ESTUDIOS DE ECONOMÍA APLICADA VOL. 36-1 2018 PÁGS. 119 – 148

Measuring Spatial Price Level Differences within a Country: Current Status and Future Developments

TIZIANA LAURETI [a], D.S. PRASADA RAO [b]

[a] *University of Tuscia, Department of Economics, Engineering, Society and Business Organization, Via del paradiso, 47, 01100 Viterbo, Italy. E-mail:* laureti@unitus.it

[b] *University of Queensland, Shool of Economics, Blair Drive, St. Lucia, Queensland 4069, Australia. E-mail:* d.rao@uq.edu.au

ABSTRACT

The importance of constructing sub-national purchasing power parities (PPPs) has been acknowledged in the literature for over two decades. In countries characterised by large territorial differences in prices and quality of products and household characteristics, calculation of sub-national PPPs can improve and serve as inputs for regular compilation of major indicators as spatial price index numbers, real per capita expenditure adjusted for price level differences across regions and estimates of inequality and poverty produced. This paper provides an overview of the data requirements and index number methods specifically designed for producing consistent set of spatial price comparisons; and a review of the attempts to compile sub-national PPPs around the world. Recent efforts by researchers in Italy in cooperation with the Italian Statistical Office (ISTAT) in the compilation of PPPs for regions in Italy are reviewed and preliminary results from these studies are presented. Finally, the paper scopes future efforts in this important area by discussing the opportunities and challenges offered by the availability of high-frequence point-of-sale data in the form of scanner data for the compilation of sub-national PPPs.

Keywords: Spatial Price Indexes, Country Product Dummy, Scanner Data.

Medición de las diferencias de nivel de precios espaciales dentro de un país: Estado actual y evolución futura

RESUMEN

La importancia de construir paridades subnacionales de poder adquisitivo (PPP) ha sido reconocida en la literatura por más de dos décadas. En países caracterizados por grandes diferencias territoriales en precios, calidad de productos y características del hogar, el cálculo de PPPs subnacionales puede mejorar la compilación regular de indicadores de tan gran importancia como los índices espaciales, el gasto real per cápita ajustado por diferencias de nivel de precios regionales y las estimaciones de desigualdad y la pobreza a escala nacional; también puede servir de insumo para estos agregados macroeconómicos. Este artículo proporciona una visión general de los requisitos en cuanto a datos y métodos de construcción de números índice diseñados específicamente para producir un conjunto coherente de comparaciones espaciales de precios; y una revisión de los intentos de compilación de PPPs subnacionales que se han llevado a cabo en todo el mundo. En especial, se pasa revista a la reciente colaboración en Italia entre los investigadores y el ISTAT en lo que a compilación de PPP para las regiones italianas se refiere, y se presentan los resultados preliminares de estos estudios. Finalmente, este artículo pone de manifiesto interesantes futuras líneas de investigación en esta importante área, discutiendo las oportunidades y los desafíos que ofrece la disponibilidad de la fecha de alta frecuencia de los puntos de venta en forma de datos de escáner para la compilación de PPP subnacionales.

Palabras Clave: Índices de precios espaciales, país producto dummy, datos escáner.

Clasificación JEL: C21, D12, E31

Artículo recibido en octubre de 2017 y aceptado en enero de 2018
Artículo disponible en versión electrónica en la página www.revista-eea.net, ref. ə-36120

ISSN 1697-5731 (online) – ISSN 1133-3197 (print)

1. INTRODUCTION

Comparisons of real incomes and monitoring standards of living and measuring changes in prices over time are tasks performed by the national statistical offices. The consumer price index (CPI) which is compiled in most countries on a monthly, quarterly and annual basis is an important yardstick to measure changes in prices of goods and services paid by consumers over time. CPI is used by the national central banks to monitor inflation and to calibrate monetary policy. CPI is also used by employers and employees in assessing movements in real wages over time. The survey framework and statistical methodology for the collection of consumer price data are well-established and the index number methods for the compilation of CPI are described in the ECE-ILO (2010) *CPI Manual*.

In contrast, spatial price comparisons at the national level are quite rare and only a few countries have produced measures of spatial price differences within countries. However, spatial price comparisons across countries to measure differences in price levels and Purchasing Power Parities of currencies (PPP) have gained prominence over the last forty to fifty years. Research on international comparisons of prices started in 1968 as a collaborative research project between the United Nations Statistical Office and a small research team headed by Professor Irving Kravis at the University of Pennsylvania. The International Comparisons Project, subsequently named International Comparison Program (ICP), has expanded from a small project covering 10 countries in its Phase I to cover 177 countries in the most recently completed round of the ICP in 2011. A brief history of the ICP and its framework can be found in World Bank (2013) and Rao (2013a). The ICP is a major international statistical initiative undertaken by the World Bank, OECD, EUROSTAT, Asian Development Bank (ADB), African Development Bank (AfDB) and UNECLAC and conducted under the auspices and guidance of the UN Statistical Commission.

International comparisons of prices and real incomes are anchored on the concept of purchasing power parities of currencies which are defined as the number of currency units of a country that can purchase the same basket of goods and services that can be purchased with one unit of currency of a reference or numeraire currency. For example, according to OECD database, PPP of Euro in Italy in 2016 with respect to US dollar is 0.722 which means that 0.722 euro in Italy has the same purchasing power as 1 US dollar in the United States.

Cross-country price comparisons have several important uses. International PPPs computed at the economy level for the gross domestic product can be used in ranking countries in terms of real size of the economies and also used in comparing real per capita incomes across countries (World Bank, 2015). These PPPs are used by the IMF and the European Union for internal administrative purposes. The PPPs from the ICP are also used in measuring global inequality and global poverty showing the number of people in absolute poverty under \$1/day and \$2/day international poverty lines. The Human Development Index (HDI) makes use of real per capita income in PPP terms as one of the components of HDI. There are numerous other uses of PPPs for international comparisons which are described in detail in World Bank (2013).

The importance of spatial price comparisons across regions within a country has not been recognized to the same extent as cross-country comparisons of prices. For large countries such as the USA, China, India and Indonesia and for countries which are relatively small, it is equally important to compare prices across different states, districts or regions within countries. In countries like India, there are significant price level differences between urban and rural areas. Similarly, in countries like Italy significant price level differences are observed in the northern and southern regions of the country. Price level differences are observed across different capital cities within a country. Accurate measurement of price level differences across regions within a country is critical for assessing inequality in the distribution of real incomes of the population residing in different parts of the country. For larger countries like China and India, spatial price comparisons are important in assessing and ensuring balanced development across all the regions within these countries.

