DESCRIPTION OF *ALL THE GINIS* DATASET

(This version February 2019)

Created by Branko Milanovic
Graduate Center, City University of New York and
Stone Center on Socio-economic Inequality

Original dataset created: Summer 2004
Previous version of the dataset: November 2016

Coverage of years: 1948-2017
Coverage of countries (or territories): 201
Total number of Ginis reported: 5121
Total number of country/years covered by Gini: 2276
Overall coverage: 16-20 percent[1]

This dataset consists only of the Gini coefficients that have been calculated from actual household surveys. It uses no Ginis estimates produced by regressions or short-cut methods.

What is this database? This database represents a compilation and adaptation of Gini coefficients retrieved from nine sources. The nine sources are:

(1) *Luxembourg Income Study (LIS)* dataset that covers the period 1967-2016 and includes 50, mostly developed, countries. There are 345 Gini observations (an increase of 25% compared to the 2016 version of *All the Gini*) all calculated from direct access to household surveys and micro (unit record) data representing the status of LIS database as of December 2018.

[1] Maximum coverage is obtained as the product of the number of countries and the number of years (1948 to 2017). There are 201 countries which gives a maximum (fully-dense) coverage of 14,070 country/years. (This is however somewhat of an overestimation because more than 40 countries did not exist in all the years included here, so the effective coverage is estimated at between 16% and 20%).

(2) *Socio-Economic Database for Latin America and the Caribbean (SEDLAC)* that covers the period 1974-2014 and includes 23 Latin American and Caribbean countries. There are 319 Gini observations all calculated from direct access to micro data from household surveys. The data are provided directly by SEDLAC from their 2016 version.

(3) *World Bank's Eastern Europe and Central Asia (ECA)* database that covers the years 1990-2011 and includes 30 countries. There are 257 Gini observations all calculated from direct access to household surveys.

(4) *World Income Distribution (WYD)* dataset that covers the period 1980-2012 and includes 153 countries. There are 642 Gini observations, about 90 percent calculated from direct access to household surveys. For the years after 2000, that percentage is close to 100. It represents the database as of October 2016.

(5) *Survey of Income and Living Condition (SILC)* conducted by Eurostat that includes years 2005-2008 with 29 countries. There are 103 Gini observations all calculated from direct access to household survey data. It has recently become more difficult to get the entire annual set of SILC data and that explains why the latest year is 2009.

(6) *POVCAL*, World Bank-based dataset that covers the period 1978-2017 and includes 171 countries. There are 1711 Gini observations, most of which are calculated from direct access to household surveys. There continues to be significant expansion (70% increase compared to 2016) and improvement of POVCAL. It is the source whose coverage has expanded by most between the current and the previous version of this database. POVCAL includes the data from the advanced (rich) countries and most of the Ginis since 2000 are calculated from micro data. The data for China, India, and Indonesia are also presented separately for rural and urban areas.

(7) *Individual data sets (INDIE).* These are data taken from individual studies (listed in the Appendix) which either report or provide their own Gini estimates calculated from micro data. As with the rest of the data, Ginis from such studies have to be calculated from nationally representative household surveys. They must cover no fewer than three (ideally, successive) years. Their advantage is that they are consistently calculated, using the same type of survey and welfare aggregate. Such data however are

available for only 17 countries. INDIE data cover the period 1950-2015 and include 329 Ginis.

(8) *World Institute for Development Research WIDER (WIID)* dataset covers the period 1948-2014 and includes 166 countries. There are 1386 Gini observations compiled from various sources, some of which are based on direct access to household surveys and others to grouped data. The data are downloaded from the version WIID3.4 (January 2017). Since this version drops the observations from the previous WIID versions, the number of Gini observations is now smaller than before. WIID data are seemingly very clearly coded but the fact that WIDER uses a bewildering array of sources with e.g. very uneven household adjustments (various equivalency scales) leads to major inter-temporal inconsistencies when the series is used—even when an effort is made to standardize the definitions (as here). I was tempted to drop the series altogether, but for the reasons of consistency with the past versions of *All the Ginis* I have retained it. However I advise a high degree of caution or even skepticism when using these data.

(9) *CEPAL*. These are historical data on Latin American countries obtained from published documents by CEPAL. They cover the period 1950-1987 and include 6 countries. There are 29 Gini observations (see Annex 1. Sources of CEPAL data).

This gives a grand total of 5121 Gini observations (see Table 1).

