

DESCRIPTION OF *ALL THE GINIS* DATASET

(version November 2014)

Created by Branko Milanovic

Originally at World Bank, Research Department

Now: at Luxembourg Income Study, New York

Original date of dataset creation: Summer 2004

Previous version: Summer 2013

Coverage of years: 1950-2012

Coverage of countries: 166

Total number of “standardized” Ginis: 2218

Increase in the number of “standardized” Ginis since previous version: 8.6%

Overall coverage: around 21 percent¹

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 in order to create a single “standardized” Gini variable. The nine sources are:

(1) *Luxembourg Income Study (LIS)* dataset that covers the period 1967-2010 and includes 40, mostly developed, countries. There are 232 Gini observations all calculated from direct access to household surveys and micro (unit record) data representing the status of LIS database as of November 2014.

¹ Maximum coverage is obtained as the product of the number of countries and the number of years (1950 to 2012). There are 166 countries which gives a maximum (fully-dense) coverage of 10458 country/years. (This is however somewhat of an overestimation because some 30 countries did not exist in all the years included here.)

(2) *Socio-Economic Database for Latin America and the Caribbean (SEDLAC)* that covers the period 1974-2012 and includes 23 Latin American and Caribbean countries. There are 301 Gini observations all calculated from direct access to household surveys. The data are taken from SEDLAC Inequality LAC 2014 version.

(3) *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.

(4) *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.

(5) *World Income Distribution (WYD)* dataset that covers the period 1980-2012 and includes 152 countries. There are 631 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 November 2014.

(6) *POVCAL*, World Bank-based dataset that covers the period 1978-2011 and includes 124 countries. There are 798 Gini observations, most of which are calculated from direct access to household surveys.

(7) *World Institute for Development Research WIDER (WIIDI)* dataset that covers the period 1950-2012 and includes 159 countries. There are 1490 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 newly updated September 2014 version WIID3.0b (while maintaining the observations from the previous WIID2.0 version where the new version data are missing).

(8) *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 “Sources of CEPAL data” at the end of this document).

(9) *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 1950-2011 and include 291 Ginis.

This gives a grand total of 4132 Gini observations.

The yearly and country coverage refers to that used in the current version of *All the Ginis* database. Individual sources might have a more updated, and hence broader, coverage in both dimensions. These nine sources are used to create a new, relatively consistent, variable called `Giniall`. Deininger-Square data are not used because they were either superseded or are included in WIDER. (For completeness, however, they are displayed in the dataset.)

Variables in *All the Ginis*. Suffix “W” suffix refers to the variables taken from the WIDER dataset; 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 “CEPAL” for 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 noted that 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.

How are the “standardized” Gini coefficients in *All the Ginis* database created? 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. When there are conflicts such that two or more datasets provide Ginis for the same country/year (and these Ginis come from nationally representative household surveys), we use the approach described as “*choice by precedence*”, which in our view reflects the reliability, degree of variable standardization, and consistency of geographical coverage of each dataset, to create a “standardized” Gini. The newly created variable is called *Giniall*.

² POVCAL does not provide information on whether the welfare concept is gross or not 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.

The “choice by precedence” works as follows. We take first INDIE data which are calculated from long-term individual country studies. Next, we take the data from five databases (LIS, SEDLAC, SILC, ECA and CEPAL) whose Ginis are calculated entirely from the direct access to household surveys’ micro data, whose variables are harmonized (income or consumption defined the same way across surveys), and whose coverage is regional so that countries are mostly non-overlapping.³ We take them in the following order: first, LIS data; second, SEDLAC; third, SILC; fourth World Bank ECA, fifth, CEPAL. Then we move to the sources whose coverage is in principle worldwide. We take them in the following order: WYD, POVCAL, and finally WIDER. This essentially means, for example, that a Gini for a given country/year which is available in WIDER, but is also available in another database, will be taken from that other data base. (In other words, WIDER data will cover only those country/years that are not covered by the other eight sources.) The reverse is true for LIS: most of its Ginis (except when they conflict with INDIE) will be included. The number of such “conflicts” (same country/year) can be gauged by comparing the total number of Gini observations from all nine sources, which is 4132, and the number of Ginis provided by `GiniAll` which is 2218. As Table 1 shows, 1914 observations (46 percent of all Ginis) are discarded because of such “conflicts.”