The importance of constructing spatial price comparisons within a country, also referred to as *sub-national purchasing power parities*, has been acknowledged in literature over the last decade. However, no systematic attempts have been made to compile sub-national PPPs on a regular basis with a few exceptions which are reviewed in section 3.

The main objective of this paper is to provide an overview of the current status of compilation of sub-national PPPs and to discuss the index number methods employed in making spatial price comparisons across regions. The paper also describes the current efforts to compile sub-national PPPs in Italy and present findings from a study undertaken in collaboration with the Italian Statistical Office (ISTAT). The future of sub-national PPP compilation with increased availability of price data from non-conventional sources such as scanner data recorded at the point of sale and the opportunities and challenges posed by the availability of big data for the regular compilation and dissemination of spatial price comparisons across regions within countries are discussed.

The paper is organized as follows. Section 2 describes the steps involved and present the index number methods commonly used in the compilation of spatial price index numbers and sub-national PPPs. Section 3 provides a brief review of the state of the art with respect to compilation of sub-national spatial price indexes. Section 4 describes the Italian experience and the ISTAT project and presents preliminary results from the project. Section 5 is devoted to a discussion of new sources of data and methods and the future developments regarding sub-national PPPs. Section 6 offers concluding remarks.

2. SPATIAL PRICE INDEXES: METHODS AND DATA

Spatial price comparisons analyse the relationships between price levels and the purchasing power of money in two geographical areas at a certain point in time. The concept of purchasing power parity is used to measure the price level in one geographical area compared to that in another area.

As mentioned above, at international level, PPPs for currencies of different countries are compiled by the International Comparison Program (ICP), which is administered by the World Bank and overseen by the United Nations Statistical Commission with the collaboration of the OECD, EUROSTAT and other regional organizations (Rao, 2013a). PPPs facilitate conversion of Gross Domestic Product (GDP) and its major aggregates into a common currency thus allowing for cross-country comparisons. Likewise, sub-national PPPs allow for intra-country spatial comparisons and can serve as inputs for compiling key economic indicators produced by countries, such as regional price comparisons, real per capital incomes and poverty estimates.

The process of compiling PPPs is quite complex and is carried out in several stages in the ICP and in the Eurostat-OECD program[1]. The following is a description of the steps involved in the computation of PPPs for spatial comparisons across regions within a country.

- During the first stage, price and expenditure weights data are collected through special surveys.
- Price data are aggregated and PPPs are computed for major economic aggregates. Aggregation of price data are accomplished in two stages.
- As expenditure data are not available at the item level, PPPs are calculated for the smallest commodity groups for which expenditure data are available. The smallest level of aggregation at which expenditure data are available are known in ICP parlance as *basic headings*. For example, Uncle Ben's rice is an item that belongs to the basic heading *rice*. PPPs at the basic heading level are similar to the elementary price index numbers computed as a part of the CPI compilation. More specifically, elementary spatial price indexes are computed by aggregating the prices of items belonging to a group of similar well-defined goods or services (called Basic Heading, BH).

[1] Providing a detailed description of the process involved in the compilation of PPPs within the ICP is difficult within the space limitations for this paper. Only an overview is provided here and interested readers may refer to Rao (2013a) and also World Bank (2013) for a more complete description.

- In the final stage, elementary PPPs are aggregated using expenditure share data as weights in order to obtain PPPs for higher-level aggregates such as consumption, investment and GDP.

The following sub-sections provide an overview of the essential elements involved in the stages listed above.

2.1. Price and Expenditure Data

The first and a fairly critical step in the compilation of PPPs is to prepare a long list of goods and services that will be priced in the regions involved in spatial price comparisons. The list used in price comparisons needs to meet and balance the requirements of *comparability* and *representativity*. Comparability means that identical products with same or similar quality need to be priced across all the regions so that the PPPs based on such data reflect solely price level differences. However, in countries with significant diversity in the availability of products and consumption patterns, products that are comparable across regions may not necessarily be representative of the products that are typically consumed in different regions. Given these two competing requirements, caution is required in the use of product lists used for compilation of CPI in different regions. CPI compiled for different regions would be based on products typically consumed or sold in respective regions and the CPI products may not always be strictly comparable. It is useful to note here that not all the products may be priced in all the regions. Reliable regional price comparisons can be made as long as there is a reasonable overlap in the items priced in different regions.

Moreover, price data are generally collected in the main cities within a country (usually regional or provincial capitals) and some municipalities may not be included in the CPI survey due to organizational issues. Therefore, several cities or geographical areas of interest may not be included in price comparisons due to lack of data, thus resulting in information loss.

In order to compute spatial price indexes when data collected for CPI aims are unsuitable for achieving territorial comparability, national statistical offices (NSOs) have had to draw on other data sources such as *ad-hoc surveys*. However, *ad hoc surveys* may be very expensive for the NSOs and do not provide information on consumption expenditure as in the case of CPI data.

Several researchers have used data from household surveys, such as Coondoo *et al.* (2011) who proposed a procedure for estimating regional consumer price index numbers in India based on the estimation of item-specific region-wise Engel curves and Majumder *et al.* (2012) who suggested a demand system based approach to the calculation of spatial prices in India.

Expenditure data at the basic heading level are necessary in the process of aggregating price data leading to PPPs for regional comparisons. Expenditure data for consumption items can be obtained from household expenditure surveys which are regularly conducted in almost all countries in the world. It is possible to make use of expenditure share or weights data for consumption from the CPI and more generally from the national accounts for aggregates other than consumption.

Both price and expenditure data need to be validated using data editing procedures described in World Bank (2013) before embarking on the next stage of PPP computation.

2.2. Aggregation methods for spatial price comparisons

Spatial price comparisons across different regions are essentially multilateral comparisons based on price data from different regions. In this section, the focus is on price comparisons involving *M* regions. This means that price data are collected from these *M* regions. Let P_{jk} represent the price index for region *k* compared with the base or reference region *j*. This is essentially a binary index comparing two regions. In the case of spatial comparisons, comparisons between all pairs of regions are of interest and, therefore, the main aim is to compile price indexes in the following price index matrix.

$$\mathrm{P} = \begin{bmatrix} P_{11} & P_{12} & P_{1k} & P_{1M} \\ P_{21} & P_{22} & P_{2k} & P_{2M} \\ P_{j1} & P_{j2} & P_{jk} & P_{jM} \\ P_{M1} & P_{M2} & P_{mk} & P_{MM} \end{bmatrix} \tag{1}$$

where $P_{jk}\ (j,k = 1,2,...,M)$ measures the price level in region *k* relative to the base region *j*. In spatial comparisons, comparisons between all regions are considered important and hence the term, *multilateral comparisons*.

