 period to 16-20%.[35] It is therefore clear that per year and more often than not the underlying distributional data will be coming from a different year. Of course this is a major source of error, as discussed in section 4, and it is rather ignored in the global inequality literature when estimating confidence intervals.

The second issue also picks on a point that the literature is consistent about, but perhaps in a way that leaves certain room for improvement. In the course of estimating the global distribution of income, which is naturally required before one gets the estimate of its inequality by any measure, the distributional data from all countries are converted typically into 100 percentiles or 1000 permille,

[34] The authors also use other initial functional forms for estimating the rough distribution in their first stage. Those include lognormal, General Quadratic, Beta, Generalized Beta and Singh-Maddala distributions, but the log-normal distribution appears to have less difference with compared with the actual data (Shorrocks and Wan, 2008, tab.2).

[35] For the period 1820-2017 we investigate here the ratio is a priori expected to be lower, although I use secondary and synthetic sources as well. More on this in the data section that follows.

and then overlaid to form the overall income distribution, e.g. Milanovic (2007).[36] This technique would work perfectly in producing a concept 3 type of global inequality measure if all countries had the same population. In that case each person from each country would carry the same weight in the final distribution. However, countries vary substantially in terms of their population. Therefore, if we split distributions into percentiles and then overlay the various country level distributions we would be assigning very different weights to individuals from different countries. The solution to this problem is to estimate the income per population unit of say 100 or 1000 persons, in order to have all persons assigned the same weight in the final distribution considered. Such is the procedure followed here.

3 Data

This section briefly browses through the data that provide the groundwork for the estimates found in the results section. On the outset there are a number of different elements that need to be combined to estimate global inequality in any meaningful sense. First are income data on a country level and money exchange rates used to convert all currencies to a common denomination, second within country income distributions, and finally population data.

Given that, for conciseness and tractability, the results presented below focus on concepts 2 and 3, along with their variants, only data related to the calculations of these concepts are discussed here. Thus income data from household surveys are utilized. The problem with this choice is that not all income inequality information is given with an average of the income captured in the source. In those cases I use the national account data discounted by the ratio between those household surveys with available averages, and their corresponding national account statistic (preference is given to the use of HFCE from the national accounts as noted above (from the World Bank data provided in 2011 PPP rates), and in its absence GDP data from Maddison).

With respect to the distributional data, the first inequality dataset with a global reach was the seminal Deininger and Squire (1996) dataset that combined almost all available inequality data point available at the time.[37] Later it became rather obsolete as it was surpassed by the World Income Inequality Database (UNU-WIDER, 2018). This source contains Gini values, in many cases accompanied with quantile and decile shares, as well as average or median income values from the underlying sources, which vary between national statistical authorities and research studies. In addition, data from OECD (2019), CEDLAS (2019) focusing on Latin America, EU-SILC EUROSTAT for European countries, and the Luxembourg Income Study (LIS, 2018) are used, according to the rules discussed in the Monte Carlo method (see section 4).

According to United Nations (2018) "note on Income Inequality Data" the LIS data are "the only source that provides inequality statistics calculated using a uniform set of assumptions and definitions on the basis of thoroughly harmonized microdata to maximize comparability". This is

[36] An equivalent way to go about it is to estimate the number of people in predefined income brackets, e.g. Sala-i Martin (2006); Moatsos et al. (2014). The point made here applies the same in this case as well.

[37] For its latest form see Deininger and Squire (2013).

true for data coming directly from household data. There are also at least three important sources of synthetic data, two for the relatively recent period, and one for the more historical one. The historical one comes from the work of Zanden van et al. (2013), that use a couple of econometric approaches, namely the Williamson hypothesis[38] and the link between inequality in heights and income inequality, to estimate income inequality in historical context. In turn, they provide a more complete dataset in terms of coverage for use in historical research, compared to Bourguignon and Morrisson (2002).

The first of the two contemporary ones is the SWIID database by Solt (2016) which is the synthetic version of UNU-WIDER (2018). SWIID is synthetic in the sense that Solt is using econometric techniques to estimate the probable indices for all countries in WIID and for all years according to the information contained in WIID. Although this approach may not please all audiences, Solt is careful in providing his estimates with the appropriate error term, which combined with the underlying income concept that he uses (gross income), is very useful in our approach as it supplements the more historical dataset from Zanden van et al. (2013) with the same income concept.

The second contemporary synthetic source on inequality is the Estimated Household Income Inequality Data Set (EHII) produced by the University of Texas Inequality Project, Galbraith and Halbach (2016). It is built upon the econometric relationship between United Nations Industrial Development Organization (UNIDO) statistics on inequality of industrial wages and the distributional data from Deininger and Squire (2013) marked as high quality. The EHII data are provided in the form of gross income household Gini indices.[39] The advantage of EHII compared to SWIID is that the former is using within country inequality information for all countries and all the years it provides estimates for.

A number of additional secondary sources and research studies where used to deepen the reach of the data on inequality: Atkinson et al. (2017) for 25 developed countries, data on Spain by Prados de la Escosura (2008)[40], Botswana from Bolt and Hillbom (2016), historical data based on social tables for USA from Lindert and Williamson (2016), as well as data from Jones and Weinberg (2000) and the US census for post WWII period, Alfani and Tadei (2017) for Central African Republic, Ivory Coast and Senegal, Kenya from Aboagye and Bolt (2018) and Bigsten (1986), South Korea data from Kang (2001), Brazil from Bértola et al. (2006) and Mizoguchi (1985) for East and South East Asia.

Another, and unique in its nature, source for income inequality data is the top incomes data made available by the World Inequality Database (WID) at `WID.world`, see for example Atkinson et al. (2011) and Alvaredo et al. (2013).[41] The contribution of this source in understanding the

[38]The Williamson hypothesis states that if the basis labour wage increases faster than GDP per capita then inequality is being reduced and vice versa.

[39]With the econometric relationship at hand, one can also produce other flavors of income concept such as net income. This approach is used for some global inequality concepts as discussed in the following section on the Monte Carlo procedure.

[40]An updated version of those estimates were kindly communicated to me by the author.

