MEMORANDUM

No 11/2011

Inequality and growth in the very long run: inferring inequality from data on social groups

Jørgen Modalsli

ISSN: 0809-8786

Department of Economics
University of Oslo

This series is published by the
University of Oslo
Department of Economics

P. O.Box 1095 Blindern
N-0317 OSLO Norway
Telephone: + 47 22855127
Fax: + 47 22855035
Internet: http://www.sv.uio.no/econ
e-mail: econdep@econ.uio.no

In co-operation with
The Frisch Centre for Economic Research

Gaustadalleén 21
N-0371 OSLO Norway
Telephone: +47 22 95 88 20
Fax: +47 22 95 88 25
Internet: http://www.frisch.uio.no
e-mail: frisch@frisch.uio.no

Last 10 Memoranda

No 10/11	Eilev S. Jansen, Tord S. H. Krogh *Credit conditions indices: Controlling for regime shifts in the Norwegian credit market*
No 09/11	Karin Jacobsen , Kari H. Eika, Leif Helland, Jo Thori Lind, Karine Nyborg *Are nurses more altruistic than real estate brokers?*
No 08/11	Michael Hoel *The supply side of CO_2 with country heterogeneity*
No 07/11	Anne Line Bretteville-Jensen, Erik Biørn, Randi Selmer *Quitting Behaviour of Cigarette Smokers* *Are there direct effects of a screening program?*
No 06/11	Kari H. Eika *Near and Generous? Gift Propensity and Chosen Emotional Distance*
No 05/11	Finn R. Førsund *Weight Restrictions in DEA: Misplaced Emphasis?*
No 04/11	Vladimir Krivonozhko, Finn R. Førsund, Andrey V. Lychev *Terminal units in DEA: Definition and Determination*
No 03/11	André K. Anundsen, Tord S. H. Krogh, Ragnar Nymoen, Jon Vislie *Overdeterminacy and endogenous cycles: Trygve Haavelmo's business cycle model and its implications for monetary policy*
No 02/11	Nils Chr. Framstad *Portfolio Separation Properties of the Skew-Elliptical Distributions*
No 01/11	Karine Nyborg , Tao Zhang *Is corporate social responsibility associated with lower wages?*

Previous issues of the memo-series are available in a PDF® format at:
http://www.sv.uio.no/econ/forskning/publikasjoner/memorandum

Inequality and growth in the very long run: inferring inequality from data on social groups

Jørgen Modalsli, University of Oslo*

March 23, 2011

Abstract

Income distribution data from before the Industrial Revolution usually comes in the shape of social tables: inventories of a range of social groups and their mean incomes. These are frequently reported without adjusting for within-group income dispersion, leading to a systematic downward bias in the reporting of pre-industrial inequality. This paper suggests a correction method, and applies it to an existing collection of twenty-five social tables, from Rome in AD 1 to India in 1947. The corrections, using a variety of assumptions on within-group dispersion, lead to substantial increases in the Gini coefficients. Combining the inequality levels with data on GDP, a robust positive relationship between income inequality and economic growth is confirmed. This supports earlier proposals, based on fewer data points, of a "super Kuznets curve" of increasing inequality over the entire pre-industrial period.

Keywords: Pre-industrial inequality, social tables, Kuznets curve, history

JEL codes: D31, N30, O11, C65

*E-mail: j.h.modalsli@econ.uio.no. This paper is part of the research activities at the centre of Equality, Social Organization, and Performance (ESOP) at the Department of Economics at the University of Oslo. ESOP is supported by the Research Council of Norway. I am grateful to Branko Milanovic, Rolf Aaberge, Erik Sørensen, Halvor Mehlum and Kalle Moene for comments and suggestions.

1 Introduction

Not much is known about inequality in the very long run. The lack of data has been addressed by Milanovic *et al.* (2011), who collect a large set of social tables. The social tables give data on the size and average income of social classes in many pre-industrial societies, with the catch that the income distribution within each class is unknown. This paper uses these social tables to draw inference on the long-run development of inequality, as well as the relationship between inequality and growth, while explicitly allowing for different levels of within-group inequality. The dimension of within-group inequality is missing in Milanovic *et al.*, leading to too low reported Gini coefficients.

1.1 Inequality in the very long run

The seminal contribution on the long-run evolution of inequality is Kuznets (1955). Using a few observations from the United States, England and Germany, Kuznets argues that inequality goes up with the industrial revolution and then decreases with modernization. While Kuznets treats the Industrial Revolution as a rather specific process (he dates the possible "widening phase" in England as going from 1780 to 1850, and postulates even shorter periods for the other countries), more recent views on industrialization stress the changes as being more gradual.

Kuznets based his conclusions on a very small data set. Over the years, better estimates of inequality through the Industrial Revolution has emerged; a macroeconomic picture of the entire post-1820 period is given by Bourguignon & Morrisson (2002). However, data on the period before 1820 remains sparse. Van Zanden (1995) uses data on European cities and argues that the period of increasing inequality started before the Industrial Revolution.[1] He documents a positive correlation between growth and inequality in European cities after the mid-1500s, with the growth-inequality relationship switching sign some time between 1870 and 1900. Lindert (2000) finds weak evidence of increasing inequality in Britain and the United States from the 1700s, again with a peak in inequality some time after industrialization. Hoffman *et al.* (2002) adjust for changing consumption baskets in several European countries and find this make the increasing-inequality trends even stronger, in particular before 1650 (their analysis starts in 1500).

The most comprehensive analysis of pre-industrial inequality so far is given by Milanovic *et al.* (2011). The authors collect a comprehensive set of social tables - listing social groups, their sizes and incomes for 25 country-time points. The main body of their paper discusses the relationship between economic activity and

[1]The term "super Kuznets curve", meaning a positive relationship between growth and inequality going further back than proposed by Kuznets, is due to van Zanden.

feasible inequality levels, but the data is publicly available and ready to be used for other purposes.[2]

Social group	Share of pop.	Per capita income (nomisma per year)	Income in terms of per capita mean
Tenants	0.37	3.5	0.56
Urban "marginals"	0.02	3.51	0.56
Farmers	0.52	3.8	0.61
Workers	0.03	6	0.97
Army	0.01	6.5	1.05
Traders, skilled craftsmen	0.035	18	2.90
Large landowners	0.01	25	4.02
Nobility	0.005	350	56.31

Table 1: Example of social table: Byzantium, ca year 1000. Source: Milanovic *et al.* (2007), based on Milanovic (2006)

An example of a social table is given in Table 1. It lists the social classes in Byzantium, ca year 1000. The data set used in this paper consists of 25 such social tables, with a varying number of groups and class definitions. Though far from a balanced panel (only a few countries have observations for more than one period), this is the first comprehensive cross-region data series on pre-industrial inequality, as opposed to the more country- or region-specific discussions of the other studies.

1.2 Interpolating inequality: Limitations of existing approaches

Common for all elaborations on pre-industrial inequality is the need for some type of interpolation. Often a combination of techniques is used, as the data available can be of many types. For example, Lindert (2000) uses a combination of social tables, factor prices, wage data, and land holdings, as well as more detailed data on wealth and income for the richer parts of the population. In most cases, information on the distribution among the poor is particularly hard to find.

[2]Milanovic *et al.* have a total of 28 observations, but for three of these (Tuscany 1427, Holland 1561 and Japan 1886) they do not appear to have access to the underlying data. For the remaining 25 observations, based on a wide range of studies described in their paper, I thank Branko Milanovic for supplying the dataset; most of the observations are also available online at `http://gpih.ucdavis.edu/`. The working paper version of their paper (Milanovic *et al.*, 2007) has a fuller exposition of the data and methodology.

For the social tables collected by Milanovic *et al.* (2011), we have the advantage of a comprehensive table for the entire population.[3] For each social class, we have an estimate of mean income of the group, as well as the relative size of the group. The distribution within each group, however, is not known. For this reason, analyzing inequality using social tables data requires additional assumptions on the characteristics of the social groups.

A natural starting point is to consider a distribution where the entire group is concentrated at its mean income. Taking the "farmers" in Table 1 as an example, this would mean that all farmers had an income of 3.8 nomisma per year. With this, it is straightforward to calculate an inequality measure such as the Gini coefficient. Milanovic *et al.* (2011) describe this as the lower bound of the Gini coefficient, and denote it as "Gini1". In the following, this will be referred to as a "point distribution", as the population is concentrated at a finite number of points.[4]

Going one step further, we can think of a distribution where all the members of group i are poorer than all members of group $i+1$; in the terms of Table 1, all "tenants" are poorer than the poorest farmer. This will be referred to as a population being *perfectly sorted* by groups; in other words, there is no overlap between the population ranges. For group borders at midpoints between group means, Milanovic *et al.* (2011) denote this as "Gini2", but we could also conceive a situation where we set the group borders so as to *maximize* the inequality consistent with the assumption of perfect sorting.

For most social table distributions, the assumption of perfect sorting greatly limits the possible Gini coefficients. An illustration of this is shown in Figure 1, which shows the Lorenz curve for a population of four groups.[5] The Lorenz curve plots cumulative population against cumulative income, and the area between the Lorenz curve and the 45-degree line is equal to the Gini coefficient of the population. When groups are perfectly sorted, the points $(0,0)$, (P_1, Z_1), ... are known; (P_i, Z_i) refers to the cumulative population and income of all groups up to group i. If there is no dispersion within groups, the Lorenz curve is given by the solid line, and the minimum Gini is the shaded area in the figure.

Now consider a set of within-group dispersions that preserves the perfect ordering of incomes by groups. The points (P_i, Z_i) still have to be on the Lorenz curve. Moreover, by the definition of the Lorenz curve, it must always be weakly

[3] There are of course substantial uncertainty inherent in compiling the tables. This goes for any pre-industrial data series, including wage and other price series, and will not be discussed further here.

[4] Analytical expressions will be detailed below; the "point distribution" Gini is equal to the between-group Gini, given in Equation (7).

[5] A related analytical proof for the case when group interval borders are given is found in Gastwirth (1972).

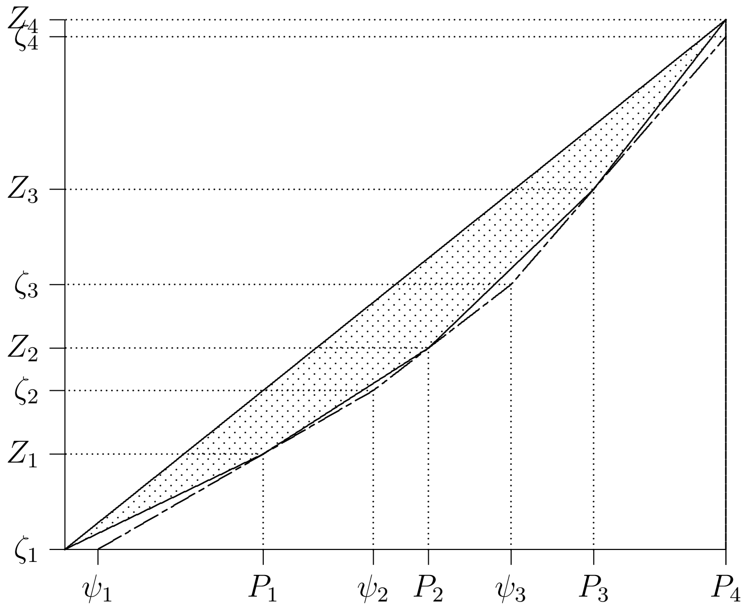


Figure 1: Lorenz curve and Gini coefficients for two restrictive assumptions

convex — the Lorenz curve plots population sorted by income, and the slope of the curve corresponds to the income of an individual at that point. It follows that the most outward-lying Lorenz curve is a series of straight lines going through the points (P_i, Z_i) with kinks somewhere between these points; an example of such a line is the dotted line in the figure. Correspondingly, the Gini coefficient can only go up by the area between the solid and dotted line.

