

On the Assessment and Use of Cross-National Income Inequality Datasets

Frederick Solt
frederick-solt@uiowa.edu

June 2, 2015

Abstract

Researchers should ensure the data they employ are fit for their purpose, and they should maximize the quality of the data they choose. In this paper, I review how this advice applies to broadly cross-national research on income inequality. I demonstrate that the guidance offered in Jenkins (2015) to those pursuing cross-national research runs completely counter to the recommendations found in Atkinson and Brandolini (2001, 2009), the source of the aforementioned advice and the works upon which Jenkins (2015) claims its own is based. I then show how the Standardized World Income Inequality Database (SWIID) incorporates Atkinson and Brandolini's recommendations to provide the most comparable data available for those engaged in broadly cross-national research on income inequality.

The assessment of the Standardized World Income Inequality Database (SWIID) offered in this issue (Jenkins 2015) suffers from bad timing. It was written over the same period that a major revision of the SWIID and a new article of record for the dataset were being prepared. Although I was able to alert its author that on several important points the work's observations had already been superseded, most notably regarding top income shares, he justifiably concluded that the piece should critique the dataset as it was then available. (Fortunately, he could and did incorporate my advice that simply ignoring the SWIID's documentation and the uncertainty associated with the its estimates was not justifiable.) This has an entirely understandable but undeniably regrettable consequence: the bulk of the details in the paper's

description and assessment of Version 4.0 of the SWIID simply do not accurately describe Version 5.0, released at the start of October 2014, or even Version 4.1 that preceded it.

Still more regrettable—and unrelated to its timing—is that there is a great deal in the paper that does not even accurately describe Version 4.0.¹ A point-by-point correction of the many erroneous statements regarding the SWIID found in the now outdated paper, however, would serve little purpose; readers should simply refer to the SWIID’s article of record (Solt forthcoming) for details on the current dataset and its construction.

Here, I take up instead the recommendations the paper offers to those pursuing cross-national research on the causes and consequences of income inequality. Given its author’s admitted unfamiliarity with cross-national income inequality datasets (Jenkins 2015, 3) and his inexperience conducting cross-national research on income inequality (none of his many and deservedly influential publications fall into any of the three categories of studies that employ cross-national inequality datasets listed at pp.2-3), it is hardly surprising that the task of providing an assessment of cross-national income inequality datasets and how they should be employed proved to be a difficult remit. The advice for researchers that the paper repeats from Atkinson and Brandolini (2001, 2009) is indeed excellent. On the other hand, the paper’s own guidance—though undoubtedly well intentioned—is badly mistaken. When pernicious recommendations are so thoroughly intermingled with sound advice they only become more dangerous. In what follows, I do my best to thresh the chaff from the wheat.

¹A particularly egregious—and obvious—example is the unsupported assertion that the SWIID imputes *all* observations (Jenkins 2015, 27), when in fact it imputes only observations for which no data is available from the LIS (see Solt 2009 and forthcoming). Making the assertion even more bizarre is the later acknowledgement that the LIS data is perfectly reproduced in the SWIID (Jenkins 2015, 31). One possible explanation is that its author, like many other producers of inequality statistics from microdata, overlooked the fact that such statistics have associated measurement error due to being based on survey samples rather than the actual population of interest. It is this measurement error, not multiple imputation, that accounts for the variation across the simulated series that make up the SWIID in the observations for which LIS data are available (see Solt forthcoming, 10-11). It is too bad that the authors of the datasets reviewed in this issue were not invited to participate in the project from the outset. It is likely that, if they had, this and many of the other errors that mar Jenkins (2015) and the other assessments would have been avoided.

Some *Very* Bad Advice

The course the paper advocates for cross-national researchers begins with selecting data from the somewhat dated UNU-WIDER (2008a) dataset, which collects and presents information on income and consumption inequality from a wide array of sources available at the time of its release.² Researchers should then ‘systematically employ some sort of selection “algorithm”’ (Jenkins 2015, 14), preferably one that includes discarding all but those observations rated 1 on the 1 to 4 *Quality* variable included in the dataset (Jenkins 2015, 11-12), by carefully investigating the documentation to choose the best observation for a particular country-year from among the multiple possibilities presented by the UNU-WIDER (2008a) dataset. The piece justifies this approach by quoting Atkinson and Brandolini (2009, 399), ‘[o]ne has to *look at the data*, exercising judgment as to whether they are fit for purpose. Data quality *does* matter’ (Jenkins 2015, 38).

