ANNEX C

Indexes and estimation techniques

Environmental revealed technological advantage index

Definition: The index of revealed technological advantage is defined as:

$$RTA_i = \frac{P_{ij}/P_j}{P_i/P}$$

Where P_{ij} is the number of patents in the technology field *i* in region *j*, P_j is total number of patents in region *j*, P_i is the national patents in technology field *i*, and P is the total national patents of all fields

Interpretation: The index of revealed technological advantage is defined as the region's share (over national value) of patents in a particular technology field divided by the region's share (over national value) in all patents fields. The index is equal to zero when the region holds no patents in a given technological field; is equal to 1 when the region's share in the technological field equals its share in all fields (no specialisation); and above 1 when a positive specialisation of the region is observed within its country.

Gini index

Definition: Regional disparities are measured by an unweighted Gini index. The index is defined as:

$$GINI = \frac{2}{N-1}\sum_{i=1}^{N-1}\left|F_i - Q_i\right|$$

where N is the number of regions, $F_i = \frac{i}{N}$, $Q_i = \frac{\sum_{j=1}^{i} y_j}{\sum_{i=1}^{n} y_i}$ and y_i is the value of variable y (e.g. GDP per capita, unemployment rate, etc.) in region *j* when ranked from low (y_i) to high (y_N) among all regions within a country.

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

Interpretation: 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. Only countries with more than four regions are included in the computation of the Gini index

Specialisation index

Definition: Specialisation is measured according to the Balassa-Hoover index, which measures the ratio between the weight of an industry in a region and the weight of the same industry in the country:

$$BH_i = \frac{Y_{ij}/Y_j}{Y_i/Y}$$

where Y_{ij} is total employment of industry *i* in region *j*, Y_j is total employment in region *j* of all industries, Y_i is the national employment in industry *i*, and Y is the total national employment of all industries. A value of the index above 1 shows specialisation in an industry and a value below 1 shows lack of specialisation.

Interpretation: The value of the specialisation index decreases with the level of aggregation of industries. Therefore, the specialisation index based on a 1-digit industry (e.g. manufacturing) would underestimate the degree of specialisation in all 2-digit industries belonging to it (e.g. textile, chemistry, etc.).

Urban sprawl index

Definition: The index measures the evolution of sprawl over time in a metropolitan area:

$$SI_i = \frac{\left[urb_{i,t+n} - \left(urb_{i,t} * \left(\frac{pop_{i,t+n}}{pop_{i,t}}\right)\right)\right]}{urb_{i,t}} *100$$

where *i* refers to a particular metropolitan area; *t* refers to the initial year; *t+n* refers to the final year; *urb* refers to the built-up area in square kilometres; *pop* refers to the total population of the metropolitan area. The built-up area (or urbanised land) is computed as the land within the boundaries of the metropolitan area covered by private and commercial buildings, infrastructure and major transportation infrastructure.

Interpretation: The urban sprawl index measures the growth in built-up area adjusted for the growth in population. When the population is stable, the urban sprawl index is basically the growth of built-up area. When the population changes, the index measures the increase in the built-up area relative to a benchmark where the built-up area would have increased in line with population growth. The urban sprawl index is equal to zero when both population and built-up area are stable over time. It is bigger (lower) than zero when the growth of built-up area is greater (smaller) than the growth of population, i.e. the density of the metropolitan area has decreased (increased). Similarly, the index could be computed to compare the sprawl for a given year over a set of metropolitan areas.

Computation of typologies of land cover and changes in land cover

To measure the different uses of land and its changes with respect to small portions of territory, data from the earth's surface collected through remote sensing and geographic information systems are used. Despite recent progress in earth observation, remote sensing and techniques for processing large datasets, there is not a unique global dataset recording land cover change. The sources of data are the following: MODIS (Moderate Resolution Imaging Spectroradiometer) Land Cover Data to measure land cover in one year (2008) for all countries. Corine Land Cover for Europe (developed by the European Environmental Agency and the European Space Agency), the Japan National Land Information, and the National Land Cover Database for the United States are used to capture land cover in different years and therefore measure changes in land uses. For Canada, Chile, Korea and Mexico, it was not possible to measure changes in land uses.

These land cover datasets, however, differ in many aspects such as the spatial resolution (though they all get down to 0.5 km cell size) classification systems, and the definitions of land cover classes; therefore, it was necessary to reclassify the typologies of land cover in order to produce the same classes regardless of the source dataset. The final classification used to calculate the statistics for regions and metropolitan areas consists of six classes:

1. Water (lakes, river, lagoons, etc.).
2. Agriculture (annual crops, rice fields, orchards, pastures etc.).
3. Forest (coniferous, broad-leaved, mixed, etc.).
4. Non-forest natural vegetation (natural grasslands, shrub lands, sparsely vegetated areas, etc.).
5. Urbanised area (residential and industrial buildings, major transportation, land for urban uses, etc.).
6. Other (bare lands, wetlands, glaciers).

