Standardizing the World Income Inequality Database — **Source link** ⧉

Frederick Solt

Institutions: Southern Illinois University Carbondale

Related papers:

- Economic Growth and Income Inequality

- A Rational Theory of the Size of Government

- A New Data Set Measuring Income Inequality

- Inequality and Growth in a Panel of Countries

- The Standardized World Income Inequality Database

Luxembourg Income Study
Working Paper Series

Working Paper No. 496

Standardizing the World
Income Inequality Database

Frederick Solt

September 2008

Luxembourg Income Study (LIS), asbl

Standardizing the World Income Inequality Database[*]

Frederick Solt

Assistant Professor of Political Science and Sociology

Southern Illinois University

Carbondale, Illinois 62901

fsolt@siu.edu

July 29, 2008

Abstract

Cross-national research on the causes and consequences of income inequality has been hindered by the limitations of existing inequality datasets: greater coverage across countries and over time is available from these sources only at the cost of significantly reduced comparability across observations. This article presents the Standardized World Income Inequality Database (SWIID), which standardizes the United Nations University database (UNU-WIDER 2008) while minimizing reliance on problematic assumptions by using as much information as possible from proximate years within the same country. The resulting series of gross and net income inequality data maximize comparability for the largest possible sample of countries and years and so are better suited to broadly cross-national research than other sources.

---

[*]The SWIID data, along with replication materials, are available at the author's website: <http://www.siu.edu/~fsolt>.

Introduction

Economic inequality is an enduring focus of inquiry in the social sciences. Economists, sociologists, and political scientists have long sought to explain why incomes are relatively equal in some countries and times and much larger disparities between rich and poor are found in others.[1] The deleterious effects of income inequality on transitions to democracy have similarly received sustained attention, and in recent years there have been repeated calls for additional scholarship on whether income distribution influences other political and social phenomena (e.g., APSA Task Force on Inequality and American Democracy 2004; Neckerman and Torche 2007).[2] Although progress is being made even on these last questions (see, e.g., Anderson and Singer 2008; Petrova 2008; Solt 2008), research on inequality's causes and consequences has been greatly hampered by data issues, namely the limited number and often questionable comparability of the observations available for quantitative cross-national analysis (e.g., Neckerman and Torche 2007, 349).

This article introduces the Standardized World Income Inequality Database (SWIID), which maximizes the comparability of income inequality statistics for the largest possible sample of countries and years and so is better suited than existing income inequality datasets for use by scholars engaged in broadly cross-national research. Existing datasets are strong in either comparability or breadth of coverage; to date, broader coverage is available only at a substantial loss of comparability. The SWIID employs a transparent procedure to increase the comparability of available cross-national inequality data. Although it will not be ideal for all research on economic inequality, its advantages over other cross-national datasets will make it an invaluable resource for those interested in ascertaining the causes and effects of income inequality cross-nationally and over time.

---

[1]Classic works include Kuznets (1955); Lenski (1966); and Meltzer and Richard (1981). Some of the most recent contributions are Huber et al. (2006); Lee, Nielsen, and Alderson (2007); and Bergh and Fink (2008).

[2]Early quantitative studies of inequality's effect on democratization include Russett (1964) and Dahl (1971). For a recent work, see Reenock, Bernhard, and Sobek (2007).

Existing Cross-National Inequality Datasets

There have been many efforts to compile cross-national datasets on income inequality over the last half-century (for a review, see Atkinson and Brandolini 2001). In the past decade or so, two projects have been particularly influential: the Luxembourg Income Study (LIS) and the dataset assembled by Deininger and Squire (1996) for the World Bank. However, both have limitations that have impeded broadly cross-national research. The LIS has generated the most-comparable income inequality statistics currently available but covers relatively few countries and years. The Deininger and Squire dataset and its successors, on the other hand, can be used to provide many more observations, but only at a substantial loss of comparability. These datasets and the tradeoffs involved in using them are discussed in turn below.

The LIS has earned a reputation as the best data available for making cross-national comparisons of income inequality (see, e.g., Smeeding 2005). The LIS team acquires reliable microdata from national household income surveys, carefully harmonizes and standardizes them, and calculates income inequality statistics using a uniform set of assumptions and definitions.[3] Unfortunately, LIS data are at present available for only thirty countries, almost all of which are among the world's richest. On average, inequality in each of these countries is observed in just five years, with most of the observations dating from after 1993. Scholars interested in examining inequality's causes or consequences in a broader spatial or temporal sample are forced to rely on other data.

