Income Measures in Cross-National Surveys: Problems and Solutions*

MICHAEL J DONNELLY AND GRIGORE POP-ELECHES

Comparable household income measures are crucial for most social science analyses of cross-national public opinion survey data. However, income questions in many cross-national surveys suffer from comparability and interpretability limitations that have not been adequately addressed by the existing literature. In this article, we examine the income measure in one major survey, the World Values Survey (WVS), arguing that a variety of problems arise when drawing inferences—descriptive or causal, individual or aggregate—using the standard ten-category measure. We then propose and implement a number of corrections to these potential biases and present a series of diagnostics that confirm the importance of our proposed corrections. We conclude by documenting some of the same challenges in the income measures used in other cross-national surveys. The accompanying data set can be merged with the WVS to make better use of the income measure.

INCOME MEASUREMENT IN CROSS-NATIONAL SURVEYS

The comparability of responses in cross-national surveys has triggered significant methodological debates (King et al. 2004; Stegmueller 2011). In this article, we address a question that should at least in theory be more easily comparable across different cultural contexts: household income. Income is a key explanatory variable for many prominent theoretical questions in political behavior (ranging from economic voting to post-materialist values) and represents a standard demographic control variable for most other analyses. Most theoretical approaches in comparative politics assume that income matters, that it shapes interests, values, ideologies, or capabilities.

Unfortunately, because of inconsistencies in how the questions are administered and coded, responses to household income questions are often not meaningfully comparable across different countries and time periods. While in the conclusion, we show that these problems also affect several other widely used cross-national surveys, we here focus primarily on the World Values Survey (WVS), an important source for cross-national public opinion analyses.[1] First, we document some of the inconsistencies in how household income was measured across surveys. Second, we discuss the implications of these measurement choices cross-national survey analyses and illustrate some of the ways in which the variable has been improperly used and interpreted in previous studies. In the final section, we propose three alternative cross-nationally comparable income measures, describe the methodological approaches for calculating them, and explain how scholars can use our data to conduct more empirically sound tests of their theories. A few researchers have noted measurement and coverage problems with the WVS income variable (Layard, Mayraz and Nickell 2008) and others have opted not to use

* Michael J. Donnelly, Assistant Professor of Political Science and Public Policy, Department of Political Science, School of Public Policy and Governance, University of Toronto, 100 St. George St, Room 3018, Toronto, ON M5S 3G3 (mj.donnelly@utoronto.ca). Grigore Pop-Eleches, Professor of Politics and Public and International Affairs, 220 Bendheim Hall, Princeton University - Woodrow Wilson School and Department of Politics, Princeton, NJ 08544 (gpop@princeton.edu). To view supplementary material for this article, please visit http://dx.doi.org/10.1017/psrm.2016.40

    [1] A Google Scholar search for "World Values Survey" produces >23,000 results.

the WVS income question at all and have instead relied either on alternative individual-level survey questions like social class to capture relative income (Andersen and Fetner 2008) or on country-level gross domestic product (GDP)/capita to capture cross-country income differences (Pop-Eleches and Tucker 2012). However, none of the critiques have focused on the problem posed by uneven distributions of income scales across countries and, more importantly, none have provided a systematic solution to facilitate the valid analysis and meaningful interpretation of the cross-country effects of household income.

The WVS income question is administered in three forms. The standard approach, which is similar to other surveys, including the International Social Survey Programme (ISSP), the European Social Survey (ESS), and the Eurobarometer, is to give respondents a showcard with ten income brackets, each labeled with a letter, and to ask them to place themselves in one of the brackets.[2] The income brackets (in the local currency) are defined by the country's principal investigators (PIs), and are intended to represent the deciles of the household income distribution. This approach was followed and the values of the bracket cut-points are available for 152 of the 245 surveys included in the cumulative file of the five-wave WVS. The top right panel of Figure 1 shows a typical result from this approach. Egyptian respondents from the 2000 survey are spread across all the income brackets, and while none of the categories match 10 percent exactly and at least two of them are significantly different, most come close. The flaws could be caused by poor sampling, poor response rates, or a misspecification of the income brackets.[3]

The second approach (used in Norway in 1996 and Romania in 2005) is to ask respondents to write in their household income, and to recode this into brackets later.[4] The bottom right panel in Figure 1 shows the resulting income distribution in one such survey.

