Meeting of the International Comparison Program (ICP)

Technical Advisory Group (TAG)

May 17-19, 2021

ICP Research Agenda (item07): SubNational PPPs

Estimations of the Subnational Spatial Consumer Price Indices (SN-SCPIs) by using scanner data

(main results of the experiments conducted)

ASESD-"Camilo Dagum" Centre: C. Giusti, M. Pratesi, L. Biggeri, T. Laureti, I.Benedetti;

ISTAT-Unit of Price Statistics: F. Polidoro, F. Di Leo, M. Fedeli

Needs to conduct more experiments

During the meeting held in October 2019, of the task Force on Sub-national PPPs, I proposed to implement some experiments on:

1. The use of scanner data to compute SN-SCPIs

- Various experiments have already been conducted but issues:
  - ✓ essentially based on partial data bases
  - ✓ pay much attention as possible to the products and services for which there are no overlapping items in all the different areas
- **Need for new experiments with more complete data set**

2. The computation of Subnational poor specific SCPIs

3. The collection of information on the main outlets where poor do the purchase (through HES)

➢ **ASEDSD and Price's Unit of Istat, conducted the experiments**

1. The use of scanner data to compute SN-SCPIs

1.1. The characteristics of the available scanner data (Strengths; Weaknesses)

Istat: **complete scanner data base** of the modern distribution (16 chains distribution) referred to the year 2018 purchased by the market research company ACNielsen

- ✓ Very **detailed information on each transaction**: sales; expenditure; quantities and the characteristics of products sold (brand, size, outlet , etc. at the barcode level GTIN (Global Trade Item Number, for thousands of products)
- ✓ **Unit Prices** are calculated taking into account turnover and quantities

Weekly prices are obtained as weekly turnover/weekly quantities

Monthly and annual prices are calculated with arithmetic mean of weekly prices weighted with quantities

- ✓ **More than two million prices for over 60,000 products**
- ✓ Deep data cleaning is needed and there are several outliers

1.1. The characteristics of the available scanner data (cont.)

✓ **Coverage: supermarkets and hypermarkets**, especially for food, beverages and personal and home care products

- **95% of modern** retail chain distribution; **55.4% of total retail trade distribution** for no perishables and seasonal products; **10,5% in terms of the total expenditures** of families for the consumption (share not uniform across Italian territory)
- It is obvious that to cover all products and services purchased by consumers, it is **necessary to use data coming from other sources** as traditional surveys on prices, other big data, etc.)

✓ No coverage outlets in rural areas; small shop, local and open market excluded

1.2. Issues coming to assure the complete ("like to like") comparability

➡ Having so much detailed data available, we started to estimate SN-SCPIs by applying the principle of complete comparability, that is of "like to like" products

➡ But we ran into various problems, due to the non-overlap of the products across different provinces; may be affected also by the sale policy of the retail companies

✓ Only **103 over 107 provinces**

✓ **Data for 72 BHs; 5 out of 12 ECOICOP Divisions covered**

✓ It has been necessary to **exclude 7 groups of products.**

1.2. Issues coming to assure the complete ("like to like") comparability

- **Difficult to find that equal products are consumed in all the sub-areas**, because the consumers of the sub-areas may have very different habits and behaviors in terms of the choice of the outlets and brand of the products, also because the sale policy of the retail companies can be different in the different sub-national areas

- ✓ To solve the problem could be useful to somewhat loosing the tight principle of comparability by **considering detailed homogeneous group of products and to take into account all the products purchased by the consumers in each group in the different areas**

- ✓ What it is important is that any product that belongs to the group (also if it is of different brands etc.) be quite similar and **satisfy the same consumer's need**

