London, 20 May 2026

ESCoE Conference on Economic Measurement 2026

Sub-National Spatial Price Indices in Italy

Experimental Results from 2024 Data

ALESSANDRO BRUNETTI, ANNA DI FRANCO, BARBARA DRAMIS, AGOSTINA ZANOLI

Istat | DCSW

Outline

- Sub-National Spatial Price Indices: the Italian project
- Data sources: scanner data, CPI data, ad hoc surveys, web-scraping
- Methodology
- Some results for five Ecoicop divisions
- Concluding remarks and perspectives

Sub-National Spatial Price Indices: the Italian project

In Italy there is a high socio-economic heterogeneity among different geographical areas.

Sub-national Spatial Price Indices or Regional Purchasing Power Parities are essential to compare economic well-being aggregates and cost of living at local level.

Previous experience in Italian Sub-National PPPs:

- Italian National Institute of Statistics (ISTAT) carried out official experimental Sub-national PPPs computations by using price data from Consumer Price Index (CPI) and ad hoc surveys and focusing on comparing consumer prices across the 20 Italian regions (in 2008 and 2010).
- Some studies analysed the use of scanner data in Sub-national PPPs compiling (Laureti T., Ferrante C., Dramis B., 2017; Ferrante C., Laureti T., Polidoro F., 2019).

In 2020, ISTAT planned to implement the regular production of Sub-national SPIs, for the 12 Ecoicop Divisions, by including a specific project in the Italian National Statistical Plan.

Sub-National Spatial Price Indices: Data requirement

Comparability:

Pricing products that are identical, or at least equivalent, to ensure that differences in prices will not be influenced by differences in quality

Representativeness:

The product selected must represent the products purchased in each area within the country. Representative products are those products frequently purchased by resident households and are likely to be widely available within each region

Data sources

A multi-source approach was used for the Sub-National Spatial Price Indices (SSPIs) estimation of the five divisions of the Ecoicop classification, presented in this paper:

- **Scanner data**
- **CPI data** for fruit, vegetables and fish, for some catering services and for accommodation services for which CPI definition is accurate enough to guarantee comparability
- **Ad hoc surveys** (Fresh food products, clothing and footwear, catering services etc.)
- **Web-scraping procedure** (Food delivery)

The tobacco products, whose prices have no variability across the Italian territory (but have different weights in the regions), have been included in the calculation with parity equal to 100.

Data source: scanner data for grocery products

Scanner data provided by Nielsen (based on an agreement between Istat and ADM, the Association of large-scale retail trade distribution) for processed food, including non-alcoholic and alcoholic beverages and household maintenance products, year 2024.

They come from a sample of outlets of twenty-one retail trade chains covering 60% of the total turnover for the hard-discount channel and over 90% for each other distribution shop type at the national level.

The outlets sample, widespread in all the municipalities within the 107 Italian provinces, is the same sample extracted for the inflation estimates.

The detailed information contained in these data (turnover and quantities sold for each GTIN in each outlet) can be used to compile weighted SSPIs at a highly detailed territorial level and allow for the fulfillment of both requirements for spatial comparison of prices: comparability and representativeness.

GTIN included in the analysis:
166.995 for the 1st Ecoicop division (01 Food and Non- Alcoholic Beverage)
28.569 for the 2nd Ecoicop division (02 Alcoholic Beverages and Tobacco)
31.301 for the 5th Ecoicop division (05 Furnishing, household equipment and routine household maintenance)

Scanner data: advantages and disadvantages

- ✓ Data for all the 107 Italian provinces.
- ✓ Both requirements for spatial prices comparison, comparability and representativeness, are fulfilled.

PROS

CONS

- ✗ Careful data classification work between the barcode description and the Ecoicop classification is necessary.

Data source: CPI data

CPI data (2024) for products belonging to classes:

01.1.3 Fish and other seafood

01.1.6 Fruits

01.1.7 Vegetables

05.6.2 Household services

11.1.1 Restaurants, cafés and the like

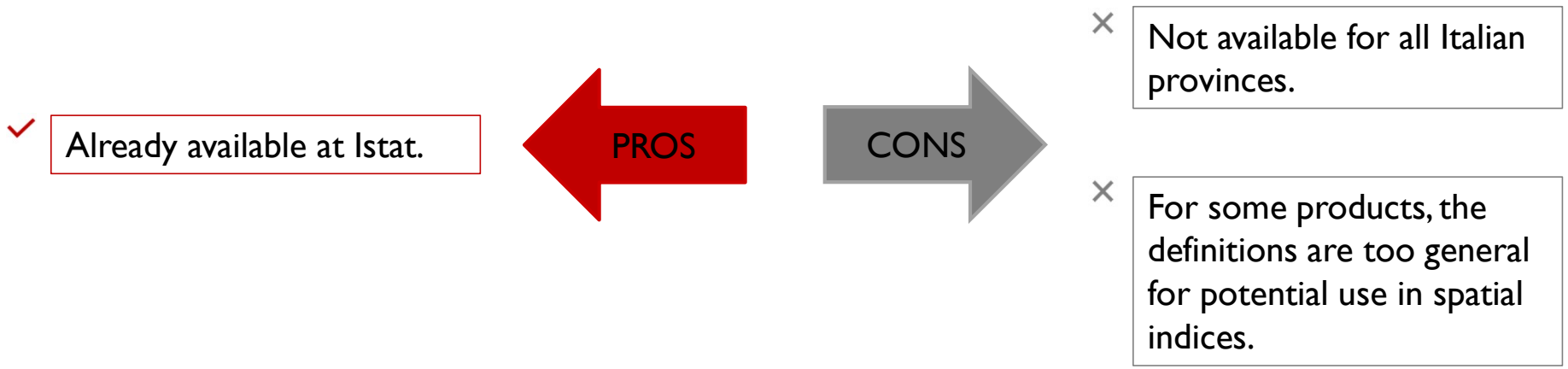
11.2.0 Accommodation services

Collected in the survey of consumer price indices are accurate enough for SSPIs

Data characteristics:

- 80 provinces
- Prices not estimated
- Outlier detection: percentiles 1° and 99° or 5° and 95° according to the product
- The total number of references (price per product within a given store) after eliminating outliers was 90.766

CPI data: advantages and disadvantages



Data source: ad hoc surveys

Ad hoc surveys have been implemented to cover products prices not available from other sources (Fresh food products, clothing and footwear, catering services, etc.).