The most frequently employed alternative to the LIS has been the inequality dataset compiled by Deininger and Squire (1996). Deininger and Squire combined many earlier datasets and evaluated the quality of their observations. On the basis of this evaluation, they identified a series of 682 observations that they labeled 'accept' for use by other researchers. These data have in fact been used extensively: hundreds of cross-national studies have drawn on the Deininger and Squire dataset. Unfortunately, as Deininger and Squire themselves pointed out, the observations are rarely comparable across countries or even over time within a single country because many

---

[3]The LIS also allows interested researchers to calculate their own inequality statistics using different definitions or for specialized purposes—for example, comparing inequality among only those of retirement age—by submitting commands electronically. For more information on the LIS, see its website at http://www.lisproject.org.

are based on different income definitions (e.g., gross or net) and different reference units (e.g., households or persons).

Although often entirely overlooked by researchers using their data, Deininger and Squire recommended two strategies for dealing with this issue. Their first recommendation was to use only those observations that are based on the same type of underlying data. Although likely to provide the most comparability possible, few users of their dataset adopted this strategy because it dramatically reduces the number of observations available for analysis. The most common basis for calculation in the 'accept' series, household gross income, was used by only about one-third of the total observations.

Their alternate recommendation was to calculate the average difference in inequality between observations that varied in their income type or reference unit and then adjust observations by this difference as needed. For example, they found that the Gini indices of observations based on net income were on average 3 points lower than those based on gross income. They therefore recommended adding 3 points to net-income-based inequality observations to make them comparable with the gross-income-based observations. This approach, however, is also problematic. Consider again the difference between gross and net income inequality in a given country and year. This difference depends on the degree to which taxes are progressive and the extent to which government transfers redistribute income to poorer members of society. As a result, it varies greatly across countries and to a lesser extent also over time (see, e.g., Bradley et al. 2003; Bergh 2005). A constant adjustment across all countries and years will therefore underestimate inequality for some observations and overestimate it for others. Similar problems arise with other such constant adjustments.

The successor to the Deininger and Squire dataset is the World Income Inequality Database (WIID), created by the World Institute for Development Economics Research of the United Nations University (UNU-WIDER 2008). The WIID provides the most comprehensive set of income inequality statistics available, incorporating data from both of the datasets described above as well as many additional sources, providing in its current version (2.0c) a total of 5314 observations in 160 countries. The sources of the data are documented, and the income definition, area and population covered, and reference unit for each observation are noted. This information allows researchers to maximize comparability by choosing only those observations with identical values on these criteria. The tradeoff between comparability and coverage, however, remains intact: the most common combination in-

4

cludes just 508 different country-years in only 71 countries and so discards the vast majority of the information in the dataset.[4] Although the WIID makes mistakenly combining non-comparable observations less likely, researchers are faced "with the reverse problem of not knowing how to piece together the information in a meaningful way" (Atkinson and Brandolini 2001).

A recent attempt to address this last issue is the Standardized Income Distribution Database (SIDD) created by Babones and Alvarez-Rivadulla (2007). Using version 1.0 of the WIID, Babones and Alvarez-Rivadulla calculated the average differences between various income definitions and reference units. They then used these findings as constant adjustments for the original WIID data, resulting in a single series representing household per capita gross income inequality. This process does lead to 1218 observations in 143 countries, but, as discussed above, constant adjustments fail to capture the substantial variation across countries and over time in the differences between one income definition or reference unit and another. For this reason, the greater coverage of the SIDD still comes at a significant cost of reliability and comparability.

Constructing a Standardized World Income Inequality Database

The goal of the SWIID, therefore, is to meet the needs of those engaged in broadly cross-national research by maximizing the comparability of income inequality data while maintaining the widest possible coverage across countries and over time. The approach, in brief, is to standardize income inequality observations using as much information as possible from proximate years within the same country. This process is spelled out in detail below.