1.3 The experiments are conducted by using two approaches

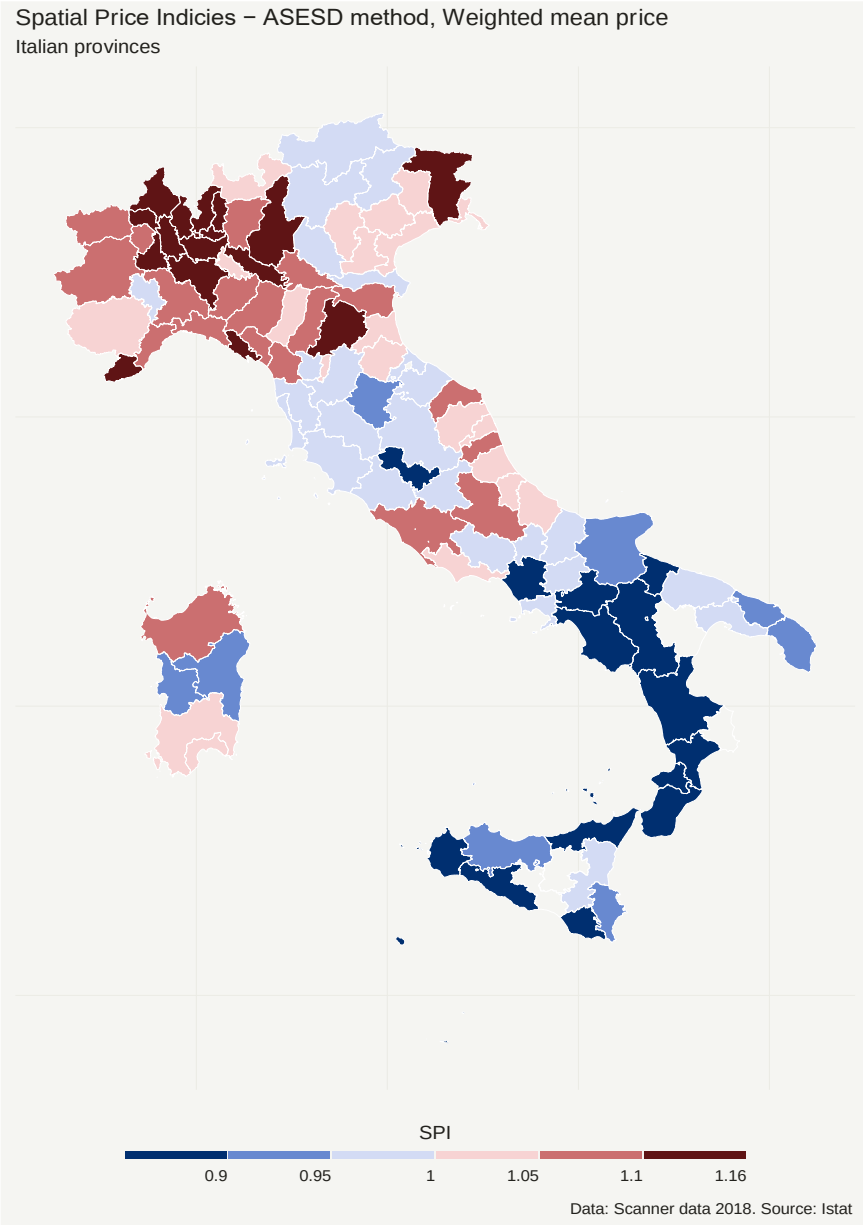
- Therefore to verify our hypothesys we do the Estimations of the Subnational Spatial Consumer Price Indices (SN-SCPIs) by using two different definitions of products' comparability
- The data base used is the same for food and beverage categories and for each of the 103 provinces