The max-inequality Lorenz reflects a distribution where the population of a group is concentrated at the two extremes of the income groups' range; the richest individuals in group i have the same income as the poorest in group $i+1$. The position of these income and population points, denoted (ψ_i, ζ_i) in the figure, that gives the highest possible Gini is in general not easy to find in closed form. However, as is evident from the figure, for most distributions the scope of increasing the area between the solid and dotted lines is very limited, and becomes more so as the number of groups goes up.

The limitation of assuming perfectly sorted groups, if this does not correspond to known characteristics of the underlying population, is the main motivation for imposing within-group distributions that have overlaps between the income ranges of groups. This will be the topic of the next section.

2 Social tables and log-normal group distributions

2.1 The distribution of income within groups

To put some structure on the within-group dispersion of income, it will be assumed for the remainder of this paper that income within each social class is log-normally distributed. The log-normal distribution is commonly used to model income inequality. For a stochastic process with a given population, where relative changes in incomes are random, the central limit theorem yields a log-normal distribution for this population (see, for instance, Crow & Shimizu (1987, chap. 1), citing Gibrat (1930, 1931)). If group incomes are log-normally distributed, the corresponding theoretical justification is that while the conventional stochastic processes operate within groups, there is no mobility between groups. While somewhat stylized, this is a reasonable and easily understood assumption, in particular on historical data.

With log-normal distributions within groups, the aggregate distribution will not itself be log-normal. Rather, it captures the salient features of a presumably stratified society; the distribution shape will reflect the group data and its smoothness will depend on within-group dispersion. The log-normal distribution has mass along the entire positive income range; correspondingly, there will be overlap between groups and the Lorenz curve will pass to the right of the points (P_i, Z_i) in Figure 1.

The log-normal distribution is most conveniently expressed in terms of μ, the mean of log income, and σ, the standard deviation of log income.Denoting the mean income of a group as y_i and the standard deviation of the income as s_i, the expressions for these parameters are

$$\mu_i = \log(y_i) - \frac{1}{2}\log\left(1 + \left(\frac{s_i}{y_i}\right)^2\right) = \log(y_i) - \frac{\sigma_i^2}{2} \tag{1}$$

$$\sigma_i^2 = \log\left(1 + \left(\frac{s_i}{y_i}\right)^2\right) \tag{2}$$

The cumulative distribution function (cdf) is

$$F^L(x; \mu, \sigma) = \Phi\left(\frac{\log(x) - \mu}{\sigma}\right) \tag{3}$$

where $\Phi(\cdot)$ is the standard cumulative normal distribution, $\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} \exp\left(\frac{-t^2}{2}\right) \mathrm{d}t$.

Denoting the relative size of each group (social class) as p_i and the total number of groups as N, it follows that the aggregate cumulative income distribution function of the population is defined as

$$F(x) = \sum_{i=1}^{N} \left[p_i F^L(x; \mu_i, \sigma_i) \right] \tag{4}$$

where μ_i and σ_i are defined by (1) and (2).

2.2 Calculating Gini coefficients from group data

Following Aitchison & Brown (1957), the expression for the Gini coefficient for a log-normal distribution is given by $G = 2\Phi(\sigma/\sqrt{2}) - 1$. Extending their procedure to the case of many groups, the expression for the Gini coefficient is

$$G = \sum_{i=1}^{N} \sum_{j=1}^{N} p_i p_j \frac{y_i}{\bar{y}} \left(2\Phi \left(\frac{\mu_i - \mu_j + \sigma_i^2}{\sqrt{\sigma_i^2 + \sigma_j^2}} \right) - 1 \right) \tag{5}$$

where $\bar{y}$ is the population mean income, $\sum_{i=1}^{N} p_i y_i$.[6]

This expression has N^2 terms; two for each combination of i and j. Each of the terms considers a separate part of the Lorenz square;[7] group i's share of income $p_i y_i / \bar{y}$ (on the vertical axis) is multiplied with group j's share of population p_j (on the horizontal axis). If there was no overlap, these parts would be separate rectangles and constitute a grid; however, in this case, the areas should be considered as density functions over the entire square. Each of these areas are weighted by a number between -1 and 1, depending on the corresponding values of μ and σ for the two groups. The sum of these weighted squares is a measure of the distance between all individuals; the Gini coefficient.

The relative simplicity of the equation comes from two features of the log-normal distribution. First, multiplying a constant with a log-normally distributed variable returns another log-normally distributed variable. Second, the convolution of two log-normally distributed variables is itself log-normally distributed. Combining this with the definition of the Gini coefficient from the Lorenz curve, we find (5) as described in the Appendix.

[6] The calculation of Equation (5) is given in the Appendix, section A.1.1.

[7] The term "Lorenz square" refers to the square on which the Lorenz curve is plotted; the horizontal axis represent aggregate population, sorted from poorest to richest, while the vertical axis represent cumulative aggregate income.

As the expression (5) has many more terms than the number of groups, and some of the terms are negative, it is not straightforward to interpret the effect of different parameters on the resulting Gini coefficient. For this reason, it is more convenient to work with a re-formulated expression. First, replace the parameter μ with the group means, using (1).[8] Second, add each ij term to the corresponding ji term to get the preferred expression for the Gini coefficient

$$G = \underbrace{\sum_{i=1}^{N}\sum_{j=i+1}^{N} p_i p_j \left(\frac{y_j}{\bar{y}} \left[2\Phi\left(\frac{\log\left(\frac{y_j}{y_i}\right)}{\sqrt{\sigma_i^2+\sigma_j^2}} + \frac{\sqrt{\sigma_i^2+\sigma_j^2}}{2} \right) - 1 \right] - \frac{y_i}{\bar{y}} \left[2\Phi\left(\frac{\log\left(\frac{y_j}{y_i}\right)}{\sqrt{\sigma_i^2+\sigma_j^2}} - \frac{\sqrt{\sigma_i^2+\sigma_j^2}}{2} \right) - 1 \right] \right)}_{\text{Across-group inequality } (G_A = G_B + G_R)} + \underbrace{\sum_{i=1}^{N} p_i^2 \frac{y_i}{\bar{y}} \left[2\Phi\left(\frac{\sigma_i}{\sqrt{2}} \right) - 1 \right]}_{\text{Within-group inequality } (G_W)} \tag{6}$$

which is decomposed into across-group and within-group inequality.[9]

The first term of (6) is the sum of inequality across groups; all pairwise comparisons between individuals in group i and individuals in group j. We can contrast this to the Gini coefficient for no within-group dispersion, which is the population-weighted sum of all pairwise differences between the groups

[8] One could also substitute in s for σ, but this does not add clarity; as the Gini coefficient is a relative measure, the standard deviation only enters scaled, as s/y, and this can just as well be summarized in the σ measure.

The Gini coefficient expressed only in means and standard deviations is

$$G = \sum_{i=1}^{N}\sum_{j=1}^{N} p_i p_j \frac{y_i}{\bar{y}} \left(2\Phi \left(\frac{\log\left(\frac{y_i}{y_j}\right)}{\sqrt{\log\left[\left(1+\frac{s_i^2}{y_i^2}\right)\left(1+\frac{s_j^2}{y_j^2}\right)\right]}} + \frac{\sqrt{\log\left[\left(1+\frac{s_i^2}{y_i^2}\right)\left(1+\frac{s_j^2}{y_j^2}\right)\right]}}{2} \right) - 1 \right)$$

[9] The decomposition is analogous to that in Lambert & Aronson (1993) with the "between" and "residual" terms merged. It also corresponds to the classification suggested by Ebert (2010), who denotes G_A as the "between" component.

The analysis here is also related to Yitzhaki & Lerman (1991), who study the relationship between stratification and inequality. The aggregate group data can be construed as giving stratification but not inequality, and the Gini coefficients presented here measure stratification-induced inequality differences between populations.

$$G_0 = \underbrace{\sum_{i=1}^{N}\sum_{j=i+1}^{N} p_i p_j \left(\frac{y_j}{\bar{y}} - \frac{y_i}{\bar{y}}\right)}_{\text{Between-group inequality } (G_B)} \tag{7}$$

and see that the expressions are closely related. G_A differs from G_B in that the group means are modified by a number between -1 and 1; the evaluation of the $2\Phi(\cdot) - 1$ function.

The values for y and p in a given population is known from the social tables. The dispersion, however, is not. It is therefore of interest to know how the inequality of a population changes when dispersion changes - how G changes with s_i, or σ_i. From Equation (6), increases in G can be decomposed into increases in across-group inequality and increases in within-group inequality.

2.3 De-composing inequality effects

The across-group Gini is always increasing with group dispersion. Formally, this effect can be evaluated by taking the derivative of the across-group Gini by the dispersion measure of one or both groups. The derivative is always positive; an increase in dispersion will always increase the across-group Gini coefficient.[10] Because the log-normal distribution has positive mass across the entire income range, there is always *some* overlap; this is why the across-group term depends on σ even for small dispersions.

Milanovic (2002, p. 82-83) discusses the relationship between group means, group dispersions and income overlaps. He shows that for the overlap to be small, countries must either be very homogeneous internally (low within-group dispersion), or their mean incomes must be very far apart. Equation (6) allows for a formal discussion of this. Consider an increase in the dispersion of group j, and the mean pairwise income difference between individuals in group j and (the poorer) group i. If the groups did not overlap; there would be no change; the lower distance resulting from a decrease in the income of the poorer individuals would be exactly offset by the increase in the income of the richer individuals, as the mean

[10] The derivative with respect to $\sigma_i^2 + \sigma_j^2$ is

$$\frac{\partial G_A}{\partial\sqrt{\sigma_i^2+\sigma_j^2}} = \frac{y_j}{\bar{y}}\phi\left(\frac{\log\left(\frac{y_j}{y_i}\right)}{\sqrt{\sigma_i^2+\sigma_j^2}} + \frac{\sqrt{\sigma_i^2+\sigma_j^2}}{2}\right) + \frac{y_i}{\bar{y}}\phi\left(\frac{\log\left(\frac{y_j}{y_i}\right)}{\sqrt{\sigma_i^2+\sigma_j^2}} - \frac{\sqrt{\sigma_i^2+\sigma_j^2}}{2}\right)$$

The derivative with respect to σ_i or $c_i = s_i/y_i$ can then be found by the chain rule; this will not change the sign.

of group j is unchanged. With overlap, however, some of the poorest j-individuals are moving *away* from the richest i-individuals; the overlap makes the effect of increased dispersion greater. The degree of overlap is again influenced by the distance between groups ($\log\left(\frac{y_j}{y_i}\right)$) and the dispersion level ($\sigma_i^2 + \sigma_j^2$). This means that lower distance between groups increases the effect on the overlap term from increasing dispersion; groups that are close will have larger overlaps. The effect of changing dispersion is smaller for very large or very small dispersions; this reflects the bounding of the Gini coefficient to be between 0 and 1.

The last term in (6) is the sum of within-group Gini coefficients; a weighted sum of the Gini coefficients for log-normal distributions as reported by Aitchison & Brown (1957). It is straightforward to see that the within-group Gini increases with dispersion. As within-group pairs constitute a relatively small part of all possible pairs, the weights are low; for small groups, the squaring of the population share means that the resulting inequality contribution is low.

Returning to the aggregate Gini coefficient, it is useful to verify that Equation (6) takes on familiar values at the extremes of dispersion. First, consider a situation where within-group dispersion approaches zero: $\sigma_i \to 0$; in that case, the across-group Gini collapses to the between-group Gini (7) as both Φ functions are evaluated at plus infinity. Similarly, we can consider a situation where dispersion approaches infinity; in that case, as $\sigma \to \infty$, the Φ evaluations on y_j and y_i are evaluated at plus and minus infinity, respectively. The Gini coefficient approaches $\sum_{i=1}^{N}\sum_{j=1}^{N} p_i p_j y_i/\bar{y}$, which sums to 1; full inequality.