The starting point is Version 2.0c of the WIID data, released in May

---

[4]Note that "filling out" a variable from the WIID with standard multiple imputation routines, which predict missing values using the other variables within the dataset to be analyzed, does not solve this problem. Because multiple imputation allows more observations to be included, it avoids discarding information about the *other* variables in the analysis and so is preferable to simply excluding observations with missing data in the inequality variable. Inequality values imputed in this way, however, do not add any information to the dataset: uncertainty in the estimates of inequality's effects—and any bias in these estimates due to the sample actually observed—remain just as large (see, e.g., King et al. 2001; Gelman and Hill 2007, 529-543).

2008. The measure of income inequality employed is the Gini index. As scaled in the WIID, the Gini index has a theoretical range from zero, which indicates that each reference unit receives an equal share of income, to one hundred, indicating that a single reference unit receives all income and all others receive nothing. Next, two series of inequality observations—providing information about inequality in gross and net income, respectively—from the LIS are added to the dataset. As the quality and comparability of these data are unparalleled, these observations serve as the baseline to which the WIID data are standardized.

Once the LIS data are added, the first step in standardizing the inequality data is to eliminate those observations that do not provide coverage of all or nearly all of a country's population. Many of the WIID observations cover only urban or rural residents or otherwise omit significant parts of the population. These observations were generally excluded. However, in the absence of any WIID observations with complete coverage for Argentina or Uruguay, and in light of their very high rates of urbanization (approximately 90%), I follow Babones and Alvarez-Rivadulla (2007) in retaining the urban-only observations for these two countries. Historical inequality data predating 1960, which are often based on unreliable surveys, were also removed from the sample.

Next, the data were sorted according to their reference unit and income definition. The WIID dataset contains over two dozen different reference-unit codes, but, as previous researchers have noted (e.g., Babones and Alvarez-Rivadulla 2007, 11), many of these are essentially equivalent. Five distinct reference units can be identified: (1) household per capita, (2) household adult equivalent, (3) household without adjustment, (4) employee, and (5) person.[5] Similarly, although the WIID data are classified into 26 income definitions, these are easily grouped into just four: (1) net income, (2) gross income, (3) expenditures, and (4) unidentified. Rather than assume con-

---

[5] Observations using undocumented country-specific reference units, such as the social assistance household or national scale household equivalent, were disregarded. It is also worth mentioning here that several different definitions of "household adult equivalent" appear in the WIID dataset, including the square root of household size (the definition preferred by the LIS) and the OECD scale. The differences in the Gini indices based on these different definitions of adult equivalent, however, are typically quite small, less than one point on the zero to one hundred scale. I have therefore opted to treat them as a single group to facilitate the standardization process, although at the cost of slightly greater uncertainty.

stant differences across reference units for various income definitions and vice versa, the data were classified according to the *combination* of reference unit and income definition. This yields nineteen categories (no observations provide information about the distribution of consumption per employee). Due to their superior quality, the two series of LIS data, which are based on household adult-equivalent net and gross income, are considered separate categories, bringing the total number of categories of data to twenty-one. Rather than choose among sources, when more than one observation was available within a category for a particular country and year, these observations were averaged.

This provides a dataset of country-year observations, each of which has data on inequality in one or more of the twenty-one categories. What is needed to generate a series with data on all countries and years from the incomplete inequality variables in twenty-one categories are the ratios between each pair of variables. If the ratio ρ_{ab} between the Gini index data in categories a and b were known, missing observations in a could be replaced simply by multiplying available data in b by ρ_{ab}. But as noted previously, the relationship between Gini indices with different reference units and income definitions will vary considerably from country to country and also over time depending on the extent of redistributive policies, details of tax law, patterns of consumption and savings, family structure, and other factors. In other words, ρ_{ab} is not constant but varies across countries i and years t. Further, ρ_{abit} is only directly calculable for those pairs of categories in those countries and years for which it is not immediately useful, that is, only when data is already available in both categories for that observation.

Those ratios ρ_{abit} that are directly calculable are valuable nevertheless because they provide information about what the ratios that are missing are likely to be. Because the factors that affect these ratios—redistributive policies, patterns of consumption, and so on—tend to change only slowly over time within a given country, the best prediction for a missing ratio will be based on available data on the same ratio in the same country in proximate years, thereby minimizing any differences in these factors. With this in mind, the ratios ρ_{abit} were predicted from the results of a series of models.