But the piece’s advocated approach does serious violence to the meaning of the quoted passage. First, looking at the data and ‘exercising judgment as to whether they are fit for purpose’ in the original work does not refer to picking and choosing individual observations but rather to ensuring a match between theory and measure. One should not unreflectively employ whatever data may be most readily available, Atkinson and Brandolini (2009, 389) explain: one must carefully consider whether the relevant variable is inequality in post-tax, post-transfer net income, as with theories linking inequality and growth, or whether it is instead pre-tax, pre-transfer market income, as when considering median-voter theories of taxation. The data employed must be fit for the theory examined. This is a fundamental tenet of research design that is too often neglected.

Second, ‘data quality’ in the quotation does not concern the quality ratings found in the UNU-WIDER (2008a) dataset and its predecessors. Atkinson and Brandolini (2001, 790) specifically argue against using such quality assessments as a criterion (as does the dataset’s user guide, see UNU-WIDER 2008b, 15). The *Quality* variable describes mainly the quality of the *documentation* of an observation; this in turn has much more to do with the timing and location of the underlying survey than with the validity of its data. What actually ‘*does* matter,’ according to the quoted article, is avoiding the

²Here too the timing was unfortunate, as this dataset has since begun to be revised and brought up to date (see UNU-WIDER 2014).

abrupt ‘breaks in consistency’ that occur in series composed of data drawn from different sources and calculated on different bases and the ‘breaks in continuity’ that occur even in series from a single source when that source’s underlying methodology changes (Atkinson and Brandolini 2009, 389-393).

That the paper’s advice completely contradicts that offered in the works of Atkinson and Brandolini (2001, 2009) is most easily seen in the context of its own example, an exploration of the relationship between unemployment, inflation, and income inequality within OECD countries since 1980. The process of arriving at a dataset for analysis (1) starts with the UNU-WIDER (2008a) data, using only those with the highest *Quality* scores.³ As already mentioned, Atkinson and Brandolini (2001, 790) counsel *against* selecting observations on these grounds.

The next step (2) retains those observations ‘in which the income definition referred to gross or disposable income’ (p.36). Atkinson and Brandolini (2001) point to two problems with the practice of mixing observations calculated on differing bases such as pre-tax gross income and post-tax disposable income. The first problem they note is that the differences between ‘gross and disposable income can . . . be substantial’ and ‘[d]ifferences across countries, and across time, are to be expected as a result of differences in government fiscal policies and in tax incidence’ (Atkinson and Brandolini 2001, 788, 789). As a result, mixing the two fails to meet their standard of generating ‘a dataset in which the observations are as fully consistent as possible’ (Atkinson and Brandolini 2001, 790). The second problem arises from the fact that one should expect taxation and the progressivity thereof to influence the relationships under examination: mixing gross- and disposable-income observations therefore suggests a simple failure to consider what is the theoretically relevant variable, a step that Atkinson and Brandolini (2009, 389) remind us is crucial.

Finally, the selection proceeds by (3) inspecting the remaining sources for each country, employing judgment to choose among them when multiple observations for a given country-year are available. ‘Wherever possible,’ the author writes, ‘I chose for each country the ones providing the longest series according to a particular definition’ (Jenkins 2015, 36). This fails to correspond to Atkinson and Brandolini’s exhortation to ensure that data do

³Although this first screen in the selection algorithm is left implicit, I confirmed it in the replication materials. Although these materials are not made available online, they can be procured via request to the author (Jenkins 2015, i).

not exhibit breaks in consistency and continuity. Those authors maintain that ‘where two consecutive observations of the Gini index differ by four percentage points or more’ investigation is ‘required’ (Atkinson and Brandolini 2009, 392). Such events occur in two of the twenty series obtained. There is an increase of 6.1 points in Belgium between 1992 and 1995, which is largely (if not completely) mirrored in other available data, and an even sharper increase of 4.1 points in a single year in Italy, which is not. ‘Looking at the data’ reveals that the 4.1 Gini-point difference between the European Commission’s data for 2001 and Brandolini’s (2004) data for 2002 is clearly a break in consistency rather than an actual rise in income inequality.