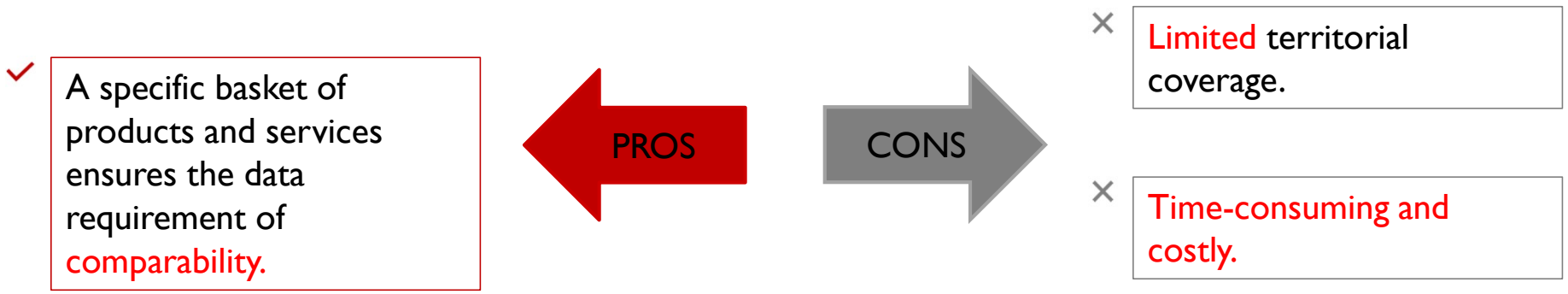
Survey's framework:

- Six surveys: two surveys a year in April/June and October/December, each survey repeated every 3 years
- RPPPs survey basket of products highly comparable between the municipalities defined for each survey
- 21 capital cities involved
- CPI outlets sample and price collectors
- Temporal Adjustment Factors (TAFs) were calculated based on HICP (Harmonized Index of Consumer Price) data to adjust data collected in a specific month to annual prices

Number of price quotations used:

- Division 01 Food products: 9.167
- Division 03 Clothing and footwear: 14.299
- Division 05 Furnishing, household equipment and routine household maintenance: 6.553
- Division 11 Catering services: 9.033

Ad hoc surveys: advantages and disadvantages



Data source: web scraping

A web-scraping procedure has been implemented to collect price data for Food delivery:

- Each month, we collect data from an online food delivery platform. We simulate the purchase of four Margherita pizzas and four soft drinks
- Data collection for 24 cities
- Outlier detection: percentiles 1° and 99°
- Observations after outlier detection: 16.812

✓ The number of quotations collected through the web-scraping procedure is significantly higher than that of ad hoc surveys.