The third approach asks respondents to rank their household's income on a scale from 1 to 10 but without providing specific income bands.[5] Despite the similarity in question wording, this approach yields a measure with a very different interpretation, since it taps subjective rather than objective income assessments. Given these fundamental differences, we argue that the income measures from these surveys[6] should not be merged into a unified WVS income measure but rather treated as a separate subjective income indicator.[7] Besides the use of subjective incomes and the lack of documentation for some country-waves, there is a substantial problem with many surveys that asked the income question in the standard format. While the PIs were instructed to choose brackets that correspond to income deciles, it appears that some

---

[2] The question wording, with minor variations, was as follows: "Here is a scale of incomes. We would like to know in what group your household is, counting all wages, salaries, pensions and other incomes that come in. Just give the letter of the group your household falls into, before taxes and other deductions."

[3] The online supplementary information includes a figure showing the distributions of income for all country-waves.

[4] The wording for the Romania (2005) survey was "Taking into account all the income in your household, from all members, in the past month, what was the total income?"

[5] The complete question wording was as follows: "Here is a scale of incomes on which 1 indicates the 'lowest income decile' and 10 the 'highest income decile' in your country. We would like to know in what group your household is. Please, specify the appropriate number, counting all wages, salaries, pensions and other incomes that come in."

[6] From the documentation, we can confidently determine that 136 surveys used the first (objective categories) method, two used the second (fill in the blank) method, and 22 used the third (subjective) method. Of the remaining 81 country-waves, Donnelly and Pop-Eleches (2016) estimate that eight were conducted using the subjective measure.

[7] Based on our inquiries, which included a consultation of the WVS archives and documentation, and e-mail inquiries to the principal investigators for surveys without explicitly listed income bands, we determined that 22 surveys have used this subjective income question. In a separate analysis, we discuss the methodological implications of this problem in greater detail (Donnelly and Pop-Eleches 2016).
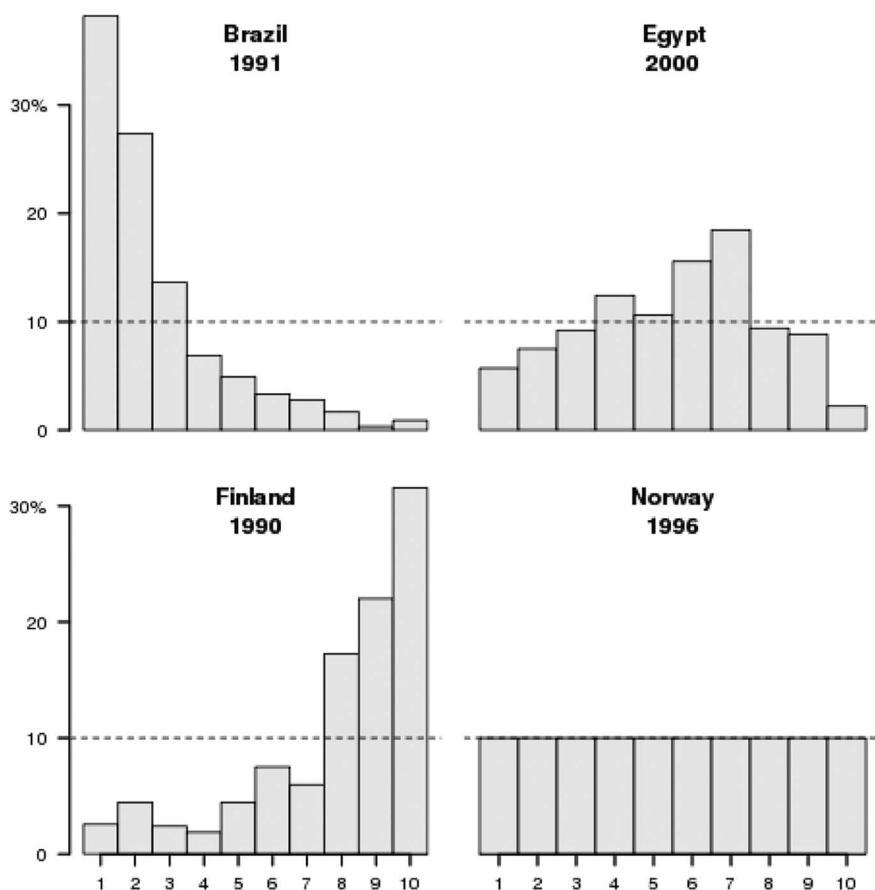
Fig. 1. *Weighted distribution of incomes in four World Values Survey*

organizers used other criteria for choosing brackets or relied on faulty (perhaps outdated) income distribution data. Therefore, as illustrated by the Brazilian and Finnish survey in Figure 1, income distributions often deviate substantially from the deciles envisioned by the WVS research design.