First, in those countries with sufficient data, predictions were generated by loess regression, which incorporates the maximum amount of information from proximate years by fitting a smooth curve point-by-point through the available data. Next, predictions were generated through multilevel modeling (Gelman and Hill 2007, 272-275). In order of increasing availability—

but also increasing uncertainty as reflected in larger standard errors—$\hat{\rho}_{abit}$ was predicted as a function of (1) country-decade, (2) country and region-decade, (3) country, (4) region-decade, (5) region, and (6) advanced or developing world. The predictions of all of these models were then combined for each ratio ρ_{abit}, assigning each country-year the available prediction with the smallest standard error.

These predictions $\hat{\rho}_{abit}$ alone, however, do not take advantage of all of the information available in the WIID data. An additional prediction of each conversion factor can be generated in a two-step process through other categories of data. That is, the ratio of the LIS net-income data (labeled category 1) to the data in category b can be calculated as the product of the ratio between data in category a and category b and the ratio of the LIS net-income data to data in category b:

$$\hat{\rho}_{1bit} = \hat{\rho}_{abit} \times \hat{\rho}_{1ait} \tag{1}$$

These two-step predictions improve upon the conversion factors predicted in one step in two ways. First, for some combinations of a and b, few or no observations of both categories of the Gini index are available, making modeling $\hat{\rho}_{abit}$ in one step impractical or impossible. Second, the uncertainty in the predicted conversion factor can often be reduced by averaging the one-step prediction with one or more two-step predictions.

Once all of the predicted ratios $\hat{\rho}_{abit}$ were calculated, twenty series of estimates comparable with the LIS net-income data were gained by multiplying these predicted ratios by the available data in each of the twenty other categories. Because each of these comparable series is incomplete, they were combined into a single variable by assigning each observation with the estimate with the least uncertainty or, when the average of some or all of the available estimates yielded an even smaller standard error, this average.

A final piece of information about the income inequality in a particular country and year is gained by noting that the distribution of income within a country typically changes only slowly over time: contemporary levels of inequality should generally be very similar to levels observed in the preceding year. With two exceptions discussed below, dramatic differences in the estimates of inequality for a given year and those preceding and following it likely reflect persisting errors in measurement. Allowing observations to

be informed by the estimates for surrounding years works to minimize such errors. This was achieved by using the five-year weighted moving average algorithm presented in Equation 2:

$$G_{it} = \frac{1}{6} \times (G_{it-2} + G_{it-1} + (2 \times G_{it}) + G_{it+1} + G_{it+2}) \qquad (2)$$

The first exception to the foregoing regards the data from the Luxembourg Income Study. Because of the very high quality of the LIS data, differences from one year to the next are unlikely to be caused by persistent measurement error, so observations from this source were therefore not adjusted with the moving average algorithm.[6] The second exception involves the countries of eastern Europe and the Soviet Union during the collapse of communist rule. The sharp increases in inequality observed in most of these countries from 1990 to 1991 would appear to be due to the profound restructuring of these countries' societies and economies rather than measurement error. Applying the moving average algorithm to this region results in overestimates of inequality in 1989 and 1990 and underestimates in 1991 and 1992; therefore the algorithm was not used in these countries during these years.

Simply applying the moving-average algorithm to the net-income inequality variable, however, would lose the estimates of uncertainty associated with each observation. Therefore, the variable was re-generated one thousand times through Monte Carlo simulation and the moving-average algorithm applied to each simulation. Values for all missing data between observations for years after 1975 were also interpolated for each simulation. Finally, the thousand simulations were averaged to generate a final series of point estimates of LIS-comparable net-income inequality and associated standard errors. The entire process was then repeated to generate a series standardized on the LIS household adult equivalent gross-income data. The final dataset covers 153 countries, with 3331 country-year observations on net inequality and 3273 country-year observations of gross inequality.

---

[6]The measurement errors in the LIS data (which average 0.38 points in the net-income series and 0.44 points in the gross-income series) were retained, however. On the uncertainty in the LIS data, see http://www.lisproject.org/keyfigures/standarderrors.htm.