[41]One of the limitations as recognized by a few of the key WID contributors (Atkinson et al., 2011, p.4-5), is that the underlying concept it pre-tax gross income. Another mismatch with househod surveys comes from the covered population, where it is individuals older than 20 years old in the case of WID, while HHS typically cover all ages. In

evolution of inequality around the world and bring the topic of inequality back on the research and societal agenda is indispensable, such as the seminal contribution by Piketty (2014). A number of authors choose to complement the HHS distributional data with information about the top income from WID in order to correct for under-reporting (see Lakner and Milanovic (2013), and Anand and Segal (2014)). I follow their example here, albeit using a different set of–rather more intermediate–assumptions (see next sub-section for details).[42]

Finally, for conducting the calculations a few more data sources are required. Population data are taken from the United Nation's World Population Project. And for the standard errors of the PPP exchange rates I use the preliminary estimates by Deaton (2012). These estimates are for the 2005 PPP exchange rates and express the implied uncertainties due to the variance in relative prices structure and the expenditure habits among countries. I will use the standard deviation for each country as it is for all PPP exchange rates rounds before and including 2005, while for the latest and most methodologically complete round of 2011 I will use half the SD from this source. Do note that the error terms calculated by Deaton are considerable with most countries having a CV of around 15-17%.

3.1 Making top ends meet

In their attempt to combine HHS and top income data, Lakner and Milanovic (2013, p.4) make a rather extreme assumption, as they themselves also acknowledge (Anand and Segal, 2014). They assume that the gap between the mean income in the HHS and the mean HFCE in the National Accounts is the missing income from the top decile into which they further fit a Pareto distribution. They use this approach as a "proxy for the extent of 'missing' top incomes". In characterizing this method as implausible, (Anand and Segal, 2014, p.27) use the example of India in 2008 where the survey mean is only 53% of the HFCE from the national accounts. Which means that they attribute the remaining 47% of the total HFCE to the top decile. I find it hard to argue against the point by Anand and Segal in this case.

In their own calculations Anand and Segal (2014, p.26) take a different approach, and assume that surveys are representative of all households, but "under-report the incomes of the top percentile in the national distribution". They account for this discrepancy by multiplying the the population in each income group from the HHS with 0.99, "and append the top percentile with its income share

addition, some serious concerns regarding the WID approach have been raised by Galbraith (2018, e.g. in p.11-2), however investigating this dimension is far beyond the scope of this paper.

[42] Another tentative approach to correct for the under-reporting by rich households is the one proposed by Korinek et al. (2006). Korinek et al. statistically estimate that adjusting for the under-reporting of high income households, also known as "selective compliance", in the United States for 2004 would add about 5 points to the Gini index. However, this approach has not been tested to a wider set of countries. In addition, as discussed in (Anand and Segal, 2014, p.20), Székely and Hilgert (1999) most often in Latin American countries the richest individuals found in household survey data had "an income no higher than what would be expected of a mid-level manager in an international firm". This is a strong indication that the very rich households simply are not present in the surveys. In turn, and with respect to the method of Korinek et al. (2006), Ravallion (2018b, p.7) note that the underlying assumption of the method would fails "if (for example) none of the super-rich participate in the survey". Therefore it seems that for a wide range of countries, this method would not be applicable.

from the tax data (assuming that its share of 'control' income is equal to its share of survey income)". They also make an econometric estimation from the pool of countries for which they have data on top deciles from HHS and top 1% from WID, in order to impute the top 1% also for countries and years for which WID does not have data on (Anand and Segal, 2014, p.21). This procedure increases the mean income as identified by the HHS.

Given that the top income data are available from the level of top 10% and that under-reporting is not only confined in the very top of the distribution, a method that combines elements from the approaches by Lakner and Milanovic (2013) and Anand and Segal (2014) would have a few important merits. In the procedure I propose here one has to work in multiple steps starting by increasing the income of those at the top 10% in the survey so that its share of income with respect to the HHS total to match the share of income from the top income data.[43] Then, in the next step, part of this additional income is moved toward the upper half of the decile to match the share of top 5% from top income data, the same for the top 1%, 0.1%, etc, until all data from top incomes are exhausted.

Since, as mentioned above, in the estimation of the distribution I work by calculating the income per group of 100 people, when this process is applied, an additional smoothing step is required in order to avoid having income at the threshold points of the 90th, 95th etc. percentiles to jump up abruptly. This step takes the form of fitting a Pareto distribution. Do note that in this procedure the income of the HHS is kept as a reference point and no addition of tax or NAS income upon income from surveys takes place. It is only the shape of the distribution from the tax data that is used, along with a proportional increase in the survey mean income (this is the same assumption as the one by Anand and Segal (2014) mentioned above).

4 The Monte Carlo procedure

[This quite a technical section can be skipped, it will probably be moved to an appendix.]

As discussed above, on the data side there are two key issues with the estimation of global inequality. One is fragmentary data, since we only have quantile income shares or summary inequality indices and not the entire distribution of income at the country level. The other is that the available data are sparse, which means that for most years we do not have distributional data at all. On this Milanovic (2002a, p.16) resides with Samuelson's quip: "yes, you can draw them as straight lines, but only with a very thick chalk". In effect the Monte Carlo procedure allows us to estimate just how thick this chalk must be when one wants to draw the evolution of global income inequality.