<u>*Transitivity and base invariance*</u>

The multilateral price index numbers P_{jk} are required to satisfy two basic properties in addition to a range of axiomatic properties expected of binary or bilateral price index numbers (see Balk, 2008).

Transitivity of price index numbers in (1) is an internal consistency requirement which stipulates that a direct comparison between any two regions must equal any indirect comparison between the two regions obtained through a third region. Formally stated, *transitivity* requires that for any three regions *j, k* and *l* the index number method used in computing P_{jk} should be such that

$$P_{jk} = P_{jl} \cdot P_{lk} \tag{2}$$

For example, transitivity ensures that a direct comparison of prices between Rome and Venice would be the same when Rome is compared to Milan and Milan compared to Venice and Rome and Venice compared through Milan as a link city.

From transitivity requirement in (2), it is easy to see that elements of the matrix in (1) are inter-related. It is easy to prove that multilateral price comparisons, $P_{jk}\ (j,k=1,2,...,M)$, are transitive if and only if there exist *M* positive numbers $\{\Pi_1,\Pi_2,...,\Pi_M\}$ such that

$$P_{jk}=\frac{\Pi_k}{\Pi_j}\quad for\ all\ \ j,k=1,2,...,M \tag{3}$$

where Π_j is interpreted as price level in region *j*. In international comparisons these Π_j's are interpreted as purchasing power parities and are denoted by PPP_j. Equation (3) implies that *PPPs* can be determined only up to a factor of proportionality and, therefore, *PPPs* are expressed relative to a reference region. If region 1 is the reference region then $PPP_1=1$ so that all the other *PPPs* are interpreted relative to region 1.

It is useful to note that some of the commonly used price index numbers such as the Laspeyres, Paasche, Fisher and Tornqvist indexes do not satisfy transitivity. Therefore it has become necessary to devise aggregation methods that satisfy transitivity. Rao (2013b) and Diewert (2013) provide an overview of multilateral aggregation methods used in international price comparisons.

The requirement of *base invariance* simply requires that all the regions are treated symmetrically in making price comparisons. This means that transitive comparisons based on a star region whereby each region is compared with the designated star region and comparisons between regions are all linked through the star region are not desirable. In this case, the region designated as the star region is treated asymmetrically compared to the remaining regions.

2.3. Methods for aggregation to basic heading level

This is the first stage of aggregation of price data at the item level leading to price comparisons at the basic heading level. For example, prices of different types of *pasta products* are aggregated leading to regional price level comparisons for the commodity group, *pasta*. Let p_{ij} represent price of *i-th* item in region *j* $(i=1,2,...,N;\ j=1,2,...,M)$. Essentially there are *N* products in the basic heading. Note that not all items may be priced in all the regions.

Jevons index based methods

If all the items in the basic heading are priced in all the regions, then transitive price comparisons between regions are given by the Jevons index which is the simple unweighted geometric average of price ratios in the regions. The index is given by

$$P_{jk}^{Jevons} = \prod_{i=1}^{N} \left[\frac{p_{ik}}{p_{ij}} \right]^{1/N} \tag{4}$$

It is easy to see that the Jevons index in (4) satisfies transitivity and base-invariance properties expected of multilateral price comparisons.

A more common scenario is the case where not all commodities are priced in all the regions. As long as the price data exhibits connectivity, then the following approach in two steps can be used. In the first step, make price comparisons between two regions using prices on items that are priced in both regions. Let N_{jk} represents the set and number of commonly priced items in regions *j* and *k*. Then

$$P_{jk}^{Jevons} = \prod_{i \in N_{jk}} \left[\frac{p_{ik}}{p_{ij}} \right]^{1/N_{jk}} \tag{5}$$

The Jevons index in (5) is clearly not transitive since different pairs of regions have different levels of overlap of products priced. A set of transitive comparisons can be produced using the binary Jevons indexes in (5) as building blocks and the procedure suggested by Gini (1931), Elteto and Koves (1964) and Szulc (1964) referred to as the GEKS index.[2] The GEKS index applied to Jevons index in (5) leads to:

$$P_{jk}^{GEKS-Jevons} = \prod_{l=1}^{M} \left[P_{jl}^{Jevons} \cdot P_{lk}^{Jevons} \right]^{1/M} \tag{6}$$

Using simple multiplications it is easy to show that price comparisons derived using (6) are transitive and base invariant. This is the aggregation method used by

[2] The GEKS approach derives a transitive index that deviates the least from a given matrix of binary comparisons. If non-transitive binary indexes are represented by $P_{jk}\ (j,k = 1,2,...,M)$ then the GEKS transitive indices are obtained by minimizing $\sum_{j=1}^{M} \sum_{k=1}^{M} \left[\ln P_{jk}^{GEKS} - \ln P_{jk} \right]^2$ subject to $P_{jk}^{GEKS} = P_{jl}^{GEKS} \cdot P_{lk}^{GEKS} \quad for\ all\ \ j,k$ and l.