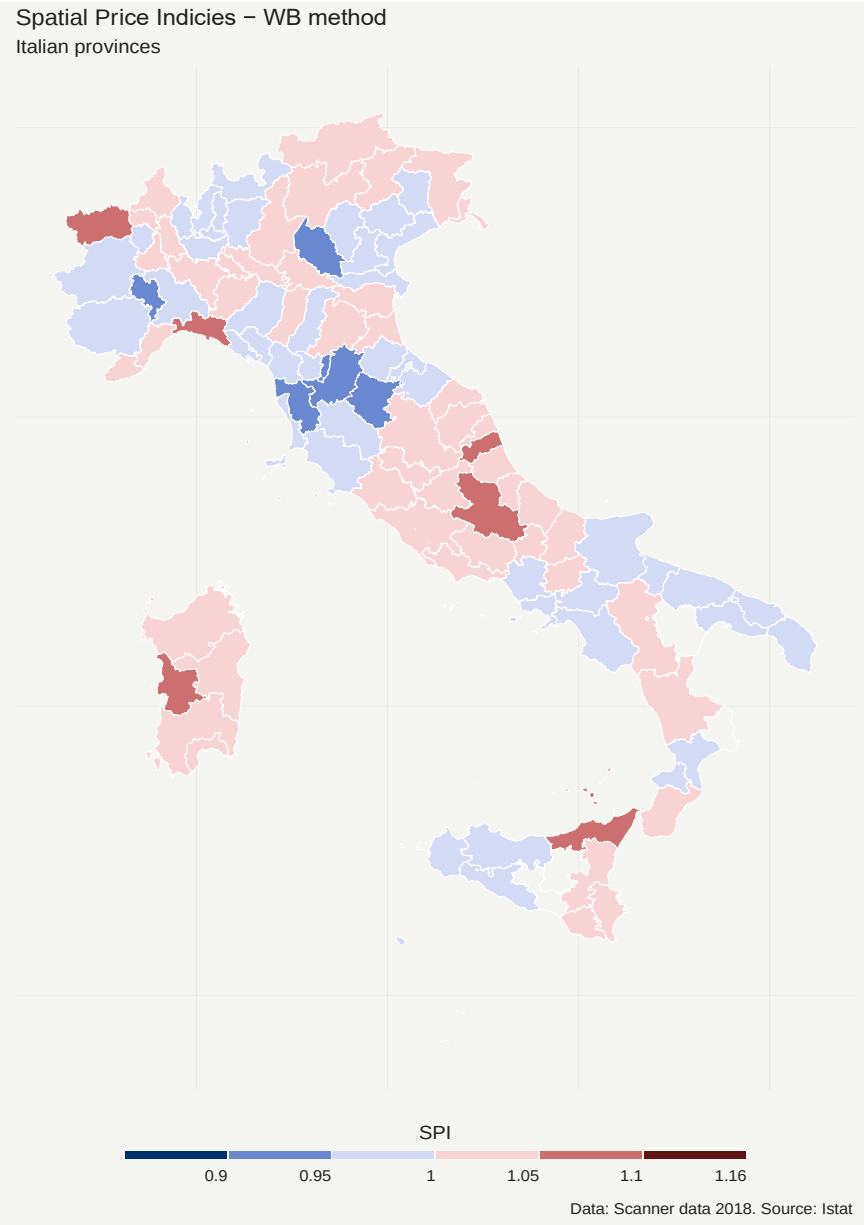
WB approach: the principle of comparability is applied in a very tight way by considering the **comparisons of the "like to like" items** (products) for the different sub-national areas. Computation of the SCPIs possible for each provincial **53 food BHs** by using a weighted CPD model and then their aggregation at the provincial SCPIs

ASESD approach: the **principle of comparability is applied at the level of group of products**, by loosing the tight specifications at the level of the elementary products. The **hypothesis** is that the elementary products (items) belonging to each group satisfy in any case the same consumer needs. The **groups of product** chosen have been at level of the **ECOICOP 8-digit classification (that are 102**). For each group the weighted average price is computed and the provincial SCPIs is obtained by the application of weighted CPD model.