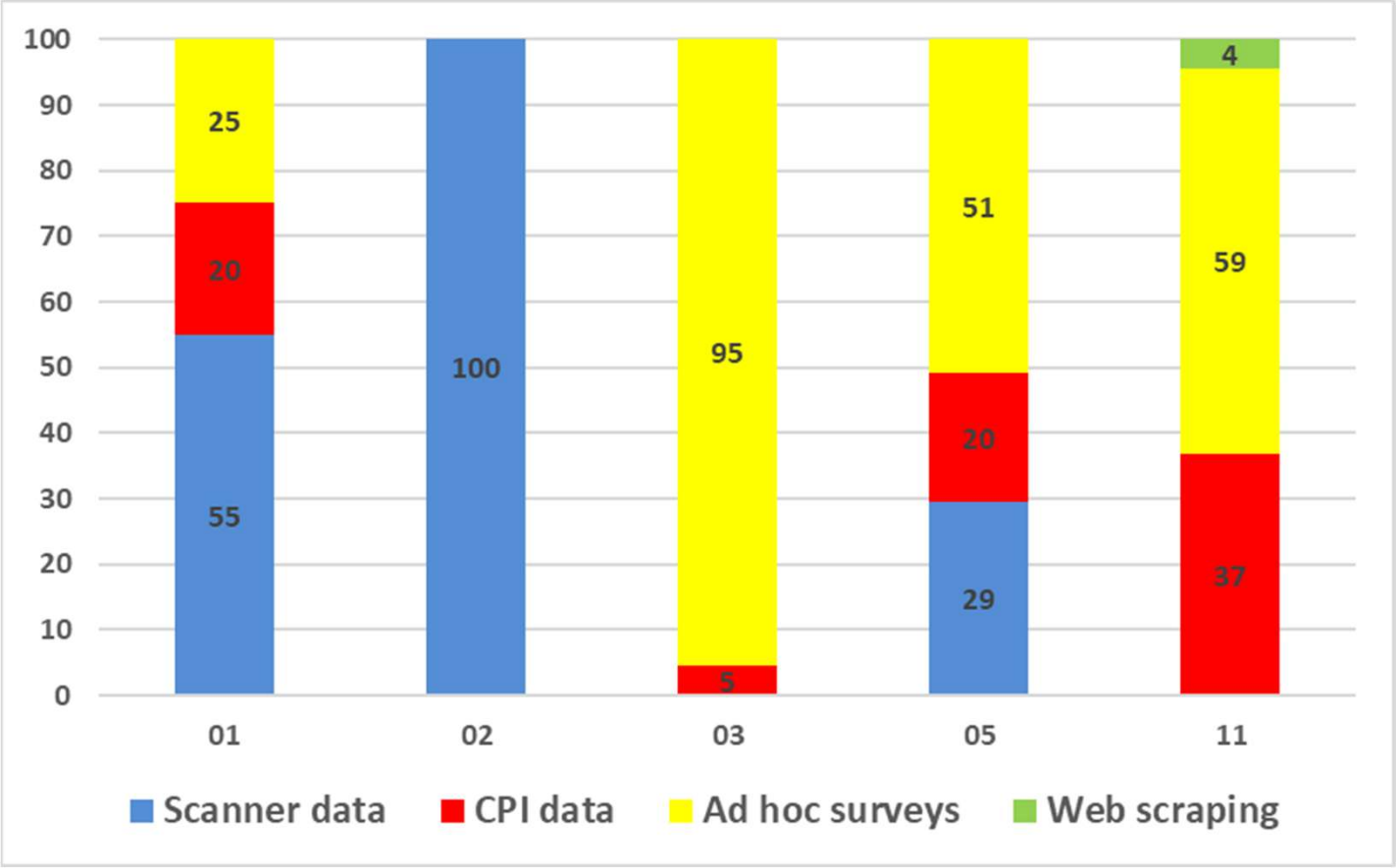
PROS

CONS

× Continuous maintenance of the procedure is required to address changes on the website

× It is crucial that the website owner grants consent for the scraping process

Data sources: relative relevance by division



Methodology

According to International Comparison Program guidelines the methods used are:

- **RPD (Regional product dummy)** model at Basic Heading (BH) level (lower level for which we compute RPPPs)
  - Weighted RPD for scanner data
  - Unweighted RPD for CPI data, ad hoc surveys data and web-scraped data

- **GEKS (Gini – Èltetö – Köves - Szulc)** for aggregations above BH level (using household expenditures as weights)

Methodology: Regional product dummy

N items belonging to a particular Basic Heading (n=1,2,..N)
R regions (r=1, 2..R)
p_{nr} price of the n-th item in region r

The model underlying the **RPD approach** is :

$$\boldsymbol{p_{nr} = P_n * SSPI_r * u_{nr}}$$

where P_n is the average price of the n-th item, SSPI_r is the specific regional factor (parity or general price level of the area considered with respect to other areas) for the basic heading for the region r, u_{nr} is a random error. Considering the logarithms:

$$\boldsymbol{ln\, p_{nr} = \sum_{r=1}^{R} a_r D_r + \sum_{n=1}^{N} b_n D_n^* + v_{nr}}$$

where D_n^* are dummies for the items, D_r are dummies for the areas (regions) $a_r \text{and } b_n$ are, the differences in the effects associated with the regions and the item type, v_{nr} are random errors N(0, σ^2).

Having expenditure or quantity weights for each item, the weighted RPD model can be written as:

$$\boldsymbol{\sqrt{w_{nr}}\, ln\, p_{nr} = \sum_{r=1}^{R} a_r \sqrt{w_{nr}} D_r + \sum_{n=1}^{N} b_n \sqrt{w_{nr}} D_n^* + \sqrt{w_{nr}} v_{nr}}$$

Methodology: Regional product dummy

Parameters of this model can be estimated using ordinary least squares, imposing a restriction, that a coefficient corresponding to a specific area is set equal to zero (a_1=0) or equivalently $SSPI_1$=1 thus considering it as reference area to which the coefficient estimates are referred.

The SSPIs between an area r and the reference area is given by $\widehat{SSPI}_{\mathrm{r}}=\exp(\widehat{a_r})$.

Methodology: GEKS for aggregation above BH

The **GEKS (Gini- Èltetö-Köves-Szulc) method** is used for SSPIs aggregation above Basic Heading. Weights calculated on household expenditures.

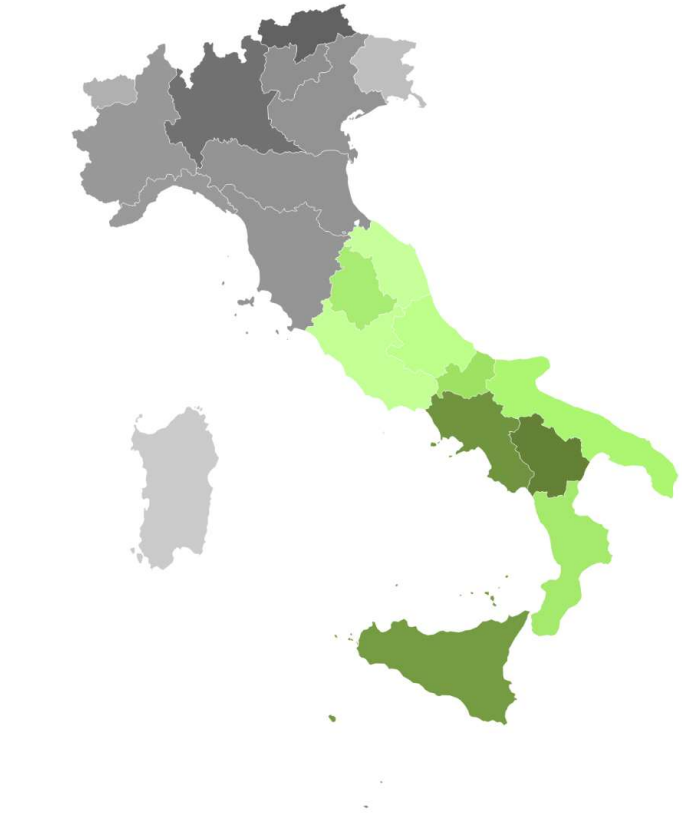
$$SSPI_{jk}^{GEKS} = \prod_{l=1}^{R} (F_{jl} \cdot F_{lk})^{1/R}$$