Analytical Challenges

The 10-point income scale reported by WVS cannot be interpreted as a straightforward indicator of household income in a cross-national regression context for several reasons. First, to the extent that we are interested in cross-national comparisons of income levels, the 10-point scale corresponds to very different income intervals in different countries: thus, a US citizen in category 4 has a significantly higher income than a Moldovan respondent with the same rating on the 10-point WVS income scale. Nor do these problems disappear in fixed-effects models given that an income shift from 2 to 4 in the United States corresponds to a much greater increase in household income than a corresponding shift in Moldova.

Even among countries with similar income levels the differences between categories is likely to be affected by differences in income inequality. To draw meaningful inferences about income levels, we would need to either restrict our analysis to a single country or to countries with very

similar income levels and income distributions[8] or—better yet—transform the survey-specific income intervals into a common cross-nationally and cross-temporally comparable income scale (as proposed in the following section). The choice of scale—relative or absolute, logged or otherwise transformed—should be theory driven.

Even if we are interested in the effects of a respondent's income relative to her compatriots, the interpretation of the WVS income measure is not straightforward. Most studies have used the original WVS 10-point scale as a continuous measure and have interpreted it as capturing *income deciles* (Mishler and Rose 2001; Beramendi and Anderson 2008; Inglehart et al. 2008). Some analysts have collapsed deciles into quintiles (Solt 2008) or have used a three-category income measure which combines varying combinations of the 10-point income scale into low- middle-, and high-income groups approximating terciles (Alesina and Giuliano 2009; Donnelly and Lefkofridi 2014).[9] The problem with this interpretation is illustrated most starkly by the income distribution from the Brazilian and Finnish surveys in Figure 1: a respondent in the fifth income category is relatively well off in the Brazil 1990 survey (88.6 percentile) but rather poor in the Finland 1991 survey (13.5 percentile). Even for less extreme examples many of the income categories do not come close to capturing actual deciles, which means that even for fixed-effects models a one-unit change is not a meaningful metric across surveys. Indeed, we find that the vast majority of WVS do not have the uniform distribution necessary to interpret them as deciles: running a set of weighted χ^2 tests, we failed to reject the null that the income scale is uniform—that is, that the categories represent deciles and respondents are sampled randomly from the population—in only seven of the 241 surveys in the WVS cumulative file.[10]

While this problem could be addressed by using the 10-point scale to calculate the cumulative distribution of incomes in a given survey and thereby determine the relative position of respondents in their country's income distribution, only one of the studies we surveyed used this approach.[11] Such an approach is useful for analyzing how relative income affects attitudes and behavior—for example, with respect to redistribution—but it is less useful for capturing the impact of overall income differentials. To address the challenges inherent in the multiple interpretations of income, in the next section we provide a menu of options to allow researchers to choose the most appropriate approach for testing their particular theory.

SOLUTIONS

Our first step toward creating cross-nationally comparable income measures is to select country-waves where documentation on income brackets is complete. This is the case in 137 of the 241 surveys of the first five waves of the WVS. Using International Monetary Fund (IMF) data on exchange rates, purchasing power parity (PPP) conversion factors, and inflation, we calculate the value of each bracket cut-point for every country-wave. After assigning the first nine

---

[8] While analyses focusing on a small sample of Organisation for Economic Co-operation and Development (OECD) democracies in a short time period (e.g. in De La O and Rodden 2008) approximate this condition, they are still limited by the fact that the distribution of WVS income categories differs significantly across surveys.

[9] While this measure deviates less from a uniform distribution because it takes into account the varying size of the original categories, the size of these "terciles" still varies significantly across countries (from around 15 to over 50 percent) thereby creating similar problems as the use of deciles.

[10] We used a 95 percent threshold. Some deviation is to be expected even in surveys with carefully designed categories, since response rates and other deviations may not be taken into account by the survey weights.

[11] The exception is Weakliem, Andersen and Heath (2005), which used regression-based income percentile estimates.