The issue with the estimation of the error term in global inequality calculations is that the "global distribution is not estimated from a stratified random sample of the world population, so standard methods are not applicable to calculate sampling errors or confidence intervals for estimates of global inequality" (Anand and Segal, 2014, p.18). In such cases the Monte Carlo procedure comes

[43] Naturally, with this procedure the HHS total is also increased. So this also needs to be taken into account in the calculations.

particularly handy (Sosa Escudero and Gasparini, 2000; Efron, 1979). In essence, to conduct a Monte Carlo simulation, one needs to provide distributions of possible values for the data sources that are treated as stochastic, and repeats the calculation a large number of times.[44]

Table 1 shows the sources of error considered in the calculation of global inequality here, and non-sampling errors in particular.[45] This table adapts on the one offered by Atkinson (2016, p.52) originally constructed for global poverty. There are 14 sources of sampling and non-sampling errors listed. Six of them are accounted for in a considerable extend, while two more are treated in a clearly partial way.[46] Six more are entirely ignored, which implies that the error terms identified here are more likely to be low band estimates, depending on the direction of the bias that those ignored error sources are introducing.[47]

The ignored sources of uncertainty are: (I) the issue of international comparability of top income data (item 3 on the table) as identified by the Alvaredo and Londoño Vélez (2013) study of top incomes in Colombia;[48] (II) bias due to the design of the household surveys or the social tables where parts of the population are typically missing (item 5), for example HHS do not consider individuals that are imprisoned or in elderly care facilities; (III) when income data are missing (item 7) I do not consider that country (this choice is consistent with the literature, and so is the following choice); (IV) as Atkinson (2016) show there are considerable inaccuracies in population totals (item 8), which both the literature in global poverty and the literature in inequality ignore; (V) countries at different stage of development probably require the use of different equivalence scales (item 13) according to the SEDLAC methodology guide, and while in the available inequality data various equivalence scales are used, this is not done systematically and with the purpose of global comparability, and it therefore adds to the uncertainty of the estimates[49]; and (VI) intra-household inequality (item 14) is not captured in the sources, which is also a reason why we cannot directly estimate gender based poverty rates.[50]

The error sources which are considered, along with the methods used for their consideration, can be better discussed in an integrated way within the diagram of figure 4. Figure 4 shows the Monte

[44] 10.000 iterations is a standard figure for this purpose.

[45] Interested readers can find a wealth of detailed information about sampling errors in household surveys in the USA CENSUS guide. For EU-SILC specific cases a detailed repository can be found in the website of Tim Goedeme and in Goedeme (2013).

[46] Those are comparability of HHS in time as various methodological factors change, and intra country price differences. For the former factor, only the component of a possible income concept mismatch is treated by the way of adding a regression based parameter with its uncertainty. For the latter factor, the rural/urban split is considered for the three countries for which the household surveys in the contemporary period are providing the relevant information (that include the two most populous countries in the world), while for the other countries and years this problem is ignored. Ravallion (2018b, p.8) discuses evidence from research which shows that “ignoring spatial price differences overestimate[s] inequality measures, though the trends over time are little affected”.

[47] This is an important topic that falls outside of the scope of this first attempt to ballpark the size of uncertainty in global inequality estimates.

[48] There are various definitions of the control income that can be used to construct the top incomes as a share of total income. Depending on which of these options is taken, a different estimates is produced as a result. On this point Anand and Segal (2014) note that “international comparisons of these top income shares may suffer from inconsistencies” and therefore “these non-comparabilities do add uncertainty to the [global inequality] estimates.”

[49] See page 24 in SEDLAC “Methodological Guide” version April 2014.

[50] For some pioneering work in the field see Wright (1995).

Table 1: Checklist of Sampling and Nonsampling Errors in global inequality estimates

.

	Source of error	Approach/Comment
1	Incomplete measurement of income	Adjust for top income under-reporting; by applying average rate with its uncertainty when no top income data is available.
2	International comparability of top income sources	ignored
3	Income inequality measurement error	Use available standard deviation or its imputed value
4	Income concept mismatch	Add a regression based constant (see text)
5	Population missing from sampling frame (e.g. those incarcerated)	ignored
6	Population in countries or territories without distributional data	imputed by randomization
7	Population in countries or territories without income data	ignored
8	Inaccurate or out-of-date population totals	ignored
9	Standard error of PPP indexes	first estimates included
10	Surveys not comparable over time (e.g. due to methodological changes)	Partially corrected for income concept; other issues are ignored.
11	Extrapolation of out-of-date survey data	A correction is applied based on the history of HHS and NAS growth rates.
12	Rural/Urban and other geographical differences	Rural/Urban split only for the countries with available data, namely: China, India and Indonesia.
13	Appropriate equivalence scale	ignored
14	Intra-household inequality	ignored

Carlo procedure as a set of various steps of calculations. In this diagram three dark gray filled boxes are more clearly identifiable indicating a starting point for each different data source. The bottom most right box represents the final estimation of a global inequality measure.[51] The purpose of this Monte Carlo experiment is to derive a numerically estimated distribution of probable values composed of 10,000 estimated points for each global inequality measure we are interested in, and for each of the years we are interested in. This means that the procedure needs to run 10,000 times for each country that in turn after the aggregation step of country level distributions to a global distribution get the desired 10,000 estimations for the global inequality measure. The part of the diagram shown, outside of the gray background at the bottom right, should be read as the section that is being executed 10,000 times for each country in a given year, while the area with a gray background is the aggregation that is executed also 10,000 times in a given year consolidating upon the acquired country level distributions.

The overall starting point, at the top left box, is the test for the existence of an income inequality index (Gini) which can be used estimate the income distribution by inversion using the log-normality assumption (see section 2.4.3). If such a Gini index is available for a given country year then the diamond to the immediate right denotes the random selection of a specific value using the standard error term–which represents the sampling error or the estimation error in case of a synthetic value (e.g. for EHII or SWIID)–from that Gini estimate.[52] If such an error term is not available, then one is imputed based on the relationship between the quality characterization available on the WIDER dataset and the available standard error in income inequality indices (Belgium, USA or SEDLAC). When the WIDER dataset does not have information for that country year, then the average standard error from all available sources is used (along with the standard error of its average, so in each iteration the standard error to be used for the Gini index is drawn from the distribution of available standard errors). This drawn Gini is then inverted using the log-normality assumption, and a rough distribution obtains (corresponding to the rectangle to right of the first diamond on the diagram). Over to the next box to the right where other (possibly) available inequality information, such as income shares, is used to adjust the distribution in a way that it both corresponds to the drawn Gini value and to the additional inequality information (by minimizing the Euclidean distance among them).