Table 1. Number of Gini observations and their characteristics by database

Database	Number of observations	Mean Gini	Standard deviation of Gini	Minimum Gini	Maximum Gini
LIS	345	35.4	9.0	20.2	69.8
SEDLAC	319	51.4	5.2	34.5	67.6
World Bank ECA	257	31.9	5.7	17.5	45.7
World Income Distribution (WYD)	642	38.7	10.1	17.8	77.4
SILC	103	30.8	3.9	23.1	38.9
POVCAL	1711	38.7	9.7	16.2	74.3
Independent Gini country series (INDIE)	329	39.0	8.0	23.7	62.5
WIDER (WIID)	1386	38.3	9.9	16.2	67.6
CEPAL	29	52.2	6.4	35.0	61.0
Total	*5121*	*38.8*		*16.2*	*77.4*

Variables in *All the Ginis*. Suffix "WY" to the data from the World Income Distribution database; suffix "SEDLAC" to the data obtained from the SEDLAC dataset; suffix "LIS" to the data from LIS; suffix "EE" to the data from World Bank ECA database; suffix "SILC" to the data obtained from SILC; suffix "POVCAL" to the data obtained from World Bank POVCAL database; Suffix "W" suffix refers to the variables taken from the WIDER dataset; suffix "CEPAL" to the data obtained from UN Economic Commission for Latin America and the Caribbean, and suffix "INDIE" to the data obtained from independent individual inequality studies. Thus, for example, `Dhh_LIS` indicates a dummy variable such that it takes the value of 1 if income recipient is household, and 0 if it is individual. The variable is taken from LIS (as shown by the "LIS" suffix), "hh" stands for household, and the prefix "D" denotes a dummy variable.

There are three kinds of variables: (a) country and year, (b) Gini value (in percent) which must come from a nationally-representative household survey covering

the entire resident and non-institutionalized population, and (c) information on the welfare concept and recipient unit to which the reported Gini refers. The last point is addressed by three dummy variables: `Dhh_database` which denotes whether the Gini refers to households (value=1) or individuals, as in household *per capita* income (value=0); `Dinc_database` which denotes whether the concept used is income (value=1) or consumption/expenditures (value=0); `Dgross_database` which denotes whether the concept used is gross (value=1) or net, as in disposable income (value=0). [2] The most common concept used, household net per capita income, will therefore be characterized by the following combination of dummy values: `Dhh`=0, `Dinc`=1, `Dgross`=0.

It should be kept in mind that the Ginis shown here, even if full correction were made for the three observable characteristics of surveys (namely, `Dinc, Dgross` and `Dhh`) may still differ for at least two reasons. First, even if the observable characteristics are coded the same, there could still be differences as, for example, in the way benefits from owner-occupied housing or home-consumption are imputed, for which we cannot adjust. Second, the Ginis may be calculated from micro or grouped data; they may be calculated using slightly different formulas or using geometrical approximations to the Lorenz curve. Thus, there could be differences in the Gini values that are due to the apparently small but important differences in the formulas used by different authors, or type of data (micro or grouped) they had access to. The user should keep in mind that, like every compilation, this one suffers not only from the bias of the final compiler (which may be thought fixed across the observations) but from the bias of individual earlier producers or compilers of the data.

[2] POVCAL does not provide information on whether the welfare concept is gross or net and hence that variable is absent. SEDLAC sources are coded as "gross" although they are in-between gross and net income. Namely, income data provided by SEDLAC are net of the wage tax, but do not deduct other direct taxes. Given that wage taxes are often greater than other direct taxes, one may be justified in treating it as net income, although to be on the conservative side we stick with "gross" label. For some WIDER data there is no information whether the welfare concept is income or consumption/expenditures. In such cases, *Dinc_W* variable is missing.

Gini estimates obtained using equalized household income and assigning such income to either households or individuals are not included in the data base. The main reason is lack of between country comparability of such Ginis. Different countries use different equivalence scales, and consequently equivalent income and its distribution will differ in function of equivalence scale used. It would be misleading to treat them as comparable just because they use an equivalence scale.

The Gini coefficients from each of the nine sources are downloaded and presented in (or transformed into) the format given above (points (a)-(c)). If, for example, the original dataset provides more information on additional Gini characteristics (as WIDER often does) that information is not used.

A significant change in the 2018 version. In the past, the nine sources were used to create a new variable called `Giniall` that summarized all the sources using all the available data and would display only one Gini value per country/year. In the (quite numerous) cases where there were duplicate or triplicate observations, coming from different sources, for any one country/year, the choice was done by what I called "the order of precedence" using as first choice the data judged best, and then going further down the list. Thus the final `Giniall` variable had the advantage of covering as many cells (country/years) as possible but the disadvantage that the data were not consistent: a country series might have several values from one source, followed by several values form another, and then yet by a third. This led to some confusion: researchers tended to use, for simplicity, `Giniall` despite the caveat that it was not a homogeneous variable (neither within countries, nor between). In order to avoid such confusion, this version of the dataset drops the summary variable.