Assessing the SWIID

Before this dataset can be useful to researchers, however, its reliability and validity must be assessed. First, we examine the uncertainty in the SWIID estimates. For a particular country and year, the size of the standard error depends largely on how much data is available for that observation and for its country and region in other years, or more precisely on how frequently the original source category or categories share observations with the other categories in the dataset. About 30% of the observations have associated standard errors of 1 point or less on the 0 to 100 scale of the Gini index. Over 60% of the standard errors are less than 2 points, and more than 85% are less than 3 points. On the other hand, 99 observations, or slightly fewer than 3%, have standard errors greater than 5 points, and 11 standard errors (0.3%) are greater than 10 points. Figure 1 displays a box plot of these standard errors broken down by region.[7] The largest standard errors are concentrated in the developing world, especially in the countries of Latin America and the Caribbean and of Africa; this is not surprising, given the relative paucity of data for many countries in those regions.[8] The relatively small uncertainty in most observations suggests that, although the SWIID standardization process is not perfect, it is nevertheless quite good. Moreover, the inclusion of standard errors in the SWIID allows researchers to take the remaining uncertainty into account by excluding the least reliable observations or, ideally, by performing their analyses using multiple Monte Carlo simulations of the SWIID data and averaging the results (see King et al. 2001).

---

[7]Consistent with convention, the top and bottom of each box indicate the 75% and 25% percentiles of the standard error in the region and the band within the box marks the median; the whiskers extend to the standard error farthest from the median but not farther from the box's top or bottom than 1.5 times the height of the box. The open circles mark all observations that lie beyond the whiskers.

[8]The largest standard errors are also concentrated in the earlier years of the period covered by the SWIID: about 80% of the standard errors over 5 (and all of the standard errors over 10) date to 1980 or before. The countries with the most observations in the SWIID net inequality series with standard errors over 5 points are Morocco (5 observations: 1960, 1965, 1975-1976, 1980); South Africa (5 observations: 1960, 1970, 1975-1976, 1980); Kenya (6 observations: 1960-1961, 1964, 1967, 1971, 1974); Malawi (8 observations: 1969, 1977-1979, 1982-1985); Jamaica (10 observations: 1968, 1973, 1975-1982); and Sierra Leone (11 observations: 1968, 1976-1985). Sierra Leone accounts for 4 of the 11 observations with standard errors over 10; no other country has more than one observation in this category. The largest standard error, 17.8 points, is for Swaziland in 1974.

Figure 1: Standard Errors in SWIID Net Inequality by Region
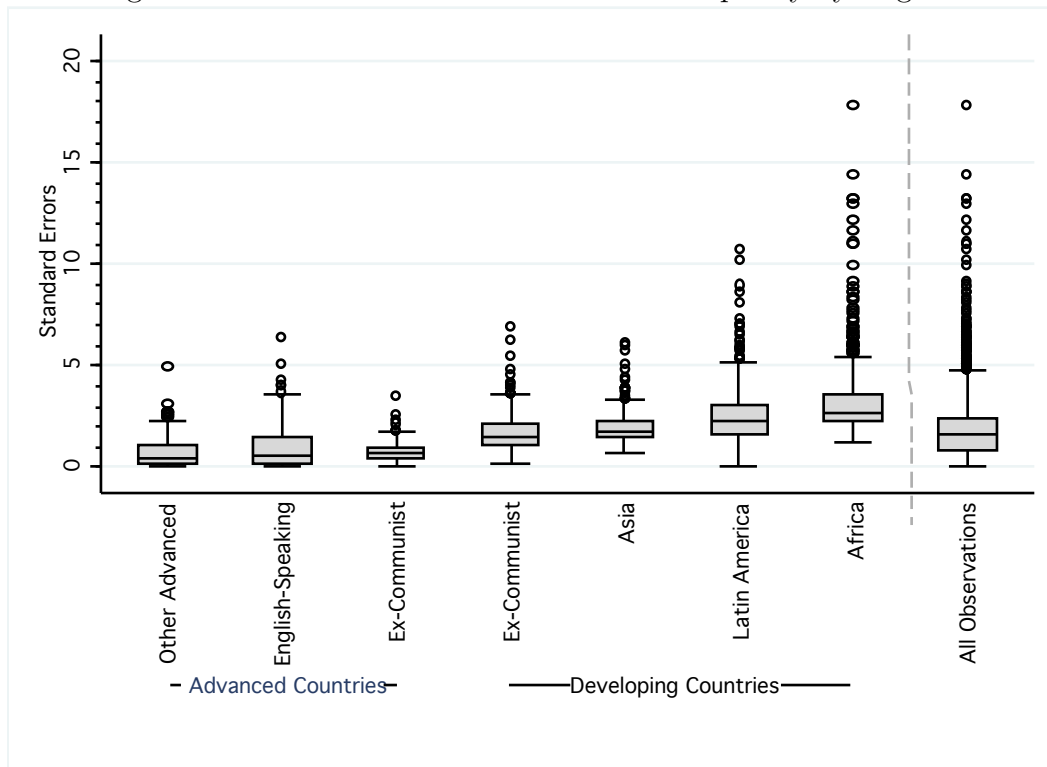
Another simple assessment of the SWIID is a comparison of the relationship of gross and net income inequality between, on the one hand, the countries of the developing world and, on the other, the advanced industrial countries. In the developing world, where taxes are quite uniformly low and effective policies to redistribute income are very rare, gross and net income inequality should be very highly related. Some advanced countries, however, engage in substantially more redistribution than others, attenuating the relationship between gross and net income inequality within this group of countries. The correlation between gross and net income inequality, therefore, should be considerably lower among the advanced countries than among the developing countries. This is in fact the case in the SWIID data. Among the countries of the developing world, the correlation between gross and net income inequality is .967; this correlation is only .749 among the advanced industrial countries. Put differently, in the developing world, differences in gross income inequality explain 93.5% of the variance in net income inequal-