In short, the advice—and the example offered as a model for researchers to follow—provided by Jenkins (2015) is very nearly opposite to that found in Atkinson and Brandolini (2001, 2009). I would further add that this advice also flies in the face of the concerns raised by the ongoing crisis of replication roiling the social sciences and beyond.⁴ The exercise of researcher judgment on which observations to include in an analysis, even when this judgment is ostensibly delimited by explicit conditions, has been the cause of major errors with severe real-world consequences (see, e.g., the documentation by Herndon, Ash, and Pollin 2014 of the problematic data selection in Reinhart and Rogoff 2010). Relying on judgment creates additional researcher degrees of freedom (Simmons, Nelson, and Simonsohn 2011) that make analyses less trustworthy even in the complete absence of ‘fishing expeditions’ and ‘*p*-hacking’ aimed at finding a particular result (Gelman and Loken 2014). This, too, is made clear by reference to the example put forward as a model for researchers in Jenkins (2015).

Why opt to use, for the United States, only the five observations of disposable-income inequality in the UNU data from the LIS (1986, 1991, 1994, 1997, 2000) rather than the complete 17-year-long disposable-income inequality series provided by Brandolini (1998) that is also included in the dataset? There’s no hint in the terms of the declared algorithm: one can only conclude that it must have been a matter of judgment that choosing the longest series was not possible in this case. The sample also excludes the still longer series from the U.S. Census Bureau, which is perhaps justifiable in terms of the algorithm on the grounds that the welfare definition

⁴Atkinson and Brandolini (2001, 2009), of course, were written well before the eruption of this crisis (their later work was accepted for publication on January 3, 2006), and so these concerns played no role in their suggestions.

employed was ‘Monetary Income, Gross’ rather than ‘Income, Gross’ (a distinction that makes very little difference in the U.S. case, as observations from the LIS using both definitions included in the UNU-WIDER (2008a) dataset confirm). Lest one think the U.S. case an aberration, the sample excludes *all but one* observation for Australia.⁵ The UNU-WIDER (2008a) dataset includes 35 observations with complete geographic and population coverage—and top *Quality* scores, no less—for Australia in the time period examined, including a series calculated on the same basis by the Australian Bureau of Statistics in eight of ten years from 1995 to 2004. This series may have been discarded because the welfare definition was ‘Monetary Income, Disposable’ rather than ‘Income, Disposable,’ but in this case too this is a distinction with little difference. Moreover, these decisions are completely invisible to readers (and to reviewers, had the piece been subject to peer review); they only come to light when one is carefully perusing both the original dataset and Jenkins’ replication materials. One could point to still other questionable decisions made in this application of judgment, e.g., regarding Norway. This underscores my point: even if the possibility occurs to them and the necessary replication materials have been made public, neither reviewers nor readers are likely to have the time and inclination to wade through all of the data and confirm that every decision an author makes in applying his or her own judgment is in fact defensible.

Not surprisingly, how one exercises one’s judgment affects the results reached. Table 1 presents replications of Models 3 and 4 of Table 4—that is, the models that employ the hand-selected WIID data and cluster-robust standard errors by country—along with additional models that replace the data selected for the United States and Australia in that piece with the longer Brandolini (1998) and ABS series. The statistically significant results found in Model 3 disappear in Model 3A, while the estimated coefficient for unemployment that is not statistically significant in Model 4 becomes statistically significant in Model 4A. In other words: the advice offered in Jenkins (2015) makes possible virtually any result, and so it makes all results untrustworthy (see Simmons, Nelson, and Simonsohn 2011).

I turn finally to one last bit of ill-considered advice the paper offers as a ‘complementary’ approach for cross-national researchers: the use of so-called

⁵This is the origin of my earlier reference to 20 series though 21 countries are included in this analyses. There could be no break in the data for Australia as it consists of only a single observation.

Table 1: A Paradise for p -Hackers

	Model 3	Model 3A	Model 4	Model 4A
Inflation	0.253** (0.078)	0.187 (0.110)	0.201* (0.098)	0.236** (0.078)
Unemployment	0.361* (0.152)	0.325 (0.174)	0.281 (0.161)	0.316* (0.149)
Time	0.229*** (0.064)	0.110 (0.140)	0.313** (0.115)	0.338** (0.104)
DV Adjustment	<i>no</i>	<i>no</i>	<i>yes</i>	<i>yes</i>
N	242	261	242	261
R^2	0.145	0.068	0.374	0.523

Notes: * $p < .05$, ** $p < .01$, *** $p < .001$; OLS estimates with country-level cluster-robust standard errors. Models 3 and 4 replicate those models in Jenkins (2015), Table 4; Model 4 includes ‘dummy variable adjustments,’ discussed further in the text below. Models 3A and 4A are identical but replace the data selected for the United States and Australia with the longer Brandolini (1998) and Australian Bureau of Statistics series available for those countries (see text). The statistically significant results found in Model 3 disappear in Model 3A, while the estimated coefficient for unemployment that is not statistically significant in Model 4 becomes statistically significant in Model 4A.