✓ The results are showed in the following cartograms

Italian Provincial SCPIs computed by using scanner data- Average Italy =1



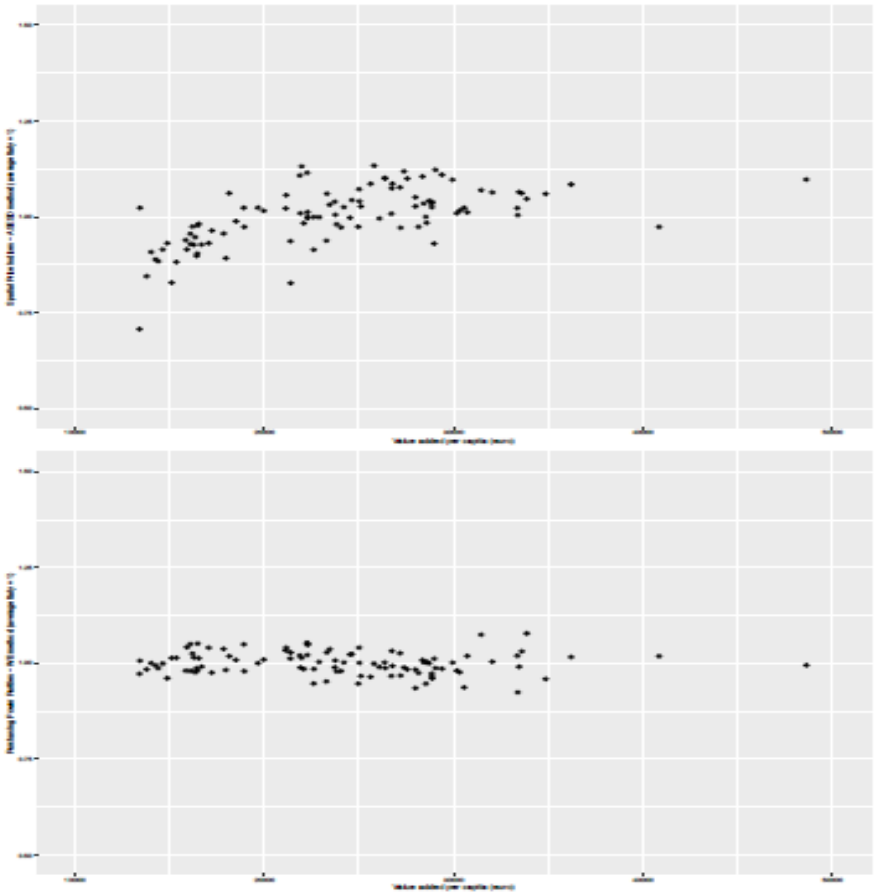


1.3 The experiments conducted (cont.): General comments

- The **results** obtained are undoubtedly **interesting** and shown the substantial different level of food's prices in the Italian Provinces.
- The **results of the two computations are different** and those with WB approach are smoother (without any clear North/ South divide as was expected), than those obtained by ASESD approach.
- It is important to take into account that the procedures followed are different too, and that the first estimations **could** be **more influenced by the characteristics of the modern retail trade** (and of their sale policy) which is not uniformly distributed across Italian territory.

✓ From simple computation on the **correlation between the SCPIs and the value added per capita by Italian province**, it is clear that the **SCPIs computed according, ASESD approach has higher internal consistency** (in fact the correlation coefficient is about +0,60 against -0,10)

➢ In conclusion the **approach followed by ASESD**, to compute the SN-SCPIs, **looks useful and promising**. It might be interesting to choose finer products groups.



