NEW DATA SOURCES FOR THE VALORISATION OF THE ABSOLUTE POVERTY THRESHOLDS[1]

Alessandro Brunetti, Stefania Fatello, Orietta Patacchia, Rosabel Ricci

Abstract. Inter Institution Scientific Commission on Absolute Poverty (IISCAP) was established at the National Institute of Statistics (Istat) with the aim of reviewing and improving the methodology used to estimate the incidence of poverty. Several issues were addressed by the Commission, among which the advancements in the valorisation of the absolute poverty lines that were made possible by the availability of new data sources for the scope of official statistics, and specifically by the development of data collection of the survey on consumer prices. These advancements are at the core of the present paper. On the one hand, the assessment of the minimum average prices of food items included in the poverty basket has been greatly improved through the use of transaction data on grocery products coming from the outlets of the modern distribution, already in use for the purpose of the inflation estimate. Moreover, the information provided by the administrative data on rental contracts, which also contributes to the compilation of the consumer price index, allowed the abandonment of the former model approach, based on Household Budget Survey. As a result, the housing component of the poverty line now includes exogenous estimates of the monetary value of rents, based on the Tax Office database, that gives almost a complete picture of all the lease contracts existing in a given year in Italy.

1. Introduction

The review of the methodology used to estimate absolute poverty in Italy was the main goal of the Inter Institution Scientific Commission on Absolute Poverty (IISCAP), established at the National Institute of Statistics (Istat). In the final reports of the Commission subgroups, several innovations were suggested with the aim to update the basket of absolute poverty, that dates back to 2005, and to improve the procedures used to measure the incidence of poverty, taking into account the

[1] The authors are grateful to an anonymous referee for useful comments to a preliminary version of the paper. Of course, they remain solely responsible for any errors or omissions that may have remained. In particular, Rosabel Ricci contributed to the writing of Section 2.1, Stefania Fatello contributed to Section 2.2 and Orietta Patacchia contributed to Section 3. Finally, Alessandro Brunetti contributed to the remaining sections of the paper.

potential of the new data sources that became available in the last decades for the scope of the official statistics[2].

The present paper focuses on the advancements in the valorisation of the absolute poverty lines that were made possible by the development of data collection of the survey on consumer prices. Indeed, the acquisition of transaction data on grocery products coming from the outlets of the modern distribution, already in use for the purpose of the inflation estimate, offers the opportunity for an important enhancement in the assessment of the minimum average prices of the large part of food items included in the poverty basket. In addition, the information provided by the administrative data on rental contracts, which also contributes to the compilation of the consumer price index, allow to replace the former econometric approach based on Household Budget Survey data. Specifically, the monetary value of rents included in the housing component of the poverty line are now exogenously estimated basing on the information contained in the Tax Office database that gives almost a complete picture of all the lease contracts existing in a given year in Italy.

The paper is organized as follows: Section 2 addresses the issue of the estimation of the minimum average prices of the food items in the poverty basket. In particular, in Section 2.1 we will firstly provide a short resume of the methodology used so far to estimate the average minimum price of food products, basing on the traditional price collection, which, with only minor innovations, has still been used for a small subset of products (i.e., fresh food sold in variable weight packages), for which scanner data are not available at present.

Section 2.2 describes in some detail how transaction data have been used for the valorisation of the largest part of food items, focusing on the new approach to the estimation of minimum average prices. The approach, that has been developed within IISCAP aims to make the most efficient use of the new data source. As made explicit, one of the main innovations concerns the selection of references for the estimation of the minimum average price, which is based on the analysis, item by item, of the most representative formats, in terms of packaging type and size.

A synthetic illustration of the Tax Office dataset used for reviewing the monetary value of the housing component of the poverty line is provided in Section 3, in which we sketch the criteria applied to select the subset of lease contracts for the calculation the minimum rental cost per square meter.

The method used to update the food and rental thresholds of the poverty lines in the upcoming years is briefly described in the concluding Section 4.

[2] For a general view on the methodological approach used by Istat to estimate absolute poverty and the innovations introduced following the recommendations of the Inter Institution Scientific Commission, see Freguja and Polidoro (2024).

2. Estimating minimum average prices for food items.

The food component of the poverty basket has undergone major revisions made by the experts of the IISCAP[3] resulting in the enlargement of the basket, which now includes 100 items. However, for the valorization of the food thresholds, the subset of items, for which the minimum average price has to be estimated, is 96[4].