³ By “non-overlapping”, we mean that the same country is (generally) not included in two different databases. LIS coverage is mostly limited to rich countries, SILC covers European Union and candidate countries, SEDLAC and CEPAL cover Latin America and the Caribbean, and World Bank ECA database covers Eastern Europe and Central Asia. The country overlaps are few. The most obvious one is between LIS and SILC, but in addition LIS includes about a dozen surveys from Latin America, thus overlapping with SEDLAC, and surveys from Eastern Europe, thus overlapping with ECA.

Table 1. Number of Gini observations by dataset

	(1) Number of Gini observations in the original data set*	(2) Number of Gini observations used in Gini _{all}	Percentage of observations used: (2)/(1)
INDIE	291	291	100
LIS	232	181	78
SEDLAC	301	259	86
SILC	103	80	78
ECA	257	227	88
CEPAL	29	22	76
WYD	631	350	55
POVCAL	798	183	23
WIDER	1490	625	42
<i>Total</i>	<i>4132</i>	<i>2218</i>	<i>54</i>
Memo: Deininger-Squire	700	0	0

* As used here. WIDER, for example, has many more observations but they do not fulfill the conditions (a)-(c). See the text below.

However, the fact that the Ginis from the original sources are provided in the database gives flexibility to the users to decide on a different “precedence” approach and to use or not the data from the sets they choose. In my opinion, the key issue is whether LIS and 4 other “regional” sources should be given precedence over INDIE or not.

The composition of the final variable *Gini_{all}* by the welfare aggregate (income or expenditures; net or gross) and recipient (household or person) is shown in Table 2.

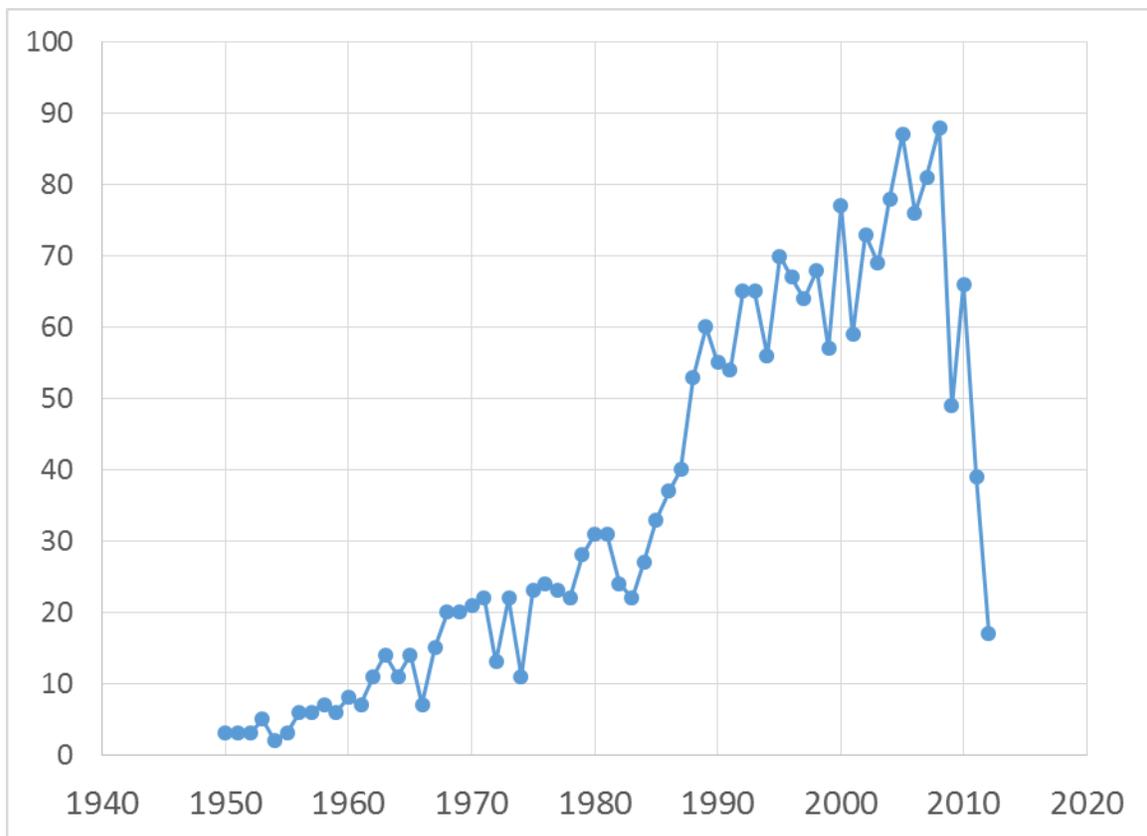
Table 2. Composition of variable *Gini_{all}* by welfare aggregate and recipient (number of observations)