Regional coverage. In terms of the five big regions, namely Africa, Asia, Latin America and the Caribbean, former transition countries of Eastern Europe and the USSR, and WENAO (Western Europe, North America and Oceania), the representation is relatively uniform (see Table 2). Of course, when one takes into account the number of countries per region, the real difference in representation becomes apparent. Africa has

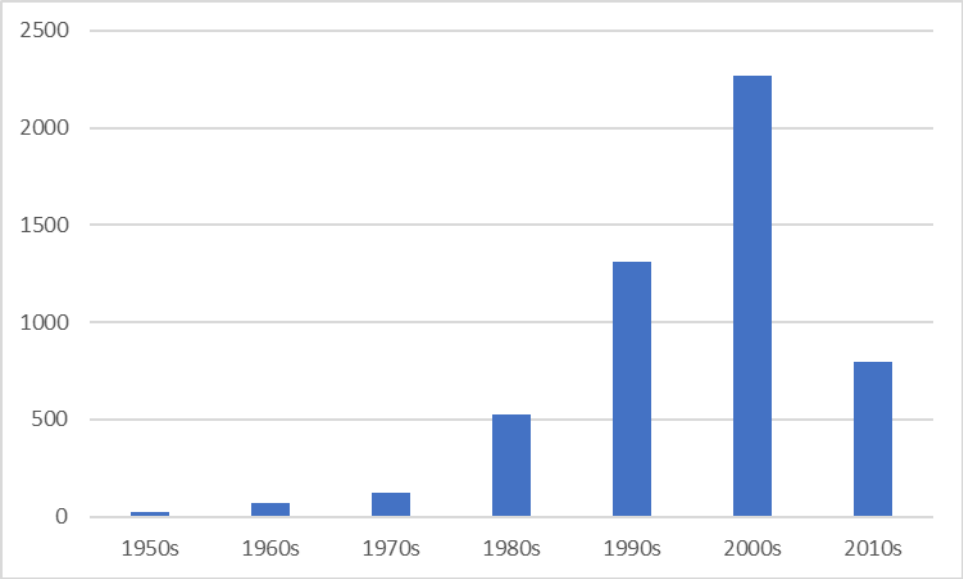
53 countries included and only less than 11 Gini observations on average per country. WENAO, on the other hand, has 26 countries included and an average of 47.6 observations per country. (All the countries included in the data base are listed in Annex 2.)

Table 2. Number of Gini observations by geographical area

	Number of Gini observations from all sources (country/years)	Number of countries or territories	Average number of observations per country or territory
Africa	571	53	10.8
Asia	770	57	13.5
Latin America and the Caribbean	1193	33	36.2
Former transition countries	1350	32	42.2
WENAO	1237	26	47.6
Total	*5121*	*201*	*25.5*

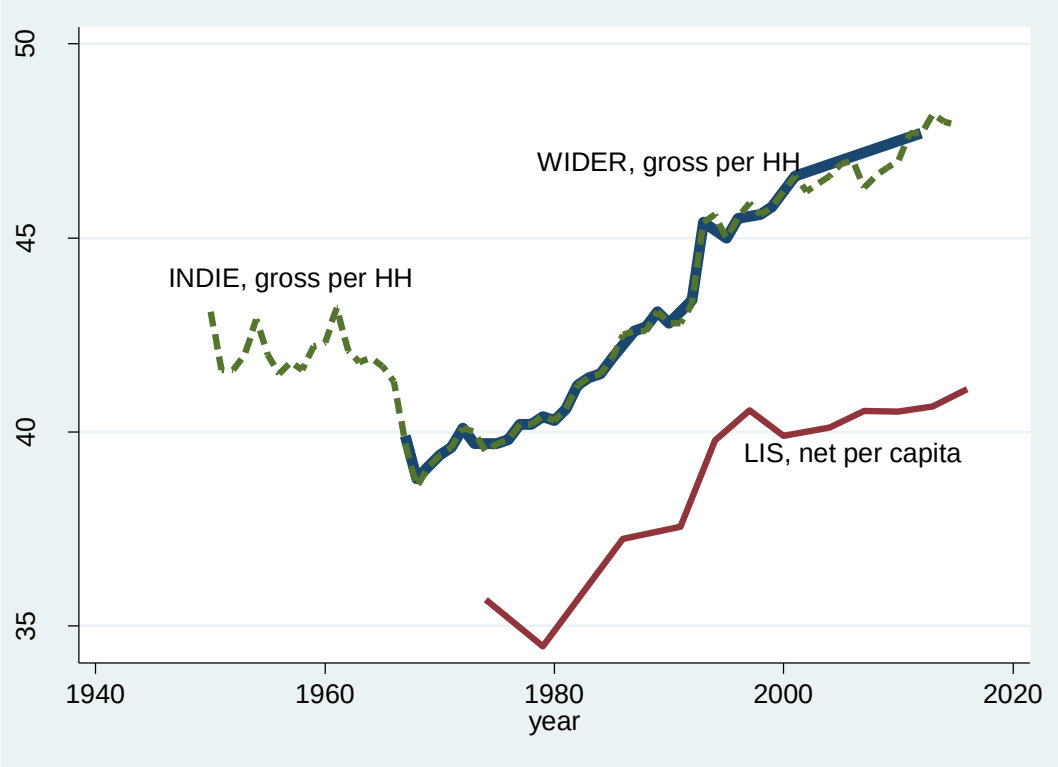
Time coverage. Not surprisingly, the number of Gini observations increases as we come closer the present (see Figure 1). In the 1950s, there are only 27 observations; for the last complete decade (2001-10), there are 2266 observations. Given the average delay of some two to three years between the time when the surveys are conducted and the synthetic data such as Gini coefficient become available, there is little doubt that the decade 2011-2020 will have even more observations. It should be noted however that if more sustained effort were made to collect the published statistics from the 1950s and 1960s, the under-representation of these years could be diminished.