2.4 Finding within-group dispersions

From the discussion above we now know that when group distributions are log-normal, we can calculate aggregate and composite inequality measures in closed form, given group sizes, means and standard deviations. The standard deviations are not in the social tables. Because of this, we have to make a case for the "correct" level of within-group dispersion in each case to calculate aggregate inequality. In the following, dispersions will be described in terms of coefficients of variation, $c = s/y$. It will be assumed that coefficients of variations are constant across groups; that standard deviations are proportional to group income.[11]

The true level of within-group inequality is not known. The following paragraphs discuss two possible ways of inferring reasonable ranges for c. In Section 3, a wide range of parameters will be examined.

[11] Alternatively, one could assume a linear relationship between (scaled) standard deviations and group means of the form $(s_i/\bar{y}) = \alpha(y_i/\bar{y})^{\beta}$ (the method used here corresponds to $\beta = 1$). This is discussed in the Appendix, and does not affect the results.

Within-group dispersion in modern societies

The first approach for finding within-group dispersion parameters is to look at modern data. Census or other survey data often include income data, as well as several characteristics that makes it possible to group the population into "social classes" corresponding to the social tables. Using data from the International Integrated Public Use Microdata Series (Minnesota Population Center, 2010), the coefficient of variation of income can be calculated for groups based on occupation, industry and employment class. For nine developed and developing countries between 1970 and 2007, a summary of the group data is given in Table 2.[12]

Classification	Mean of c_{min}	Mean of c_{median}	Mean of c_{max}	Mean # of groups
Occupation	1.0	1.3	3.1	9.4
Industry	0.9	1.5	2.9	13.9
Empl.classification	1.5	2.0	6.0	2.7
Empl.class (detailed)	1.1	1.7	6.1	5.8

Table 2: Within-group inequality (coefficient of variation) in modern societies

The range of variation coefficients is not large. Comparing the dispersion in the most and least diverse groups, for less than half of the country-years is the former more than three times the latter. Moreover, the the mean and minimum of the dispersion of groups are quite similar. The median within-group coefficient of variation is between 0.7 (Canada, 1981) and 4.8 (Mexico, 2000), with most being around 1. There is no clear relationship between development status and dispersion, though the groupings by "employment class" consistently yield higher dispersions than the other two groupings. In any case, Gini coefficients of pre-industrial inequality should be calculated for dispersions (coefficients of variations) somewhere in the range between 1 and 2. Both of these will be used in the following section.

Well-apportioned groups

In addition to inference from modern data, we can draw a restriction on the coefficient of variation from the group structure in the social tables, by saying that the weighted sum of within-group Ginis should not be larger than the between-group Ginis. This could be justified by saying that groups should be "well-apportioned"; for a group to have a separate identity when tabulating incomes, the differences within the group should be less than the differences across the groups. The level of dispersion consistent with this well-apportioned assumption will be denoted c_w;

[12] The countries are Brazil, Canada, Colombia, Mexico, Panama, Puerto Rico, South Africa, United States and Venezuela. A fuller exposition is given in the Appendix, table A.9.

it will differ across societies (it is derived from the group means and sizes). To calculate c_w, insert for the definition of σ (2) and the dispersion structure in the expression for within-group inequality in (6), and equate the average within-group dispersion with the between-group Gini.

The standard deviation of logs becomes $\sigma_w = \sqrt{2}\Phi^{-1}\left(\frac{G_0+1}{2}\right)$. Inserted in (5), we get the expression for the upper bound on the Gini coefficient consistent with well-apportioned groups:

$$G_{\text{“well-apportioned”}} = \sum_{i=1}^{N}\sum_{j=1}^{N} p_i p_j \frac{y_i}{\bar{y}} \left[2\Phi\left(\Phi^{-1}\left(\frac{G_B+1}{2}\right) + \frac{\log\left(\frac{y_i}{y_j}\right)}{2\Phi^{-1}\left(\frac{G_B+1}{2}\right)} \right) - 1 \right] \tag{8}$$

where G_B is given by Equation (7); that is, the expression depends only on the means and group sizes in the original data. For a simple back-of-the envelope calculation of inequality comparison across societies, Equation (8) is a good candidate. It can be seen as an upper bound of dispersion by making the following claim: if within-group dispersion was really bigger than c_w, the compiler of the table would not have chosen the groups in this way, as they do not add to the "structuring" of information about the society.

3 Re-evaluating pre-industrial inequality

With the methodology in place, pre-industrial inequality can be re-evaluated using the social table data compiled by Milanovic *et al.*. The overall level of inequality goes up by a large amount when within-group inequality is accounted for. In addition, changing dispersion also affects how we rank the various societies in terms of inequality.

Seven different sets of assumptions on within-group dispersion will be illustrated. The first and second set are the measures used by Milanovic *et al.*. Their "Gini1" assumes no within-group inequality — this is the "point distributions" discussed above — and is equal to the between-group Gini coefficient.[13] The "Gini2" variable is the inequality associated with within-group inequality and perfect group sorting, for given group interval borders, as described by Kakwani (1980, chap. 6). While Gini1 corresponds to $c = 0$, Gini2 does not map into the methodology used in this paper.

For the groupwise log-normal distributions, the coefficient of variation will be

[13] The between-group Gini, G_B, can be calculated by Equation (7).

assumed constant across groups.[14] The values for c shown here will be $0.1, 0.5, 1$ and 2, covering most of the range discussed above. There will also be an assumption set with "well-apportioned" groups, where the within-group Gini coefficients are equal to the between-group coefficients. These differ between populations, as the estimates are calculated from group means and sizes, but are still constant across groups within each population.[15] The assumption sets used are summarized in Table 3.

#	Within-group dispersion	Var. coeff c	Var. of log $\sigma^2 = \log(1+c^2)$	Gini within groups $G_i = 2\Phi(\sigma/\sqrt{2}) - 1$
1	None (MLW "Gini1")	0	0	0
2	Perfect sorting (MLW "Gini2")	-	-	-
3	Very low	0.1	0.01	0.06
4	Low	0.5	0.22	0.26
5	Intermediate	1	0.69	0.44
6	High	2	1.61	0.63
7	"Well-apportioned"	c_w	-	-

Table 3: Assumptions on within-group dispersions

3.1 The level of inequality in pre-industrial societies

The Gini coefficients increase significantly when within-group dispersion is accounted for. Figure 2 shows how the calculated Gini coefficients are sensitive to assumptions on within-group dispersion. The Gini estimates used by Milanovic *et al.* ("Gini1" and "Gini2") span only a small range of the possible values. Even the low coefficient of variance assumption of $c = 0.1$ gives higher Gini estimates for all but eight populations; increasing c to 0.35 leaves only Moghul India with higher Gini2. Like other populations with few groups, Moghul India has a large group containing the majority of the population; unlike the other populations, however, this group is not the poorest, and the income distance to the richer and poorer groups is relatively high. This allows for high inequality while preserving the assumption of no overlap. In the terms of Figure 1, the data points for Moghul India allow a large distance between the solid and dotted line, while for the other populations, this space is very small.

From Section 2.4 above, we know that the most coherent modern-day social groups have coefficients of income variations between .5 and 1. Using the still

[14] Most results hold up to other linear relationships between s_i and y_i. This is detailed in the Appendix, Table A.5 and Figure A.1.

[15] See Equation (8) for the calculation of the well-apportioned groups.

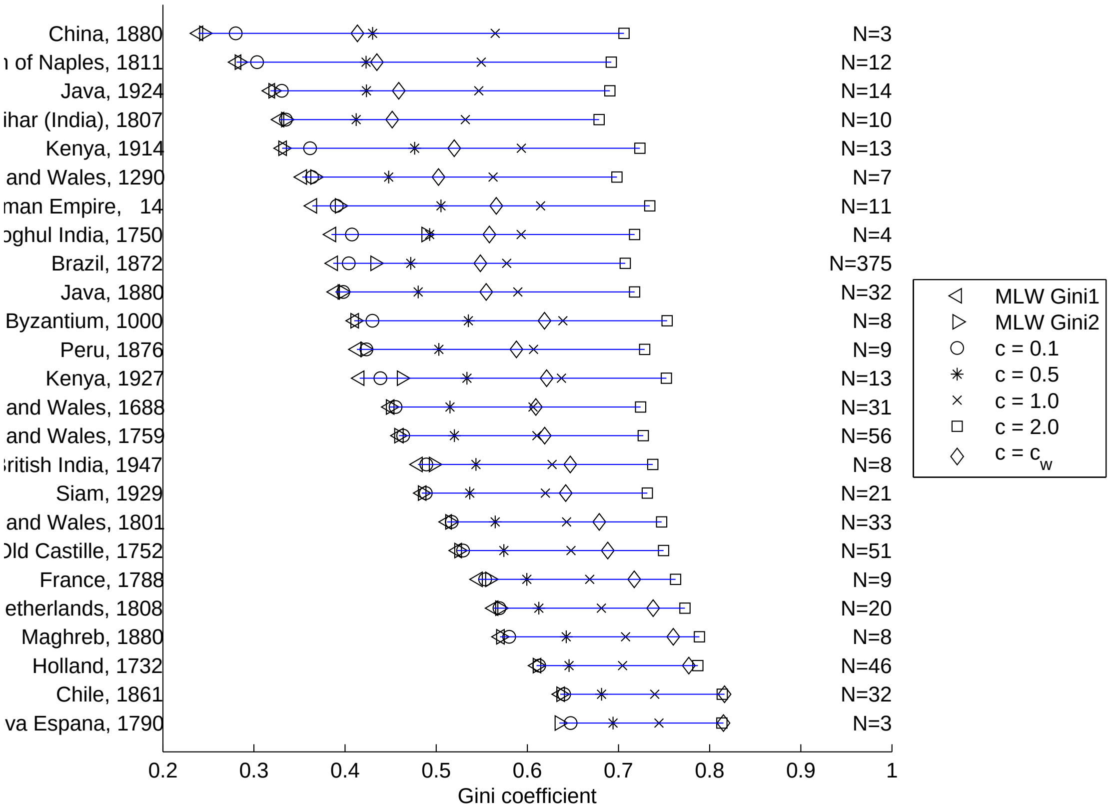


Figure 2: Comparison of Gini coefficients for the seven assumption sets

low value of $c = 0.5$, the calculated Gini coefficients for all the pre-industrial populations are higher than the Gini2 value. Further increasing within-group dispersion to $c = 2$, all Gini coefficients are higher than 0.7; very high inequality by any standard.

There is some change in sorting as c increases. At $c = 0.5$, around 7 per cent of all pairwise comparisons of societies change; at $c = 2$ this number has increased to 13 per cent. Above $c = 2$ the re-shuffling does not increase much more.[16] For the societies with higher between-group inequality, that is, the lower half of Figure 2, the sorting of societies is almost perfectly preserved — for example, by all measures, England and Wales in 1759 was just a little bit more unequal than in 1688. Hence, we can conclude that while the level of inequality is very sensitive to assumptions on within-group dispersions, the ranking is not.

With a large within-group dispersion measure, $c = 2$, calculated Gini coefficients are in some cases more than twice as large as the benchmark values. If inequality in these societies was this high, the value of the social tables data is low, as we would expect there to be variation in dispersion between populations, making it harder to rank the societies with respect to each other.

It could be a source of concern if the Gini coefficient of a population was highly dependent on the number of groups in that population. On the one hand, a high number of recorded groups could reflect a highly stratified society with corresponding inequality. On the other hand, we must assume that the number of recorded groups also reflects some pragmatism on the associated (often contemporary) researcher's part, with respect to how much data it is possible to collect. In any case, there is not a high correlation between the number of groups and the Gini estimates; for all estimation sets, linear OLS regression does not yield a significant slope parameter.[17]

To sum up, there are two main messages from Figure 2. First, the level of pre-industrial Gini coefficients is in general sensitive to assumptions on within-group dispersions. Second, the ordering of societies with respect to each other experiences some changes; around 10% of all compared pairs change order when the coefficient of variation within groups goes from 0 to 1. This is a relatively low number, and as will be seen in Section 4, it does not affect the relationship between growth and inequality.