The SSPI for region k with reference to region j, chosen as the basis, is given by the geometric mean of the Fisher type indices of all direct comparisons between region j and region k, and indirect across all possible links between the R regions (l, k, j ∈R).

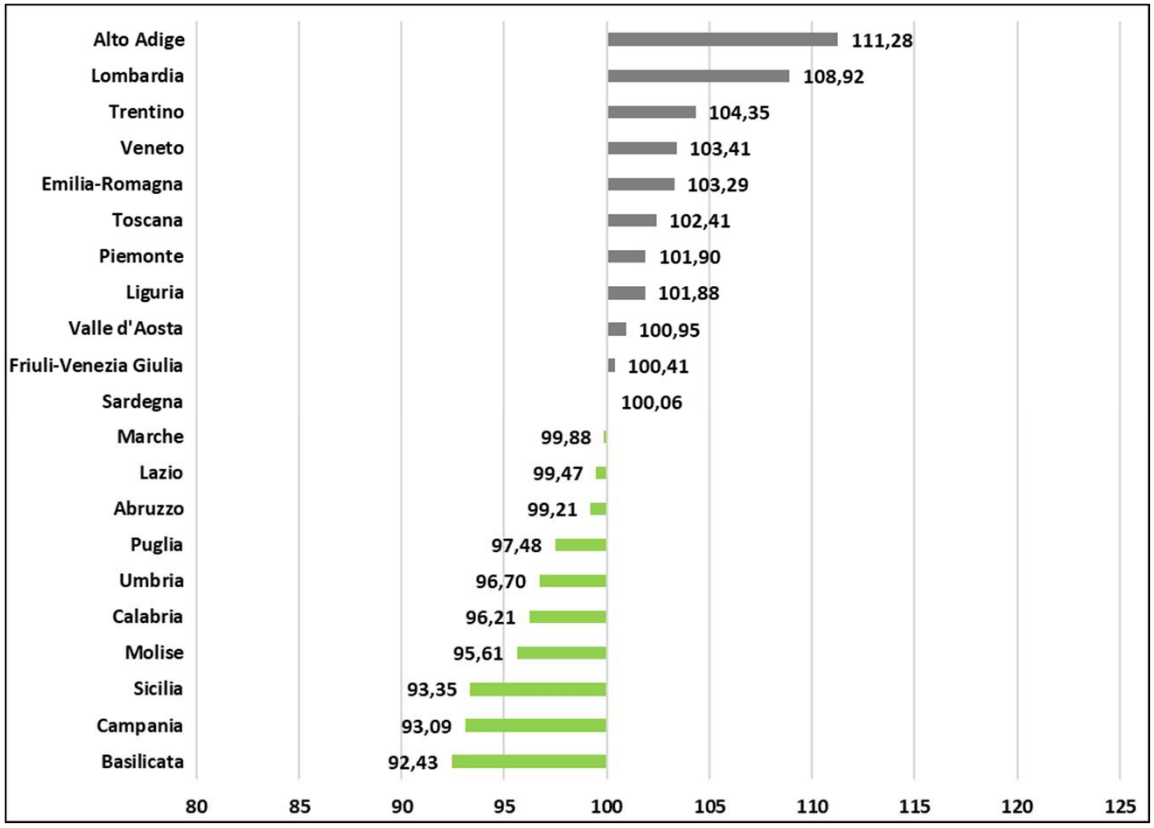
The Laspeyres type and Paasche type indices used for the Fisher type indices are calculated on the basis of Basic Heading SPIs weighted by base region and partner region expenditures, respectively.

To express the parities referring to the national average, each SSPI is divided by the geometric mean of the price level indices of the participating regions, and conventionally, multiplied by 100.

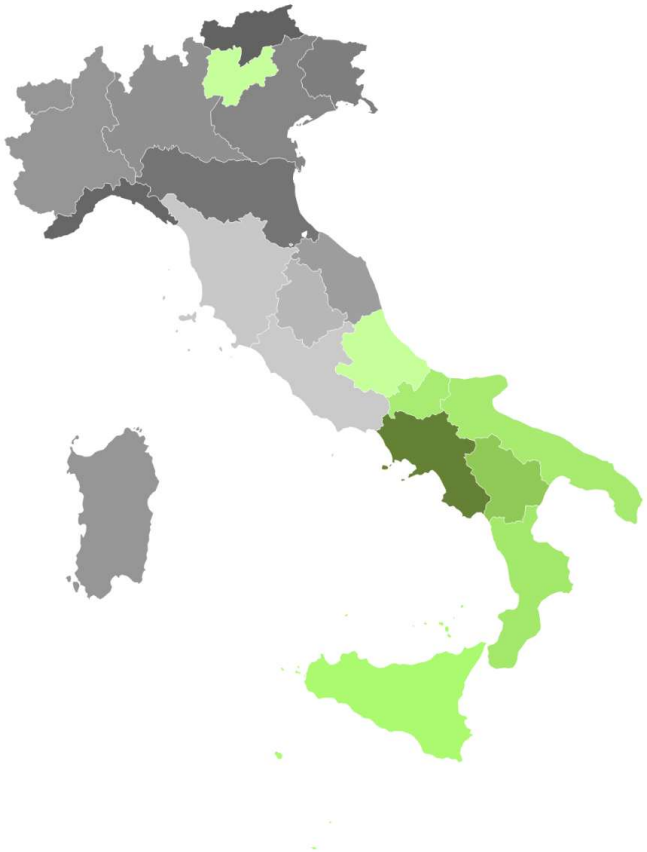
Results: Five aggregated divisions (2024) - Italy=100