‘dummy variable adjustments’ to account for the incomparability in the Gini indices algorithmically selected from the UNU-WIDER (2008a) dataset in accordance with its other recommendations (Jenkins 2015, 38). The piece reiterates (and indeed references as if original) the point that simply including dummy variables for each of the different welfare definitions and equivalence scales upon which these statistics were calculated makes the implausible assumption that the differences between these various bases are constant across countries and over time (see, e.g., Atkinson and Brandolini 2001, 788-790; Atkinson and Brandolini 2009, 388; Solt 2009, 233; Solt forthcoming, 4-6). In light of this, the piece advises that researchers ‘should use a number of carefully-defined interaction variables to account for variations in Gini differences across time and space’ (Jenkins 2015, 38). It offers no further guidance for how one might actually implement this last vague recommendation. This is just as well: such dummy variables, with or without ‘carefully-defined’ interaction terms, introduce bias by comparing the difference between definitions *across different country-years* (see Atkinson and Brandolini 2001, 788).

This is easily seen by imagining two countries, Ruritania and Megalomania. Megalomania has high income inequality and only measures of gross income inequality; Ruritania has low income inequality and only measures of net income inequality. A dummy variable identifying those observations calculated using gross income will capture not only the difference between inequality in gross and net income but also the difference in the level of inequality between Ruritania and Megalomania. Here, again, the recommendations of Jenkins (2015) contradict the advice of Atkinson and Brandolini (2001, 788), who pointed out, ‘A more satisfactory procedure . . . compares *paired* estimates, i.e, when both gross and net incomes are available for *the same country at the same date.*’ One simply should not proceed in the manner recommended in Jenkins (2015). Its guidance is wrong at every step.

Good Advice, and Its Application

I turn now to the *actual* recommendations provided by Atkinson and Brandolini (2001, 2009) and describe how the SWIID incorporates them. Atkinson and Brandolini (2009, 389) enjoin researchers to begin by looking at the data they employ to be sure that it is fit for their purposes, that is, to first identify the relevant variable for their theory rather than allowing data availability to decide for them. The SWIID includes estimates of the Gini index of net-income inequality and of market-income inequality, as well as estimates of absolute redistribution (the difference between those two income-inequality measures) and relative redistribution (the proportional reduction in market-income inequality), thereby meeting the needs of most researchers (see Solt forthcoming, 21).⁶ The sample to be analyzed should likewise be chosen to match the universe implied by theory as closely as possible rather than be dictated by data availability. The SWIID provides the broadest possible sample of countries and years—more than twice as many observations as in the

⁶Most, but certainly not all: those whose theories speak directly to income shares or ratios have no alternative but to instead look to the quintile and decile data provided by UNU-WIDER (2014), carefully selecting only those statistics based on the same welfare definition and equivalence scale. Despite appearances to the contrary, the World Top Incomes Database (Alvaredo et al. 2014) does not provide cross-nationally comparable data for this purpose; substantial differences across countries in the definitions of taxable income and the tax unit as well as in the prevalence of tax avoidance and evasion across incomes led that dataset’s compilers to use it only to compare trends over time across countries, not levels (see Atkinson, Piketty, and Saez 2011, 4-5 and *passim*).

next largest dataset—so as to provide data for testing theories of wide scope (see Solt forthcoming, 13); regional and national sources may be available for tests of theories of more narrow application.⁷

To maximize what Atkinson and Brandolini (2009) refer to as data quality, the SWIID is designed to avoid ‘breaks in consistency’ that occur when series are drawn from different sources and ‘breaks in continuity’ that occur with changes of methodology within the same source. Its first step in addressing these issues is to use *all* of the available data with full geographic and population coverage and for which the welfare definition and equivalence scale are known—drawing on the UNU-WIDER (2008a) dataset, regional sources, national statistical offices, and the academic literature. This follows Atkinson and Brandolini’s (2001, 790) observation—echoed by UNU-WIDER 2008b, 15—that the quality of documentation should not to be used to select data; it also eliminates problematic researcher degrees of freedom (cf. Simmons, Nelson, and Simonsohn 2011). It uses these data to estimate missing values in inequality data available from the Luxembourg Income Study (LIS 2014*a, b*), the source of the most comparable cross-national income inequality data available.