OECD and EUROSTAT in their international comparisons of prices.[3]

Region-product-dummy (RPD) Method

The region-product-method is the regional version of the country-product-dummy (CPD) method used in international comparisons. This is a regression based method proposed by Summers (1973)[4] for aggregation of price data at the basic heading level when expenditure weights are not available. Details of the CPD method and its properties are discussed in Rao (2004). The RPD method suggests that price levels are estimated by regressing logarithms of prices on region and product dummy variables and the model is given by:

$$\begin{aligned} \ln p_{ij} &= \ln PPP_j + \ln P_i + \ln u_{ij} \\ &= \pi_j + \eta_i + v_{ij} \\ &= \sum_{k=1}^{M} \pi_k D^k + \sum_{i=1}^{N} \eta_i D^i + v_{ij} \end{aligned} \tag{7}$$

where D^k is a region-dummy variable that takes value equal to 1 if the price observation is from *k-th* region; and D^i is a commodity-dummy variable that takes value equal to 1 if the price observation is for *i-th* commodity. The random disturbance is assumed to satisfy the standard assumptions of a multiple regression model. Parameters of this model can be estimated once one of the parameters of the model is set at a specified value. For example, if region 1 is taken as the reference or numeraire region then π_1 is set at zero and remaining parameters are estimated. If $\hat{\pi}_j \; (j = 1, 2, ..., M)$ are estimated parameters then required transitive price comparisons based on RPD method are given by:

$$P_{jk}^{RPD} = \frac{\exp(\hat{\pi}_k)}{\exp(\hat{\pi}_j)} \quad \textit{for all} \;\; j, k = 1, 2, ..., M \tag{8}$$

The RPD method based price comparisons in (8) are transitive and base-invariant. Further it is easy to show that when all the items are priced in all the regions then the RPD based price comparisons are identical to the Jevons-based comparisons in equation (4) (see Ferrari *et al.,* 1996 or Rao, 2004 for a proof of this statement).

[3] The GEKS-Jevons methods has certain limitations especially in the case where there are regions which have no commonly priced items, not a common occurrence but nonetheless important. See Rao (2013b) for further discussion on these aspects and other related issues.

[4] Summers proposed this as a method of imputing price data for missing observations.

2.4. Aggregation methods for aggregation above basic heading level

The next and final step in the compilation of spatial price comparisons is to aggregate the results from basic heading level comparisons to higher level aggregates such as food, consumption, investment and regional gross domestic product. A distinguishing feature of aggregation at this stage is that expenditure data are available making is possible to make use of standard price index number formulae.

Suppose there are N basic headings[5] and let $p_i^j \ and \ e_i^j$ represent price and expenditure for *i-th* basic heading in region *j*, respectively. Data available for making comparisons above the basic heading level can be represented by the following price and expenditure matrices.

$$P = \begin{bmatrix} p_1^1 & p_1^2 & p_1^j & p_1^M \\ p_2^1 & p_2^2 & p_2^j & p_2^M \\ p_i^1 & p_i^2 & p_i^j & p_i^M \\ p_N^1 & p_N^2 & p_N^j & p_N^M \end{bmatrix} \quad E = \begin{bmatrix} e_1^1 & e_1^2 & e_1^j & e_1^M \\ e_2^1 & e_2^2 & e_2^j & e_2^M \\ e_i^1 & e_i^2 & e_i^j & e_i^M \\ e_N^1 & e_N^2 & e_N^j & e_N^M \end{bmatrix} \tag{9}$$

If region 1 is used as the numeraire in making comparisons for all the basic headings, then $p_i^1 = 1 \ for \ i = 1, 2, ..., N$ for region 1 and for all other regions $p_i^j = P_{1j}^{GEKS-Jevons} \ or \ P_{1j}^{RPD}$ for *i-th* basic heading for all the remaining regions. Implicit quantities and expenditure shares are computed using:

$$q_i^j = \frac{e_i^j}{p_i^j} \quad and \quad s_i^j = \frac{e_i^j}{\sum_{i=1}^N e_i^j} = \frac{p_i^j \cdot q_i^j}{\sum_{i=1}^N p_i^j \cdot q_i^j}$$

Three index number formulae used in international and sub-national price comparisons are presented below.

<u>Fisher-based GEKS Index</u>

The Fisher price index is known to have a range of axiomatic and economic theoretic properties (Diewert, 1976 and Balk, 2008). The Fisher index, known as the Fisher Ideal index number, satisfies all the axiomatic properties expected of a price index number formula with the exception of *circularity* test[6]. Fisher index is also known to be exact and a superlative index number. The Fisher

[5] Note that N is used here to denote the number of basic headings even though it was used to denote the number of items within a basic heading. This is used to keep notation simple and there is no scope for confusion.

[6] Circularity test along with identity test implies transitivity.

index is the geometric average of the Laspeyres and Paasche index numbers and is given by:

$$P_{jk}^{Fisher} = \sqrt{P_{jk}^{Laspeyres} \cdot P_{jk}^{Paasche}} \tag{10}$$

where

$$P_{jk}^{Laspeyres} = \frac{\sum_{i=1}^{N} p_i^k q_i^j}{\sum_{i=1}^{N} p_i^j q_i^j} = \sum s_i^j \left(\frac{p_i^k}{p_i^j} \right),$$

$$P_{jk}^{Paasche} = \frac{\sum_{i=1}^{N} p_i^k q_i^k}{\sum_{i=1}^{N} p_i^j q_i^k} = \left[\sum_i s_i^k \left(\frac{p_i^k}{p_i^j} \right)^{-1} \right]^{-1}$$

As the Fisher binary index in (10) is not transitive, it is possible to use GEKS procedure to generate transitive multilateral price comparisons across different regions. The resulting index is given by:

$$P_{jk}^{GEKS-Fisher} = \prod_{l=1}^{M} \left[P_{jl}^{Fisher} \cdot P_{lk}^{Fisher} \right]^{1/M} \tag{11}$$

The GEKS-Fisher based formula in (11) is used in cross-country comparisons undertaken within the ICP at the World Bank (2015) and the OECD-EUROSTAT (2012) comparisons.

It is useful to conclude this description on GEKS-Fisher index with two comments. The GEKS method can be used on any other binary index that satisfies time/region reversal test. For example, if Fisher index is replaced by the Tornqvist index which is another exact and superlative index it is possible to obtain GEKS-Tornqvist index number formula[7]. The second point to note is that the GEKS approach has now been generalized and a weighted GEKS (WGEKS) has been proposed in Rao (2009) and properties of this method are discussed further in Hajargasht *et al*. (2017).