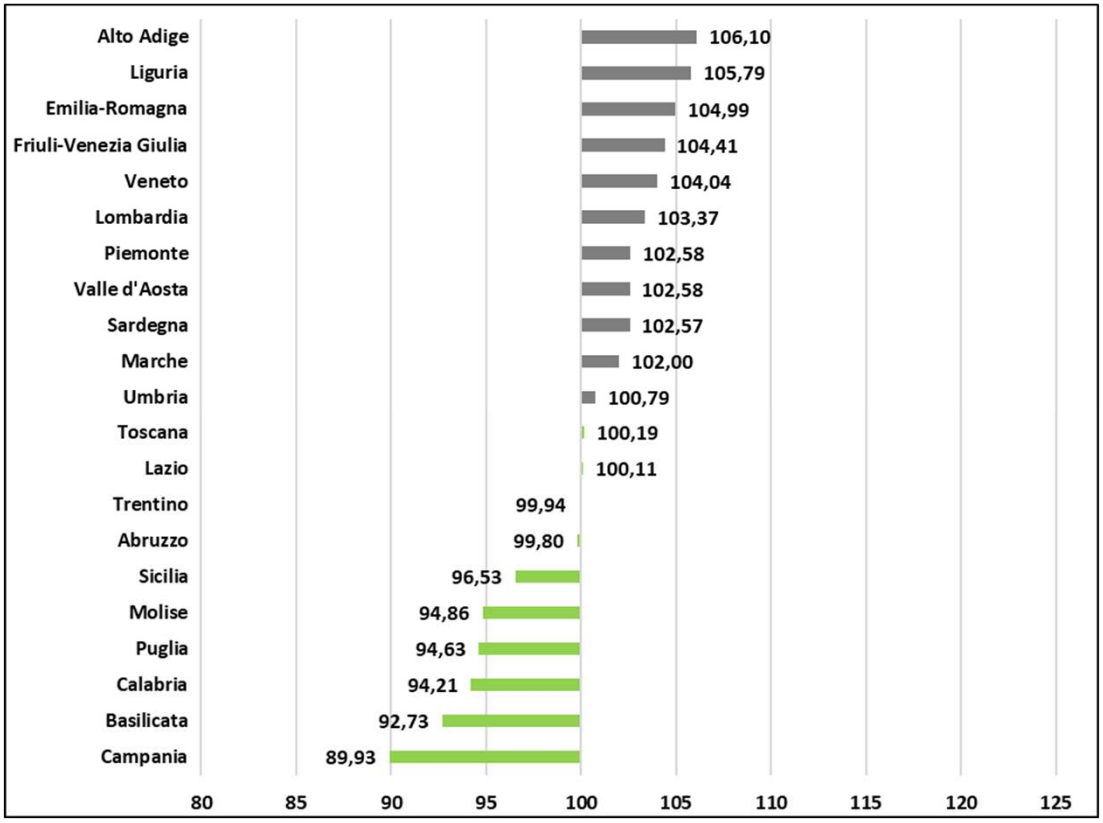


Created with Datawrapper



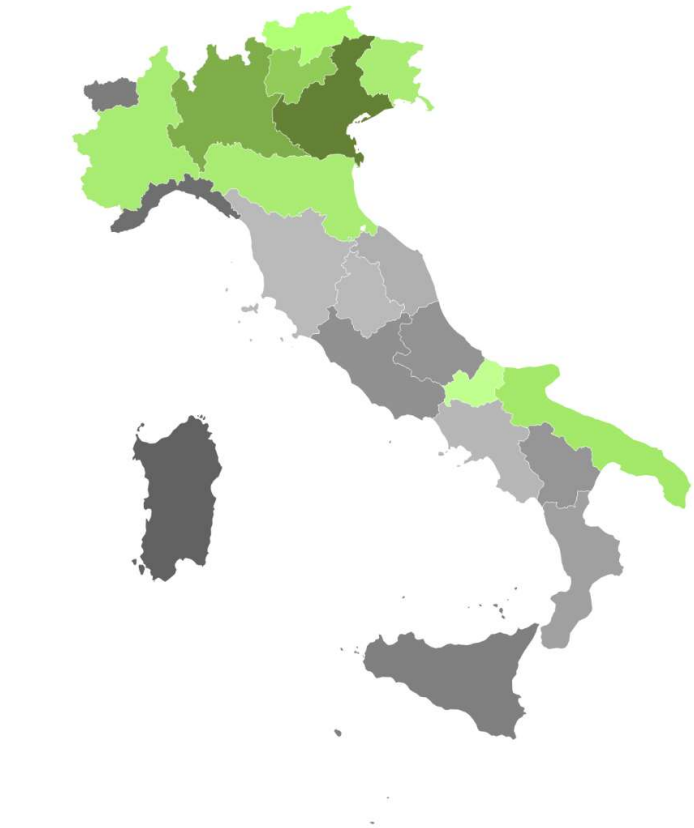
Results: Food and non-alcoholic beverages (2024) - Italy=100