Figure 1. Total number of Gini observation by decade



All the Ginis database gives the user full flexibility, whether she wishes to use the data from only one source, or to combine various sources, or to use various sources but to keep the definitions of the aggregates and recipients the same. A simple illustration is provided in Figure 2 which shows US Ginis from three different sources. INDIE and WIDER that use gross household income while LIS uses net income per person (household disposable per capita income). The Gini levels of the first two sources, in the years when both are available, coincide. Their inequality is some 5 to 7 Gini points higher than inequality expressed in terms of disposable per capita income. One could go on listing similar examples for practically every country for which several data sources exist.

Figure 2. Ginis for the United States: different sources, different welfare aggregates and income recipients



To see what data exist for a given country it is useful to list the data as shown below for the Unites States (where data sources from which US is omitted by construction, namely, SEDLAC, SILC. World Bank ECA, and CEPAL are omitted).

```
. list year gini_LIS gini_WYD  gini_POVCAL gini_INDIE gini_W if contcod=="USA"
```

	year	gini_LIS	gini_WYD	gini_PO~L	gini_I~E	gini_W
9289.	1950	.	.	.	43.1	.
9290.	1951	.	.	.	41.6	.
9291.	1952	.	.	.	41.6	.
9292.	1953	.	.	.	42	.
9293.	1954	.	.	.	42.9	.
9294.	1955	.	.	.	42	.
9295.	1956	.	.	.	41.5	.
9296.	1957	.	.	.	41.8	.
9297.	1958	.	.	.	41.6	.
9298.	1959	.	.	.	42.2	.
9299.	1960	.	.	.	42.3	.
9300.	1961	.	.	.	43.2	.
9301.	1962	.	.	.	42.1	.
9302.	1963	.	.	.	41.8	.
9303.	1964	.	.	.	41.9	.
9304.	1965	.	.	.	41.7	.
9305.	1966	.	.	.	41.3	.
9306.	1967	.	.	.	39.7	39.9
9307.	1968	.	.	.	38.6	38.8
9308.	1969	.	.	.	39.1	39.1
9309.	1970	.	.	.	39.4	39.4
9310.	1971	.	.	.	39.6	39.6
9311.	1972	.	.	.	40.1	40.1
9312.	1973	.	.	.	40	39.7
9313.	1974	35.68055	.	.	39.5	35.9
9314.	1975	.	.	.	39.7	39.7
9315.	1976	.	.	.	39.8	39.8
9316.	1977	.	.	.	40.2	40.2
9317.	1978	.	.	.	40.2	40.2
9318.	1979	34.4759	.	.	40.4	40.4
9319.	1980	.	.	.	40.3	40.3
9320.	1981	.	.	.	40.6	40.6
9321.	1982	.	.	.	41.2	41.2
9322.	1983	.	.	.	41.4	41.4
9323.	1984	.	.	.	41.5	41.5
9324.	1985	.	.	.	41.9	41.9
9325.	1986	37.25031	37.04661	37.529999	42.5	.
9326.	1987	.	.	.	42.6	42.6
9327.	1988	.	.	.	42.6	42.7
9328.	1989	.	.	.	43.1	43.1
9329.	1990	.	.	.	42.8	42.8
9330.	1991	37.56624	.	38.240002	42.8	.
9331.	1992	.	.	.	43.3	43.4
9332.	1993	.	.	.	45.4	45.4
9333.	1994	39.78604	39.38738	40.220001	45.6	.
9334.	1995	.	.	.	45	45
9335.	1996	.	.	.	45.5	45.5
9336.	1997	40.56313	39.88	40.759998	45.9	.
9337.	1998	.	.	.	45.6	45.6
9338.	1999	.	.	.	45.8	45.8
9339.	2000	39.90026	39.9	40.380001	46.2	.
9340.	2001	.	.	.	46.6	46.6
9341.	2002	.	.	.	46.2	.
9342.	2003	.	.	.	46.4	.
9343.	2004	40.10884	40.28456	40.529999	46.6	40.62
9344.	2005	.	.	.	46.9	.
9345.	2006	.	.	.	47	.
9346.	2007	40.53691	.	41.080002	46.3	41.75
9347.	2008	.	43.17992	.	46.6	.
9348.	2009	.	.	.	46.8	.
9349.	2010	40.52726	.	40.41	47	40.46
9350.	2011	.	.	.	47.7	47.6
9351.	2012	.	.	.	47.7	47.7
9352.	2013	40.65271	.	41.040001	48.2	.
9353.	2014	.	.	.	48	.
9354.	2015	.	.	.	47.9	.
9355.	2016	41.10104	.	41.5	.	.