The Geary-Khamis Method

The Geary-Khamis (GK) method proposed by Geary (1958) and its properties studied by Khamis (1972) was the principal aggregation method used for cross-country price comparisons since the inception of the ICP in 1968. It is

---

[7] The Tornqvist index is given by $P_{jk}^{Tornqvist} = \prod_{i=1}^{N} \left[p_i^k / p_i^j \right]^{\frac{s_i^j + s_i^k}{2}}$.

only in the 2005 and 2011 ICP comparisons, that the GK method was replaced by Fished-based GEKS formula discussed above. The GK method is designed to provide measures of spatial price differences through PPPs and international or inter-regional average prices for each of the commodities. The GK system is defined through the following inter-related system of equations:

$$
\begin{aligned}
PPP_j &= \frac{\sum_{i=1}^{N} p_i^j q_i^j}{\sum_{i=1}^{N} P_i q_i^j} \qquad \forall j = 1, 2, \dots, M \\
P_i &= \frac{\sum_{j=1}^{M} \left(p_i^j \cdot q_i^j \right) \big/ PPP_j}{\sum_{j=1}^{M} q_i^j} \qquad \forall i = 1, 2, \dots, N
\end{aligned}
\tag{12}
$$

The GK system compares prices in each region *j* with average prices of the items obtained by averaging observed prices in different regions after adjusting for differences in the purchasing power or price levels in different region. PPP for each region *j* is obtained by using quantities purchased in the region at observed prices p_i^j .

The GK spatial price comparisons are obtained by solving the simultaneous system of (*M*+*N*) equations in as many unknown PPPs and Ps. The GK method has been a popular choice for cross-country comparisons given its additive consistency property which ensures that real consumption obtained by converting nominal consumption using PPPs would be the same as the value of quantities purchased at international or regional average prices. The GK method guarantees, for each country *j*, that:

$$
\frac{\sum_{i=1}^{N} p_i^j \cdot q_i^j}{PPP_j} = \sum_{i=1}^{N} P_i \cdot q_i^j
$$

The main criticism against GK method is that the PPPs are obtained using a Paasche-type index which does not allow for substitution when faced with average prices P_i *'s* and therefore likely to yield biased estimates of PPPs. This criticism has been addressed in the works of Rao (1976), Neary (2004), Feenstra, Ma and Rao (2009) and Feenstra *et al.* (2013). Despite the drawbacks associated with the GK method, it is still used when additively consistent comparisons are required. Aten (2006) and Aten and Figueroa (2015) use GK method for higher levels of aggregation.

Weighted Region-Product-Dummy (RPD) Method and Stochastic Approach

The RPD method described in equation (7) in section 2.3 is designed for the purpose of aggregating price data in the absence of expenditure data. Price level differences at item level are treated as of equal importance when they are no indicators of importance of the products. At the basic heading level, it is possible to accord weights to individual price data by running weighted least squares on the model:

$$\ln p_i^j = \sum_{k=1}^{M} \pi_k D^k + \sum_{i=1}^{N} \eta_i D^i + v_{ij}$$

with weights proportional to expenditure shares, s_i^j . Weighted RPD is equivalent to applying ordinary least squares to the following model:

$$\sqrt{s_i^j} \ln p_i^j = \sum_{k=1}^{M} \pi_k \sqrt{s_i^j} D^k + \sum_{i=1}^{N} \eta_i \sqrt{s_i^j} D^i + \sqrt{s_i^j} v_{ij} \qquad (13)$$

Properties of the weighted RPD method are discussed in Rao (2009). The use of RPD/CPD method is also referred to as the stochastic approach to multilateral comparisons. Hajargasht and Rao (2010) and Rao and Hajargashat (2016) discussed properties of the weighted CPD model and showed how most of the multilateral index numbers can be derived using weighted CPD model.

At sub-national level various methods have been used to calculate spatial price indexes, including the RPD model and GEKS method, as we shall see in the next section.

3. SUB-NATIONAL SPATIAL PRICE INDEXES: THE STATE OF THE ART

Though not as prominent as international comparisons of prices through ICP, there have been research projects and studies, conducted by NSOs and individual researchers, on the compilation of spatial price indexes in various countries including the United States, Brazil, India, Indonesia, China, Italy, Australia, New Zealand and the United Kingdom.

There are two principal approaches used in sub-national price comparisons. The first and foremost is the approach whereby data collected for the purpose of compiling CPI are used in constructing spatial price indexes. The second approach is to use unit-value prices derived from household expenditure data. In addition, spatial price comparisons undertaken by individual researchers rely on price data obtained from a variety of official published sources.

3.1. Sub-national PPPs based on CPI data

Early research on sub-national price comparisons was mostly conducted in the United States. Indeed, the first official measures of sub-national differences in cost of living were the Bureau of Labor Statistics (BLS) standard budgets, developed in the 1940s in response to a Congressional mandate to determine: "what costs a worker's family to live in the large cities of the United States". Under this program, basic living standards and income requirements were developed for a specified type of family living in urban areas of the country. These standards were then translated into hypothetical market baskets of goods and services. The rationale of using standard budget indexes as measures of price differences required that all urban families in the United States have approximately the same standards of living, even if it was not essential to collect prices for identical market baskets of goods and services in all locations (Sherwood, 1975).

In early 1990s a number of studies were carried out by the BLS and the Bureau of Economic Analysis (BEA) in order to determine whether CPI price data could be used in making inter-area comparisons across the United States (Primont and Kokoski, 1990; Kokoski, 1991). The country-product-dummy (CPD) method was employed at the item level in order to accommodate the diversity of items in the sample. Tornqvist binary indexes and the Gini-Elteto-Koves-Szulc (GEKS) method were then used to aggregate price relatives over items in order to obtain inter-area of price indexes for higher level commodity aggregates. Following this line of research, Kokoski *et al.* (1994) and Kokoski *et al.* (1999) further improved US inter-area price comparisons using the hedonic CPD model at item level by including quality characteristics of the products and the outlet information. Moulton (1995) constructed hedonic inter-area indexes of the cost of shelter using a new source of detailed microdata from the CPI Housing Survey.

These pioneering efforts at the BLS and BEA were later continued by Aten (2005, 2006) finally leading to the regular compilation of spatial price differences in the United States through the compilation of Regional Price Parities (RPP) for 38 metropolitan and urban areas of the U.S. (representing about 87% of the total population) for the years 2003 and 2004. This work was later expanded to cover the remaining non-metropolitan portions of each state (Aten, 2008). Since 2014, the RRPs and the price-adjusted estimates of regional personal income have become official statistics of the United States BEA and are now being published annually (Aten and Figueroa, 2015).

The Australian Bureau of Statistics (ABS) compiled and disseminated experimental indexes of cost of living in eight Australian capital cities using existing price data collected as a part of CPI compilation and calculated spatial price indexes using GEKS (Waschka *et al.,* 2003). In a similar vein, Statistics

New Zealand (SNZ) have been evaluating the possibility of carrying out spatial price comparisons of prices since 2005 and two experts have been assigned to develop a methodology for constructing spatial cost of living indexes. The expert report (Melser and Hill, 2005) deals with all the issues and provides advice on constructing sub-national PPPs thus providing reference text for implementing PPPs in other countries. However, to our knowledge, SNZ has never disseminated estimates of sub-national PPPs.