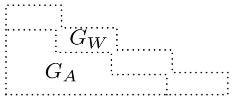
[16] For comparison, the expected change in pairwise sorting for random data sets is around 1/2 (50%).

[17] This holds regardless of whether Brazil 1872, with 375 groups, is included in the regression. See Appendix, section A.2.6.

3.2 The contributions of subgroups to inequality

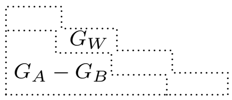
As discussed in the previous section, the increase in inequality comes both from inequality within and across groups. Using Equation (6), we can look at the contributions of group pairs to inequality. From each pair of groups, we get the weighted sum of pairwise income differences between individuals of the groups. As an example, consider the social table for Byzantium, AD 1000, as given in Table 1. A Gini decomposition based on group pairs, with within-group dispersion at $c = 1$, is given in Table 4.

All Gini components ($G_A + G_W$)



	$i=1$	$i=2$	$i=3$	$i=4$	$i=5$	$i=6$	$i=7$	$i=8$
$j=1$	3.4							
$j=2$	0.4	0.0						
$j=3$	10.1	0.5	7.3					
$j=4$	0.8	0.0	1.2	0.0				
$j=5$	0.3	0.0	0.4	0.0	0.0			
$j=6$	3.1	0.2	4.4	0.2	0.1	0.2		
$j=7$	1.3	0.1	1.8	0.1	0.0	0.1	0.0	
$j=8$	10.3	0.6	14.5	0.8	0.3	0.9	0.3	0.1

"Within" and "overlap" terms ($G_A - G_B + G_W$)



	$i=1$	$i=2$	$i=3$	$i=4$	$i=5$	$i=6$	$i=7$	$i=8$
$j=1$	3.4							
$j=2$	0.4	0.0						
$j=3$	9.1	0.5	7.3					
$j=4$	0.4	0.0	0.6	0.0				
$j=5$	0.1	0.0	0.2	0.0	0.0			
$j=6$	0.1	0.0	0.2	0.0	0.0	0.2		
$j=7$	0.0	0.0	0.0	0.0	0.0	0.1	0.0	
$j=8$	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.1

Table 4: Example of group pair contributions, Byzantium, AD 1000.

The upper panel shows the entire Gini coefficient. The diagonal is the within-group Gini components; these would all be zero if there was no within-group dispersion. The other cells in the upper panel are the across-group components. Because groups are weighted by products of group sizes and incomes, small groups only add to inequality if differences *between* groups are very big. The lower row

($j = 8$) gives the contributions from the "nobility" group with very high income; because the difference from other groups is so big, interactions with this group contribute greatly to inequality. The most sizable contributions come from the interaction of the very small, very rich mobility group ($j = 8$) with the two poor, very big tenant and farmer groups ($i = 1, i = 3$). The sum of all the cells in the upper panel is the total Gini coefficient for this population, given a within-group coefficient of variation of 1.

Most of the large effects from group income differences come from the differences between group means, and are as such contained in the between-group Gini (G_B). The lower panel subtracts the between-group components,[18] giving the additions to inequality that arise solely from within-group dispersions.

When the between-group inequality is subtracted, nearly all contributions to inequality from the upper groups disappear. Within-group Gini coefficients, in particular for $i = 1$ and $i = 3$, the largest groups, contribute a total of 11 Gini points to the total Gini.[19] In this case, however, the across-group contribution is even more important. Inequality across farmers (group 1) and tenants (group 3) - large groups that have means close together - is particularly evident. This combination adds 9.1 points to a total Gini coefficient of 64 — nearly half the increase from the between-group Gini of 41. This highlights the restriction an assumption of perfect sorting places on inequality. As the means are so close, any perfectly sorted within-group distribution would have both these groups compressed over a very short income range.

Table 5 shows the decomposition of the increase in inequality for all the societies. For no within-group dispersion ($c = 0$), by construction, the within-group Gini is zero and the across-group component is equal to the between-group component. As c increases, both components go up; with many groups, more of the increase is in across-group inequality, as more of the possible pairs of people are in separate groups. Some populations are clear outliers. For example, the social table for China has nearly all the population in the poorest group, and hence the "within" term of this group accounts for nearly the entire increase in G for high c. For Chile, the difference between group means is so big that increasing within-group dispersion has a less pronounced effect on both components. And for Naples, where group means are close, nearly all the increasing inequality is from increases in the across-group component.

[18] G_B is given in Equation (7).

[19] Throughout the text, Gini coefficients will be scaled to be between 0 and 100; a "Gini point" refers to a change of 1 in this measure.

		$c = 0$		$c = 1$		$c = 2$	
		G_A	G_W	G_A	G_W	G_A	G_W
Roman Empire, 14	$(N = 11)$	36	0	41	21	44	29
Byzantium, 1000	$(N = 8)$	41	0	53	11	60	16
England and Wales, 1290	$(N = 7)$	35	0	49	8	59	11
England and Wales, 1688	$(N = 31)$	45	0	58	3	68	4
Holland, 1732	$(N = 46)$	61	0	69	1	77	2
Moghul India, 1750	$(N = 4)$	39	0	42	17	47	25
England and Wales, 1759	$(N = 56)$	46	0	60	2	71	2
Old Castille, 1752	$(N = 51)$	52	0	63	2	72	3
France, 1788	$(N = 9)$	55	0	62	5	69	7
Nueva Espana, 1790	$(N = 3)$	63	0	64	10	67	14
England and Wales, 1801	$(N = 33)$	51	0	61	3	70	4
Bihar (India), 1807	$(N = 10)$	33	0	49	5	61	6
Netherlands, 1808	$(N = 20)$	56	0	64	4	71	6
Kingdom of Naples, 1811	$(N = 12)$	28	0	52	3	64	5
Chile, 1861	$(N = 32)$	64	0	71	3	78	4
Brazil, 1872	$(N = 375)$	40	0	56	2	68	3
Peru, 1876	$(N = 9)$	41	0	53	7	63	10
China, 1880	$(N = 3)$	24	0	24	32	25	46
Java, 1880	$(N = 32)$	39	0	53	6	63	9
Maghreb, 1880	$(N = 8)$	57	0	63	7	68	10
Kenya, 1914	$(N = 13)$	33	0	36	24	39	33
Java, 1924	$(N = 14)$	32	0	49	6	61	8
Kenya, 1927	$(N = 13)$	42	0	46	17	51	25
Siam, 1929	$(N = 21)$	48	0	60	2	70	3
British India, 1947	$(N = 8)$	48	0	56	7	64	9

Table 5: Gini coefficients decomposed for different levels of within-group dispersion

3.3 Robustness checks: Removing inequality at the top and bottom

The standard deviations used in the discussion above are based on conjecture, and until better data is available, this will continue to be the case. Other objections could be raised to the use of log-normal distributions. The next paragraphs address two of these.

Richer groups

For the richer income groups of historical inequality data (the upper social classes), we often have more detailed information on group structures. Hence, imposing the log-normal distribution, with positive mass across the entire income spectrum and a left-skewed distribution, might be harder to accept for these groups.

However, these upper groups are typically small, and it turns out that the contribution to aggregate inequality from dispersion within these groups is also small. As an example, consider the decomposition illustration of Table 4.

As is seen in the left column of the upper panel, the contributions to overall Gini from the richest group ($j = 8$) are substantial, even though it only consists of one per cent of the total population. However, all of this contribution comes from the difference in group means, which is present before the within-group dispersion is introduced. If we remove the between-group inequality, and move to the lower panel, it is clear that the contribution of the upper group is very low. As there is almost no overlap with the other groups, and the population of the richest group is low, the contribution of the richest group to the increased dispersion is almost zero.

Similar exercises can be conducted for the other social tables. Counting the "inequality contribution" from a group as all terms in (6) that include the group, we can check how much the richer groups contribute to overall inequality. Taking as the threshold any groups with a mean income of more than five times the population mean, and using the assumptions of $c = 1$, the result of this accounting exercise shows that there are no large contributions by the rich groups.[20] Even for the cases where these groups make up a considerable size of the population (they are largest in France and New Spain), the contribution from these groups only make up a small factor of the inequality that is added by within-group dispersion. It follows that removing the assumption of log-normal distributions within groups for the richer groups would not significantly alter the results in this paper.

[20] The table is given in the Appendix: Table A.6

Poorer groups and subsistence minima

Log-normal distributions have positive mass across the entire positive income range. Hence, by assuming such distributions within groups, we postulate that many people are very poor. However, some positive income level needs to be fulfilled in order to survive - the *subsistence income*. If we believed that everyone, at all times, lived at or above subsistence, we would have to revise our assumptions on within-group distributions. Inequality-limiting subsistence is one the key messages of Milanovic *et al.* (2011). As an example, the mean income of "Agricultural day laborers and servants" in France 1788 was 312 PPP dollars a year. With subsistence income at 300 dollars (as assumed in their paper), most people in that group (covering 36 per cent of the population) must have had incomes very close to the mean.

There is no need to assume that the subsistence border holds with absolute certainty; indeed, there is ample historical evidence to suggest that large groups have been living below subsistence level for long periods of time. A notable example is given in Clark (2008, chapter 6), where the Malthusian period is described as a situation with "social mobility and the survival of the richest". In pre-industrial England, according to Clark, poor families on average did not replace their population, while rich families did; consequently, there was continuous social mobility downward. However, it is not unlikely that subsistence income plays *some* role in truncating income distributions at the bottom, and it is useful to see how the results presented would change if the income of everyone was above subsistence minimum. In order to explore the effect on inequality on imposing subsistence minima, the setup of Section 2 is altered in three ways, using the assumption $c = 1$.

The first two adjustments keep the same log-normal distributions, but alter them at the tails. For the first adjustment, any population below the subsistence minimum is simply shifted up to the subsistence minimum. This reduces inequality at the lower end, but skews group means, as the same group-wise log-normal distributions are kept for the rest of the population. The second adjustment addresses this by also shifting the richest part of the population in each group down to a "group upper bound", in such a way as to keep group means at the pre-adjustment levels.

The final adjustment is of a different type. Instead of defining the log-normal distribution on the entire positive income scale (starting at 0), it is defined over the scale starting at y_{min}. This means that there is no population mass below y_{min}. In practice, this amounts to subtracting y_{min} from all group means before calculating the log-normal distributions, and then right-shifting these distributions by y_{min}.

For each of these three adjustments, the aggregate Gini coefficients are recalculated. The calculation is done using numerical methods, calculating all pairwise differences in a discrete (but very fine-grained) population space.[21]

An adjustment by minimum incomes does shift the Gini estimates down for several populations, while others are virtually unchanged. Three populations stand out with large corrections: Byzantium and the two Kenya observations. All of these three have rather low population mean incomes, making the minimum income more quantitatively important; the population mean in Kenya 1914 is only 50% above minimum. Here, the same subsistence income is used for all populations; one could argue that the subsistence level is lower in tropical areas. If subsistence income in Kenya is actually lower, the downward revision of the Gini coefficient would be less.

A strong downward change in the Gini is expected across the line, as assumptions of no population mass below minimum income correspond directly to assumptions of very low within-group inequality at the bottom of the income distribution. The fact that substantial inequality (inequality above G_B) remains even after such an extreme revision shows that group overlap always needs to be accounted for when using group data, even if one adheres strongly to limiting subsistence incomes.

4 Inequality and economic growth

4.1 The Kuznets curve

Milanovic *et al.* (2011) do not discuss the evolution of inequality with economic activity except for the hypothesis on the relationship between subsistence income and feasible inequality. However, with the data available, and the framework for interpolating inequality in place, we have the opportunity to re-visit the broad sweeps of Kuznets and van Zanden: that inequality should increase with economic activity in pre-industrial societies. Moreover, the differences between pre-industrial and modern inequality can be assessed.

