ANNEX C

Indexes and estimation techniques

Gini index

Definition: Income inequality among individuals in regions and metropolitan areas are measured by an unweighted Gini index. The index is defined as:

$$GINI = \frac{\sum_{i=1}^{N}(2i - N - 1)y_i}{N\sum_{i=1}^{N}y_i}$$

where N is the number of households, and y_i is the value of variable y (e.g. equivalised household disposable income) in household j when ranked from low (y_1) to high (y_N) among all households within a metropolitan area.

The index ranges between 0 (perfect equality: y is the same in all households) and 1 (perfect inequality: y is zero in all households except one).

Theil entropy index

Definition: Regional disparities are also measured by a Theil entropy index, which is defined as:

$$Theil = \sum_{i=1}^{N} \frac{y_i}{\bar{y}} ln\left(\frac{y_i}{\bar{y}} \right)$$

Where N is the number of regions in the OECD, y_i is the variable of interest in the i-th region (i.e. household income, life expectancy, homicide rate, etc.) and $\bar{y}$ is the mean of the variable of interest across all regions.

The Theil index can be easily decomposed in two components: one is the disparities within subgroups of regions – where for example is subgroup is identified by a set of regions belonging to a country; another one is the disparities between subgroups of regions (i.e. between countries). The sum of these two components is equal to the Theil index.

In order to decompose the Theil index, let's start by assuming *m* groups of regions (countries). The decomposition will assume the following form:

$$Theil = \sum_{j=1}^{M}\sum_{i=1}^{N} s_j \frac{y_{ij}}{\bar{y}_j} ln\left(\frac{y_{ij}}{\bar{y}_j} \right) + \sum_{j=1}^{M} s_j ln\left(\frac{\bar{y}_j}{\bar{y}} \right)$$

Where the first term of the formula is the *within* part of the decomposition it is equal to the weighted average of the Theil inequality indexes of each country. Weights, s_i, are computed as the ratio between the country average of the variable of interest and the OECD

average of the same variable. The second term is the between component of the Theil index and it represents the share of regional disparities that depends on the disparities across countries.

Interpretation: The Theil index ranges between zero and ∞, with zero representing an equal distribution and higher values representing a higher level of inequality.

The index assigns equal weight to each region regardless of its size; therefore differences in the values of the index among countries may be partially due to differences in the average size of regions in each country.

Methodology to estimate GDP at the metropolitan level

The proposed methodology uses GDP per capita values in TL3 regions and Metropolitan Statistical Areas (MSA) as data inputs (MSA are used only for Australia, Canada, and the United States) and the distribution of the population based on the GHS population grids.

Using ArcMap 10.2.2, the suggested methodology is composed of three main steps:

- Convert the TL3 and MSA polygons (that contain the GDP per capita values) to a raster – use the GHS population grid to define the cell size of the new GDP per capita raster;
- Using the previously created GDP per capita raster and the GHS population grid, multiply the GDP per capita of a square km by the corresponding population in the same square km; this would result in a GDP raster (a grid that shows GDP per square kilometre);
- With the use of the polygons of the metropolitan areas, calculate the sum of the square km's GDP values lying within the metropolitan boundaries.

It has to be noted that the estimates of GDP in the metropolitan areas do not adhere to international standards; the comparability among countries relies on the use of the same methodology applied to areas defined with the same criteria.

From:

OECD Regions and Cities at a Glance 2018

Access the complete publication at:
https://doi.org/10.1787/reg_cit_glance-2018-en