	Income			Expenditures			Missing	<i>Total</i>
	Per person	Per household	Missing info	Per person	Per household	Missing info		
Net	641	140	0	279	17	0	0	1077
Gross	452	250	0	279	16	0	0	997
Missing	23	0	9	104	0	0	8	144
<i>Total</i>	<i>1116</i>	<i>390</i>	<i>9</i>	<i>662</i>	<i>33</i>	<i>0</i>	<i>8</i>	<i>2218</i>
	<i>1515</i>			<i>695</i>			<i>8</i>	<i>2218</i>

Note: all POVCAL welfare aggregates treated as “net”.

Figure 1 shows the number of Gini all observations by year. There is a steady increase until around 2008. After that point, the number of observations falls because the most recent surveys are only gradually becoming available. The average time lag between the time a survey is fielded and its results are included here is about 3 years. Between 1990 and 2010, the yearly number of observations is almost always over 60 and reaches as high as 88.

Figure 1. Number of Gini all observations by year



In terms of 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 3). Of course, when one takes into account the number of countries per region, the real difference in representation becomes apparent. Africa has 50 countries included and only

5.5 Gini observations on average per country. WENAO has 25 countries included and on average almost 22 observations per country.

Table 3. Number of `Giniall` observations by geographical area

	Number of Gini observations	Number of countries	Average number of observations per country
Africa	273	50	5.5
Asia	432	33	13.1
Latin America and the Caribbean	471	28	16.8
Former transition countries	496	30	16.5
WENAO	546	25	21.8
<i>Total</i>	<i>2218</i>	<i>166</i>	<i>13.4</i>

The new Gini variable and the caveats. As explained, the key new variable provided here is `Giniall` that gives values of the Gini coefficients from nationally representative household surveys for 2218 country/years.⁴ In principle, `Giniall` observations should be comparable, but two important caveats need to be made.

First, the dummy variables indicate whether the welfare concept used to calculate `Giniall` is income or consumption (`Dinc`), whether it is on a net or gross bases (`Dgross`), and whether the recipient unit is household or individual (`Dhh`). Thus, in the empirical work, an adjustment for each of these characteristics is desirable. `Giniall` should not be displayed or run in regressions, except in special circumstances, alone, that is, without any adjustment or awareness of the underlying concepts.

Second, one must keep 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

⁴ In several instances, “nationally representative” is interpreted more “leniently” as to allow us to include, for example, surveys from Argentina and Uruguay which used to be strictly speaking urban only but where urban population accounted for the quasi totality of country’s population.