To this aim, the data coming from the consumer price survey have been used. As the first step, a mapping was made which traces each food item included in the poverty basket to one or more elementary products of the consumer price basket[5]. In particular, almost one third of the items (33), relating mainly to fresh products such as fruits, vegetables, meat and some foods usually sold in packages of variable weight, were linked to elementary products for which price information is still coming from the traditional price collection carried out at the provincial level by the Municipal Statistical Offices (MSOs)[6]. For this subset of items, the minimum average prices are estimated according to the former methodology[7], described in Section 2.1.

For the remaining 63 food items of the poverty basket (linked to package products), the minimum average prices were assessed using the new source of the consumer price survey, namely the scanner data source.

The scanner data, indeed, is a powerful source of information, allowing for a wider coverage of stores as compared to the traditional price collection, both from the point of view of channels and the number of outlets (about 4,000), and offering immeasurably much more granular information about the products actually sold in the reference period. In order to make the most efficient use of the information coming from transaction data, a new methodology for the estimation of minimum average prices has been developed within the Commission, which is discussed in Section 2.2.

It is important to stress that, regardless the source of information, the minimum average prices have been compiled at the regional level, using provincially collected prices, in order to allow the estimation of poverty thresholds for the NUTS 2 subdivisions. All estimates are referred to year 2022.

---

[3] de Martino *et al.* (2024) discuss in details the meticulous analysis that has been carried out in order to define the new food component of the basket.

[4] See de Martino *et al.* (2024).

[5] A detailed review of the methodology used to compile the consumer price indices and the corresponding basket of products is contained in Istat (2023).

[6] In 2022, the traditional price collection took place in 79 provinces participating to the consumer price survey.

[7] Indeed, a couple of innovations were introduced that improve the approach used to estimate the minimum average prices of the traditionally collected products. Firstly, sales and temporary price reductions were taken into account (they were not in the former approach) and secondly, as stated below, estimates were produced at a more detailed territorial level (i.e., regional level). For an exhaustive illustration of the former methodology, see Istat (2009).

2.1. The traditional price collection for the valorisation of the food sub-component of the poverty basket.

For the fresh food products sold in in packages of variable weight, the information collected at local level by MSOs with traditional techniques was used to estimate the corresponding minimum average prices. In particular, this sub-group consists of 33 food items generally referred to: milk and derivatives; fresh meat; preserved meat; cereals and derivatives; fish and fish products; vegetables and fruit. As a whole, the 33 food items were linked to 113 elementary products included in the Harmonized index of consumer price (HICP) basket. In few cases, the link between the elementary products of the HICP basket and the food items of the poverty basket is one-to-one (this is the case for example of Dried, salted or smoked meat). At the opposite, for Fresh or chilled fruit, 37 elementary products were considered, and 51 for Fresh or chilled vegetables other than potatoes and other tubers.

In order to cover both the modern and the traditional outlets, the minimum average prices were estimated, at the provincial level, by keeping separated the modern retail channel (which includes hypermarkets, supermarkets, outlets with surface between 100 and 400 s.m. and discounts) from the traditional channel (mini markets, traditional shops, local markets).

Formally, let $\{a_i\}$ be the set of the elementary products linked to the food item A and $\{p^m_{j,h,k}(a_i)\}$ the set of quotations of the j references collected in month m in outlets of type h of the municipality k for the elementary product a_i.

The minimum price $\hat{p}^m_{h,k}(a_i)$ is given by:

$$\hat{p}^m_{h,k}(a_i) = \min_j \{p^m_{j,h,k}(a_i)\} \quad (1)$$

provided that at least two price offers have been collected (imputed prices were not taken into account).

The aggregation of the minimum prices $\hat{p}^m_{h,k}(a_i)$ involves a series of further steps:

1. The monthly minimum price of the elementary product a_i of the municipality k is calculated by aggregating $\hat{p}^m_{h,k}(a_i)$ across outlet types. Specifically, $\hat{p}^m_k(a_i)$ is obtained as the weighted arithmetic mean of $\hat{p}^m_{h,k}(a_i)$ with weights proportional to the importance of the channels in terms of turnover[8] (Table 1).

[8] The 5-digits level of the ECOICOP classification (i.e. sub-classes) is the lowest level for which the turnover weights of different outlet types are available. Therefore, for all the elementary products included in the same ECOICOP sub-class (or class, in case more detailed information on turnover is not available), the corresponding monthly minimum average price is calculated using the same weighting coefficients.