More on WIDER dataset. The past and current WIDER datasets are much broader than the data included here. We have extracted from WIDER only the observations that are conceptually the same as those contained in the other datasets used here. This means that they are derived from nationally representative household surveys, provide information on a "complete" welfare concept whether it is for income or

expenditure or on a net or gross basis, with household as the basic statistical unit, and with household or person as the recipient unit. We have included only Ginis, not quintile and decile shares that are also often available in WIDER. But, in addition to those data, WIDER dataset includes also observations on the distribution of earnings. Earnings are obviously only one component of income (hence, not a "complete" concept) and individual workers (not households) are the basic statistical units. Such data are not included here.

More on *World Income Distribution* (WYD) dataset. WYD database is an original database created as part of the work on global income distribution. The objective of the work is to gather and analyze detailed household surveys for as many countries as possible for several benchmark years and come up with estimates of global inequality. The currently available data exist for seven benchmark years (1988, 1993, 1998, 2002, 2005, 2008 and 2011). Some of the data for the forthcoming benchmark year 2013, which is still not complete, are also included.

World Income Distribution approach is as follows. If a country does not have a household survey for a given benchmark year, then a year as close to the benchmark as possible is selected, provided it is not more than 2 years apart from the benchmark year.[3] This explains: (i) the clustering of Gini observations around the years 1988, 1993, 1998, 2002, 2005, 2008 and 2011 and (ii) that the earliest observations are from 1985.[4] The objective of WYD data base was to create as "rich" (numerous in terms of countries) and "dense" (ventiles or percentiles for each country) coverage for the benchmark years, not to maximize the number of Gini observations, or provide longer-term series for individual countries.

The household survey data provided by LIS, SILC, World Bank ECA and SEDLAC were all used in creating World Income Distribution dataset. However, Gini observations, coming from LIS, SILC, SEDLAC or ECA are listed under their respective

[3] There are just a few exceptions to this rule.
[4] Other than three observations.

original data sources, not as part of WYD. WYD thus includes only the Ginis from the surveys that do not originate from LIS/SILC/SEDLAC/World Bank ECA. For example, micro data for Thailand or Indonesia are not part of other databases used here and are thus listed under WYD. For the exact origin and information on these surveys, the user needs to consult the documentation provided by World Income Distribution database (see the Web links given below).

World Income Distribution (WYD) database was used in several publications, in particular Branko Milanovic, *Worlds Apart: Measuring International and Global Inequality*, Princeton: Princeton University Press, 2005; Branko Milanovic, "True world income distribution, 1988 and 1993: First calculation based on household surveys alone", *Economic Journal*, vol. 112, No. 476, January 2002, pp. 51-92; and Branko Milanovic, "Global inequality recalculated and updated: The effect of new PPP estimates on global inequality and 2005 estimates", *Journal of Economic Inequality*, volume 10, issue 1, 2012, pp. 1-18.

How to refer to *All the Ginis* database? Simply as *All the Ginis* or by abbreviation *ATG* database (version February 2019); and the web reference
https://www.gc.cuny.edu/Page-Elements/Academics-Research-Centers-Initiatives/Centers-and-Institutes/Luxembourg-Income-Study-Center/Branko-Milanovic,-Senior-Scholar/Datasets
http://econ.worldbank.org/projects/inequality

Where to find the original (source) databases? The data, descriptions and explanations regarding how the source databases were constructed can be found on the following Websites.