In case there is no available inequality index from the particular year, but there is one from a different year (second box on left column), then this is used with a penalty on the standard error used. This penalty is estimated based on the percentage deviation of Gini from those countries with available data. In case of an interpolation the weighted average of the available information prior and after the year of interest is used, with the closest available datapoint getting higher weight.[53] In case of an extrapolation only one value is used. One way or another, a Gini value is drawn to be used as before for estimating the rough distribution. If on the other hand, there is no Gini index available on a particular country at any year (first diamond on left column), then a random value is drawn from the uniform distribution between the lowest and the highest Gini values in the entire

[51] Several inequality measures are used. See the results section, as well as the methodological discussion above.

[52] In terms of flow-chart notation, all diamonds denote a choice at random based on an available distribution.

[53] See the appendix for the calculation procedure (forthcoming).

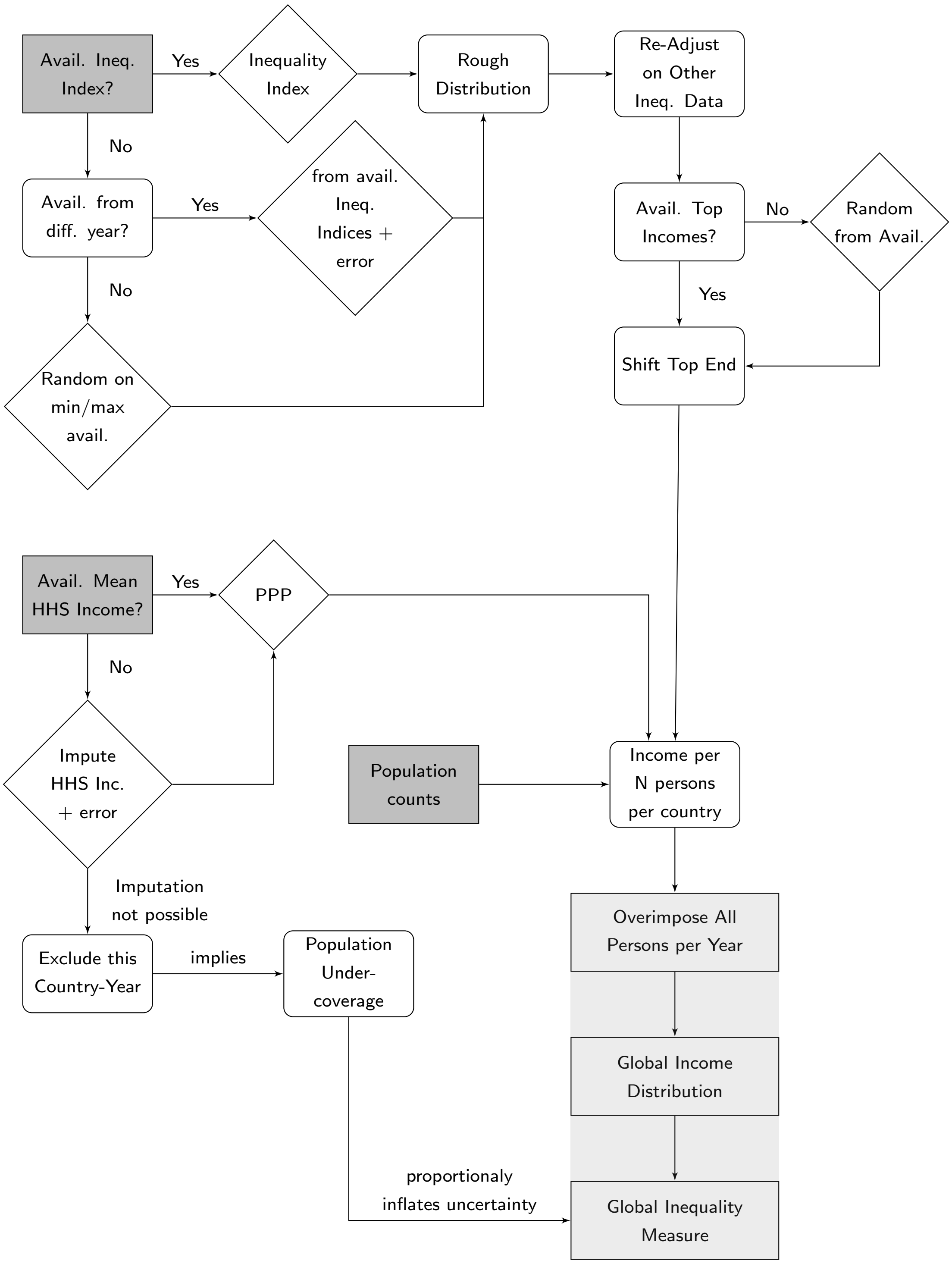

Figure 4: Architecture of Data and Monte Carlo method for estimating global income inequality. The diagram represents the procedure repeated for each country in a given year; with the last two boxes representing the aggregation on a per year basis to derive the global inequality estimation. Diamond shaped objects mark a draw from a distribution along the Monte Carlo simulation.

dataset (with the same income concept–gross income). The procedure then hooks back to the point of estimating the rough distribution as before. A difference now is that if there is some additional inequality information this now gets a higher weight in adjusting the rough distribution.[54]

One important remark on the above procedure has to do with the possibility of a mismatch, in a given country-year combination, between the income concept we choose to use and that of the available Gini. In the results presented below the income concept is that of gross inequality, and for the cases that such a Gini is not available, a drawn parameter is added to account for the uncertainty introduced by this mismatch.[55]

After the re-adjusted distribution has been established the next step is to account for the top incomes under-reporting using the method described above (see subsection 3.1). When information on top incomes is not available then the distribution of the available top income shares for that year is used. If less than 10 countries have available top income shares for that year, then the countries with available data in the closest possible year are used as well in order to reach 10 countries at minimum. Therefore at the end of this procedure an estimated top-income adjusted distribution obtains.

Next gray box checks whether or not a mean household survey income is available. If there is available, then the PPP conversion rate is applied to convert the local currency units to PPP dollars. The exact PPP value is drawn from the normal distribution implied by the available standard error for the particular country. The convention has as follows: (i) when the PPP rates are from 2005 or prior then the SD value from Deaton (2012) are used, and when there are none available then the distribution of available values is used at random (uniformly); (ii) when PPP rates are from 2011 then half of the SD available from Deaton (2012) are used, and in case that the country is not in the estimates half the–as above–randomized value is used instead.

