**Redistribution, Inequality, and Growth**

**Online Appendix**[[1]](#footnote-2)

This file contains data documentation and STATA data and replication do files used in “Redistribution, Inequality, and Growth” IMF Staff Discussion Note 14/02 by J.D. Ostry, A. Berg, and C.G. Tsangarides (2014).

## Data

The data used in the analysis comes from the following sources.

1. Version 7.1 of the Penn World Tables: *Growth rate, initial income (rgdpch), investment (ki), population growth (pop), and openness (openk).*

Link: <https://pwt.sas.upenn.edu/php_site/pwt71/pwt71_form.php>

1. Version 3.1 of the “Standardizing the World Income Inequality Database (SWIID)” constructed by Solt (2009): *Inequality (market and net) and redistribution.*

Link to data and replication codes: <http://myweb.uiowa.edu/fsolt/swiid/swiid.html>

1. IMF World Economic Outlook: *Terms of trade and interest rate.*
2. World Bank World Development Indicators (2013): *Tax revenue (percent of GDP), Subsidies and other transfers (percent of expense).*
3. OECD National Accounts database (2012): *Social security benefits paid by general government (percent of GDP), Current transfers received by households (percent of GDP), Subsidies (percent of GDP), Social expenditure (percent of GDP), Total tax revenue (percent of GDP).*
4. Updated and extended version of the Lane and Milesi-Ferretti (2007) dataset covering 1970-2010: *Debt liabilities to GDP.*
5. Polity IV (2011): *Political institutions (polity2).*

Link: <http://www.systemicpeace.org/polity/polity4.htm>

1. Barro and Lee educational attainment dataset (version (1.2)): *Average years of total schooling.*

Link: <http://www.barrolee.com/>

The SWIID database uses a “custom missing-data algorithm” to standardize the United Nations University’s World Income Inequality Database (WIID) and additional observations. The final dataset covers 173 countries over the period 1960-2010. We construct the following three samples for our analysis (summary statistics below).

1. The “full” sample includes all observations where net and market inequality data are available in Solt (2009), version 3.1.
2. The “baseline” sample starts with the “full’ sample and removes a set of specific observations where Solt concludes that the raw surveys are unreliable. In addition, we require that one of two conditions be satisfied in order to include an observation, each designed to ensure that the redistribution measure is informative: either (a) that the country contain at least one survey of some sort of net concept (e.g. disposable income or expenditure) and one market concept (which includes definitions of either market or gross income), so that there is country-specific information on redistribution itself from the survey data for that country; or (b) that uncertainty associated with estimated redistribution is very small relative to the size of redistribution. We require that the calculated ratio of the difference between market and net inequality to the associated standard error (constructed from the standard errors of net and market inequality provided by Solt) is greater than 1.96.
3. The “restricted” sample strikes out all observations for developing countries prior to 1985 and developed countries prior to 1975. It also requires that there be three surveys, not one as in the “baseline” sample, for each income concept in criteria (a), and it does not use criterion (b) above.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Various samples of redistribution | | | | | |
| Sample | Observations | Mean | St.Deviation | Min | Max |
| Baseline | 3667 | 7.5 | 7.2 | -14.2 | 30.3 |
| Full | 4396 | 6.9 | 6.9 | -14.2 | 30.4 |
| Restricted | 2158 | 7.7 | 7.3 | -11.9 | 30.3 |

## STATA replication files

There are 3 data files and one STATA do file to reproduce the Tables included in the SDN.

|  |  |  |
| --- | --- | --- |
| Table | Data source | Do-file |
| Table 1 | transfers\_corr.dta | SDN\_Tables\_APR2014\_final.do |
| Table 2 | SDN\_5year\_average\_dataset.dta | SDN\_Tables\_APR2014\_final.do |
| Table 3 | SDN\_5year\_average\_dataset.dta | SDN\_Tables\_APR2014\_final.do |
| Table 4 | SDN\_5year\_average\_dataset.dta | SDN\_Tables\_APR2014\_final.do |
| Table 5 | SDN\_spells\_dataset.dta | SDN\_Tables\_APR2014\_final.do |
| Table 6 | SDN\_spells\_dataset.dta | SDN\_Tables\_APR2014\_final.do |

1. This online appendix is a companion to SDN/14/02, originally posted in February 2014, and revised in April 2014. The revised version corrects a few issues with the data and estimations. None of these changes is consequential for the results. [↑](#footnote-ref-2)