Turning to Europe, it is worth noting that only two countries, viz., the United Kingdom and Italy, carried out experimental estimation of sub-national price indexes involving their respective statistical offices and unofficial research studies have reported results for Germany and the Czech Republic. Roos (2006) reported on intra-country regional price level differences in Germany using a sample of price levels in 50 German cities in 1993. Kocourek *et al.* (2016) estimated regional price levels in districts of the Czech Republic using data from price surveys carried out by Czech Statistical Office and found significant regional differences in market prices of goods, services, as well as housing and rentals.

In 2000, the UK Office for National Statistics (ONS) carried out a one-off exercise to produce indicative figures concerning price level differences across regions. Subsequently, in 2004, the ONS published estimates of regional price level comparisons for the year 2003. These two exercises were carried out as a by-product to provide data for EUROSTAT PPP program (Fenwick and O'Donoghue, 2003). A year later, the ONS produced regional consumer price level estimates based on CPI price data mainly for food items, tobacco and drinks; supplemented with administrative data sources and more importantly a purpose-designed price level survey for items such as clothing, furniture, electrical goods and travel (Wingfield *et al.*, 2005). The ONS then produced Relative Regional Consumer Price Levels (RRCPLs) for 2010 following more suitable procedures (ONS, 2011) which involved data on price observations from the existing monthly CPI collection and regional price surveys which were required for computing Spatial Adjustment Factors for EUROSTAT-OECD PPP Program. The methodology employed by ONS for constructing RRCPLs is consistent with the approach used by EUROSTAT for calculating the PPPs (ONS, 2011).

The Italian Statistical Office (ISTAT) carried out two experimental studies for computing regional PPPs and disseminated the results on two occasions: in 2008 with reference to price data from 2006; and in 2010 with reference to 2009 data. In 2005, ISTAT in cooperation with Unioncamere and Istituto Tagliacarne, launched a research project aimed at testing the possibility of using and integrating the statistical information currently supplied by CPI surveys.

In order to calculate sub-national PPPs for consumer prices for 20 Italian cities (the regional chief towns), ISTAT used the same procedures as that used in the

ICP which followed the principle of strict comparability of products. Moreover, ad-hoc surveys were designed and carried out for the "Clothing and Footwear" and "Furniture" product groups. In the first experiment, sub-national PPPs, obtained using the GEKS formula, were computed for three expenditure dividions (Food and Beverages, Clothing and Footwear, Furniture), which represented approximately 34% of the total consumer expenditure (ISTAT, 2008). In the second experiment, sub-national PPPs were compiled for all the COICOP[8] expenditure divisions. The source of data and methodology remained the same with the exception of rents for which spatial comparisons were carried out using CPD models and Household Budget Survey data that includes some detailed information on the characteristics of the dwellings (ISTAT, 2010). A more comprehensive account of spatial price comparisons in Italy is provided in section 4.

3.2. Spatial price comparisons using unit-value prices

Since the early-2000s there have been studies carried out by researchers focusing on spatial price comparisons in large countries where significant differences in price levels and standards of living are observed in sub-regions within these countries. Aten (1999) and Aten and Menezes (2002) focused on spatial differences in food prices in 10 Brazilian cities. These measures are based on unit value prices obtained from household expenditure surveys. Besides reviewing practices for price collection, the Geary-Khamis, GEKS and Fisher methods were employed in compiling spatial price indexes. The initial analysis focusing on food prices was extended to include the price of services.

The use of unit-value price data from household expenditure data provided a framework for measuring spatial price comparisons in countries like India and Indonesia. Spatial price differences in India were measured and analysed by Coondoo *et al.* (2004) who estimated regional price index numbers from the *National Sample Survey (NSS)* data on household expenditure surveys conducted regularly in India. The method used, based on a specific demand system, is closely related to the CPD model. Covering 45 food items and data for 25 states/provinces in India, spatial price indexes were computed using variants of the CPD model and standard multilateral methods (see section 2 for a review of the index number methods used in making spatial price comparisons). Majumdar, Ray and Sinha have further refined the methodology for compiling RRPs for India, Indonesia and Vietnam, thus broadening and improving the analyses with the aim of proposing a unified framework for estimating intra and inter-country spatial price indexes (Mishra and Ray, 2014, Majumder *et al.*, 2015a, 2015b,

[8] The Classification of individual consumption by purpose, abbreviated as COICOP, is a classification developed by the United Nations Statistics Division to classify and analyze individual consumption expenditures incurred by households and non-profit institutions.

and Majumder *et al.*, 2016). Deaton and Dupriez (2011) also used unit values from household survey data to investigate spatial differences in prices for India and to obtain more reliable urban and rural price differentials to be used in official poverty estimates (Deaton and Tarozzi, 2005).

The use of unit value prices from household expenditure surveys has its limitations, the main problem is that meaningful unit values are available only for broadly classified food items. Spatial differences in food prices need to be augmented to include other consumption goods and services in order to provide a comprehensive measure of price level differences. In this respect, these attempts to measure spatial price level differences using unit values fall well short of the RRPs compiled by the BEA in USA.

3.3. Comparisons using prices from alternative sources

Several studies for computing sub-national price level differences have been carried out for the Chinese provinces with coverage extended beyond food prices as is the case with studies on India. The most important study was conducted by Brandt and Holz (2006), who estimated spatial price indexes for the base year 1990 using the cost basket method; and by Li *et al.* (2005) who carried out a study for computing sub-national PPPs in China for the year 2002. Shanghai was taken as the base province while average price data were obtained from 31 cities and districts within the provinces and referred to 127 goods and services. More recently, Biggeri *et al.* (2017) computed regional PPPs for the year 2014 using price data on a sample of 62 goods and services. Finally, Menggen *et al.* (2017) make use of official price data on 140 goods and services for 31 regions in China and compile spatial-temporal price comparisons covering the period 2000 to 2015.

Recently, the Asian Development Bank carried out a specific experimental research study on data from the Philippines (Dikhanov *et al.*, 2011) and compiled sub-regional PPPs by integrating data from the ICP in the Asia-Pacific region, of which Philippines is a part, and the consumer price index (CPI) data from the national statistical office of the Philippines.