To avoid the biases that result from the use of dummy variables, Atkinson and Brandolini (2001, 788) direct that the adjustments between data using different welfare definitions and equivalence scales should be calculated from comparisons of observations that differ in these ways but describe the same country-year. The SWIID’s estimates of observations missing from the LIS are therefore based on adjustments calculated only from data that share a country and year (see Solt forthcoming, 9). Suppose, for example, that data on Ruritania are available from both the LIS Key Figures and the Ruritanian National Statistics Bureau (RNSB) in 2010 but only from the RNSB in 2011. The SWIID’s adjustment to the 2011 RNSB data would be based on the relationship between the LIS and RNSB data in 2010, not on a comparison of the 2011 RNSB data to the 2010 LIS data.

Atkinson and Brandolini (2009, 389) point out that the theoretical relationship between Gini indices computed on the basis of varying welfare definitions suggests these adjustments should be multiplicative rather than

⁷I note, however, that regional sources are often surprisingly insufficient, even for fairly recent years in data-rich parts of the world. For example, using only Eurostat data to provide information about the context of inequality in which the first four waves of the European Social Survey were conducted (from 2002 to 2009) would result in missing data for more than one-fourth of the country-years in the sample (see Solt 2015).

additive, so the SWIID adjustments are calculated as ratios rather than differences. Because, as Atkinson and Brandolini (2001, 790) observe, welfare definitions and equivalence scales interact, the SWIID’s adjustments are calculated for each distinct *combination* of definition and scale (see Solt forthcoming, 8-9).

Atkinson and Brandolini (2001, 788-790; 2009, 388) document that differences between these welfare definitions and equivalence scales vary across countries and over time and therefore advise that global fixed adjustments are insufficient for making comparable observations that were calculated on varying bases. Therefore, the SWIID completely avoids making such global fixed adjustments for these differences. Rather, its adjustments vary over space and time as much as possible given the available data (see Solt forthcoming, 9-10 for details). Despite the fact that the SWIID includes a great many observations for relatively data-poor parts of the world, over half of its observations are based on adjustments calculated from available data within the same country as the observation, and more than a third vary over time as well; among more advanced countries, these proportions are considerably higher (see Solt forthcoming, 14-15). Adjustments that encompass more time and space are consequently less certain, and indeed the uncertainty can be considerable in relatively data-poor countries. The SWIID therefore incorporates this uncertainty—and the uncertainty due to sampling error present in even the very high quality data provided by the LIS—to ensure that researchers draw correct inferences, and it is pre-formatted for use with the tools available in both Stata and R for data measured with error. Finally, to take into account Atkinson and Brandolini’s (2009, 393) point that ‘income distribution data do not move sharply’ and that observations that indicate otherwise mark breaks in consistency or continuity rather than actual change, a moving-average smoother is applied to the SWIID estimates.

As a result of this careful attention to the sound advice of Atkinson and Brandolini (2001, 2009), the SWIID makes sense of what those authors called the ‘bewildering variety of estimates’ (Atkinson and Brandolini 2001, 784) found elsewhere. Its success in this regard can be seen by comparing its estimates with later releases of data from the LIS. The exercise reveals an impressive record of out-of-sample prediction: for only 5 of the 71 observations, or 7%, are the differences between the LIS Key Figures and previously released data from the SWIID substantively and statistically significant (see Solt forthcoming, 16-18). This record is all the more impressive given that the difference between the LIS data and observations carefully selected from the

UNU-WIDER (2014) dataset to maximize their comparability was substantively and statistically significant in 21% of country-years (Solt forthcoming, 18). The SWIID provides the most comparable data possible for the broadest sample of country-years of any cross-national income-inequality dataset.