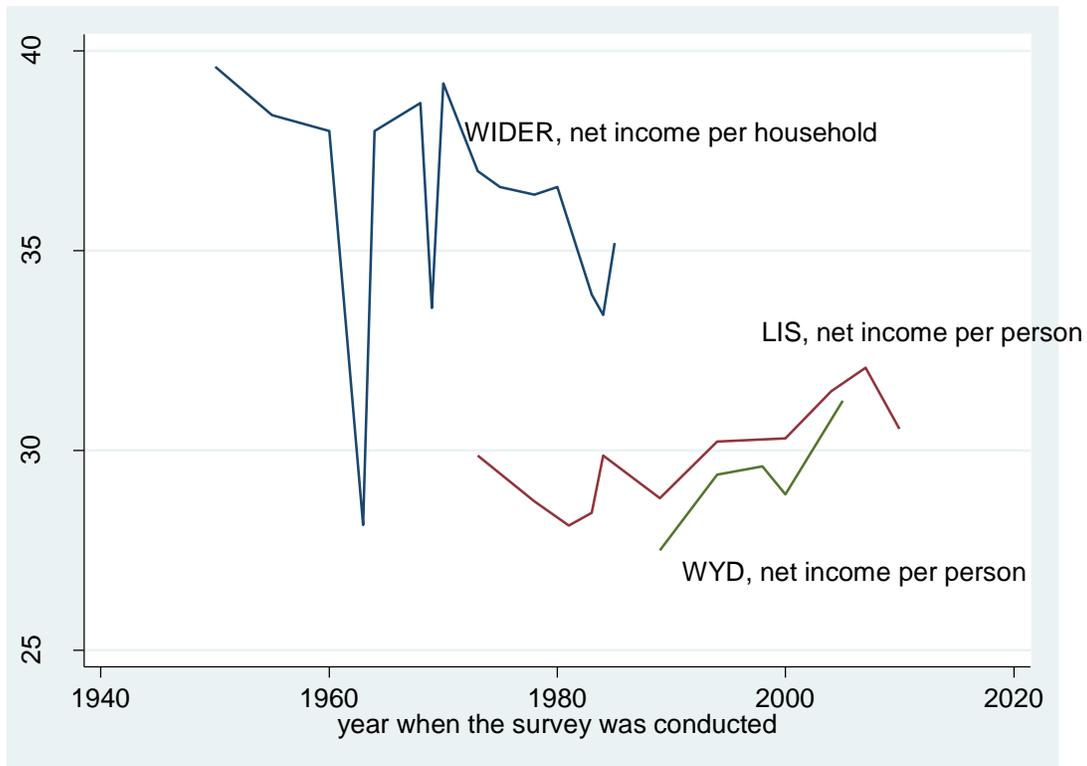
from owner-occupied housing or home-consumption are imputed, for which we cannot adjust. Second, the Ginis may be calculated from micro or grouped data; they may be calculated using slightly different formulas or using geometrical approximations to the Lorenz curve. Thus, there could be differences in 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.

All the Ginis database gives the user full flexibility, whether she wishes to use the data from only one source, or to arrange the sources differently than here (by creating a different “choice by precedence”), or to use various sources but to keep the definitions of the aggregates and recipients the same. There is thus a huge variety of choices one can make. A very simple illustration is provided in Figure 2. `Giniall` obtained for Germany come from five sources: 13 observations from an independent study, 7 observations from LIS, 1 observation from SILC, 2 observations from WYD, and 11 observations from WIDER (a total of 34 yearly observations).⁵ But the difference in income recipients, even when the welfare concept is the same (net income), implies a difference in the Gini levels for the same or adjacent years. Consequently, the use of `Giniall` without adjustment for welfare aggregate and type of recipient is not recommended.⁶

⁵ Prior to the unification, West Germany is treated as “Germany”.

⁶ Note also the clear outlier in WIDER data for 1963.

Figure 2. Gini^{all} values for Germany: different sources and different income recipients



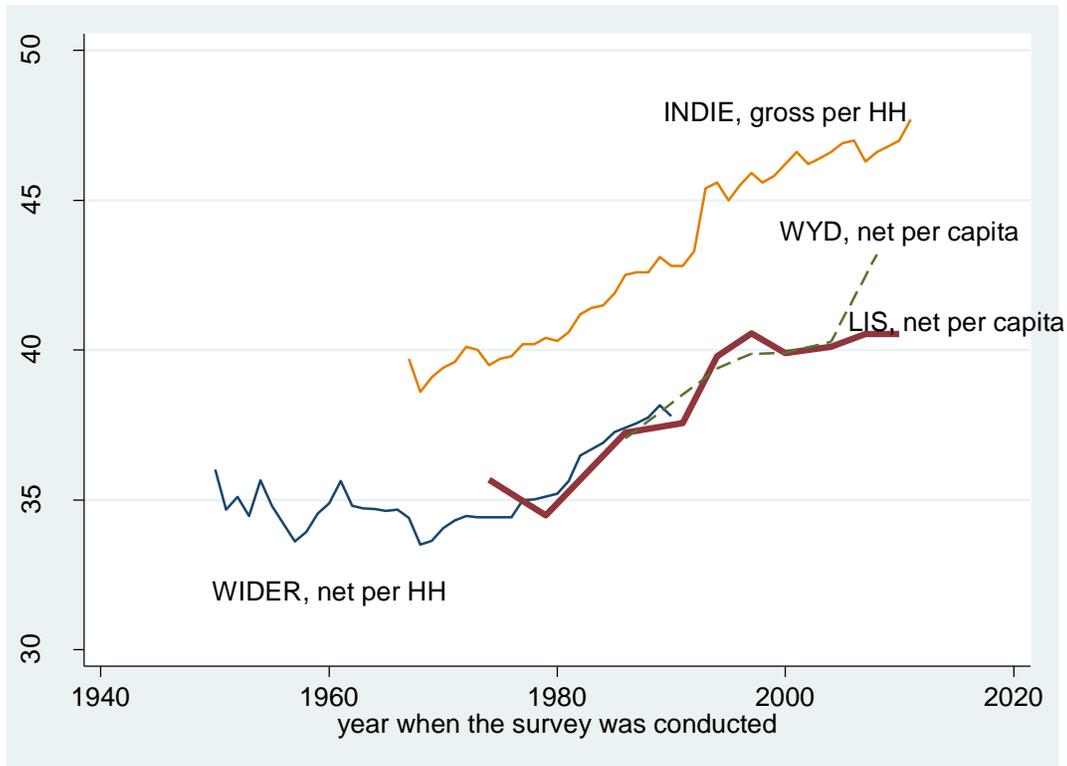
Note: for simplicity the one data point from SILC not shown.