4. SPATIAL PRICE COMPARISONS AND ISTAT PROJECTS

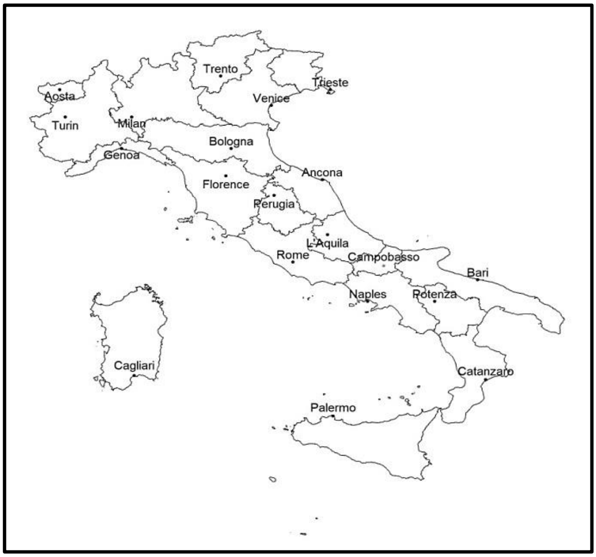
Spatial price indexes measuring differences in price levels across regions within a country are essential for comparing real income and standards of living. Several players of the economic and social debate recognize the need for compiling sub-national PPPs for Italy due to the high socio-economic heterogeneity across its geographical areas[9] (Figure 1).

[9] The second territorial level of the Nomenclature of territorial units for statistics (Nuts2) divides Italy into 19 regions and the two autonomous provinces of Trento and Bolzano, which make up the

Compared to other OECD countries, the Italian regions vary widely in terms of household economic conditions. Indeed, Italy has a dualistic economy with all Southern regions attaining a lower level of per capita income on average than the Centre-Northern regions. More specifically, in 2014 the autonomous province of Bolzano/Bozen showed the highest average annual household income (37,424 euros), which was over 15,000 euros higher than in Sicilia, the region with the lowest values (21,804 euros).

In 2015, the South and the Islands were the areas of the country with the highest level of poverty, affecting approximately one fifth of all resident households (compared to 10% in terms of national averages). In the same year, households in the Northern Italy spent more than households in the South and in the Islands. The highest expenditure was observed in the North-west (2,836.32 euros per month), approximately 1,000 euros more than the average expenditure in the Islands (1,891.78 euros per month).

Figure 1
Italian regional capitals



Source: Own elaboration.

Trentino-Alto Adige region. More specifically, Italy is subdivided into the following 20 regions, each of which has its own regional capital (specified in brackets): Aosta Valley (Aosta), Piedmont (Turin), Liguria (Genoa), Lombardy (Milan), Adige Trentino-Alto Adige (Trento), Veneto (Venice), Friuli-Venezia Giulia (Trieste), Emilia-Romagna (Bologna), Tuscany (Florence), Umbria (Perugia), Marche (Ancona), Lazio (Rome), Abruzzo (L'Aquila), Molise (Campobasso), Campania (Naples), Apulia (Bari), Basilicata (Potenza), Calabria (Catanzaro), Sicily (Palermo), Sardinia (Cagliari).

In order to investigate whether and to what extent consumer price levels differ across the Italian geographical areas, the ISTAT carried out official experimental sub-national PPP computations by using price data from CPIs and ad hoc surveys and focusing on comparing consumer prices across the 20 Italian regional capitals.

The results of the latest experiment carried out by ISTAT in 2009 showed significant differences in the level of consumer prices across the regional capitals (Istat, 2010). Consumer price levels in the Northern cities are generally higher than those in the Centre and especially in Southern Italy. Bolzano (105.6) and Milano (104.7) showed the highest prices compared with the Italian average (100) while the least expensive city proved to be Napoli (93.8).

The sub-national PPP results obtained from these analyses have encouraged ISTAT to go ahead with the project of regularly producing spatial indices of consumer prices at regional level. To this aim further research studies have been implemented by ISTAT in cooperation with the University of Florence and the University of Tuscia providing interesting results (Biggeri *et al.*, 2017; Laureti and Polidoro, 2017; Laureti *et al.*, 2017).

More specifically, within the European Multipurpose Price Statistics project, which is an important EUROSTAT project to which the NSOs are contributing, ISTAT is attempting to build up a database suitable for constructing sub-national PPPs as well as temporal comparisons of consumer prices.

Within a data warehouse approach which integrates various sources of data, the compilation of sub-national spatial price indices will be based on four main kinds of data: i) traditional CPI data (already representative and in some cases also comparable), ii) scanner data obtaining from retail trade chains of the modern distribution, iii) ad hoc collected data for certain groups of products by means of electronic devices and a dedicated software[10] data collected on the web also through web scraping techniques.

In order to use traditional CPI data an in-depth analysis of the basket and the micro data was carried out. This analysis was aimed at selecting all the products which are comparable by definition and therefore can be used for compiling spatial price indexes. Indeed, products whose prices are collected for compiling CPIs may not be comparable or representative across different areas due to differences in consumer behavior.

[10] Since the end of 2013, Istat has started the design and implementation of a new type of software (PPPJ) that will translate, in the field of the spatial comparison, the technical steps achieved in the field of temporal comparison (a software dedicated to manage all the statistical issues of the consumer price data collection, called P1J, was designed and developed for territorial data collection).

Within the frame of the research project Biggeri *et al.* (2017) focused on 7 Basic Headings (groups of products) belonging to the Food and non-alcoholic CPI group for carrying out experiments with the aim of understanding whether to use elementary price quotes or average prices obtained by averaging individual price quotes across outlets and investigating the performance of various CPD models estimated using data characterized by various levels of aggregation. Fresh meat, all fresh fish species, all types of fresh fruit and vegetables (Table 1) were selected, which accounts for 5.2 % of the entire basket.

Table 1

Product groups (BHs), number of products and monthly price quotes

BH	Description	Num. of products in CPI survey	Num. of price quotes
11.01.13.1	Fresh, chilled or frozen fish and seafood (Fish)	29	55,276
11.01.16.1	Fresh or chilled fruit (Fruit)	73	64,655
11.01.17.1	Fresh or chilled vegetables other than potatoes (Vegetables)	90	63,917
11.01.12.1	Beef and veal (Beef)	4	16,884
11.01.12.2	Pork	2	8,424
11.01.12.3	Lamb, mutton and goat	1	3,552
11.01.12.5	Other meats and edible offal	2	5,520
	Total	201	218,228

Source: Own elaboration.