2. The minimum average price of the month m in municipality k of product A, $\hat{p}_k^m(A)$, is obtained as the weighted arithmetic mean of the monthly minimum prices $\hat{p}_k^m(a_i)$ of the elementary products a_i, using the HICP weights. In a number of instances, however, no explicit weights were available: in this cases, the monthly provincial minimum average of food product $\hat{p}_k^m(A)$ is calculated using the geometric mean of $\hat{p}_k^m(a_i)$.
3. The annual minimum food prices of the province are calculated as the arithmetic mean of the provincial monthly minimum prices.
4. The annual minimum food prices for each region are calculated as the weighted arithmetic mean of the provincial annual minimum prices, with weights proportional to the resident population of the provinces.

For the valorisation of the poverty basket, about 500,000 traditionally collected quotations were handled to calculate the minimum average prices of the 33 food products. The smaller number of quotations was used for Fresh or chilled fish (7,968 quotations), while the largest number for Fresh or chilled vegetables other than potatoes and other tubers (more than 200,000).

Table 1 – *Weights of the modern and traditional retail trade channels used to compile the minimum average prices of traditionally collected food products, by ECOICOP (year 2023).*

ECOICOP	Level	Description	% Modern distribution	% Traditional distribution
01.1.2	Class	Meat	51,6	48,4
01.1.4.5	Subclass	Cheese and curd	51,6	48,4
01.1.6.1	Subclass	Fresh or chilled fruit	41,8	58,2
01.1.7.1	Subclass	Fresh or chilled vegetables other than potatoes and other tubers	62,3	37,7
01.1.7.4	Subclass	Potatoes	62,3	37,7

2.2. The use of scanner data to estimate minimum average prices of processed food products

Since 2018, Istat has been using scanner data of grocery products in the production process of the consumer price indices. In 2022, Istat received data for a sample of about 4,000 outlets, including hypermarkets, supermarkets, discounts,

outlets with surface between 100 and 400 s.m. and specialist drug[9]. The sample of outlets is stratified by provinces and retail trade channels and cover the entire national territory.

Istat receives scanner data on a weekly basis and at item code level (GTIN[10]) for each outlet in the sample. For each GTIN, elementary prices are calculated taking into account turnover and quantities (weekly price = weekly turnover/weekly quantities). Every GTIN identifies a specific product through a series of attributes (manufacturer, brand, possible sub-brand, size, packaging, variety) and all GTINs are classified within the ECR classification[11] which groups homogeneous products in specific markets.

Starting from the database available for the calculation of consumer price indices, Istat has been implemented a new methodology to estimate the average minimum price of the processed foods included in the poverty basket. In fact, the granularity of the information coming from the transaction data allowed a great improvement in terms of temporal and territorial coverage as well as of product offering, as compared to traditional sources of price collection. It also made it possible to include all retail trade channels where households purchase food products.

As the first step, the 63 food products of poverty basket have been linked with transaction data at the lowest level of the ECR classification used for scanner data (i.e., markets). All in all, 171 markets were linked.

The large amount of information provided by scanner data, however, raises the issue of the selection of the references that contribute to the estimation of the minimum average price. As a matter of fact, transaction data refer to products that are generally sold in a wide number of different packaging of different size. For example, the evidence coming from scanner data shows that canned tuna is sold in 14 different types of packaging (single or multi pack, containing up to 24 pieces) and 114 different sizes (ranging from 33gr to 4kg). With the aim of selecting the most representative formats, an analysis has been carried out to identify, product by product, the most sold types of packaging. The selection criterion was based on annual data on quantities sold. As for canned tuna, five different typologies have been selected: three single packs (80gr, 160gr and 200gr) and two multi-packs (3x80gr and 2x160gr).

For the selected typologies of packaging of each food product, the minimum average price has been estimated on a monthly basis at the provincial level, separately for hypermarkets, supermarkets, discounts and outlets with surface between 100 and 400 s.m.. Moreover, unlike the case of the traditionally collected

[9] Small traditional local shops are not covered by scanner data.
[10] The Global Trade Item Number (GTIN) is a unique product identifier that is recognized internationally.
[11] ECR markets are the lowest level of the ECR classification (classification shared by industrial and distribution companies).

price, the use of scanner data allowed us to consider the price distribution of GTINs as the basis for the estimation of the minimum average price[12].

Specifically, the minimum average price (for a given province, month, outlet type and packaging), is obtained as the weighted arithmetic mean of the prices of the lower tail (first fifth) of the price distribution, with weight given by quantities of GTINs sold[13].