For regions in other countries than the EU, United States and Japan, the MODIS Land Cover product was used to estimate the proportion of urban (class 13 in IGBP classification) and forest land (classes 1-5 in IGBP classification). The MODIS Land Cover is released each year, and 2008 data were used for estimation. The urban class refers to 2001-02 MODIS data since updated estimates of urban land are still not available for later years. For Europe, Japan and the United States, it was possible to compute also the change in urbanised land, agricultural land and forested land. Changes are expressed as net rates: for example, the rate of change of urban area is calculated as the amount of land converted to urban land cover minus the urban land converted to other classes, as a fraction of the urban land in the starting year.

Once the six classes of land cover are defined, a raster was produced with each cell being classified according to one of the six classes; by superimposing the layer of regional boundaries, we can compute the percentage of regional area covered by forest or the percentage of urbanised area in a metropolitan area, etc.

Methodology to adjust GDP, total employed and unemployed at metropolitan level

The proposed methodology uses as data inputs the values of GDP in TL2 or TL3 regions and the distribution of population on a small grid (1 km^2 cell). It is composed of four main steps, each of which is carried out using GIS software:

- Take the GDP at TL3 level and intersecting with the population grid obtained by the dataset LandScan 2000.
- Attribute each 1 km^2 cell a GDP value by weighing for the population in each cell.
- Intersect the layer of GDP in each cell with the boundaries of metropolitan areas. Cells that are not entirely included in one metropolitan area can be aggregated proportionally to the share of their area that falls within each metropolitan area (proportional calculation criteria) or, alternatively, by using a maximum area criterion.
- Calculate the sum of cells' GDP values belonging to each metropolitan area.

An improved method would be to use employment data rather than population in step 2. For example, the United Kingdom Office for National Statistics provides income estimates at ward level downscaling the regional values through various variables including household size, employment status, proportion of the ward population claiming social benefits, and proportion of tax payers in each of the tax bands, etc. A similar method is used by the U.S. Bureau of Economic Analysis to estimate the GDP for U.S. Metropolitan Statistical Areas. The Federal Statistical Office of Switzerland used CLC-Data-Classes urban continuous fabric, urban discontinuous fabric and industrial or commercial units for all neighbouring countries by calibrating with other data to estimate data for jobs in grid cells. However these types of data input are not available in most OECD countries; therefore a simpler solution was adopted.

A similar technique is applied to estimate employment and unemployment in metropolitan areas. Due to the lack of labour market data in TL3 regions, employment and unemployment in metropolitan areas are derived from TL2 regions. As such, caution should be taken in comparing these values at metropolitan level.

It has to be noted that the estimates of GDP, employment and unemployment 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.

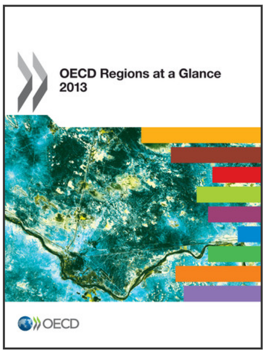
Methodology to disaggregate CO_2 emissions at regional level and metropolitan areas

Generally, emission data are available at country level from the Intergovernmental Panel on Climate Change (IPCC). To facilitate estimation of the emission levels for geographic areas like OECD regions or metropolitan areas, the EDGAR global emission database, developed by the Joint Research Centre of the European Commission, was used. The EDGAR database version 4.1 provides country emission levels separately by each compound and sector of origin (e.g. CO_2 emission from fuel production) allocated (disaggregated) to gridded maps on the basis of spatial data such as location of energy and manufacturing facilities, road networks, shipping routes, human and animal population density and agricultural land use. The spatial resolution of the grid is 0.1 by 0.1 degrees and the gridded estimates are currently available for the years 2000-08.

The methodology employed essentially sums the EDGAR estimated values for the 0.1 by 0.1 degrees grids over the relevant boundaries of the regions or metropolitan areas. The raster of total CO_2 emissions were averaged over a three-year period to smooth out potential extreme values that might occur in the yearly data.

The emissions from the energy sector include public electricity, heat production and other energy industries; while the emissions from transport include road, rail and ground transportation.

While these estimates have the advantage of using a common methodology for all metropolitan areas, they are based on sectoral energy use and GHG at the national level and population and/or sectoral shares at the local level. As a result, they cannot capture changes in energy use or GHG emissions due to local policies. They also cannot capture the eco-efficiency of cities, as the estimation assumes that all sectors use energy or produce GHG at the same rate in the entire country. The absence of a global protocol for quantifying GHG emissions attributable to urban areas limits the international comparability and it should be taken into account when using these estimates.



From:

OECD Regions at a Glance 2013

Access the complete publication at:

https://doi.org/10.1787/reg_glance-2013-en

Please cite this chapter as:

OECD (2013), “Indexes and estimation techniques”, in *OECD Regions at a Glance 2013*, OECD Publishing, Paris.

DOI: https://doi.org/10.1787/reg_glance-2013-49-en