Detailed sources and explanations of how WYD dataset was created can be found on the same Website where *All the Ginis* is,
http://econ.worldbank.org/projects/inequality (pl. go under "Datasets" and then "World Income Distribution").

For WIDER, see: https://www.wider.unu.edu/project/wiid-world-income-inequality-database

For SEDLAC, see http://sedlac.econo.unlp.edu.ar/eng/.

For Luxembourg Income Study, see http://www.lisdatacenter.org/.

For POVCAL, see http://iresearch.worldbank.org/PovcalNet/index.htm?2.

Additional information. Please contact me at bmilanovic@gc.cuny.edu or branko_mi@yahoo.com.

APPENDIX: Sources of INDIE data (17 countries)

Russia, 2001-2009 (9 data points): Irina Denisova, "Income distribution and Poverty in Russia", OECD Social, Employment and Migration Working Papers, No. 132, OECD Publishing. 2012. Page 9, Table 1. Gini of net per capita disposable monetary income calculated from the official annual national Household Income and Expenditure Survey.

China, 1985-2001 (17 data points): Ximing Wu and Jeffrey Perloff, "China's income distribution and inequality 1985-2001", *Review of Economics and Statistics*, 87 (2005): 763-775. Calculations by the authors based on published official urban and rural fractiles of the income distribution. Chinese annual surveys have been (until 2013) conducted separately for rural and urban areas, and here the results are put together to generate distribution for the entire country.

China, 2003-2015 (13 data points). Official National Bureau of Statistics estimates (without any micro data provided) as reported in Juzhong Zhuang and Li Shi, "Understanding recent trends in income inequality in the People's Republic of China", ADB Working Paper No. 489, July 2016.

USA, 1950-2015 (66 data points) "Income, poverty and health insurance coverage in the United States:2009", US Census Bureau, September 2010, Table A.2, pp. 40-43, and the subsequent such publications "Income and poverty in the Unites States" (Table A.2) from years 2010 to 2015 (all based on Current Population Survey); plus "Income, poverty and health insurance coverage in the United States 2012, US Census Bureau, September 2012, Table A.2. p. 38. Almost exactly the same data are given in "The Changing Shape of Nation's Income Distribution, 1947-1998", Current population report, June 2000, Table 4 by Arthur F. Jones Jr. and Daniel H. Weinsberg. Data are for household gross income across households, both based on March Current Population Survey (conducted every year).

Brazil 1981-2004 (with interruptions) (21 data points), "The rise and fall of Brazilian inequality 1981-2004", World Bank Working Paper No. 3867, March 2006 by Francisco H. Ferreira, Philippe D. Leite and Julie Litchfield, Table 1, p. 6. Data are from PNAD survey (Pesquisa Nacional por Amostra de Domicílios) conducted annually by the Instituto Brasileiro de Geografia e Estatística.

Italy, 1967-2008 (with interruption) (29 data points) from Giovanni Vecchi and Andrea Brandolini, published in Gianni Toniolo (ed.) *The Oxford Handbook of the Italian Economy since Unification*, Oxford University Press, 2013. Tables kindly provided by Giovanni Vecchi. Data from household surveys conducted annually (with a few interruptions) by Banca d'Italia.

Great Britain (UK), 1961-2014 (54 data points). Data calculated especially for Branko Milanovic by Jonathan Cribb from the Institute for Fiscal Studies, using micro data from Family Expenditure Surveys and Family Resource Surveys.

Japan, 1981-1990 (with interruptions) (4 data points). Based on Income Redistribution Survey (IRS) conducted at three-year intervals. From Toshiaki Tachibanaki and Tadashi Yagi, "Distribution of economic well-being in Japan: towards a more unequal society", Table 6.3, p. 113 in *Changing patterns in the distribution of economic welfare: an international perspective*, ed. By Peter Gottschalk, Bjorn Gustafsson, and Edward Palmer, Cambridge University Press, 1999

Ireland, 1973-1987 (with interruptions) (3 data points). From Tim Callan and Brian Nolan, "Income inequality and poverty in Ireland in the 1970s and 1980s", Table 10.4, p. 224 in *Changing patterns in the distribution of economic welfare: an international perspective*, ed. By Peter Gottschalk, Bjorn Gustafsson, and Edward Palmer, Cambridge University Press, 1999. Data are from the annual Household Budget Surveys (income and expenditures) conducted by the Central Statistical Office.

Poland, 1985-1997 (13 data points). Unpublished calculations by Branko Milanovic from individual data from the official annual Household Budget Surveys supplied by the Polish Central Statistical Office.