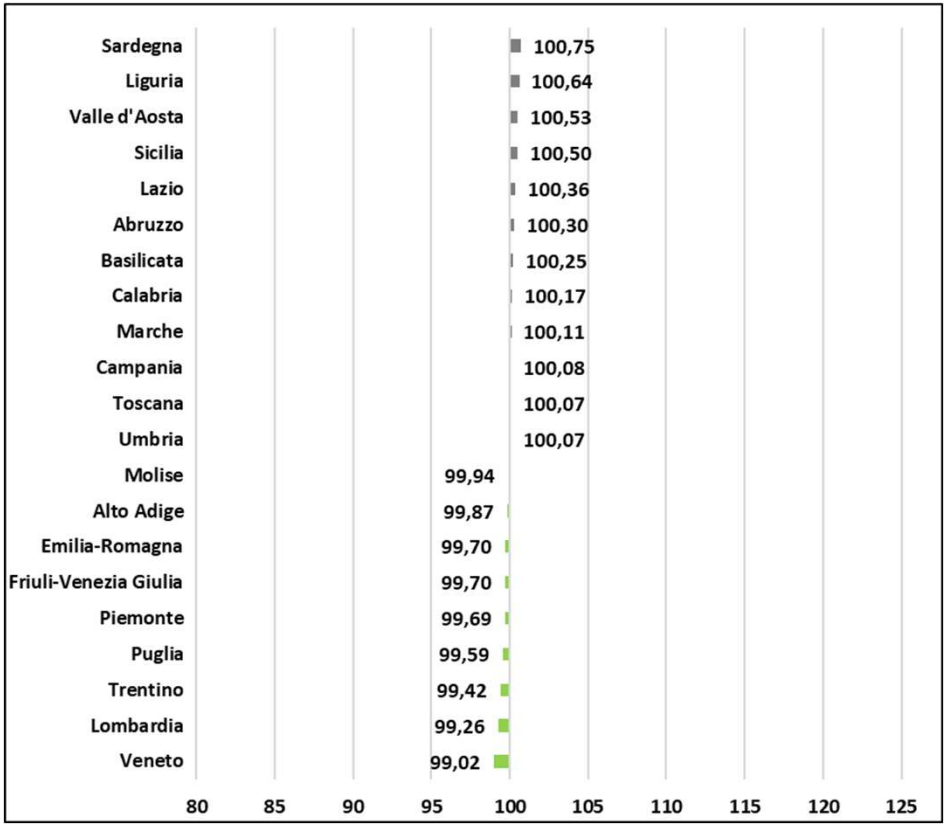




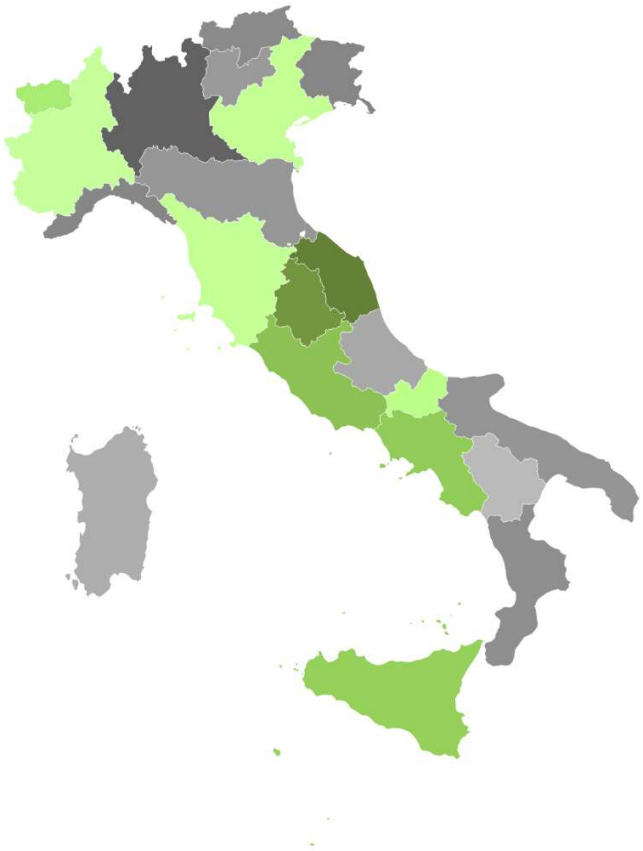
Created with Datawrapper

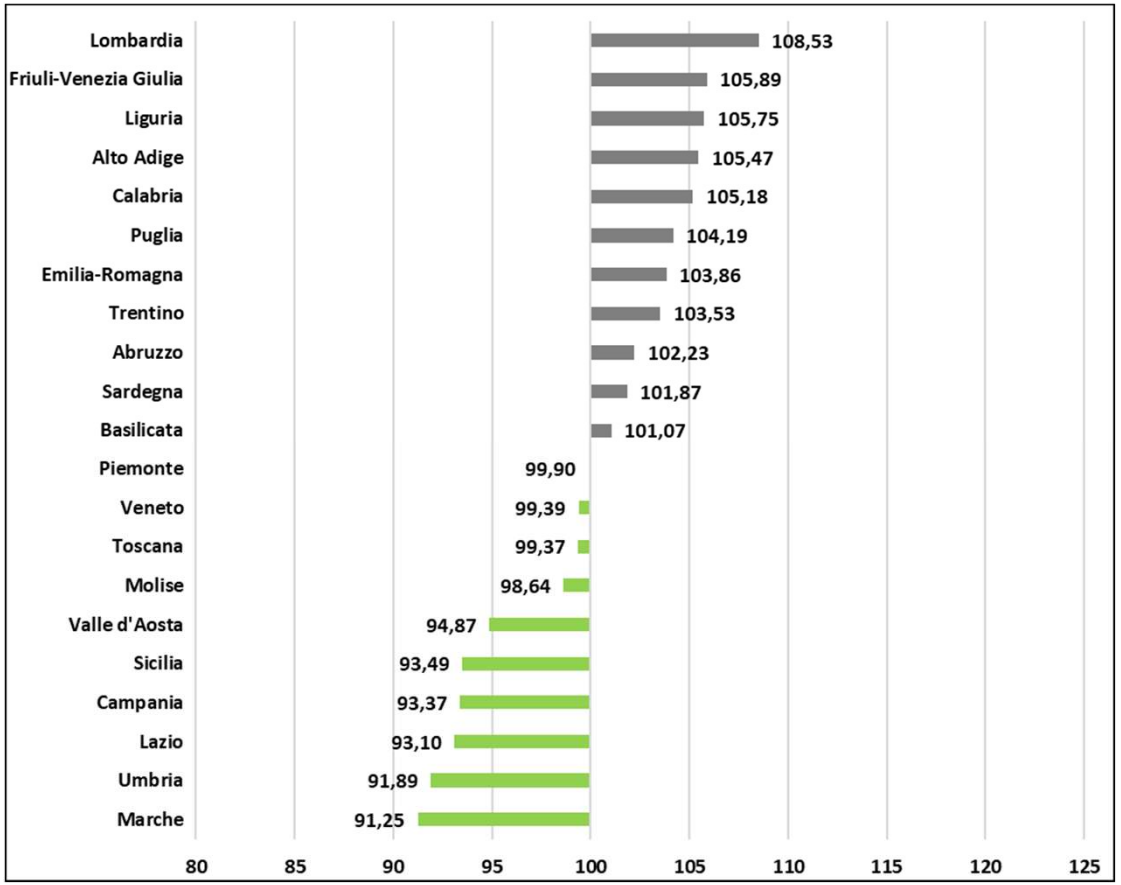
