

## #7 – The Standardized World Income Inequality Database (SWIID)

### Short description

**Type of object:** Dataset.

**Source (organisation):** Harvard Dataverse. The Harvard Dataverse Repository is a free data repository open to all researchers from any discipline, both inside and outside of the Harvard community, where you can share, archive, cite, access, and explore research data. Each individual Dataverse collection is a customizable collection of datasets (or a virtual repository) for organizing, managing, and showcasing datasets.

**Issues:** This dataset is focused on income inequalities (data on Gini index). These are the four measures analysed by SWIID.

- **gini\_disp:** Estimate of Gini index of inequality in equivalized (square root scale) household disposable (post-tax, post-transfer) income, using Luxembourg Income Study data as the standard.
- **gini\_mkt:** Estimate of Gini index of inequality in equivalized (square root scale) household market (pre-tax, pre-transfer) income, using Luxembourg Income Study data as the standard.
- **abs\_red:** Estimated absolute redistribution, the number of Gini-index points market-income inequality is reduced due to taxes and transfers: the difference between the **gini\_mkt** and **gini\_disp**.
- **rel\_red:** Estimated relative redistribution, the percentage reduction in market-income inequality due to taxes and transfers: the difference between the **gini\_mkt** and **gini\_disp**, divided by **gini\_mkt**, multiplied by 100.

**Time span:** From 1960 until 2023 (it depends on the countries)

**Geographical coverage:** 192 countries

**Link:** <https://dataverse.harvard.edu/dataset.xhtml?persistentId=doi:10.7910/DVN/LM4OWF>

### Note/Interpretations

The goal of the Standardized World Income Inequality Database (SWIID) 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 SWIID's income inequality estimates are based on thousands of reported Gini indices from hundreds of published sources, including the OECD Income Distribution Database, the Socio-Economic Database for Latin America and the Caribbean generated by CEDLAS and the World Bank, Eurostat, the World Bank's PovcalNet, the UN Economic Commission for Latin America and the Caribbean, national statistical offices around the world, and academic studies while minimizing reliance on problematic assumptions by using as much information as possible from proximate years within the same country. The data collected and harmonized by the Luxembourg Income Study is employed as the standard. The SWIID currently incorporates comparable Gini indices of disposable and market income inequality for 198 countries for as many years as possible from 1960 to the present; it also includes information on absolute and relative redistribution. (2022-11-20). The latest version (version 9.4) of this dataset has been published in November 2022.

The data are immediately accessible.

The **swiid9\_4.dta** file is pre-formatted for use with Stata's tools for analysing multiply imputed data. Estimates of each of four inequality measures and their associated uncertainty are represented by a placeholder variable (which has the measure's name but only missing data for all observations) plus 100 separate variables (prefixed with **\_1\_**, **\_2\_**, etc.): for any given observation, the differences across these 100 variables capture the uncertainty in the estimate.

[Go back to the List](#)