Iran, 1984-2011 (28 data points). From Djavad Salehi-Isfahani, "Poverty, inequality and populist politics in Iran", *Journal of Economic Inequality*, vol. 7:5–28, 2009, Table 4. The data are from the official annual Household Income and Expenditures Surveys conducted by the Statistical Center of Iran. Period 2006-2011, based on the same original source; unpublished results from a personal communication by Djavad Salehi-Isfahani (30 August 2013).

India, 1983-1997 (with interruptions) (11 data points). From Martin Ravallion, "Should poverty measures be anchored to national accounts?" *Economic and Political Weekly*, August 26, 2000, p. 3247. Calculated from the annual National Sample Survey.

France, 1975-1990 (with interruptions) (4 data points). From Andrea Brandolini, "A bird's Eye View of Long-run Changes in Income Inequality", Table A11. Estimates by INSEE. Detailed original sources given in Brandolini.

(West) Germany, 1950-1985 (with interruptions) (13 data points). From Andrea Brandolini, "A bird's Eye View of Long-run Changes in Income Inequality", Table A10. Estimates by the Central Statistical Office. Detailed original sources given in Brandolini.

Canada, 1971-1994 (24 data points). From Andrea Brandolini, "A bird's Eye View of Long-run Changes in Income Inequality", Table A10. Data from Statistics Canada. . Detailed original sources given in Brandolini.

Netherlands, 1981-1989 (with interruptions) (6 data points). From Andrea Brandolini, "A bird's Eye View of Long-run Changes in Income Inequality", Table A9. Estimates by Statistics Netherlands. Detailed original sources given in Brandolini.

Chile, 1987-1994 (with interruptions) (4 data points). From Francisco H.G. Ferreira and Julie A. Litchfield, "Calm after the storm: income distribution in Chile, 1987-1994", World Bank Research Working Paper No. 1960, November 1998.

Indonesia, 1999-2010 (with one interruption) (10 data points). From Riyana Miranti, Yogi Vidyattama, Erick Hansnata, Rebecca Cassells, Alan Duncan "Trends in Poverty and Inequality in Decentralising Indonesia", OECD Social, Employment and Migration Working Paper No. No. 148, OECD Publishing; Figure 12, p. 31. Available at http://dx.doi.org/10.1787/5k43bvt2dwjk-en.

Annex 1, Sources of CEPAL data

	Period	Number of years used	Source
Argentina	1953-72	4 observations	Table 5.1 in *Antecedentes estatisticos de la distribucion del ingreso Argentina 1953-82*, United Nations, Santiago de Chile 1987.
Brazil	1972-87	11 observations	Table 5.1 in *Antecedentes estatisticos de la distribucion del ingreso Brasil 1970-88*, United Nations, Santiago de Chile 1990. All based on Pesquisa Nacional; por Amostra de Domicilios (PNAD)
Chile	1968-71	2 observatons	Table 5.1 in *Antecedentes estatisticos de la distribucion del ingreso Chile 1940-82* United Nations, Santiago de Chile 1987. Data processed by CEPAL and World Bank
Colombia	1970-72	3 observations	Table 5.1 in *Antecedentes estatisticos de la distribucion del ingreso Colombia 1951-82*, United Nations, Santiago de Chile 1986. Data from Encuesta Nacional de Hogares, Presupuestos Familiares
Mexico	1950-77	7 observations	Table 5 in *Antecedentes estatisticos de la distribucion del ingreso Mexico 1960-77*, United Nations, Santiago de Chile 1988. All data but one from O.Altamir's calculations based micro data (EDIGF or ENIG).
Peru	1971-81	2 observations	Table 5.1 *Antecedentes estatisticos de la distribucion del ingreso Peru 1961-82.* United Nations, Santiago de Chile 1989. Both based on national survey micro data

Annex 2. List of all countries and territories and their abbreviations in the database (in the alphabetic order according to the abbreviation)