Formally, let $B_r = \{b_{i_r}\}$ be the set of the GTINs of the selected package type r linked to the food product B, and $B^m_{r,h,k} = \{b^m_{i_r,h,k}\}$ the subset of the GTINs that belong to the first fifth of the price distribution of month m in outlets of type h of the municipality k.

The minimum price $\hat{p}^m_{h,k}(B^m_{r,h,k})$ is given by:

$$\hat{p}^m_{h,k}(B^m_{r,h,k}) = \sum_{B^m_{r,h,k}} \frac{q^m_{h,k}(b^m_{i_r,h,k})}{q^m_{h,k}(B^m_{r,h,k})} \cdot p^m_{h,k}(b^m_{i_r,h,k}), \quad (2)$$

where $q^m_{h,k}(b^m_{i_r,h,k})$ is the quantity of the GTIN $b^m_{i_r,h,k}$ and

$$q^m_{h,k}(B^m_{r,h,k}) = \sum_{B^m_{r,h,k}} q^m_{h,k}(b^m_{i_r,h,k}). \quad (3)$$

The minimum average prices of the selected packages are then expressed in a standardized unit (price/kg or price/liter) before being aggregated to calculate the monthly minimum average price of food product B sold in the outlets of the same type h and province k:

$$\hat{p}^m_{h,k}(B^m_{h,k}) = \sum_r \frac{q^m_{h,k}(B^m_{r,h,k})}{q^m_{h,k}(B^m_{h,k})} \cdot \hat{p}^m_{h,k}(B^m_{r,h,k}), \quad (4)$$

where $q^m_{h,k}(B^m_{r,h,k})$ is the standardized quantity (expressed in terms of kilos or liter) of product B sold in packages of type r and

$$q^m_{h,k}(B^m_{h,k}) = \sum_r q^m_{h,k}(B^m_{r,h,k}). \quad (5)$$

Minimum average prices $\hat{p}^m_{h,k}(B^m_{h,k})$ are then aggregated across outlet types and provinces of the same region to calculate the monthly minimum average price of food product B at the regional level. For the aggregations the weighted arithmetic

[12] For more details on the methodology see Altarocca *et al.* (2024).

[13] As a whole, for the 63 food products about 12 million quotations have been processed each month. From these, 1,8 million quotations have been monthly selected for the calculation of the 63 minimum average prices.

mean formula is used with weights proportional to the turnover of the different retail channels and, as for the traditionally collected prices, the resident population.

Finally, the regional minimum average price of food product B for the reference year (2022) is obtained as the arithmetic mean of the monthly minimum average prices.

3. Tax Office database of rents as a data source for the estimation of the rental component of the poverty line

The use of administrative data for the estimation of the rental component of the absolute poverty lines represents a crucial innovation with respect to the previous methodology based on Household Budget Survey data. In particular, the data source is the database of rents held by the Real Estate Market Observatory (OMI) of the Tax Office, which is used, starting from 2022, for the compilation of the Italian CPI and HICP indices of housing rentals index for private sector[14].

The provision of the database by the Tax Office started in 2017 in the framework of the cooperation between Istat and the fiscal agencies of the Ministry of Economy and Finance. It contains information about all new rental contracts registered in the reference quarter referring to all type of properties (residential and non-residential), covering almost the entire national territory[15], and it is provided 35/40 days after the end of the reference quarter. Specifically, the information provided by the OMI database concern:

1. the contract (the value of rental price, the tax option, the start date of occupancy, the end date of occupancy and sometimes the date of the early termination);
2. the characteristics of the property (cadastral categories, the size in terms of square meters and/or number of rooms, the location of the property in terms of cadastral code, cadastral information as map sheet and number, OMI zone, address);
3. the parties involved (information about landlords and tenants including fiscal code).

[14] As a whole, more than 1.3 million of rental prices are available monthly.

[15] For properties located in the autonomous provinces of Bolzano and Trento and in few municipalities situated in Veneto, Lombardia, and Friuli that are regulated by a different cadastral system some information are missing (specifically, dwelling area in terms of squared meters and the number of rooms).

For the valorisation of the rent sub-component of the poverty line, however, it was necessary to select the statistical units (new and existing contracts[16]) from the OMI database, according to specific criteria.

Firstly, only contracts from private landlords were considered. For this purpose, OMI database was integrated with the information coming from the Statistical Register of active Enterprises (ASIA), using the VAT code of the property owners. By combining the two sources, it was possible to identify the owner type and consequently to distinguish private from social contracts.

Secondly, short term contracts (usually used for study, work or holidays) were excluded while only long term contracts (4+4 years and 3+2 years) were taken into account.