Concluding Thoughts

Given that it is purportedly based on the recommendations of Atkinson and Brandolini (2001, 2009), that the guidance of Jenkins (2015) yields such disastrously bad outcomes could be read as a warning against undertaking any broadly cross-national research on income inequality at all. There are hints to this effect in the concluding section of (and indeed throughout) the work, and if this is in fact its author's judgment, it would perhaps be an understandable conclusion from the perspective from one whose own work has been using microdata to produce inequality statistics for single countries. But this perspective is very different from that of those researchers who *do* actually wish to make comparisons across countries and over time. Just how different is thrown into stark relief by the complaint that the SWIID data for the United States do not show the 'sharp discontinuity' between 1992 and 1993 that was caused by the major redesign of the Current Population Survey between those two years (Jenkins 2015, 31). To those for whom the survey methodology is itself an important object of study, the elimination of such a break is quite naturally viewed as a flaw. But for anyone seeking to draw inferences regarding income inequality and its relationships to other variables over space and time, the elimination of artifacts that do not reflect actual changes in inequality is unquestionably a feature, not a bug (see, e.g., Atkinson and Brandolini 2009, 392).

It is fortunate then, that as I have shown above, the guidance in Jenkins (2015) bears little resemblance at all to the advice provided by Atkinson and Brandolini (2001, 2009). The procedure used to generate the SWIID, on the other hand, does match this sound advice. Those pursuing research on income inequality across many countries and over time, as a result, will often find that the SWIID is their best choice of data source.

References

- Alvaredo, Facundo, Anthony B. Atkinson, Thomas Piketty, and Emmanuel Saez. 2014. “The World Top Incomes Database.” <http://topincomes.g-mond.parisschoolofeconomics.eu>. 23 July 2014.
- Atkinson, Anthony B., and Andrea Brandolini. 2001. “Promise and Pitfalls in the Use of ‘Secondary’ Data-Sets: Income Inequality in OECD Countries as a Case Study.” *Journal of Economic Literature* 39(3):771–799.
- Atkinson, Anthony B., and Andrea Brandolini. 2009. “On Data: A Case Study of the Evolution of Income Inequality Across Time and Across Countries.” *Cambridge Journal of Economics* 33(3):381–404.
- Atkinson, Anthony B., Thomas Piketty, and Emmanuel Saez. 2011. “Top Incomes in the Long Run of History.” *Journal of Economic Literature* 49(1):3–71.
- Brandolini, Andrea. 1998. “A Bird’s-Eye View of Long-Run Changes in Income Inequality.” Banca d’Italia.
- Brandolini, Andrea. 2004. *Income Inequality and Poverty in Italy: A Statistical Compendium*. Rome: Banca d’Italia.
- Gelman, Andrew, and Eric Loken. 2014. “The Statistical Crisis in Science.” *American Scientist* 102(6):460.
- Herndon, Thomas, Michael Ash, and Robert Pollin. 2014. “Does High Public Debt Consistently Stifle Economic Growth? A Critique of Reinhart and Rogoff.” *Cambridge Journal of Economics* 38(2):257–279.
- Jenkins, Stephen P. 2015. “World Income Inequality Databases: An Assessment of WIID and SWIID.” *Journal of Economic Inequality* .
- LIS. 2014a. “Inequality and Poverty Key Figures.” <http://www.lisdatacenter.org>. August 2014. Luxembourg: LIS.
- LIS. 2014b. “Luxembourg Income Study Database.” <http://www.lisdatacenter.org>. Multiple countries; August 2014. Luxembourg: LIS.

- Reinhart, Carmen M., and Kenneth S. Rogoff. 2010. "Growth in a Time of Debt." *American Economic Review* 100(2):573–78.
- Simmons, Joseph P., Leif D. Nelson, and Uri Simonsohn. 2011. "False-Positive Psychology: Undisclosed Flexibility in Data Collection and Analysis Allows Presenting Anything as Significant." *Psychological Science* 22(11):1359–1366.
- Solt, Frederick. 2009. "Standardizing the World Income Inequality Database." *Social Science Quarterly* 90(2):231–242.
- Solt, Frederick. 2015. "Economic Inequality and Nonviolent Protest." *Social Science Quarterly* 96.
- Solt, Frederick. forthcoming. "The Standardized World Income Inequality Database." *Social Science Quarterly* . SWIID Version 5.0, October 2014.
- UNU-WIDER. 2008a. "World Income Inequality Database, Version 2.0c, May 2008." http://www.wider.unu.edu/research/Database/en_GB/database.
- UNU-WIDER. 2008b. "World Income Inequality Database User Guide and Data Sources, Version 2.0c." <http://website1.wider.unu.edu/wiid/WIID2c.pdf>.
- UNU-WIDER. 2014. "World Income Inequality Database, Version 3b, September 2014." http://www.wider.unu.edu/research/WIID3-0B/en_GB/database.