Country or territory	Abbreviation
Afghanistan	AFG
Angola	AGO
Albania	ALB
Netherlands Antiles	ANT
United Arab Emirates	ARE
Argentina	ARG
Armenia	ARM
Australia	AUS
Austria	AUT
Azerbaijan	AZE
Burundi	BDI
Belgium	BEL
Benin	BEN
Burkina Faso	BFA
Bangladesh	BGD
Bulgaria	BGR
Bahrain	BHR
Bahamas	BHS
Bosnia & Hercegovina	BIH
Belarus	BLR
Belize	BLZ
Bermuda	BMU
Bolivia	BOL
Brazil	BRA
Barbados	BRB
Brunei	BRN
Bhutan	BTN
Botswana	BWA
Central African R.	CAF
Canada	CAN
Switzerland	CHE
Chile	CHL
China	CHN
China-rural	CHN-R
Chin-urban	CHN-U
Cote d'Ivoire	CIV
Cameroon	CMR
Congo, Dem. Rep.	COD
Congo, Rep. (Brazzaville)	COG
Colombia	COL
Comoros	COM
Cape Verde	CPV
Costa Rica	CRI
Czechoslovak Socialist Republic	CSK
Cyprus	CYP
Czech R	CZE
Germany	DEU
Djibouti	DJI
Denmark	DNK
Dominican Republic	DOM
Algeria	DZA
Ecuador	ECU
Egypt	EGY
Spain	ESP

Country or territory	Abbreviation
Estonia	EST
Ethiopia	ETH
Finland	FIN
Fiji	FJI
France	FRA
Micronesia, Federated Islands of	FSM
Gabon	GAB
United Kingdom of Great Britain & Northern Island	GBR
Georgia	GEO
Ghana	GHA
Guinea	GIN
Guadeloupe	GLP
Gambia	GMB
Guinea-Bissa	GNB
Equatorial Guinea	GNQ
Greece	GRC
Greenland	GRL
Guatemala	GTM
French Guiana	GUF
Guyana	GUY
Hong Kong, China	HKG
Honduras	HND
Croatia	HRV
Haiti	HTI
Hungary	HUN
Indonesia	IDN
Indonesia rural	IDN-R
Indonesia urban	IDN-U
India	IND
India rural	IND-R
India urban	IND-U
Ireland	IRL
Iran, Islamic Republic	IRN
Iraq	IRQ
Iceland	ISL
Israel	ISR
Italy	ITA
Jamaica	JAM
Jordan	JOR
Japan	JPN
Kazakhstan	KAZ
Kenya	KEN
Kyrgyz Republic	KGZ
Cambodia	KHM
Kiribati	KIR
St.Kitts and Navis	KNA
Korea, Republic	KOR
Kosovo	KOS
Kuwait	KWT
Laos	LAO
Lebanon	LBN
Liberia	LBR
Libya	LBY
St. Lucia	LCA
Sri Lanka	LKA
Lesotho	LSO
Lithuania	LTU
Luxembourg	LUX

Country or territory	Abbreviation
Latvia	LVA
Morocco	MAR
Moldova	MDA
Madagascar	MDG
Maldives	MDV
Mexico	MEX
Macedonia, FYROM	MKD
Mali	MLI
Malta	MLT
Myanmar	MMR
Montenegro	MNE
Mongolia	MNG
Mozambique	MOZ
Mauritania	MRT
Mauritius	MUS
Malawi	MWI
Malaysia	MYS
Namibia	NAM
New Caledonia	NCL
Niger	NER
Nigeria	NGA
Nicaragua	NIC
Netherlands	NLD
Norway	NOR
Nepal	NPL
New Zealand	NZL
Oman	OMN
Pakistan	PAK
Panama	PAN
Peru	PER
Philippines	PHL
Papua New Guinea	PNG
Poland	POL
Puerto Rico	PRI
Korea, Dem. People's Republic	PRK
Portugal	PRT
Paraguay	PRY
Palestine	PSE
Qatar	QAT
Reunion	REU
Romania	ROU
Russian Federation	RUS
Rwanda	RWA
Saudi Arabia	SAU
Sudan	SDN
Senegal	SEN
Singapore	SGP
Solomon	SLB
Sierra Leone	SLE
El Salvador	SLV
Somalia	SOM
Serbia	SRB
South Sudan	SSD
SaoTome & Principe	STP
Suriname	SUR
Slovakia	SVK
Slovenia	SVN
Sweden	SWE

Country or territory	Abbreviation
Swaziland	SWZ
Seychelles	SYC
Syrian AR	SYR
Chad	TCD
Togo	TGO
Thailand	THA
Tajikistan	TJK
Turkmenistan	TKM
Timor Leste	TLS
Tonga	TON
Trinidad & Tobago	TTO
Tunisia	TUN
Turkey	TUR
Tuvalu	TUV
Taiwan, China	TWN
Tanzania	TZA
Uganda	UGA
Ukraine	UKR
Uruguay	URY
United States	USA
Uzbekistan	UZB
Venezuela	VEN
Vietnam	VNM
Vanuatu	VUT
Palestine (West Bank and Gaza)	WBG
Samoa	WSM
Yemen, Republic	YEM
Yugoslavia, Federative Socialist Republic	YUG
South Africa	ZAF
Zambia	ZMB
Zimbabwe	ZWE