Estimates on GDP per capita have been made available by Maddison (2010) and can be used for a simple linear regression of inequality on economic activity.[22] By the conventional view of the Kuznets curve, we should expect an increasing relationship between GDP per capita and inequality as measured by the Gini coefficient. Figure 3 plots these variables against each other for the different assumption sets; the dotted lines represent results of linear regressions for each of the sets, as detailed in Table 6, and each set of symbols correspond to one set of assumptions on within-group dispersion.

[21] A full description of the adjustments, as well as a table of results, is shown in the Appendix, see Table A.7.

[22] The GDI per capita estimates are those used by Milanovic *et al.* (2011); Table 1, p. 7.

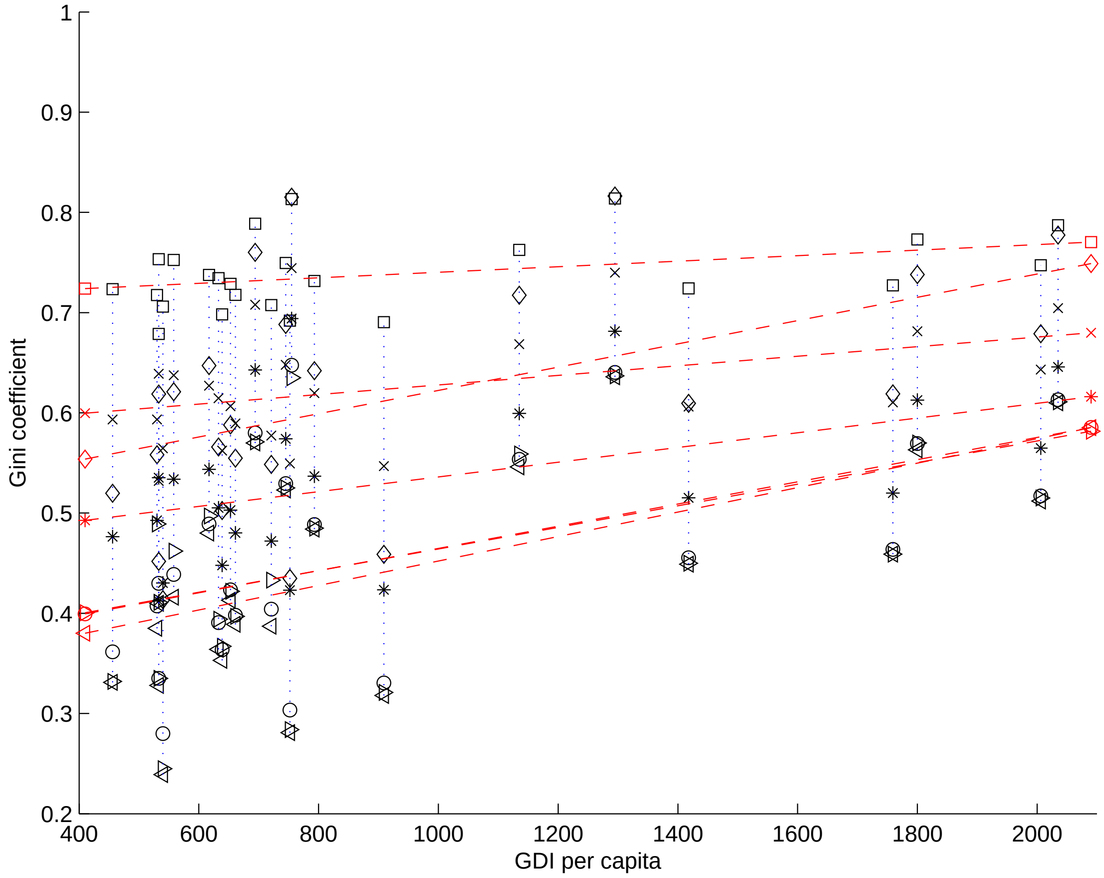


Figure 3: The relationship between Gini coefficients and GDI per capita for various assumptions

	Dependent variable: Gini coefficient						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Const	33.008***	35.630***	35.389***	46.253***	58.000***	71.276***	50.610***
	(4.087)	(4.134)	(3.918)	(3.178)	(2.350)	(1.490)	(4.467)
GDI/capita	0.012***	0.011**	0.011***	0.007**	0.005**	0.003*	0.012**
	(0.004)	(0.004)	(0.004)	(0.003)	(0.002)	(0.001)	(0.004)
R^2	0.298	0.244	0.275	0.203	0.164	0.139	0.243
N	25	25	25	25	25	25	25

Table 6: Regression results under assumption set (1)-(7); see Table 3 for definitions of the assumptions. {***,**,*}=significant at {99%,95%,90%} level (two-sided tests)

There is a positive relationship between inequality and economic activity level for all seven sets of Gini coefficient estimates. For the very high within-group dispersion (regression 6), significance does not hold at the 95% level; for all the others, it does. The results do not change when using logarithmic specifications.

Coefficients of .01 correspond to an increase of one Gini point (where the Gini is scaled between 0 and 100) when GDP increases by 100 (the ranges of Gini coefficients and GDP are shown in Figure 3). The point estimates are lower for higher dispersions. The highest estimate says that an increase in GDP per capita from 1400 to 1750 — roughly the increase in Great Britain in the century before the industrial revolution — would increase the Gini coefficient by 4.2 points.

The results confirm the hypothesis of Van Zanden (1995) that inequality was increasing well before the Industrial Revolution. The results are robust to splitting the dataset into West Europe and the rest of the world.

4.2 Pre-industrial vs. modern inequality

While the relationship between pre-industrial inequality estimates are robust to changes in within-group dispersions, comparison with modern data is not. Figure 4 compares the pre-industrial inequality ranges to modern inequality observations, as used by Milanovic *et al.* (2011). The upper two panels compare Western countries. As is evident, both for the lower and higher ranges, pre-industrial inequality was higher than modern inequality, with only a slight overlap where the United States and Italy are more unequal than some of the lower estimates for pre-industrial inequality.

Now turn to the lower two panels, comparing pre-industrial and modern inequality for today's developing countries. It is clear that now the assumptions matter much more. Milanovic *et al.* state that "inequality differences within the pre-industrial and modern samples are many times greater than are differences between their averages". Their data implies that the most unequal societies today — South Africa and Brazil — have much higher inequality than most pre-industrial societies. From the figure it is evident that such a conclusion depends on specific assumptions on the within-group dispersions. With the uncertainty implied by the lower left panel, it is hard to compare the pre-industrial developing countries to their modern counterparts without more information on within-group dispersion.

Another interesting exercise is to compare the pre-industrial Western economies (upper left panel) to modern developing countries (lower right). The GDI per capita of the poorer developing countries is in the same range as the European countries of the eighteenth century. Still, for most societies and for most sets of assumptions, the European countries were more unequal then than the developing countries are now. There are many possible explanations for this; such as more equal land ownership, better food production technologies, or improvement of

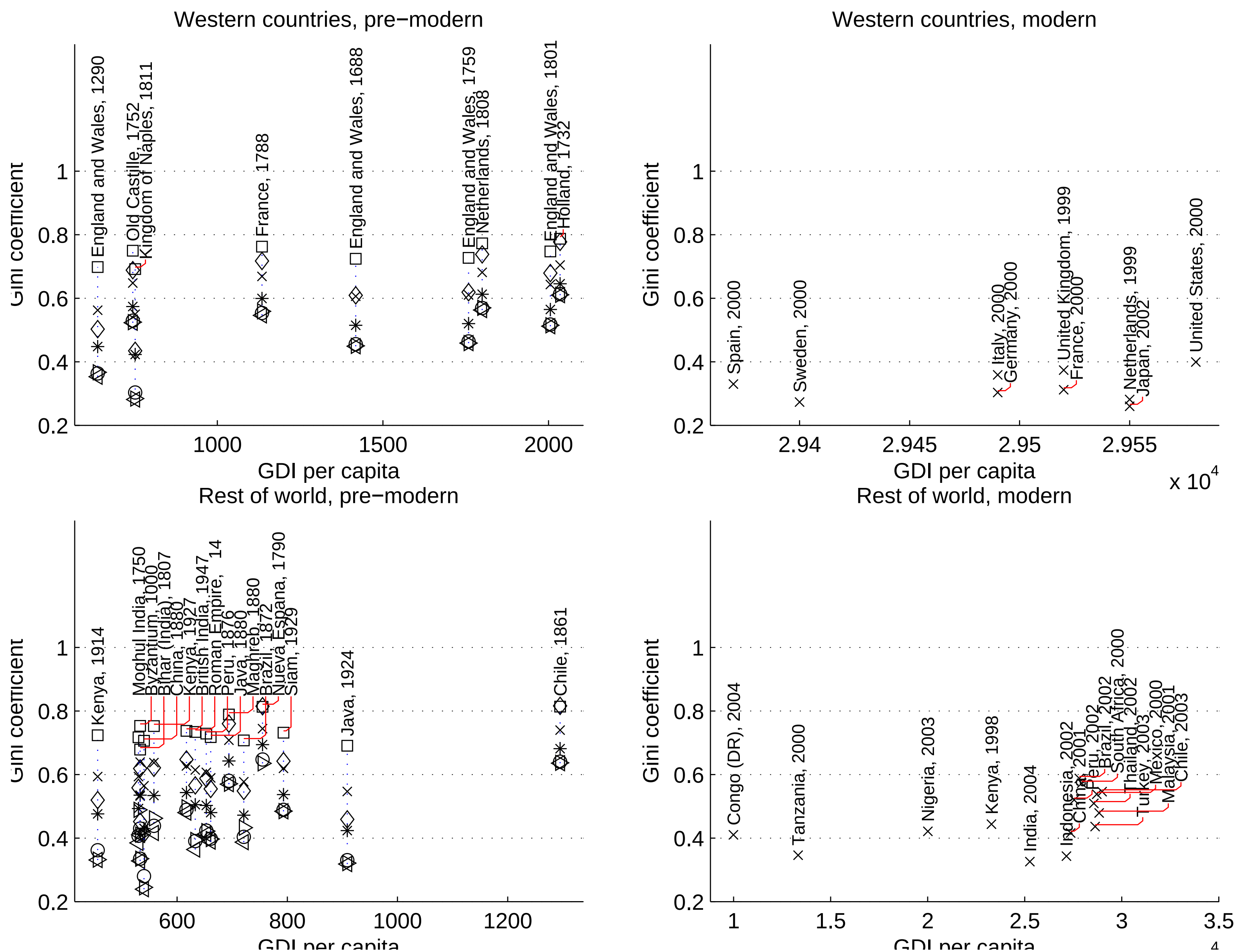


Figure 4: The relationship between Gini coefficients and GDI per capita; pre-modern and modern

democracy.

Given the wide range of variance assumptions used here, we can conclude that, barring any better country-specific data on within-group inequality, the data confirms both increasing inequality in pre-industrial societies and decreasing inequality after industrialization, at least for today's developed countries. One explanation for the lagging decrease in inequality in developing countries could be that the industrialization (or rather, modernization) process has not come far enough; they are still on the upward-sloping part of the Kuznets curve, or just past the peak.

These result are not dramatically different from previous studies on long-run inequality developments. The study of Bourguignon & Morrisson (2002) on world inequality after 1820 aims to cover the whole world by interpolating between countries and years where data is known. They find a hump-shaped profile of within-country (or rather, within "group of country") inequality in Western countries, with a peak in the late nineteenth century. For developing countries their data contains much more interpolation; the least prosperous country groups have increasing within-country inequality for the whole period. Somewhat similar results are found by Baten *et al.* (2009), using similar data but supplementing it with data on unskilled wages.

Economic historians now tend to tone down the emphasis on the industrial development on the nineteenth century and focus more on longer development arches (Crafts & Harley, 1992), which increases the need for broad empirical evidence on the "pre-industrial" period. Moreover, economists are becoming increasingly interested in the long-run interrelationships between inequality and investment. Several particular institutional developments have been postulated to be influenced by inequality; see for example Sokoloff & Engerman (2000) on the difference in development paths between North and South America and Galor & Moav (2006) and Galor *et al.* (2009) on the development of modern education systems. Getting better historical inequality data, and making better use of what is available, is critical to evaluate these claims. In that sense, it is reassuring that the main results on growth and inequality are shown to be robust. Further work should be able to explore in more detail the implications of inequality differences between the countries analyzed here.