Another illustration is provided in Figure 3 which shows US Ginis from four different sources. Here both the income aggregates and recipients differ: INDIE uses gross household income, WIDER net household income while LIS and WYD both use net income per person (household per capita income). The levels of the last three sources, in the years when they coincide, are similar. Gross income however is more unequally distributed and its Gini coefficient is throughout higher.⁷

One could go on listing similar examples for practically every country for which several data sources exist.

⁷ There is yet another, subtler, difference. WYD 2008 data are based on the directly accessed US 2008 March Current Population Survey (CPS) and thus use CPS definition of net income. LIS uses the same source of data, but “lissifies” the variables (that is, harmonizes them so they should be comparable to the variables from other countries), creating in the process its own definition of net income.

Figure 3. Gini^{all} values for United States: different sources, different welfare aggregates and income recipients



More on WIDER dataset. The original WIDER dataset is 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. This means that they are derived from nationally representative household surveys, provide information on a “complete” welfare concept whether income or expenditure (on 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 these household-level 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 six benchmark years (1988, 1993, 1998, 2002, 2005 and 2008). Some of the data for the benchmark year 2011, which is still not complete and available online, are also included here.

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.⁸ This explains: (i) the clustering of Gini observations around the years 1988, 1993, 1998, 2002, 2005, 2008 and 2011, (ii) the fact that there are at most seven Gini observations per country, and (iii) that the earliest observations are from 1986.⁹ 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 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).

⁸ There are just a few exceptions to this rule.

⁹ Other than three observations.

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 November 2014); and the web reference <http://econ.worldbank.org/projects/inequality> (pl. go under “Datasets” and then “All the Ginis”).

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: http://www.wider.unu.edu/research/WIID3-0B/en_GB/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.

USA, 1967-2011 (45 data points) “Income, poverty and health insurance coverage in the United States:2009”, US Census Bureau, September 2010, Table A.2, pp. 40-43; 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), 1961-2010 (50 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>.

SOURCES OF CEPAL DATA

	Period	Number of years used	Source
Argentina	1953-72	4 observations	Table 5.1 in <i>Antecedentes estadísticos de la distribución del ingreso Argentina 1953-82</i> , United Nations, Santiago de Chile 1987.
Brazil	1972-87	11 observations	Table 5.1 in <i>Antecedentes estadísticos de la distribución del ingreso Brasil 1970-88</i> , United Nations, Santiago de Chile 1990. All based on Pesquisa Nacional; por Amostra de Domicílios (PNAD)
Chile	1968-71	2 observatons	Table 5.1 in <i>Antecedentes estadísticos de la distribución del ingreso Chile 1940-82</i> United Nations, Santiago de Chile 1987. Data processed by CEPAL and World Bank
Colombia	1970-72	3 observations	Table 5.1 in <i>Antecedentes estadísticos de la distribución del ingreso Colombia 1951-82</i> , United Nations, Santiago de Chile 1986. Data from Encuesta Nacional de Hogares, Presupuestos Familiares
Mexico	1950-77	7 observations	Table 5 in <i>Antecedentes estadísticos de la distribución del ingreso Mexico 1960-77</i> , 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 <i>Antecedentes estadísticos de la distribución del ingreso Peru 1961-82</i> . United Nations, Santiago de Chile 1989. Both based on national survey micro data