Results: Alcoholic beverages and tobacco (2024) - Italy=100





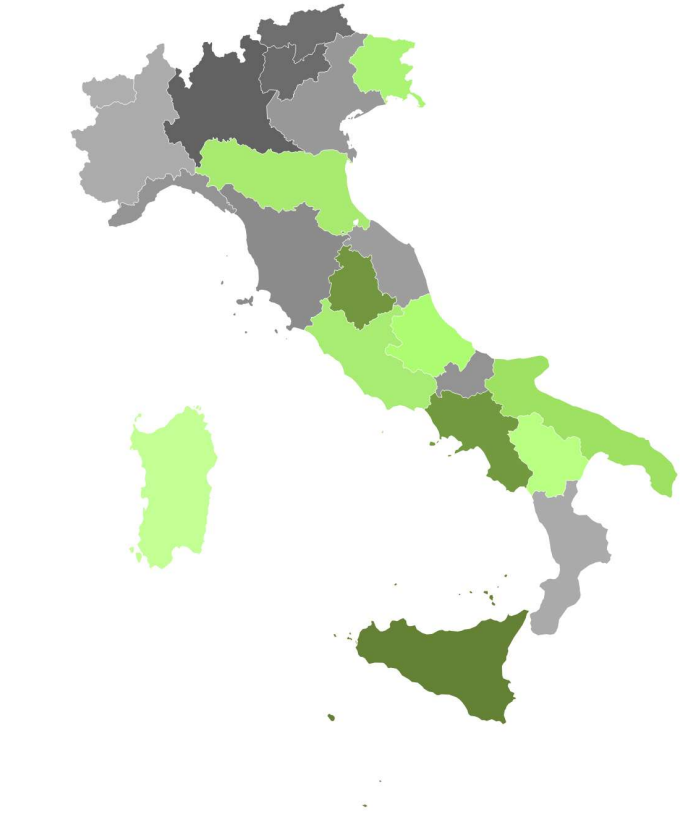
Results: Clothing and footwear (2024) - Italy=100