5 Conclusion

This article has shown that when accounting for within-group inequality in social tables, reported inequality goes up by a large amount. The increase comes from both within- and across-group inequality, and is particularly important in the case where groups are large and have means that are close to each other.

Moreover, the non-monotone link between growth and inequality is confirmed;

the data supports the long-run development of inequality as proposed by Kuznets (1955) and Van Zanden (1995). The result that inequality increases with modernization, only to decrease with development levels seen in the West over the last century, is robust to changing assumptions on within-group inequality.

With further research, we can expect to see more tabulations of income and wealth data from pre-industrial societies. For statistics of a social table format, where within-group dispersion is not given, this paper presents a straightforward, transparent way of calculating inequality. The method can also be useful for modern data. While nation-wide distribution data now exist for most countries, within-group data is frequently missing for subnational entities or social classes. The approach presented in this paper can be used in these cases, to put structure on and properly evaluate any type of incomplete data on income or wealth distributions.

References

Aitchison, J., & Brown, J. A. C. 1957. *The Lognormal Distribution, with special reference to its uses in economics*. 1st edn. Cambridge University Press.

Baten, Joerg, Foldvari, Peter, van Leeuwen, Bas, & van Zanden, Jan L. 2009. World Income Inequality 1820-2000. *Mimeo*.

Bourguignon, François, & Morrisson, Christian. 2002. Inequality among World Citizens: 1820-1992. *American Economic Review*, **92**(4), 727–744.

Clark, Gregory. 2008. *A Farewell to Alms: A Brief Economic History of the World*. Princeton University Press.

Crafts, N. F. R., & Harley, C. K. 1992. Output growth and the British industrial revolution: a restatement of the Crafts-Harley view. *Economic History Review*, **45**(4), 703–730.

Crow, Edwin L., & Shimizu, Kunio. 1987. *Lognormal Distributions: Theory and applications*. 1 edn. CRC Press.

Ebert, Udo. 2010. The decomposition of inequality reconsidered: Weakly decomposable measures. *Mathematical Social Sciences*, **60**(2), 94–103.

Galor, Oded, & Moav, Omer. 2006. Das Human-Kapital: A Theory of the Demise of the Class Structure. *Review of Economic Studies*, **73**(1), 85–117.

Galor, Oded, Moav, Omer, & Vollrath, Dietrich. 2009. Inequality in Landownership, the Emergence of Human-Capital Promoting Institutions, and the Great Divergence. *Review of Economic Studies*, **76**(1), 143–179.

Gastwirth, Joseph L. 1972. The Estimation of the Lorenz Curve and Gini Index. *Review of Economics and Statistics*, **54**(3), 306–316.

Hoffman, Philip T., Jacks, David S., Levin, Patricia A., & Lindert, Peter H. 2002. Real Inequality in Europe since 1500. *Journal of Economic History*, **62**(2), 322–355.

Kakwani, Nanak. 1980. *Income Inequality and Poverty: Methods of Estimation and Policy Applications*. A World Bank Research Publication.

Kuznets, Simon. 1955. Economic Growth and Income Inequality. *American Economic Review*, **45**(1), 1–28.

Lambert, Peter J., & Aronson, J. Richard. 1993. Inequality Decomposition Analysis and the Gini Coefficient Revisited. *Economic Journal*, **103**(420), 1221–1227.

Lindert, Peter H. 2000. Three centuries of inequality in Britain and America. *Handbook of Income Distribution*, **1**, 167–216.

Maddison, Angus. 2010 (Feb.). *World Population, GDP and Per Capita GDP, 1-2003 AD [Computer file].*

Milanovic, Branko. 2002. True world income distribution, 1988 and 1993: First calculation based on household surveys alone. *Economic Journal*, **112**(476), 51–92.

Milanovic, Branko. 2006. An estimate of average income and inequality in Byzantium around year 1000. *Review of Income and Wealth*, **52**(3), 449–470.

Milanovic, Branko, Lindert, Peter H., & Williamson, Jeffrey G. 2007. Measuring Ancient Inequality. *National Bureau of Economic Research Working Paper Series*, **13550**(Oct.).

Milanovic, Branko, Lindert, Peter H., & Williamson, Jeffery G. 2011. Pre-Industrial Inequality. *Economic Journal*, **121**(551), 255–272.

Minnesota Population Center. 2010. *Integrated Public Use Microdata Series, International: Version 6.0 [Machine-readable database].*

Sokoloff, Kenneth L., & Engerman, Stanley L. 2000. History Lessons: Institutions, Factors Endowments, and Paths of Development in the New World. *Journal of Economic Perspectives*, **14**(3), 217–232.

Van Zanden, Jan L. 1995. Tracing the Beginning of the Kuznets Curve: Western Europe during the Early Modern Period. *Economic History Review*, **48**(4), 643–664.

Yitzhaki, Shlomo, & Lerman, Robert I. 1991. Income stratification and income inequality. *Review of Income and Wealth*, **37**(3), 313–329.

A Appendix

A.1 Calculations of expressions

A.1.1 Calculation of Equation (5)

This section shows the derivation of Equation (5), using the definition of the Gini coefficient as the area below the Lorenz curve. The calculation is an extension of Aitchison & Brown (1957)'s one-group case, and makes use of some convenient properties of the log-normal distribution.

Denote the log-normal population density functions as $f(x;\mu_i,\sigma_i^2)$ and the corresponding CDF as $F(x;\mu_i,\sigma_i^2) = \int_0^x f(u,\mu_i,\sigma_i^2)\mathrm{d}u$. Throughout this section, without loss of generality, group means will be rescaled to population means; that is, the population mean is always 1.

First, as stated by Aitchison & Brown, Theorem 2.6, page 12

$$\frac{1}{y_i}\int_0^x uf(u;\mu_i,\sigma_i^2)\mathrm{d}u = \int_0^x f(u;\mu_i+\sigma_i^2,\sigma_i^2)\mathrm{d}u \tag{9}$$

where y_i is the group mean.

Secondly, from Aitchison & Brown, Corollary 2.2b, page 11

$$\int_0^\infty F(ax;\mu_1,\sigma_1^2)\mathrm{d}F(x;\mu_2,\sigma_2^2) = F(a;\mu_1-\mu_2,\sigma_1^2+\sigma_2^2) \tag{10}$$

Now consider a piecewise log-normal distribution, with the probability density function

$$g(x) = \sum_{i=1}^N p_i f(x;\mu_i,\sigma_i^2) \tag{11}$$

The Lorenz curve plots cumulative population against cumulative income. Letting both axes run over income x, cumulative population is $G(x) = \int_0^x g(u)\mathrm{d}u$ while cumulative income is $V(x) = \int_0^x ug(u)\mathrm{d}u$.

By (9), cumulative income is

$$V(x) = \int_0^x u \sum_{i=0}^{N} p_i f(u;\mu_i,\sigma_i^2)\mathrm{d}u \tag{12}$$

$$= \sum_{i=1}^{N} p_i y_i \left(\frac{1}{m_i}\int_0^x u f(u;\mu_i,\sigma_i^2)\mathrm{d}u\right) \tag{13}$$

$$= \sum_{i=1}^{N} p_i y_i \left(\int_0^x f(u;\mu_i+\sigma_i^2,\sigma_i^2)\mathrm{d}u\right) \tag{14}$$

$$= \sum_{i=1}^{N} p_i y_i F(x;\mu_i+\sigma_i^2,\sigma_i^2) \tag{15}$$

Denote the total area below the Lorenz curve as H. It can be expressed as

$$H = \int_0^\infty V(x)\mathrm{d}\left[G(x)\right] \tag{16}$$

$$= \int_0^\infty \sum_{i=1}^{N} \left[p_i y_i \left(F(x;\mu_i+\sigma_i^2,\sigma_i^2)\right)\right] \mathrm{d}\left[\sum_{j=1}^{N}\left(p_j F(x;\mu_j,\sigma_j^2)\right)\right] \tag{17}$$

$$= \sum_{i=1}^{N} \left(p_i y_i \int_0^\infty F(x;\mu_i+\sigma_i^2,\sigma_i^2)\mathrm{d}\left[\sum_{j=1}^{N}\left(p_j F(x;\mu_j,\sigma_j^2)\right)\right]\right) \tag{18}$$

Reordering and using (10) to get

$$H = \sum_{i=1}^{N}\left(p_i y_i \sum_{j=1}^{N} p_j \left(\int_0^\infty F(x;\mu_i+\sigma_i^2,\sigma_i^2)\mathrm{d}\left[F(x;\mu_j,\sigma_j^2)\right]\right)\right) \tag{19}$$

$$= \sum_{i=1}^{N}\left(p_i y_i \sum_{j=1}^{N} p_j \left(F(1;(\mu_i-\mu_j)+\sigma_i^2,\sigma_i^2+\sigma_j^2)\right)\right) \tag{20}$$

$$= \sum_{i=1}^{N}\left(y_i \sum_{j=1}^{N} p_i p_j \left(F(1;(\mu_i-\mu_j)+\sigma_i^2,\sigma_i^2+\sigma_j^2)\right)\right) \tag{21}$$

Letting F_N denote a normal distribution and Φ its standardized variant, this can further be written as

$$H = \sum_{i=1}^{N} \left(y_i \sum_{j=1}^{N} p_i p_j \left(F_N(0; (\mu_i - \mu_j) + \sigma_i^2, \sigma_i^2 + \sigma_j^2) \right) \right) \tag{22}$$

$$= \sum_{i=1}^{N} \left(y_i \sum_{j=1}^{N} p_i p_j \left(\Phi \left(\frac{0 - (\mu_i - \mu_j + \sigma_i^2)}{\sqrt{\sigma_i^2 + \sigma_j^2}} \right) \right) \right) \tag{23}$$

$$= \sum_{i=1}^{N} \left(y_i \sum_{j=1}^{N} p_i p_j \Phi \left(\frac{-(\mu_i - \mu_j + \sigma_i^2)}{\sqrt{\sigma_i^2 + \sigma_j^2}} \right) \right) \tag{24}$$

$$= 1 - \sum_{i=1}^{N} \left(y_i \sum_{j=1}^{N} p_i p_j \Phi \left(\frac{\mu_i - \mu_j + \sigma_i^2}{\sqrt{\sigma_i^2 + \sigma_j^2}} \right) \right) \tag{25}$$

Finally, by the definition of the Gini coefficient,

$$G = 1 - 2H \tag{26}$$

$$= 2 \sum_{i=1}^{N} \left(y_i \sum_{j=1}^{N} p_i p_j \Phi \left(\frac{\mu_i - \mu_j + \sigma_i^2}{\sqrt{\sigma_i^2 + \sigma_j^2}} \right) \right) - 1 \tag{27}$$

A.1.2 Calculation of c_w

This section outlines the calculation of c_w. First, consider the more general case, where the relationship between standard deviations and group means are

$$\frac{s_i}{\bar{y}} = \alpha \left(\frac{y_i}{\bar{y}} \right)^{\beta} \tag{28}$$

α_w is defined as the α that makes the average of within-group Gini coefficients equal to the between-group Gini coefficient.