Luxury dwellings (including villas, castles and houses of historical and artistic value), which are considered out of scope, were excluded as well. On the contrary only properties for ordinary housing (that is, civil, economic, popular, ultra-popular, housing of a rural type, dwellings in cottages, corresponding to cadastral categories A2 to A7) were selected.

Finally, outliers defined according to dwelling area and rent per square meter were dropped out from calculation.

Having selected the statistical units from the Tax Office database, the cost of rents per square meter is obtained as the average price of each strata, defined according to dwelling size (surface area), geographical location of the dwelling (at NUTS 2 level) and the municipality type[17].

As for the case of food products, the estimates of the cost of rents per square meter are referred to year 2022.

4. Conclusions

The new sources of information that became available for the scope of the consumer price survey offered the opportunity to improve substantially the official indicators on absolute poverty by allowing more robust estimates of the poverty lines.

The methodological solutions that have been adopted in order to make the most efficient use of the information coming, on the one hand, from transaction data on

[16] The availability of the expiring date of the contracts allowed to estimate the stock of the existing ones in the reference year (2022). Moreover, a specific procedure was implemented in order to detect those contracts that (depending on the tax regime chosen by the owner) needed to be yearly updated and to revaluate accordingly the corresponding rents.

[17] Centre of metropolitan area, municipalities of metropolitan area suburbs and municipalities with more than 50,000 inhabitants and other municipalities. See Cutillo (2023) for more details.

food products and, on the other hand, from the administrative data on rents have been discussed in the previous sections.

As already noted, for the valorisation of the thresholds of the two sub-components, year 2022 has been considered as the reference year. It is worthwhile to emphasize that, according to the conclusions of the IISCAP, the estimates of minimum average prices will undergo a five-year revision, while, for years in between, the thresholds will be updated basing on the evolution of a specific set of sub-class indices[18] of the HICP.

Acknowledgements

The innovations discussed in the paper are the result of the joint effort of experts who cooperate within Inter Institution Scientific Commission on Absolute Poverty. Special thanks go to Luigi Biggeri, (Università di Firenze), Enrica Chiappero Martinetti (Università degli studi di Pavia), Caterina Giusti (Università degli studi di Pisa), Tiziana Laureti (Università della Tuscia), Achille Lemmi (Università degli studi di Siena), Federico Polidoro (World Bank), Stefano Toso (Università degli studi di Bologna) and to Istat colleagues: Ilaria Arigoni, Andrea Cutillo, Valeria de Martino and Federico Di Leo.

References

ALTAROCCA F., DORMI C., FATELLO S., MATTA C. 2024. Some empirical evidence from the use of scanner data to estimate prices of food products included in the absolute poverty basket, *Rivista Italiana di Economia Demografia e Statistica,* Special Issue on "New approaches for measuring poverty: studies and perspectives", Vol. LXXVIII, No. 4, pp. 118-134.

CUTILLO A., 2023. The measurement of absolute poverty in Italy: an overview and a revision based on the database of real estate leases of the Tax Agency, *Rivista Italiana di Economia Demografia e Statistica,* Vol. LXXVII, No.1, pp. 5-16.

DE MARTINO V., MISTURA L., PICCINELLI R. 2024. The new food sub-component of the basket of absolute poverty: an optimal balance of elements, *Rivista Italiana di Economia Demografia e Statistica,* Special Issue on "New approaches for measuring poverty: studies and perspectives", Vol. LXXVIII, No. 4, pp. 49-64.

[18] Corresponding to the five-digits level of the ECOICOP classification.

FREGUJA C., POLIDORO F. 2024. The concept and measurement of poverty, *Rivista Italiana di Economia Demografia e Statistica*, Special Issue on "New approaches for measuring poverty: studies and perspectives", Vol. LXXVIII, No. 4, pp. 21-48.

ISTAT 2009. La misura della povertà assoluta. *Metodi e Norme n. 39*. https://www.istat.it/wp-content/uploads/2021/01/misura_della_poverta_assoluta.pdf

ISTAT 2023, Indici dei prezzi al consumo. Aspetti generali e metodologia di rilevazione - Edizione 2022. *Letture Statistiche - Metodi.* https://www.istat.it/wp-content/uploads/2023/04/Ebook-Indici-prezzi-al-consumo-2022.pdf

---

Alessandro BRUNETTI, Istat, albrunet@istat.it
Stefania FATELLO, Istat, fatello@istat.it
Orietta PATACCHIA, Istat, patacchi@istat.it
Rosabel RICCI, Istat, roricci@istat.it