When no HHS mean income is available then one is imputed using a national account statistic to shift the closest available HHS mean income estimate (HFCE growth is preferred and GDP growth is used when that is not available). If there is no HHS mean income available, then the distribution of the ratio between HHS mean income and the NAS statistic (from the 20 closest years with available data) is used to impute the HHS value from the available NAS statistic. If no NAS statistic is available, then imputation is not possible and the particular country-year combination is excluded. This in turn means that there is a population under-coverage on a global scale, and therefore uncertainty of any final global inequality measure should be inflated to represent that. This is done by increasing the SD of the final estimate inverse proportionally to the final global population coverage rate for that particular year.

Next gray box corresponds to the population data, and this information is used, along with the final estimated distribution and its (estimated) mean income, to assign a particular income to every

[54] The Euclidean distance is now estimated with a weight on the Gini of 0.25 and a weight on the additional information of 0.75. This choice is clearly arbitrary and it is followed as a reasonable choice and better than ignoring the presence of available direct information on the distribution although not in the form of a Gini index.

[55] This parameter is estimated by the observed differences between the income concept of the available data and gross income from countries that have both.

N=100 persons per country. This leads the process to its final stage marked by the three boxes in the common gray shaded area as a background at the bottom right of the diagram. The first step here is to sort all groups of 100 persons globally–thus forming a global distribution of income–and then using this information to estimate the global inequality measure one wishes.

5 Results

5.1 Stylized Benchmark

PRELIMINARY RESULTS BASED ON 100 ITERATIONS ONLY

Since the focal point is the estimation of the confidence interval around the global inequality estimates, an instructive way to first approach this is via the use of a stylized long run example, and in the next sub-section enrich it with the actual data. In this experiment I draw information from the actual available data only with respect to the typical sampling error, the typical error of estimated (synthetic) Ginis, the present day income disparities among countries, and the broad coverage pattern of distribution data availability (which implies for which country-years combinations imputation is needed).

The experiment uses the list of 169 countries in the Maddison dataset for which there are data for both population and GDP per capita in 2015. Population and income stay fixed at those levels for the entire period.[56] What does change is the availability of distributional data, and the typical error in the overall mix of those distributional data (which represents the relative high presence of estimated data in the historical period, and the less so in the more recent years). For the start of the period in (a virtual) 1820 this error term is 8 Gini points[57], the same uncertainty as the average uncertainty in the estimation methods used in Zanden van et al. (2013)[58], and at the end of the period this drops to 1 Gini point, representing the decrease presence of estimated Ginis in the pool of data.

The error terms of the PPP exchange rates is taken to be 15% as the average in Deaton (2012), while the error term for the sampling error in the available data points is taken to be 0.5 Gini points (typical value in SEDLAC and close to the average in USA 1967-1976, which is 0.57 Gini points). All countries have at least 2 HHS observations (which is higher than in reality, since we do have countries without any observation throughout), and at maximum a country has 123 in 197 years. These values are not directly specified manually, but are the result of assumptions on the probability that in 1820s only a few countries have available data (ca. 3), while by 2010s this share rises to about 54% or 91 countries.[59] The rule is that the higher the GDP per capita and population, the

[56] This is mainly because we are interested in the broadly estimating the size of the error term. An experiment that allows population and income to change in this period is provided below as well.

[57] A reference to a Gini point assumes that Gini is measured in the scale of 0 to 100, while a reference to a Gini value assumes that the Gini is measured between 0 and 1.

[58] Refering to the Williamson Hypothesis and the relationship between the coefficient of variation in heights and inequality; see the methodological section above and the source for more details.

[59] This is the probability for a country to actually have an inequality observation, either direct or estimate, in year

more probable is that a country has an inequality observation in a given year.[60] A band-pass filter is applied not allowing the drawn Gini values to be less than 0.2 or higher than 0.8, which are the limits of the observable values throughout the 1820-2016 period. Finally, an important element is the error induced by shifting a distribution from a different year to sit-in for a missing observation. This of course depends on the number of years in between the available data and the year of interest. In next section the level of the error is based on the empirical evidence from the available sources, but here a simple rule is used: the error term added is 0.5 Gini points for 1 a year shift, and grows by half a point for each additional year in the gap up to a max of 8 Gini points, where it rests for any number of years beyond 15.

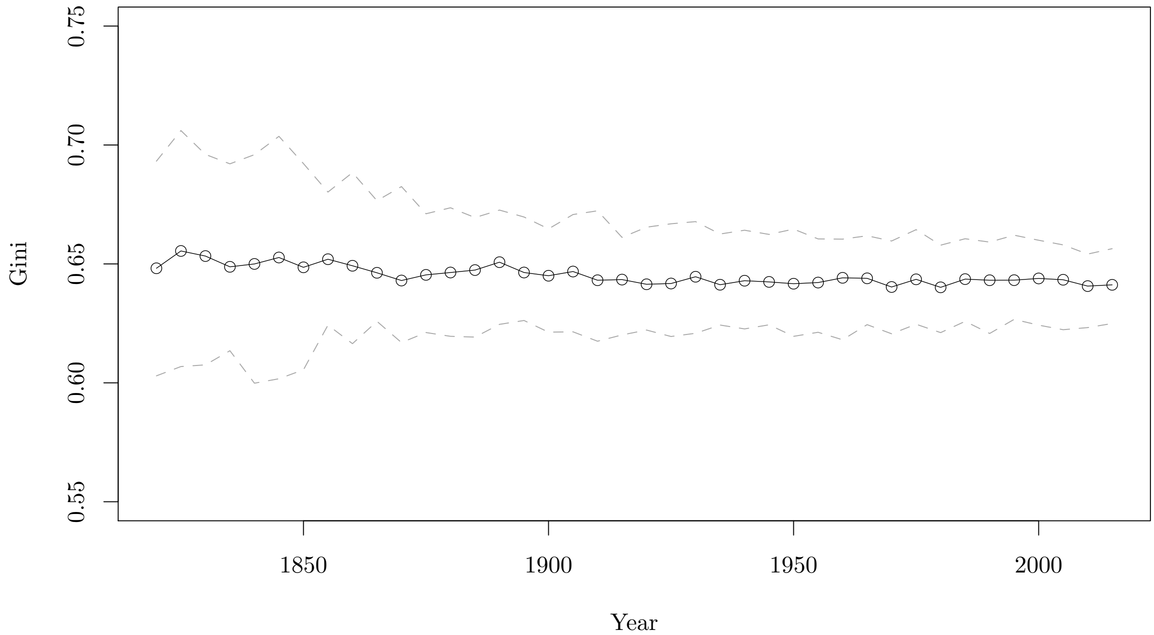

Figure 5: The stylized evolution of the concept 3 global income inequality and its 95% confidence interval, estimated for every five years within 1820-2015.