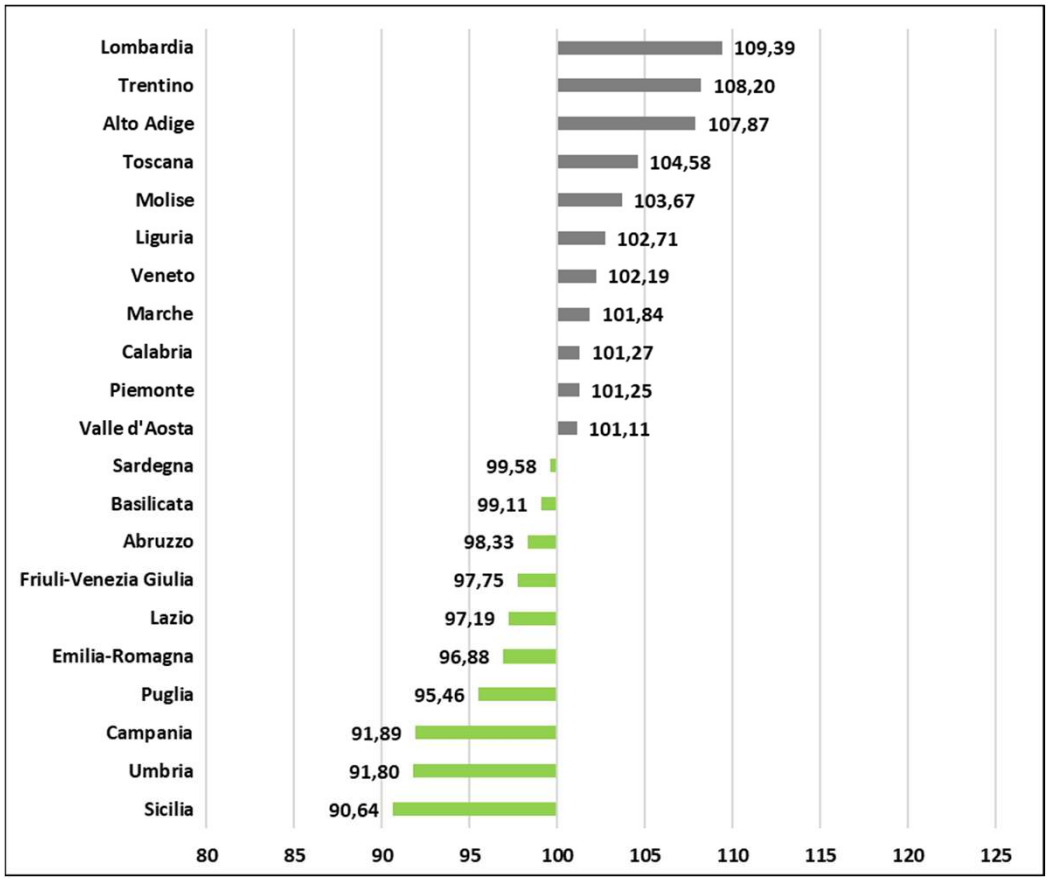




Basilicata data were imputed for failure to detect

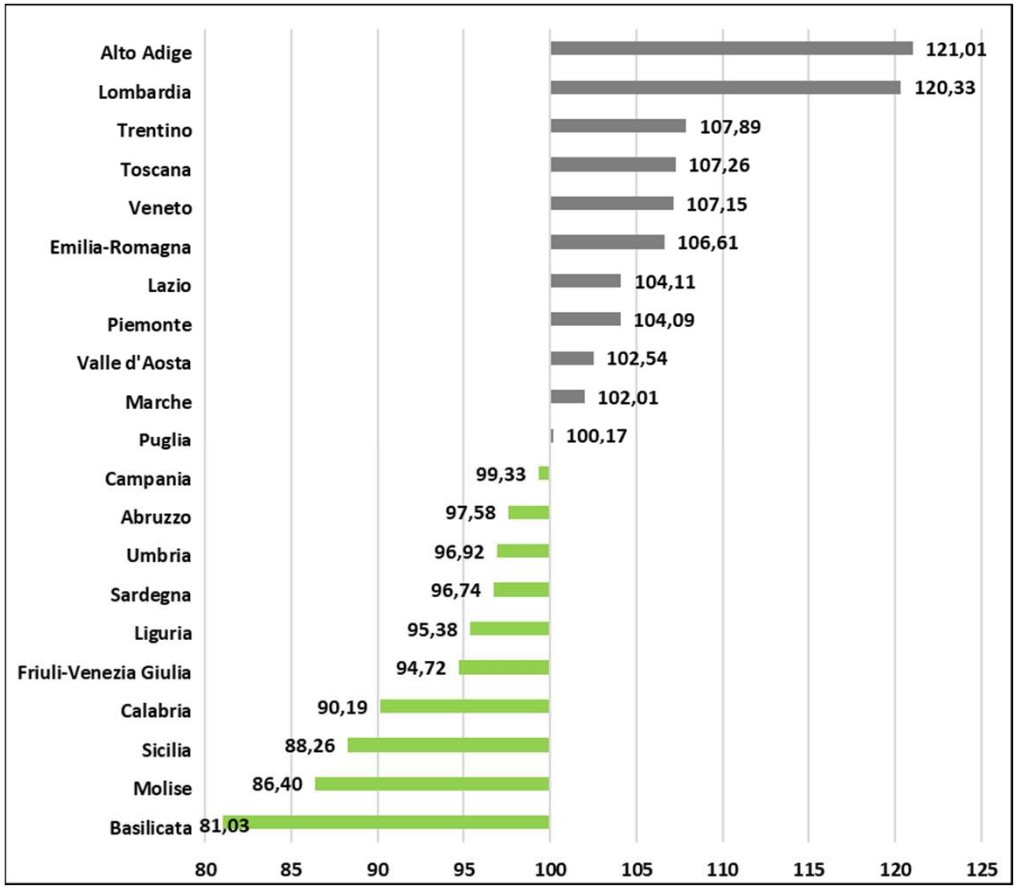
Results: Furnishing, household equipment and routine household maintenance division (2024) - Italy=100





Basilicata data were imputed due to the low number of prices collected for groups: Furniture and furnishing, carpets and other floor covering, Household textiles, Glassware, tableware and household utensils, Tools and equipment for house and garden

Results: Restaurants and hotels (2024) - Italy=100



Concluding remarks and perspectives

The main challenge of the SSPIs project is how to deal with different data sources and their aggregation.

Each source needs a specific analysis, data cleaning and process implementation in order to create a database drawing on various data sources that will be used in SSPIs calculation.

We are working on the 4th division concerning 'Housing, Water, Gas, Electricity and Other Fuel', which is important in terms of consumer expenditure (11,6% of the HICP basket). It is characterized by the presence of tariffs and rents. Rent data will come from the data base of rents provided to Istat by the Italian Tax Office.

Thank you

AGOSTINA ZANOLI | zanoli@istat.it

1926 | 2026

Istat | Istituto Nazionale di Statistica