Scatterplot of the SPIs and PPPs versus the value added per capita, Italian province

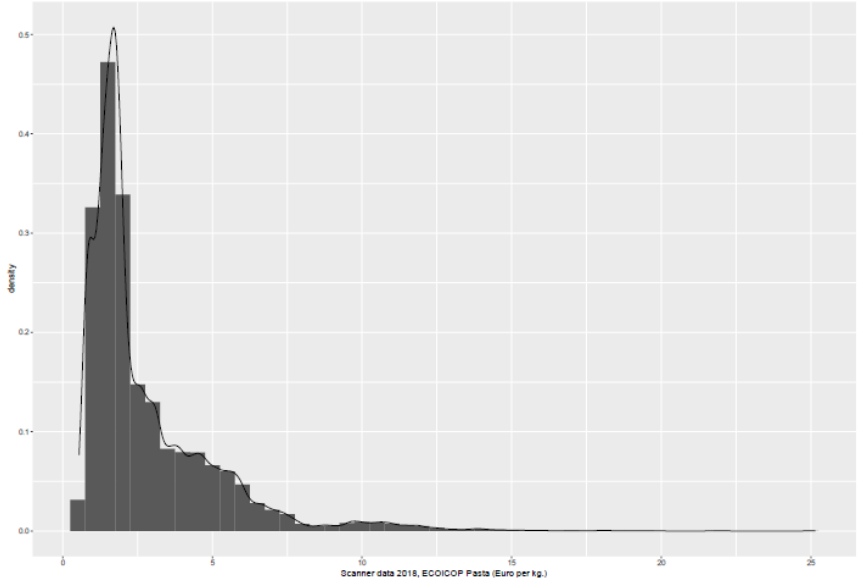
2. The computation of Subnational poor specific SCPIs

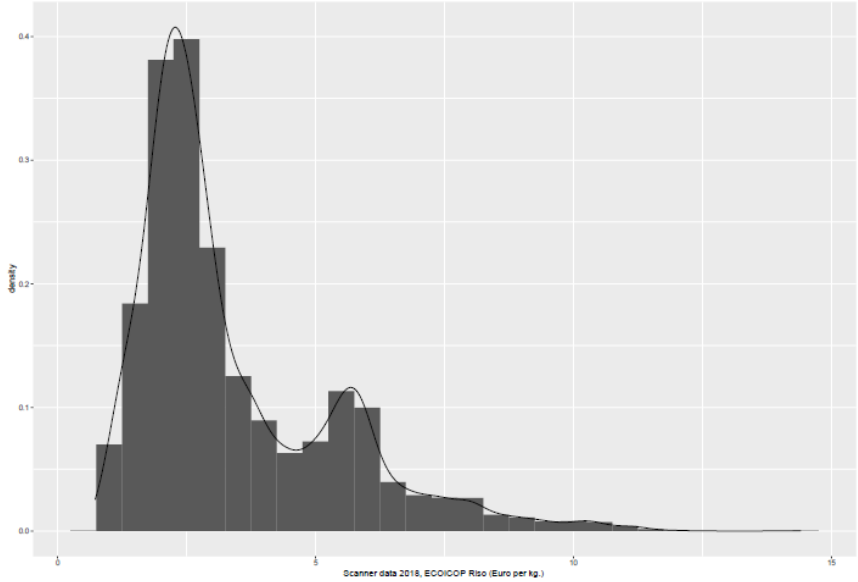
- In order to calculate SPIs closer to the prices paid by the poor, preliminary experiments have been conducted by **using the data of the first quintile of the price distributions**, assuming that the poor purchase the cheaper items of a product. (see for example the following figures)
- After the analysis of price distribution of each product the CPD model was applied to obtain the **SCPIs($Q_{0,2}$)** that is referred to the first quintile of the price distributions, while the general SCPIs(Mean) refer to the all price distributions

The results obtained are reported in the cartogram of the following slide

- The results obtained are somehow expected. Indeed, **provinces in the south of Italy show SPI smaller than 1,** while provinces in the north show values greater than 1. However, there are exceptions.
- **The value of SCPIs($Q_{0,2}$) are smoother than SCPIs(Mean**). It seems that the cheaper prices depend essentially on the kind of products purchased in the different outlets and areas.
- This experiment could be conducted also on data obtained by traditional surveys
- Need for more information on the purchases of poor?