Figure 5 shows the results. Overall the 95% confidence interval of the concept 3 global inequality estimates has a breadth of 9 Gini points in 1820, and it drops to about 3 Gini points by 2016. The confidence interval of the more contemporary period is smaller compared to the estimates by Milanovic (2012), who in 2005 estimates the confidence interval to be about 6 Gini points wide.[61]

T explicitly; and not one available from a nearby year that is shifted to 1820 or any other particular year of interest. The distribution of available data points in this stylized scenario is shown in figure 6 in the appendix.

[60] The highest ranking country in terms of both income and population will have 4 times higher probability of having an inequality observation in a given year than the lowest ranking country. This probability is linearly interpolated for countries in between these two extremes.

[61] In his estimation Milanovic assumes that all HHS data from around the world come from a single randomized global HHS, and ignoring the PPP and all other uncertainties. In the stylized scenario presented here the confidence

It therefore appears that this first stylized approximation underestimates the size of uncertainty in the modern period. At the same time, and in comparison to the estimate of a 3.5 Gini points wide confidence interval by Bourguignon and Morrisson (2002) for 1820, the stylized approximation appears to overestimate the error term. Do note however that as mentioned in the introduction the approach by Bourguignon and Morrisson (2002) makes assumptions that appear implausible for the accuracy of their data and results (Anand and Segal, 2008). For comparison for 1992 Bourguignon and Morrisson (2002) estimate a confidence interval of 0.4 Gini points, while Milanovic (2012) estimates 6.8 Gini points in 1993.

5.2 Data Based Application

TBD

6 Conclusion

TBD

References

Aboagye, P. Y. and Bolt, J. (2018). Economic inequality in Ghana, 1891-1960. *African Economic History Working Paper Series*, (38).

Ackland, R., Dowrick, S., and Freyens, B. (2013). Measuring global poverty: why ppp methods matter. *The Review of Economics and Statistics*, 95(July):813–824.

Aitchison, J. and Brown, J. (1957). *The Lognormal Distribution.* The syndics of the Cambridge University Press, Cambridge.

Alfani, G. and Tadei, F. (2017). *Income Inequality in Colonial Africa: Building Social Tables for Pre-Independence Central African Republic, Ivory Coast and Senegal.* Number 33.

Alvaredo, F., Atkinson, A. B., Piketty, T., and Saez, E. (2013). The top 1 percent in international and historical perspective.

Alvaredo, F., Chancel, L., Piketty, T., Saez, E., and Zucman, G. (2018). World Inequality Report 2018. *World Inequality Lab*, page 300.

Alvaredo, F. and Londoño Vélez, J. (2013). High incomes and personal taxation in a developing economy: Colombia. *CEQ Working Paper*, (12).

Amiel, Y. and Cowell, F. A. (1992). Measurement of income inequality. *Journal of Public Economics*, 47:3–26.

interval for 2005 is also 9 Gini points.

Anand, S. and Segal, P. (2008). What Do We Know about Global Income Inequality? *Journal of Economic Literature*, 46:1(1983):57–94.

Anand, S. and Segal, P. (2014). The Global Distribution of Income. *International Development Institute Working Paper*, 2014-01.

Aten, B. and Heston, A. (2001). Use of Penn World Tables for International Comparisons of Poverty Levels : Potential and Limitations. In *Global Poverty Workshop, Initiative for Policy Dialogue*, number 1986, pages 1–15.

Aten, B. and Heston, A. (2010). Use of Country Purchasing Power Parities for International Comparisons of Poverty Levels: Potential and Limitations. In Anand, S., Segal, P., and Stiglitz, J. E., editors, *Debates on the Measurement of Global Poverty*, chapter 6, pages 154–186. Oxford University Press, Oxford.

Atkinson, A. and Brandolini, A. (2010). On Analyzing the World Distribution of Income. *The World Bank Economic Review*, 24(1):1–37.

Atkinson, A., Piketty, T., and Saez, E. (2011). Top Incomes in the Long Run of History. *Journal of Economic Literature*, 49(1):3–71.

Atkinson, A. B. (1970). On the Measurement of Inequality. *Journal of Economic Theory*, 263:244–263.

Atkinson, A. B. (1997). Bringing income distribution in from the cold. *The Economic Journal*, 107(March):297–321.

Atkinson, A. B. (2016). *Monitoring Global Poverty: Report of the Commission on Global Poverty.* The World Bank, Washington, DC.

Atkinson, T., Hasell, J., Morelli, S., Roser, M., Hasell, J., and Morelli, S. (2017). The Chartbook of Economic Inequality. *INET at the Oxford Martin School.*

Bértola, L., Castelnovo, C., Reis, E., and Willebald, H. (2006). An exploration into the distribution of income in Brazil, 1839-1939. *XIV International Economic History Congress Helsinki-Finland 21-25 August 2006.*

Bigsten, A. (1986). Welfare and economic growth in Kenya, 1914–76. *World Development*, 14(9):1151–1160.

Bolt, J. and Hillbom, E. (2016). Long-term trends in economic inequality: lessons from colonial Botswana, 1921–74. *Economic History Review*, 69(4):1255–1284.