From Equations (2) and (28), we get

$$\sigma = \sqrt{\log\left(1 + \alpha^2 (y_i/\bar{y})^{2\beta - 2}\right)} \tag{29}$$

α_w is then defined by the α that makes the average within-group Gini coefficient (right-hand side below) equal to the between-group Gini coefficient (left-hand side below; calculated from y and p).

$$G_B = \sum_{i=1}^{N} p_i 2\Phi \left[\sqrt{\frac{1}{2}\log\left(1+\alpha_w^2 (y_i/\bar{y})^{2\beta-2}\right)} \right] - 1 \tag{30}$$

This is solved numerically when $\beta \neq 1$.
Note that when $\beta = 1$, $c_w = \alpha_w$. In this case:

$$G_B = 2\Phi\left(\sqrt{\frac{1}{2}\log\left[1+\alpha_w^2\right]}\right) - 1 \tag{31}$$

$$\alpha_w = c_w = \sqrt{\exp\left(2\left[\Phi^{-1}\left(\frac{G_B+1}{2}\right)\right]^2\right) - 1} \tag{32}$$

For $\beta = 1$, all within-Ginis will be equal to the between-Gini. For $\beta \neq 1$, the average of within-Ginis will be equal to the between-Gini. This means that alternate averages (weighting by $y_i p_i^2$ instead of p_i, for example) would produce different values for α_w if $\beta \neq 1$, but do not matter for $\beta = 1$.

A.2 Robustness checks

A.2.1 Kuznets curve regression robustness: Other function forms

More flexible Kuznets curve regressions are shown in Tables A.1 and A.2. As is evident from the tables, the result is robust to changing specifications.

A.2.2 Kuznets curve robustness: Restricting the sample

Tables A.3 and A.4 show the results of the Kuznets curve regression for limited samples, where "West European" refers to the early modern West European samples. With a sample size of only 9 for the West European countries, the slope parameters are only significant at a 10% level, though they hold systematically for all assumption sets. For the "rest" sample significance is stronger.

Though the number of observations with the split sample is too low to properly compare the effects, one can observe that the slope parameter is higher for the non-West European countries — perhaps reflecting that the West European countries are closer to the end of the era when the mechanisms of increasing inequality operate, with relatively high levels of GDP per capita.

Dependent variable: Gini coefficient							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Const	−49.853	−37.264	−39.512	−3.098	26.092	53.053***	−27.455
	(27.516)	(28.071)	(26.495)	(21.712)	(16.132)	(10.261)	(30.394)
log (GDI/capita)	14.014***	12.334***	12.674***	8.357**	5.407**	3.091*	13.219***
	(4.086)	(4.168)	(3.934)	(3.224)	(2.395)	(1.523)	(4.513)
R^2	0.338	0.276	0.311	0.226	0.181	0.152	0.272
N	25	25	25	25	25	25	25

Table A.1: Regression results under assumption set (1)-(7) (see Table 3 for definitions of the assumptions). {***,**,*}=significant at {99%,95%,90%} level (two-sided tests)

Dependent variable: Log Gini coefficient							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Const	1.604**	1.919***	1.922***	2.911***	3.553***	4.021***	2.642***
	(0.649)	(0.662)	(0.592)	(0.403)	(0.254)	(0.137)	(0.510)
log (GDI/capita)	0.321***	0.279***	0.279***	0.156**	0.086**	0.042*	0.217***
	(0.096)	(0.098)	(0.088)	(0.060)	(0.038)	(0.020)	(0.076)
R^2	0.325	0.259	0.304	0.229	0.184	0.153	0.263
N	25	25	25	25	25	25	25

Table A.2: Regression results under assumption set (1)-(7) (see Table 3 for definitions of the assumptions). {***,**,*}=significant at {99%,95%,90%} level (two-sided tests)

Dependent variable: Gini coefficient							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Const	31.321***	32.376***	33.047***	43.063***	54.989***	69.011***	46.501***
	(7.859)	(7.984)	(7.559)	(5.700)	(4.002)	(2.452)	(8.397)
GDI/capita	0.012*	0.012*	0.011*	0.008*	0.006*	0.004*	0.013*
	(0.005)	(0.005)	(0.005)	(0.004)	(0.003)	(0.002)	(0.006)
R^2	0.418	0.393	0.409	0.397	0.400	0.406	0.418
N	9	9	9	9	9	9	9

Table A.3: Regression results under assumption set (1)-(7) (see Table 3 for definitions of the assumptions). {***,**,*}=significant at {99%,95%,90%} level (two-sided tests). Restricted to only West European countries.

A.2.3 More general specification of variance structure

As noted in Footnote 14, a more general specification of the variance structure is $(s_i/\bar{y}) = \alpha(y_i/\bar{y})^{\beta}$. The specification used in the main text — with the coefficient of variation constant — corresponds to $\beta = 1$. However, cases could be made for other relationships between group mean and group dispersion; that is, other values for β. This does not change the main results.

Figure A.1 and Table A.5 show the results for $\beta = 0.5$. As shown, the results in the main text still hold up, the only difference being somewhat lower significance for some of the higher Kuznets curve estimates.

A.2.4 The inequality impact of upper groups

Results are shown in Table A.6.

As we count all terms except the within-group cells (the diagonal) as belonging to two groups, the sum of all these terms is not the overall Gini. In the table, the column "contributions of rich groups" includes all terms where the rich groups are at least one of the groups. For example, for the Byzantine example (Table 4), if the two richest groups were included as "rich", all cells that are part of the two rightmost columns and/or the two lowest rows would be included; the $(i = 7, j = 8)$ cell would not be counted twice toward the inequality contribution.

A.2.5 Adding subsistence income

This explains the numerical procedure used to calculate the values in Table A.7, discussed in Section 3.3.

A population grid X of 50 000 points is constructed, with points spaced equally apart in logs (more points at the bottom). This combines the need for high accuracy at the bottom (where there is high "population density") with the need for covering large income ranges at the top (where density is lower, and one does not need as fine a grid). The grid runs from zero to 10 000 times the mean income of the richest group. A weight is assigned to each grid point corresponding to the inverse of the spacing of points.

Adjustments 1 and 2

The log-normal PDF is then calculated for each of these points for each group, and the distributions normalized to group sizes.

As y is already normalized so that the population mean is 1, subsistence income is found by inverting the number "mean income in terms of s" found in Table 2 of Milanovic *et al.* (2011). When the lowest income group has lower mean income than this subsistence group, the lowest group mean income will be chosen, subtracting 0.0001 (the scaling is population mean) to allow for some very small dispersion at the bottom group; this does not alter the results, but simplifies the calculation.

	Dependent variable: Gini coefficient						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Const	20.144**	24.474**	23.688**	37.669***	51.780***	67.410***	38.823***
	(8.525)	(8.778)	(8.209)	(6.878)	(5.133)	(3.262)	(9.575)
GDI/capita	0.033**	0.029**	0.030**	0.021**	0.015*	0.009*	0.031**
	(0.012)	(0.012)	(0.012)	(0.010)	(0.007)	(0.005)	(0.014)
R^2	0.345	0.280	0.322	0.258	0.238	0.224	0.272
N	16	16	16	16	16	16	16

Table A.4: Regression results under assumption set (1)-(7) (see Table 3 for definitions of the assumptions). {***,**,*}=significant at {99%,95%,90%} level (two-sided tests). Restricted to only non-West European countries.

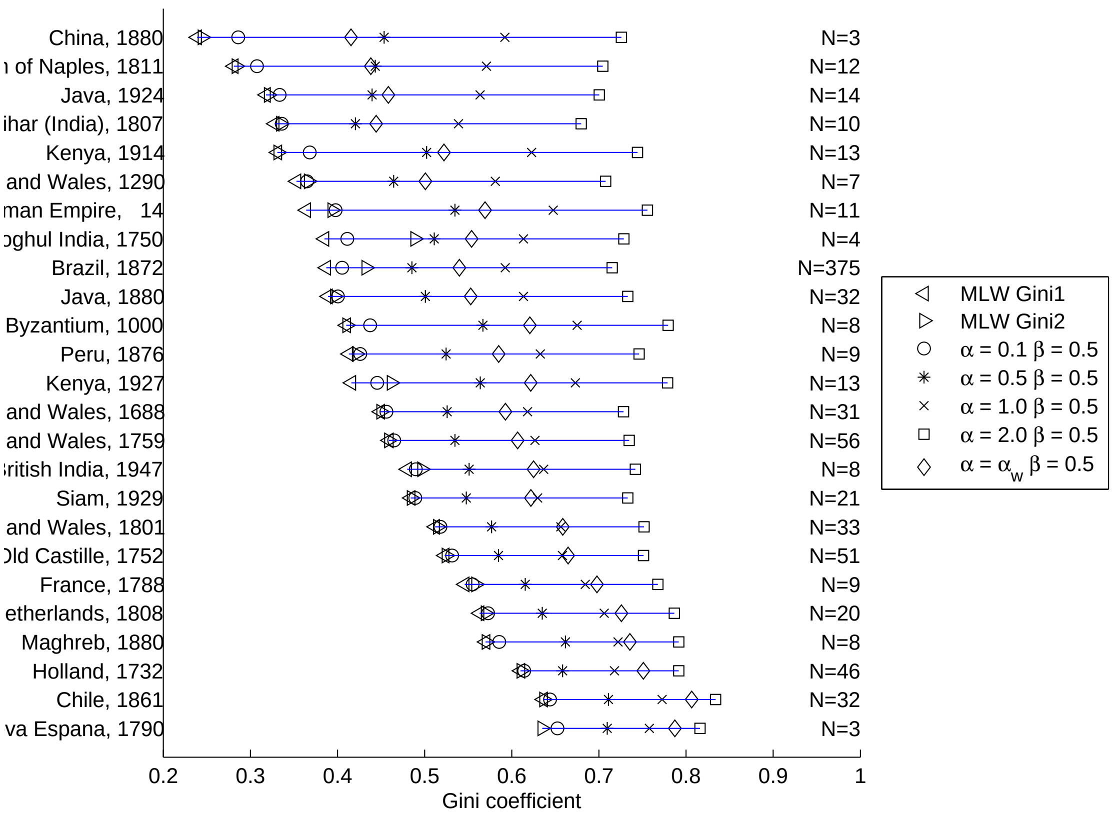


Figure A.1: Comparison of Gini coefficients for the seven assumption sets; $\beta = 0.5$

Then, for each population group, the total mass of everyone below subsistence income P is calculated, replacing the pdf values for these grid points with 0, and adding P to the distribution at the first grid point above subsistence.

For adjustment 2, in addition, a "richness line" R is introduced. Starting at the upper end of each group, move everyone above the richness line (the total mass of people in the group with income above R) down to the first grid point below R. Then decrease R until this procedure makes the mean of the group equal to the pre-adjustment mean.

Finally, all the group distributions are summed into a population distribution. Then, defining all grid points as discrete groups (ie 50 000 groups), (7) is used to calculate the overall Gini coefficient.

Adjustment 3

The log-normal distribution is now calculated on $X - y_{min}$ instead of on X, for each group (y_{min} is found in the same way as for the previous adjustments). Then, the complete distributions are right-shifted by y_{min} again, before they are added. Then, Gini coefficients can be computed on the grid points in the same manner as for adjustments 1 and 2.

Benchmark

An unadjusted Gini coefficient is also calculated by the numerical method. The largest deviations on the unadjusted Gini comparied to coefficients calculated by Equation 5 are .09 Gini points (.0009) for New Spain and .01 Gini points (.0001) for Chile; this verifies that the numerical procedure is sufficiently accurate to compare the benchmark to the adjusted values.

Detailed results

Table A.7 shows the results.

Subsistence incomes are taken from Table 2 of Milanovic *et al.* (2011); however, in many cases (denoted by an asterisk in the table) the mean income of the poorest group is lower than this subsistence level. In those cases subsistence minimum is set to the mean income of the poorest group.

A.2.6 Gini and number of groups relationship

A regression of Gini coefficients on the number of groups, for $c = 1$, is shown in Table A.8. The point estimate is very close to zero, and not significant. Brazil (with $N = 375$) is an outlier in terms of number of groups and was not included in the regression shown here. Including Brazil in the regressions does not change the sign or significance level of the coefficients.

A.3 Modern inequality data

The underlying information for Section 2.4 is in Table A.9.