Fig. 2. Relationship between GDP/capita and the per capita household income in the World Values Survey
Note: PPP = purchasing power parity.

categories their respective midpoints and assigning those in the tenth category (whose upper bound is infinite) a value based on the size of the ninth category and the value of the lower bound of the tenth,[12] we can assign a PPP (or exchange rate) income in 2005 US dollars. To test the plausibility of our calculations, we estimated the PPP and exchange rate average household incomes for each country-wave, divided by an estimate of the average household size, and compared it with GDP/capita as measured by the World Bank. Figure 2 shows that while our measurements are not the same as those calculated from aggregate economic data, they are closely related. Since household monetary income ignore business and in-kind benefits, it is unsurprising that the points generally fall below the 45° line. We have also calculated the correlations between the Gini index based on our data and Solt's (2009) measure and between economic growth between successive rounds as calculated from our data and that of the IMF. These correlations, while weaker, remain reasonably strong (0.74 and 0.45, respectively).

To calculate income levels for surveys where data on income brackets was not available in the WVS documentation, we devised an imputation approach that uses the empirical cumulative distribution, per capita GDP, and income inequality data to generate estimates of the income of individuals at a particular point in the income distribution.

The inequality data from the World Development Indicators (World Bank 2011) consist of estimates of the share of national income held by the top and bottom deciles and each of the quintiles. This provides us with an approximation of the cumulative distribution of income in those country-years. We can then map an estimate of the individual's place in the income distribution (based on a weighted ranking of the sample) onto a PPP income distribution. Our technique requires minimizing the distance between the distribution implied by the WDI income

---

[12] We use the width of the ninth category plus the lower bound of the tenth. Thus, if the ninth category ranges from $50,000 to 60,000, we impute the average income for the tenth as $70,000. Rueda, Stegmueller and Idema (2014) take a similar approach using ESS data. Following Hout (2004), they use midpoints for all but the top categories, and assume a Pareto distribution to get the estimate for the top category. We used the simpler non-distributional approach because it produced more sensible results in those countries where almost all respondents were in the top few categories, a problem that is much more common in the WVS than in the ESS.

shares and a lognormal[13] distribution by manipulating the lognormal distribution's location and scale parameters. We do this by calculating

$$(\mu^*, \sigma^*) = \underset{\mu, \sigma}{\operatorname{argmin}} \sqrt{\sum_{i=1}^{6} (q_i - F^{-1}(p_i, \mu, \sigma))^2}, \qquad (1)$$

where q is a vector of the six percentiles for which the income shares are available from the WDI (10, 20, 40, 60, 80, 90), p a vector of those corresponding cumulative probabilities, and F^{-1} the inverse cumulative distribution function (CDF) of the lognormal distribution.[14] We then use the distribution that minimizes this function—that is, the lognormal "shape" that best fits the WDI data—to calculate incomes corresponding to each percentile and divide the country's per capita GDP by the mean of the implied distribution to give the factor by which the income should be multiplied to estimate the true incomes of individuals at every percentile.

The parameters of the lognormal resulting from Expression (1) allow us to map from any given income percentile to an estimated income by inflating the value of the distribution corresponding to each percentile by a factor defined by the following:

$$\frac{GDPpc_{cy}}{\mu^* + {\sigma^{*2}}/{2}}, \qquad (2)$$

where the denominator is the standard formula for the mean of a lognormal distribution, and the numerator is GDP/capita of country c in year y.[15]

After imputing estimates for the income of individuals at 11 points in the income distribution, we calculate the ratio between the imputed and actual income for all surveys where income brackets were included in the WVS documentation. We then regress this ratio on a set of country dummies and several country characteristics (such as household size, GDP/capita, and income inequality) and the particular percentile being estimated.[16] We use those regressions to predict the same ratios in those country-waves for which brackets are missing. In other words, we use information from the survey-based income measure, the lognormal income measure, and macroeconomic data to estimate the survey-based income measure for surveys in which the income categories are unknown.

From these regressions, we obtain income-level predictions for all of the country-waves in our data set, and assign the appropriate income level (in 2005 PPP US$) to each survey respondent.

The individual-level correlation between our imputed income measure and the survey-based income (for respondents from surveys where we have the WVS documentation on the income bands) is 0.918. This correlation is generally stronger within country-waves, dropping below 0.9 in only one survey (Malta, Wave 1). It is also strong within each band on the 10-point scale,

---

[13] We use a lognormal, rather than, for instance, a Pareto distribution, because our goal is to approximate the income of the largest possible group of people. This is probably appropriate for most of the distribution, but previous work has suggested that the lognormal is a poor fit for incomes at the extremes. Thus, scholars interested in inequality or other phenomena driven by the very top or very bottom of the income distribution should take caution in using our income estimates. See Cowell (2011), for a detailed discussion of these tradeoffs.