EXAMPLES OF DISTRIBUTION OF PRICES

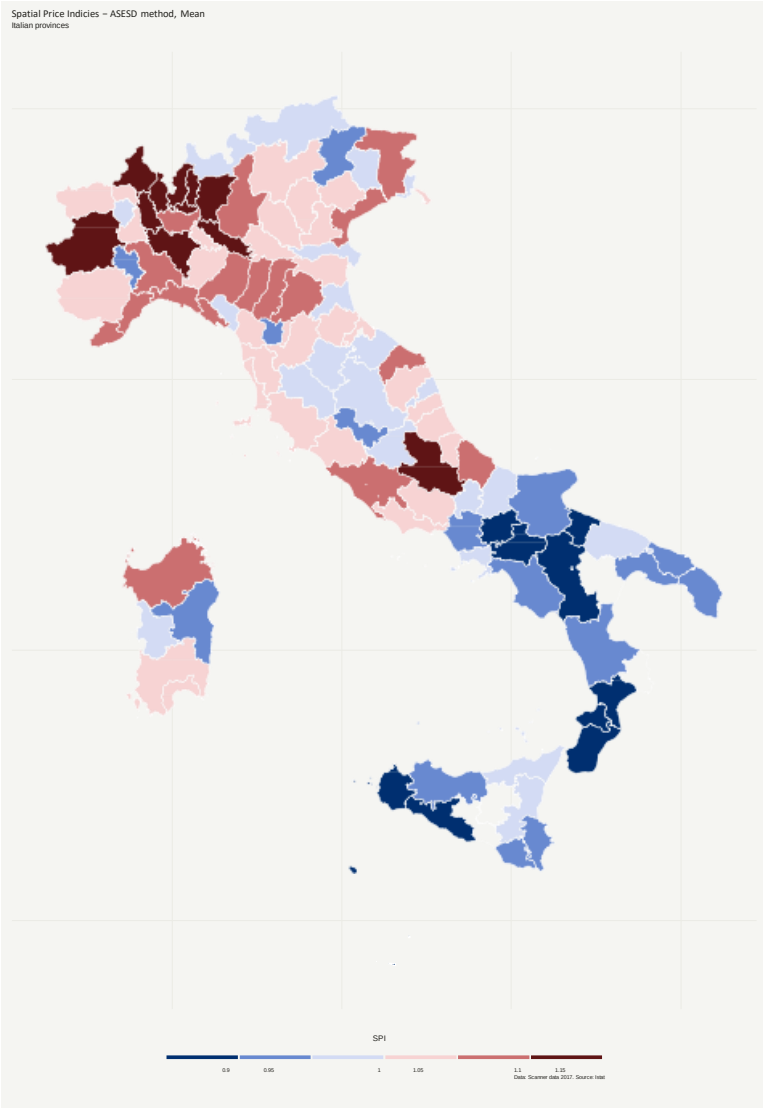


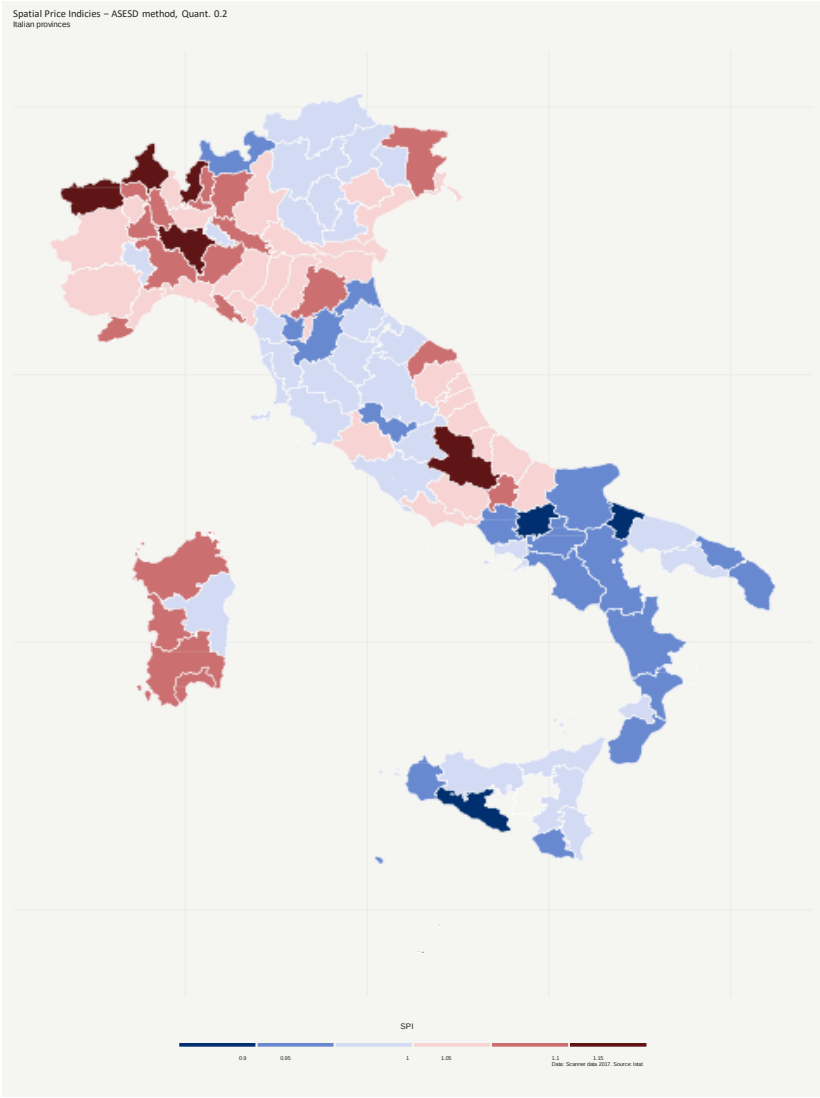


Prices for pasta (€/Kg)				
Min.	1st Q.	Med.	3rd Q.	Max.
0.54	1.42	1.9	3.6	25.14

Prices for rice (€/Kg)				
Min.	1st Q.	Med.	3rd Q.	Max.
0.73	2.1	2.7	4.43	14.4

Italian Provincial SCPIs in average and poor specific: computed by using mean unit prices (left) and quantile 0.2 of unit prices (right).





3. The collection of information on the main outlets where poor do the purchase (through HES)

- In 2019, Istat have done a specific analysis of data collected with the Italian Households Expenditure Survey (HES) **to know where people in condition of absolute poverty purchase 25 large consumption products**.
- The **households have been divided into two groups**: households over (non poor) or below (poor) the thresholds of absolute poverty.
- The results obtained from these preliminary analyses of 2019 HES data show **some interesting differences** between non poor and absolutely poor households in terms of frequency of purchase (poor do purchase with less frequency); and the choice of the type of outlet (poor use mostly traditional shop and hard discount). See the tables
- **This evidence is worth to be deepened also by breaking down the analysis at territorial level,** overcoming the problem of a too small size of the sample if we take into consideration only poor households.
- This line of research is aimed at improving the estimation of the actual prices paid by the poor families in different Italian geographical areas by taking into account their different behavior

FREQUENCY OF THE PURCHASES AND TYPES OF OTLET WHERE POOR HOUSEHOLDS MAKE PURCHASE

Table 7: Families who did not purchase in the last two weeks (%) by product - Year 2019. Source: elaboration on 2019 Istat HES data.