ity, but in the advanced countries, they explain only 56.1%. Among the advanced countries, as expected, differences in redistributive policies are much more important to explaining differences in net inequality.

A final means of assessing a cross-national inequality dataset like the SWIID is to examine its relationships with various social indicators (see Babones and Alvarez-Rivadulla 2007). For this purpose, I employ data collected by the U.S. Census Bureau (2007) on life expectancy at birth and infant mortality rates and compare the correlations of those indicators with the net and gross income inequality data from the SWIID against those obtained using the SIDD.[9] The correlations between these indicators and the 'accept' series compiled by Deininger and Squire (1996) are also included to provide an additional point of comparison. Table 1 presents the results. The net and gross income inequality data of the SWIID exhibit much higher correlations with life expectancy and infant mortality than do the SIDD or Deininger and Squire data. Further, the SWIID provides from three times to more than eight times as many observations as the other datasets and so provides a much firmer basis for drawing general conclusions about any relationships between these variables. Although not shown, the SWIID correlations are even larger when only those observations for which SIDD or Deininger and Squire data are available are considered. The SWIID is a substantial improvement over these older cross-national datasets.

Table 1: Correlations Between Income Inequality Measures, Life Expectancy, and Infant Mortality Rate

	SWIID		SIDD	D&S
	Net	Gross	Gross	Accept
Life Expectancy at Birth	−.505	−.480	−.355	−.245
	(2277)	(2265)	(590)	(267)
Infant Mortality Rate	.529	.487	.434	.353
	(2500)	(2488)	(718)	(373)

The first row presents bivariate correlations; the number of observations appears below in parentheses.

[9]On the relationship between income inequality and life expectancy, see, e.g., Moore (2006); on income inequality and infant mortality, see, e.g., Mayer and Sarin (2005).

Using the SWIID for Cross-National Research

The SWIID series for net and gross income inequality are available in formats convenient for use by researchers from the author's website: <http://www.siu.edu/~fsolt>. The command files used to generate the SWIID are also available from the same source for modification by those with specialized needs. As new data become available—the WIID continues to be revised and expanded, and the LIS has announced that it will soon initiate coverage of several new countries in East Asia and Latin America—the SWIID will be updated to incorporate them.

I conclude by emphasizing that the SWIID represents a particular choice in the balance between comparability and coverage: it maximizes comparability for the broadest available set of country-year observations. This trade-off will suit the needs of many scholars engaged in broadly cross-national research, but clearly it will not be the most appropriate option for all applications. Greater comparability can often be achieved when one's scope of inquiry is narrower. The high quality, superior comparability, and great flexibility of the data available from the LIS will continue to make it the preferred source for many cross-national studies of inequality in the advanced countries. But even the LIS data are not perfectly comparable. Those investigating the development of income inequality over time within one or a few countries are therefore best advised to seek out the original sources cited in the WIID as well as other national sources and become familiar with the exact assumptions and definitions they employ (see Atkinson and Brandolini 2001). Approaches using all of these data sources hold promise for improving our still-limited understanding of the causes and consequences of economic inequality.