[14] Sala-i-Martin and Pinkovskiy (2009) use a similar least squares approach to estimate the Gini index within countries using only income share data.

[15] This equation is simply a shape-preserving shift in the distribution.

[16] See the appendix for additional details. For countries which do not have brackets in any waves (or which only appear once), we refit similar regressions without country fixed-effects, and progressively simplify the regressions for countries with less and less available macroeconomic data.

ranging from 0.82 to 0.89. We consider this a sufficiently strong relationship to justify using the imputed measure for many applications.

USING THE DATA

To facilitate the appropriate use of World Values income data, we provide, in the online appendix, a data file containing information on the income measure for each country-wave. This information includes a mapping from the original income codes to the PPP-adjusted income levels as measured from the survey and a similar mapping for the distribution-based incomes. We also provide income percentiles and ratios corresponding to each income code. Users will be able to merge these data into the WVS and include any of these measures in their analyses. We suggest that users carefully consider which measure best tests their theory, using percentiles when the focus is on a respondent's rank in society, ratios when the relative magnitudes of income matter, and levels when the absolute differences between respondents are of concern.

CONCLUSION

In this article, we have identified several factors that limit the interpretability of the household income measure in the WVS, one of the largest and most widely used set of cross-national and cross-temporal public opinion surveys. The problems include the inconsistent implementation of the income question, and the large variations in the shares of respondents in the ten categories of the WVS income scale, which questions the validity of its widespread interpretation as representing income deciles. Therefore, we have argued that a valid cross-national and cross-temporal analysis of the effects of income on political attitudes using WVS data requires the creation of a new set of variables that allow for meaningful comparison across countries and time periods, and we have proposed and implemented three alternative income measures that capture different facets of income differences: levels, ratios, and percentiles.

These challenges are not unique to the WVS. Designing income categories is difficult, especially in a cross-national setting. Thus, many of the most widely used cross-national surveys have at least some cases where income variables deviate significantly from the uniform distributions that would be necessary to interpret them straightforwardly as interval scales. For instance, the ESS's sixth round (2012–2013) in Ireland uses carefully documented categories, adjusted for inflation, yet they produce the distribution displayed in Figure 3. It shows an undoubtedly skewed distribution, which cannot be interpreted meaningfully as capturing deciles. We found similarly non-uniform income distributions in other cross-national surveys with income scales designed to capture deciles, as in the ISSP and the Candidate Countries Eurobarometer, or quintiles, as in the Comparative Study of Electoral Systems.[17]

All of this suggests is that users of cross-national surveys should carefully check their income measures and use the techniques we describe (or an alternative solution) to adjust the data such that it is consistent and substantively meaningful across countries and/or time periods. The alternatives—using questions about social class, financial satisfaction, or possession of major consumer products—will be useful in some cases, but will prevent scholars from testing some of the key questions in comparative politics.

---

[17] In all cases, we used the demographic weights provided by the survey files in calculating the distributions, which means that the deviations from uniformity are not eliminated by reweighting.
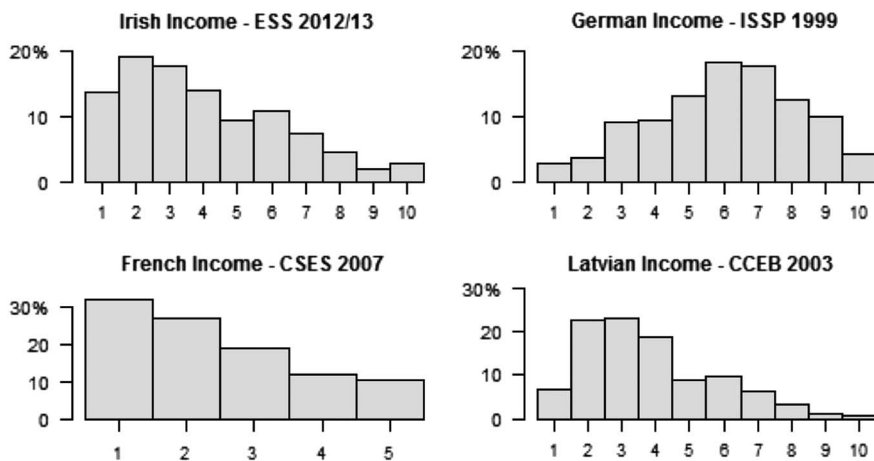
Fig. 3. Examples of income distributions from different cross-national surveys
Note: ESS = European Social Survey; ISSP = International Social Survey Programme; CSES = Comparative Study of Electoral Systems; CCEB = Candidate Countries Eurobarometer.