The data has been compiled by IPUMS (Minnesota Population Center, 2010). As reported by IPUMS, the statistical data was originally produced by

- Brazil: Institute of Geography and Statistics
- Canada: Statistics Canada
- Colombia: National Administrative Department of Statistics
- Mexico: National Institute of Statistics, Geography, and Informatics
- Panama: Census and Statistics Directorate
- Puerto Rico: U.S. Bureau of the Census
- South Africa: Statistics South Africa
- United States: Bureau of the Census
- Venezuela: National Institute of Statistics

A.4 Calculating decile shares

When a fuller knowledge of the aggregate distribution is desirable, one can calculate percentile shares. In the following, ten groups will be assumed (deciles), but any partition is possible.

Let d be the vector of population lower bounds for the groups ($d = \{0, .1, .2, .3, ..., .9\}$). Without loss of generality, rescale income so that the population mean is 1.

The lower income bounds a are then found numerically by solving

$$\sum_{i=1}^{N} \left(p_i F(a_j; \mu_i, \sigma_i^2)\right) - d_j = 0; \tag{33}$$

for each $j \in \{1, 2, 3, ..., 10\}$. (Trivially, $a_1 = 0$). As F is strictly increasing, (33) only has one solution for each j.

The upper bounds b are then the lower bounds of the group above, $b_j = a_{j+1}$; $b_{10} = \infty$.

The mean income of each decile is

$$\begin{aligned} \delta_j &= \sum_{i=1}^{N} p_i \int_{a_j}^{b_j} u f(u; \mu_i, \sigma_i^2) du \qquad (34) \\ &= \sum_{i=1}^{N} p_i \left(\int_0^{b_j} u f(u; \mu_i, \sigma_i^2) du - \int_0^{a_j} u f(u; \mu_i, \sigma_i^2) du \right) \qquad (35) \end{aligned}$$

From Equation (9) this equals

$$\delta_j = \sum_{i=1}^{N} p_i y_i \left[\int_0^{b_j} f(u; \mu_i + \sigma_i^2, \sigma_i^2) du - \int_0^{a_j} f(u; \mu_i + \sigma_i^2, \sigma_i^2) du \right] \tag{36}$$

$$= \sum_{i=1}^{N} p_i y_i \left[F(b_j; \mu_i + \sigma_i^2, \sigma_i^2) - F(a_j; \mu_i + \sigma_i^2, \sigma_i^2) \right] \tag{37}$$

From this, for each decile group j, we know the bounds (a_j, b_j) and the mean income δ_j.

Dependent variable: Gini coefficient

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Const	33.008***	35.630***	35.901***	48.419***	60.370***	72.813***	50.560***
	(4.087)	(4.134)	(3.926)	(3.246)	(2.440)	(1.562)	(4.198)
GDI/capita	0.012***	0.011**	0.011***	0.007**	0.004*	0.002	0.011**
	(0.004)	(0.004)	(0.004)	(0.003)	(0.002)	(0.001)	(0.004)
R^2	0.298	0.244	0.267	0.180	0.132	0.091	0.233
N	25	25	25	25	25	25	25

Table A.5: Regression results under assumption set (1)-(7) where the assumption set numbers refer to Figure A.1

	N	N^r	n^r	G	G_W	G_W^r	G_R	G_R^r
Roman Empire, 14	11	6	0.04	61	21	0	4	0
Byzantium, 1000	8	1	0.01	64	11	0	12	0
England and Wales, 1290	7	1	0.04	56	8	0	13	0
England and Wales, 1688	31	8	0.02	61	3	0	13	0
Holland, 1732	46	15	0.04	70	1	0	8	1
Moghul India, 1750	4	1	0.01	59	17	0	3	0
England and Wales, 1759	56	13	0.02	61	2	0	14	0
Old Castille, 1752	51	9	0.04	65	2	0	10	1
France, 1788	9	2	0.10	67	5	1	7	1
Nueva Espana, 1790	3	1	0.10	74	10	3	1	0
England and Wales, 1801	33	8	0.04	64	3	0	10	1
Bihar (India), 1807	10	0	0.00	53	5	0	16	0
Netherlands, 1808	20	10	0.03	68	4	0	7	0
Kingdom of Naples, 1811	12	1	0.01	55	3	0	24	0
Chile, 1861	32	6	0.05	74	3	0	8	1
Brazil, 1872	375	114	0.01	58	2	0	16	0
Peru, 1876	9	2	0.02	61	7	0	12	0
China, 1880	3	2	0.02	56	32	0	0	0
Java, 1880	32	22	0.01	59	6	0	14	0
Maghreb, 1880	8	1	0.01	71	7	0	7	0
Kenya, 1914	13	8	0.01	59	24	0	3	0
Java, 1924	14	2	0.01	55	6	0	17	0
Kenya, 1927	13	8	0.01	64	17	0	5	0
Siam, 1929	21	1	0.01	62	2	0	11	0
British India, 1947	8	2	0.01	63	7	0	8	0

Table A.6: Inequality contribution from the richest groups. Superscript r denotes contributions from groups with mean incomes more than five times greater than population mean

	y_{min}	Benchmark G	Adj. 1	Adj. 2	Adj. 3	Benchmark G_B
Roman Empire, 14	0.48	61	55	45	47	36
Byzantium, 1000	0.56	64	55	42	44	41
England and Wales, 1290	0.47*	56	50	44	44	35
England and Wales, 1688	0.21*	61	59	58	57	45
Holland, 1732	0.07*	70	70	70	70	61
Moghul India, 1750	0.30*	59	56	54	53	39
England and Wales, 1759	0.17	61	60	60	59	46
Old Castille, 1752	0.07*	65	65	65	64	52
France, 1788	0.26	67	63	61	62	55
Nueva Espana, 1790	0.24*	74	71	68	69	63
England and Wales, 1801	0.11*	64	64	64	63	51
Bihar (India), 1807	0.43*	53	48	44	44	33
Netherlands, 1808	0.17	68	67	66	65	56
Kingdom of Naples, 1811	0.45	55	49	43	43	28
Chile, 1861	0.16*	74	73	72	71	64
Brazil, 1872	0.23*	58	56	55	54	40
Peru, 1876	0.33*	61	57	54	53	41
China, 1880	0.56	56	48	37	39	24
Java, 1880	0.31*	59	56	54	53	39
Maghreb, 1880	0.32*	71	66	62	63	57
Kenya, 1914	0.66*	59	48	34	34	33
Java, 1924	0.33	55	52	49	48	32
Kenya, 1927	0.53	64	55	44	48	42
Siam, 1929	0.18*	62	61	60	59	48
British India, 1947	0.23*	63	60	59	59	48

Table A.7: The Gini coefficients under different assumptions on minimum incomes, with $c = 1$

	Dependent variable: Gini coefficient						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Const	0.445***	0.455***	0.458***	0.534***	0.628***	0.741***	0.616***
	(0.025)	(0.024)	(0.023)	(0.018)	(0.013)	(0.008)	(0.026)
Number of groups	−0.000	0.000	−0.000	−0.000	−0.000	−0.000	−0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
R^2	0.001	0.000	0.001	0.009	0.015	0.018	0.002
N	25	25	25	25	25	25	25

Table A.8: Lack of correlation between Gini coefficient (for $c = 1$) and number of groups.

	Occupation (ISCO)				Industry (IPUMS)				Empl.class. (IPUMS)				Empl.class. (IPUMS, detailed)			
Country and Year	$c_{\min}$	c_{median}	$c_{\max}$	N	$c_{\min}$	c_{median}	$c_{\max}$	N	$c_{\min}$	c_{median}	$c_{\max}$	N	$c_{\min}$	c_{median}	$c_{\max}$	N
Brazil 1970	0.8	1.1	1.8	10	1.0	1.4	1.9	15	1.5	2.0	23.9	3	1.1	1.4	23.9	6
Brazil 1980	1.0	1.3	6.3	10	0.8	1.6	6.2	15	1.9	3.4	34.6	3	0.7	2.3	34.6	8
Brazil 1991	1.1	1.5	3.8	10	1.2	2.0	3.8	15	1.9	2.5	20.5	3	1.0	2.1	20.5	11
Brazil 2000	1.2	1.7	11.1	10	0.8	2.1	10.6	15	1.7	4.8	7.4	3	0.8	3.5	7.2	7
Canada 1971	0.6	0.8	1.8	8	0.7	0.8	1.7	9	0.9	1.3	4.5	3	0.9	1.2	4.5	4
Canada 1981	0.7	0.7	1.4	8	0.6	0.8	1.4	14	0.8	1.1	2.6	3	0.8	1.0	2.6	4
Canada 1991	0.7	0.7	1.2	8	0.7	0.9	1.2	13	0.8	1.0	1.1	2	0.8	1.1	1.1	3
Canada 2001	0.6	0.8	1.0	9	0.6	0.8	1.1	13	0.8	0.9	1.1	2	0.8	0.9	1.1	3
Colombia 1973	1.1	1.7	3.0	9	1.3	2.0	3.0	11	2.0	2.1	2.1	2	1.6	1.8	4.1	5
Mexico 1995	0.8	1.2	2.6	9	0.7	1.8	2.6	12	1.5	2.7	9.0	3	1.0	2.0	9.0	5
Mexico 2000	1.7	3.3	9.0	10	1.5	3.3	10.2	15	3.7	4.5	13.5	3	2.6	4.8	13.5	5
Panama 1980	0.8	1.1	8.7	9	0.7	1.2	4.3	14	1.5	2.8	6.2	3	0.9	2.0	7.4	5
Panama 1990	0.8	1.2	3.7	10	0.7	1.4	2.9	14	1.3	2.7	5.0	3	0.8	1.7	5.0	5
Puerto Rico 1970	0.9	1.0	1.4	9	0.9	1.2	1.4	11	1.2	1.3	1.4	2	0.9	1.3	1.4	3
Puerto Rico 1980	0.9	1.0	1.4	9	0.8	1.1	1.4	14	1.1	1.2	1.3	2	0.9	1.1	1.4	6
Puerto Rico 1990	1.0	1.1	1.3	9	0.7	1.2	1.4	14	1.1	1.3	1.4	2	0.9	1.2	1.4	7
Puerto Rico 2000	1.6	2.0	3.7	9	0.9	2.1	4.1	14	1.9	2.3	4.5	3	1.4	2.0	4.5	8
Puerto Rico 2005	1.0	1.1	1.5	8	0.9	1.2	1.6	12	1.2	1.4	1.6	2	0.8	1.1	1.7	7
South Africa 1996	0.9	1.1	2.4	9	0.8	1.4	2.5	15	1.4	1.5	1.5	3	1.3	1.4	1.7	4
South Africa 2001	1.8	2.4	3.7	9	1.8	2.6	4.0	15	2.4	2.6	2.8	2	2.4	2.7	3.3	4
South Africa 2007	1.7	2.3	4.2	10	1.5	2.5	3.7	14	2.3	2.5	2.5	3	2.3	2.5	2.5	3
United States 1960	0.6	0.9	1.3	10	0.6	1.1	1.5	15	1.0	1.2	4.3	3	0.8	1.1	4.3	4
United States 1970	0.6	0.9	1.3	10	0.7	1.1	1.5	15	1.0	1.1	3.2	3	0.8	0.9	3.2	7
United States 1980	0.7	0.9	1.2	10	0.6	1.0	1.3	15	0.9	1.1	2.5	3	0.8	0.9	2.5	7
United States 1990	0.7	1.0	1.3	10	0.6	1.2	1.4	15	1.1	1.3	2.4	3	0.7	1.1	2.4	8
United States 2000	0.9	1.2	1.5	10	0.8	1.3	1.8	15	1.3	1.5	2.5	3	0.8	1.2	2.5	8
United States 2005	0.7	1.0	1.4	10	0.7	1.2	1.7	15	1.1	1.4	2.0	3	0.8	1.1	2.0	8
Venezuela 2001	0.9	1.1	2.6	10	0.9	1.2	1.8	15	1.2	1.4	1.6	3	0.9	1.2	1.4	6

Table A.9: Modern within-group inequality. Source data: IPUMS