References

Anderson, Christopher J., and Matthew M. Singer. 2008. "The Sensitive Left and the Impervious Right: Multilevel Models and the Politics of Inequality, Ideology, and Legitimacy in Europe." *Comparative Political Studies* 41(4-5):564–599.

APSA Task Force on Inequality and American Democracy. 2004. "American Democracy in an Age of Rising Inequality." *Perspectives on Politics* 2(4):651–670.

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.

Babones, Salvatore J., and María José Alvarez-Rivadulla. 2007. "Standardized Income Inequality Data for Use in Cross-National Research." *Sociological Inquiry* 77(1):3–22.

Bergh, Andreas. 2005. "On the Counterfactual Problem of Welfare State Research: How Can We Measure Redistribution?" *European Sociological Review* 21(4):345–357.

Bergh, Andreas, and Günther Fink. 2008. "Higher Education Policy, Enrollment, and Income Inequality." *Social Science Quarterly* 89(1):217–235.

Bradley, David, Evelyne Huber, Stephanie Moller, François Nielsen, and John D. Stephens. 2003. "Distribution and Redistribution in Postindustrial Democracies." *World Politics* 55(2):193–228.

Dahl, Robert A. 1971. *Polyarchy: Participation and Opposition.* New Haven: Yale University Press.

Deininger, Klaus, and Lyn Squire. 1996. "A New Data Set Measuring Income Inequality." *World Bank Economic Review* 10(3):565–591.

Gelman, Andrew, and Jennifer Hill. 2007. *Data Analysis Using Regression and Multilevel/Hierarchical Models.* Cambridge: Cambridge University Press.

Huber, Evelyne, François Nielsen, Jenny Pribble, and John D. Stephens. 2006. "Politics and Inequality in Latin America and the Caribbean." *American Sociological Review* 71(6):943–963.

King, Gary, James Honaker, Anne Joseph, and Kenneth Scheve. 2001. "Analyzing Incomplete Political Science Data: An Alternative Algorithm for Multiple Imputation." *American Political Science Review* 95(1):49–69.

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

Lee, Cheol-Sung, François Nielsen, and Arthur S. Alderson. 2007. "Income Inequality, Global Economy, and the State." *Social Forces* 86(1):77–111.

Lenski, Gerhard E. 1966. *Power and Privilege: A Theory of Social Stratification.* New York: McGraw-Hill.

Mayer, Susan E., and Ankur Sarin. 2005. "An Assessment of Some Mechanisms Linking Economic Inequality and Infant Mortality." *Social Science & Medicine* 60(3):439–660.

Meltzer, Allan H., and Scott F. Richard. 1981. "A Rational Theory of the Size of Government." *Journal of Political Economy* 89(5):914–927.

Moore, Spencer. 2006. "Peripherality, Income Inequality, and Life Expectancy: Revisiting the Income Inequality Hypothesis." *International Journal of Epidemiology* 35(3):623–632.

Neckerman, Kathryn M., and Florencia Torche. 2007. "Inequality: Causes and Consequences." *Annual Review of Sociology* 33:335–37.

Petrova, Maria. 2008. "Inequality and Media Capture." *Journal of Public Economics* 92(1-2):183–212.

Reenock, Christopher, Michael Bernhard, and David Sobek. 2007. "Regressive Socioeconomic Distribution and Democratic Survival." *International Studies Quarterly* 51(3):677–699.

Russett, Bruce M. 1964. "Inequality and Instability: The Relation of Land Tenure to Politics." *World Politics* 16(3):442–454.

Smeeding, Timothy M. 2005. "Public Policy, Economic Inequality, and Poverty: The United States in Comparative Perspective." *Social Science Quarterly* 86(S1):955–983.

Solt, Frederick. 2008. "Economic Inequality and Democratic Political Engagement." *American Journal of Political Science* 52(1):48–60.

UNU-WIDER. 2008. "World Income Inequality Database, Version 2.0c, May 2008." Available online at http://www.wider.unu.edu/research/Database/en_GB/database/.

U.S. Census Bureau. 2007. "International Data Base." Available online at http://www.census.gov/ipc/www/idb/idbprint.html.