The choice between the three versions of the household income measure proposed in this article—or for that matter of variations of these solutions (e.g. non-logged versions of absolute income) —needs to be driven primarily by the particular theoretical question to be tested. For example, percentiles are most appropriate for analyzing the effects of an individual's relative position on her country's socioeconomic ladder, while ratios also incorporate the magnitude of the income differences within a given survey, and levels further capture absolute differences across time and space, and are thus more relevant for scholars interested in the broader cross-national differences in household incomes. Clearly, the broad comparative scope of the income-level measure also means that its use requires the acceptance of the broader set of assumptions that we had to make in order to construct this measure.

REFERENCES

Alesina, Alberto, and Paola Giuliano. 2009. 'Preferences for Redistribution'. IZA Discussion Paper No. 4056. Leibniz: Institute for the Study of Labor.
Andersen, Robert, and Tina Fetner. 2008. 'Economic Inequality and Intolerance: Attitudes Toward Homosexuality in 35 Democracies'. *American Journal of Political Science* 52(4):942–958.
Beramendi, Pablo, and Christopher J. Anderson. 2008. *Democracy, Inequality, and Representation: A Comparative Perspective*. New York: Russell Sage Foundation.

Cowell, Frank A. 2011. *Measuring Inequality*, 3rd ed. London: Oxford University Press.

De La O, Anna, and Jonathan Rodden. 2008. 'Does Religion Distract the Poor?: Income and Issue Voting Around the World'. *Comparative Political Studies* 41(4–5):437–76.

Donnelly, Michael, and Grigore Pop-Eleches. 2016. 'Measuring Subjective and Objective Incomes in Cross-National Surveys'. Unpublished working paper, Princeton, NJ.

Donnelly, Michael, and Zoe Lefkofridi. 2014. *Unequal Responsiveness in Europe*. Florence: European University Institute Max Weber Working Paper.

Hout, Michael 2004. *Getting the Most Out of the GSS Income Measures*, GSS Methodological Report No. 101. Berkeley: University of California, Berkeley Survey Research Center.

Inglehart, Ronald, Roberto Foa, Christohper Peterson, and Christian Welzel. 2008. 'Development, Freedom, and Rising Happiness'. *Psychological Science* 3(4):264–85.

King, Gary, Christopher J.L. Murray, Joshua A. Salomon, and Ajay Tandon. 2004. 'Enhancing the Validity and Cross-Cultural Comparability of Measurement in Survey Research'. *American Political Science Review* 98(1):191–207.

Layard, R., Stephen Nickell and Guy Mayraz. 2008. 'The Marginal Utility of Income'. *Journal of Public Economics* 92(8–9):1846–857.

Mishler, William, and Richard Rose. 2001. 'Political Support for Incomplete Democracies: Realist Vs. Idealist Theories and Measures'. *International Political Science Review* 22(4):303–20.

Sala-i-Martin, Xavier, and Maxim Pinkovskiy. 2009. 'Parametric Estimation of the World Distribution of Income'. NBER Working Papers Series, Cambridge, MA, pp. 1–75.

Pop-Eleches, Grigore, and Joshua Tucker. 2012. 'Associated With the Past?: Communist Legacies and Civic Participation in Post-Communist Countries'. *East European Politics & Societies* 27(1):45–68.

Rueda, David, Daniel Stegmueller, and Timo Idema. 2014. 'The Effects of Income Expectations on Redistribution Preferences in Western Europe'. Paper presented at The Annual Meeting of the American Political Science Association (APSA), August 28–31, Washington, DC.

Solt, Frederick. 2008. 'Diversionary Nationalism: Economic Inequality and the Formation of National Pride'. *American Journal of Political Science* 52:48–60.

Solt, Frederick. 2009. 'Standardizing the World Income Inequality Database'. *Social Science Quarterly* 90(2):231–42.

Stegmueller, Daniel. 2011. 'Apples and Oranges? The Problem of Equivalence in Comparative Research'. *Political Analysis* 19(4):471–87.

Weakliem, David L., Robert Andersen, and Andrew Heath. 2005. 'By Popular Demand: The Effect of Public Opinion on Income Inequality'. *Comparative Sociology* 4(3–4):261–284.

World Bank. 2011. *World Development Indicators Database*. Washington, DC: The World Bank.