Bolt, J., Inklaar, R., de Jong, H., and Zanden van, J. L. (2018). Rebasing 'Maddison': new income comparisons and the shape of long-run economic development. *Working Paper "The Maddison Project"*, (January):67.

Bourguignon, F. (2015). *The Globalization of Inequality.* Princeton University Press, Princeton.

Bourguignon, F. and Morrisson, C. (2002). Inequality Among World Citizens : 1820-1992. *The American Economic Review*, 92(4):727–744.

CEDLAS (2019). Socio-Economic Database for Latin America and the Caribbean (CEDLAS and The World Bank).

Cowell, F. A. (2009). *Measuring Inequality*. Number December. Oxford University Press, 3rd edition.

Deaton, A. (2005). Measuring poverty in a growing world (or measuring growth in poor world). *The Review of Economics and Statistics*, LXXXVII(February):1–19.

Deaton, A. and Dupriez, O. (2011). Spatial price differences within large countries. (July).

Deaton, A. and Heston, A. (2010). Understanding PPPs and PPP-based National Accounts. *American Economic Journal: Macroeconomics*, 2(4):1–35.

Deaton, A. S. (2012). Calibrating measurement uncertainty in purchasing power parity exchange rates. *Princeton University Working Paper*, (April).

Deaton, A. S. and Dupriez, O. (2009). Global poverty and global price indexes.

Deininger, K. and Squire, L. (1996). A new dataset measuring income inequality. *The World Bank Economic Review*, 10(3):565–591.

Deininger, K. W. and Squire, L. (2013). Measuring Income Inequality Database 1890-1996.

Dikhanov, Y. and Ward, M. (2001). Evolution of the Global Distribution of Income in 1970-99. *53rd session of the International Statistical Institute, Seoul, Republic of Korea*, pages 22–29.

Dowrick, S. and Akmal, M. (2005). Contradictory trends in global income inequality : a tale of two biases. *Review of Income and Wealth*, 51(2):201–229.

Efron, B. (1979). Bootstrap Methods : Another Look at the Jackknife. *The Annals of Statistics*, 7(1):1–26.

Ferreira, F. H. G., Chen, S., Dabalen, A., Jolliffe, D., Narayan, A., Prydz, E. B., Revenga, A., Sangraula, P., and Yoshida, N. (2015). A global count of the extreme poor in 2012: data issues, methodology and initial results. *Policy Research Working Paper 7432. Washington, DC: The World Bank.*

Firebaugh, G. (1999). Empirics of World Income Inequality. *American Journal of Sociology*, 104(6):1597–1630.

Firebaugh, G. (2003). *The New Geography of Global Income Inequality*. Harvard University Press, Cambridge and London.

Galbraith, J. K. (2018). Sparse, Inconsistent and Unreliable: Tax Records and the World Inequality Report 2018. *Development and Change*, 00(0):1–18.

Galbraith, J. K. and Halbach, B. (2016). A Comparison of Top Income Shares and Global Inequality Datasets. *The University of Texas Inequality Project, Working Paper*, 73:2016.

Ghosh, J. (2018). A note on estimating income inequality across countries using PPP exchange rates. *Economic and Labour Relations Review*, 29(1):24–37.

Goda, T. and Torres Garcia, A. (2016). The Rising Tide of Absolute Global Income Inequality During 1850-2010: Is It Driven by Inequality Within or Between Countries? *Social Indicators Research*, pages 1–22.

Goedeme, T. (2013). How much Confidence can we have in EU-SILC? Complex Sample Designs and the Standard Error of the Europe 2020 Poverty Indicators. *Social Indicators Research*, 110(1):89–110.

Jones, A. F. and Weinberg, D. H. (2000). The Changing Shape of the Nation's Income Distribution 1947-1998. *U.S. Census Bureau - Current Population Reports*, P60-204:1–11.

Kang, S. (2001). Globalization and Income Inequality in Korea : An Overview. *OECD Technical Meeting*, (December).

Korinek, A., Mistiaen, J. a., and Ravallion, M. (2006). Survey nonresponse and the distribution of income. *Journal of Economic Inequality*, 4(1):33–55.

Lakner, C. and Milanovic, B. (2013). Global income distribution: from the fall of the Berlin Wall to the Great Recession.

Lindert, P. H. and Williamson, J. G. (2016). *Unequal gains: American growth and inequality since 1700.* Princeton University Press, Princeton.

LIS (2018). Luxembourg Income Study Database (LIS).

Lopez, J. H. and Servén, L. (2006). A Normal Relationship ? Poverty, Growth, and Inequality.

Maddison, A. (2010). Maddison Project, downloaded from http://www.ggdc.net/maddison/oriindex.htm.

Milanovic, B. (2002a). The Ricardian Vice : Why Sala-i-Martin's calculations of world income inequality are wrong.

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

Milanovic, B. (2006). Global Income Inequality: What it is and why it matters. *World Bank Working Paper Series*, no.3865.

Milanovic, B. (2007). *Worlds Apart. Measuring International and Global Inequality.* Princeton University Press, Princeton.

Milanovic, B. (2011). A short history of global inequality: The past two centuries. *Explorations in Economic History*, 48(4):494–506.

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

Milanovic, B. (2019). All the Ginis dataset. Technical report, Graduate Center, City University of New York and Stone Center on Socio-economic Inequality, New York.

Milanovic, B., Lindert, P., and Williamson, J. G. (2007). Measuring Ancient Inequality. *MPRA Working Papers*, (5388).

Mizoguchi, T. (1985). ECONOMIC DEVELOPMENT POLICY AND INCOME DISTRIBUTION: THE EXPERIENCE IN EAST AND SOUTHEAST ASIA. *The Developing Economies*, XXIII(4):307–324.

Moatsos, M., Baten, J., Foldvari, P., Leeuwen, B. V., and Van, J. L. (2014). Income inequality since 1820. In van Zanden, J. L., Baten, J., D'Ercole, M. M., Rijpma, A., and Timmer, M. P., editors, *How Was Life? Global Well-being since 1820*, chapter 11. OECD Publishing, Paris.