Products	No purchase - percentages of Households		
	No poor	Poor	Total
Bread	7.1	12.5	7.5
Pasta	18.8	32.1	19.7
Biscuits, rusks, snacks	17.7	31.2	18.6
Fresh meat	9.9	20.0	10.5
Frozen meat	25.6	39.9	26.5
Cured meats	16.6	41.8	18.2
Fresh fish	54.2	81.0	55.9
Frozen fish	63.8	80.6	64.9
Milk	19.6	31.1	20.3
Chesees	10.3	26.8	11.4
Yogurt	44.2	65.7	45.6
Eggs	27.8	38.8	28.5
Fres fruit	6.0	19.1	6.9
Fresh vegetables, potatoes and legumes	4.9	15.5	5.6
Dried or frozen vegetables, potatoes and legumes	44.0	63.3	45.3
Olive oil	58.5	75.9	59.6
Mineral water	29.0	48.1	30.3
Soft drinks	37.9	52.8	38.8
Wine	59.1	85.6	60.8
Coffee	42.7	65.2	44.2
Medicines	38.4	68.7	40.4
Personal hygiene products (soaps, deodorant, baby diapers, etc.)	25.1	48.7	26.6
Cleaning products	22.8	44.5	24.2
Disposable items for the kitchen (napkins, dishes, etc.)	41.1	63.1	42.5

Table 9: Types of outlet where poor households make purchases (% distribution) - Year 2019. Source: elaboration on 2019 Istat HES data.

Products	Traditional shop	Open market and street vendors	Hard discount	Hypermarkets and supermarkets	Department stores and outlet chains	Farm or direct producer	Internet
Bread	41.5	1.7	19.1	37.6	0.1	0.0	0.0
Pasta	13.9	1.0	31.3	53.3	0.5	0.0	0.0
Biscuits, rusks, snacks	10.9	1.6	32.3	55.0	0.2	0.0	0.0
Fresh meat	30.2	1.3	23.9	43.9	0.3	0.3	0.0
Frozen meat	10.9	1.6	16.4	69.3	0.0	0.0	0.1
Cured meats	15.7	1.1	29.6	53.1	0.2	0.3	0.0
Fresh fish	41.0	13.6	14.0	30.3	0.0	0.3	0.0
Frozen Fish	10.3	1.7	32.6	54.4	1.0	0.0	0.0
Milk	13.7	1.1	32.5	52.5	0.2	0.0	0.0
Chesees	14.7	0.8	30.4	53.7	0.2	0.2	0.0
Yogurt	10.7	0.9	32.5	55.6	0.3	0.0	0.0
Eggs	14.3	2.4	32.6	49.7	0.2	0.7	0.0
Fres fruit	25.1	10.8	24.9	38.8	0.1	0.2	0.0
Fresh vegetables, potatoes and legumes	25.1	9.6	24.1	40.7	0.2	0.3	0.0
Dried or frozen vegetables, potatoes and legumes	10.3	5.0	30.6	53.4	0.4	0.4	0.0
Olive oil	8.5	1.3	33.2	55.4	0.3	1.4	0.0
Mineral water	11.5	0.7	29.1	58.1	0.4	0.1	0.0
Soft drinks	8.1	2.0	31.2	58.2	0.5	0.0	0.0
Wine	13.4	8.0	32.3	53.5	0.0	0.9	0.0
Coffee	10.1	1.1	31.4	56.5	0.1	0.3	0.5
Medicines	94.0	0.0	0.7	5.4	0.0	0.0	0.0
Personal hygiene products (soaps, deodorant, baby diapers, etc.)	10.3	0.9	31.7	55.1	1.7	0.1	0.2
Cleaning products	12.8	2.3	31.1	51.7	2.1	0.0	0.0
Disposable items for the kitchen (napkins, dishes, etc.)	13.5	1.5	34.1	49.6	1.3	0.0	0.0
Toys and videogames	27.2	9.1	17.3	34.4	11.9	0.0	0.0

- The ASESD Camilo Dagum Center will continue the experiments along these lines of research

- However, to have more evidence it is necessary that other Institutions in different countries implement the same experiments

- ASESD is available to coordinate, in collaboration with ICP Global Office the activities of the involved institutions making the acquired experience available

Thanks for your attention!