Niño-Zarazúa, M., Roope, L., and Tarp, F. (2016). Global Inequality: Relatively Lower, Absolutely Higher. *Review of Income and Wealth*, (00):1–24.

OECD (2019). Income inequality (indicator).

Piketty, T. (2014). *Capital in the twenty-first century.* Belknap Press - Harvard University Press, London.

Prados de la Escosura, L. (2008). Inequality, poverty and the Kuznets curve in Spain, 1850-2000. *European Review of Economic History*, 12(3):287–324.

R Core Team (2018). *R: A Language and Environment for Statistical Computing.* R Foundation for Statistical Computing, Vienna, Austria.

Rao, D. P., Rambaldi, A. N., Huynh, L. T., Doran, H. E., and Ganegodage, K. R. (2015). UQ International Comparisons Database: UQICD Version 2.1.2.

Ravallion, M. (2016). *The Economics of Poverty: History, Measurement and Policy.* Oxford University Press.

Ravallion, M. (2017). Global Inequality When Unequal Countries Create Unequal People. *NBER Working Paper*, (w24177).

Ravallion, M. (2018a). Inequality and Globalization: A Review Essay. *Journal of Economic Literature*, 56(2):620–642.

Ravallion, M. (2018b). What might explain today ' s conflicting narratives on global inequality ? *WIDER Working Paper*, (141).

Reddy, S. G. and Pogge, T. (2010). How not to count the poor. In Anand, S., Segal, P., and Stiglitz, J. E., editors, *Debates on the Measurement of Global Poverty*, pages 42–51. Oxford University Press.

Sala-i Martin, X. X. (2006). The World Distribution of Income: Falling Poverty And...Convergence, Period. *Quarterly Journal of Economics*, 121(2):351–397.

Sen, A. (1976). Real National Income. *The Review of Economic Studies*, 43(1):19–39.

Sen, A. K. (2009). *The Idea of Justice*. Harvard University Press, Cambridge, MA.

Shorrocks, A. F. and Wan, G. (2008). Ungrouping Income Distributions, Synthesising Samples for Inequality and Poverty Analysis. *UNU-WIDER Research Paper*, (16).

Silva, A. C. and Yakovenko, V. M. (2005). Temporal evolution of the "thermal" and "superthermal" income classes in the USA during 1983-2001. *Europhysics Letters*, 69(2):304–310.

Solt, F. (2016). The Standardized World Income Inequality Database, SWIID Version 7.1, August 2018. *Social Science Quarterly*, 97(5):1267–1281.

Sosa Escudero, W. and Gasparini, L. (2000). A note on the statistical significance of changes in inequality. *Economica*, XLVI(1):111–122.

Subramanian, S. and Jayaraj, D. (2015). Growth and inequality in the distribution of India's consumption expenditure 1983 to 2009-10. *WIDER Working Paper*, (025).

Székely, M. and Hilgert, M. (1999). What' s Behind the Inequality We Measure- An Investigation Using Latin American Data.pdf.

United Nations (2018). Note on Income Inequality Data. *Department of Economic and Social Affairs*, (October).

UNU-WIDER (2018). UNU-WIDER, World Income Inequality Database (WIID4, December 19, 2018). *United Nations*.

Wilkinson, R. G. and Pickett, K. E. (2009). *The Spirit Level*. Penguin Group, London, UK.

Wright, R. E. (1995). Women and Poverty in Industrialized Countries. *Journal of Income Distribution*, 5(1):31–46.

Zanden van, J. L., Baten, J., Foldvari, P., and van Leeuwen, B. (2013). The Changing Shape of Global Inequality 1820-2000; Exploring a New Dataset. *Review of Income and Wealth*, pages 1–19.

Zheng, B. (2007). Unit-Consistent Decomposable Inequality Measures. *Economica*, 74(293):97–111.

7 Appendix

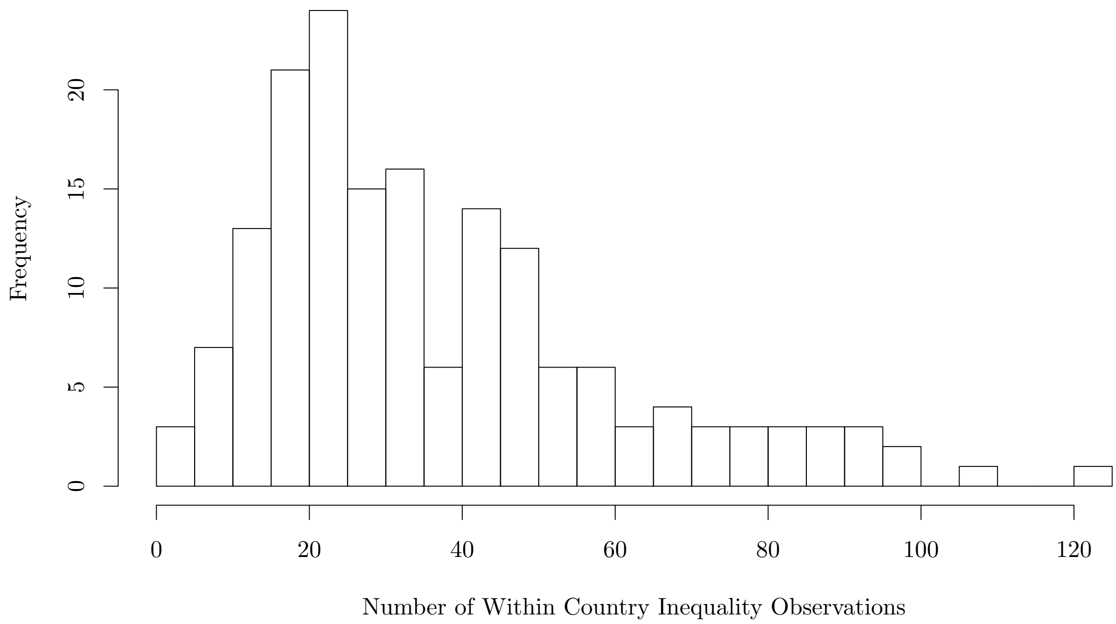

Figure 6: The distribution of available Gini data points for the stylized scenario.