Sub-national PPPs for the above mentioned BHs were estimated by using a dataset composed of 218,228 monthly price quotes collected in the 19 regional capitals considered in 2014 CPI survey. Although the varieties available in the various regional markets may vary due to distinct consumption behaviour, the products considered were characterized by a high degree of overlap across the Italian regional capitals. After having obtained annual individual price quotes for each product in the 7 BHs, hedonic CPD models, which include the characteristics of the outlet where products are sold, were estimated for compiling regional price indexes for Italy.

Table 2 reports the results obtained that show significant consumer price level differences across the various Italian regions and support the notion that price levels are higher in the Northern-Central regions than in the South, especially for the Fresh or chilled fruit BH and Fresh or chilled vegetables other than potatoes BH.

In addition, Biggeri *et al.* (2017) estimated CPD models with a spatial first-order autoregressive process for the error terms in order to explore the issue of spatial dependence in consumer prices. The spatial error CPD model, in which the spatial weights matrix W was based on an economic distance, produced point estimates of sub-national PPPs that were almost the same as those obtained using the CPD model but with lower standard errors.

As regards scanner data, it is worth noting the ISTAT project started in 2014 with the aim of carrying out a massive revision of the CPI production process in order to replace the on-field data collection for food and grocery products in supermarkets and hypermarkets with the prices obtained from this new data source.

Table 2

Estimates of sub-national PPPs for regional capitals using hedonic CPD models and CPI data, ROME=100

	Beef		Other meats		Pork		Lamb, mutton and goat		Fresh or chilled fruit		Fresh or chilled vegetables		Fresh, chilled or frozen fish and seafood	
Region		Sig.		Sig.		Sig.		Sig.		Sig.		Sig.		Sig.
North														
Aosta	106.11		109.82		111.10	**	95.59		125.17	***	125.53	***	93.67	*
Torino	112.88	**	105.91		100.09		104.73		102.35		98.80		100.67	
Genova	109.97	**	98.54		93.61	**	95.52		113.73	***	111.12	***	97.54	
Milano	87.17		105.49		96.28		97.03		150.59	***	155.02	***	125.19	***
Trento	105.44		106.09		87.65		100.35		125.16	***	119.73		92.43	**
Venezia	104.50		98.88		94.91		108.93		123.20	***	116.26		76.13	***
Trieste	102.51		123.31	*	98.14		116.89	*	121.54	***	121.44		91.07	***
Bologna	99.92		104.13		102.96		101.29		126.78	***	123.79	***	82.49	***
Centre														
Firenze	87.76	**	92.51		87.75		97.62		108.89	***	98.17		97.58	
Ancona	107.31	**	103.22		105.33		109.99	***	122.63	***	112.54	***	86.28	***
Perugia	101.62		87.28		95.88		110.45		111.09	***	101.72		111.15	***
South and Islands														
Aquila	100.66		113.95		100.14		98.37		92.12	***	87.29	***	85.20	***
Campobasso	92.47		96.25		95.82		95.55		93.97	***	89.17	***	85.32	***
Napoli	81.18	***	92.63		87.14	***	87.21	**	95.97	*	86.04	***	76.91	***
Potenza	77.92	***	86.32		87.45	**	88.34	***	99.33		96.57		74.40	***
Bari	89.13	**	81.61		96.38		101.77		89.20	***	84.40	***	74.65	***
Catanzaro	78.77	***	85.23	**	82.38	***	76.50	***	81.45	***	82.35	***	79.68	***
Palermo	87.10	**	90.38		85.08	***	80.07	***	101.98		102.49		104.09	
Cagliari	86.14	***	87.51		86.89	***	74.54	**	104.23	*	98.29		91.48	**
Obs.	177		74		89		42		1673		2018		888	
Root MSE	0.176		0.147		0.148		0.105		0.181		0.246		0.194	
AIC	-89.702		-56.262		-69.182		-57.081		-879.982		167.933		-340.517	

Source: Own elaboration.

According to an agreement with the Association of Modern Distribution, ISTAT can obtain scanner data through the market research company ACNielsen, which provide data on a monthly basis by uploading the data files on a dedicated ISTAT web portal. At present, ISTAT has received data regarding 1,470 stores belonging to 16 retail chains (Coop Italia, Conad, Selex, Esselunga, Auchan, Carrefour, Finiper, PAM, etc.), that cover approximately 93.7% of the turnover of total modern distribution and 37 Italian provinces.

Each record represents the weekly sum of turnover and quantity for a GTIN (Global Trade Item Number, formerly EAN code) sold during the week in a single store. Also provided are classification tables for mapping GTINs to ECOICOP classification and specific elementary aggregates. Moreover, additional information are provided regarding stores and GTINs, such as product descriptions, geographic location, etc. Therefore, these data can be effectively used for compiling sub-national PPPs for food and grocery groups of products since they enable to compare representative and comparable items across regions.

Laureti and Polidoro (2017) explored this new source of price data in order to compute Italian sub-national BH level. After a process of data cleaning and trimming outliers, unit value price per item code were computed by dividing the total turnover for that item by the total quantities sold over the week. Several analyses were carried out in order to understand how best to aggregate the detailed information contained in the Italian scanner data for constructing spatial price indexes. More specifically, by referring to each Italian regional capitals, ANOVA and t tests on a sample of items were carried out in order to verify if the price of the same item could reflect auxiliary services provided by the seller. Indeed, within each city, the same item is found in different supermarket chains and in different stores which belong to the same retail chain. Results showed significant differences in prices of the same items thus suggesting product differentiation which is embodied in the range or quality of services offered by different retailers across chains.

Therefore, with respect to item groupings, the authors used the finest classification of item that is available within the BH, i.e. the product code, which is identical across the Italian territory. As regards the time dimension, annual regional average prices for each GTIN code and outlet were used. In order to compile spatial price index for the 20 regional capitals at BH level, 931 points of sale of the six most important modern distribution chains (Coop Italia, Conad, Selex, Esselunga, Auchan, Carrefour), that cover approximately 57% of the turnover of total modern distribution were considered. The dataset used consists in 3,659,286 annual price quotes from the 20 regional chief towns concerning 69 BHs for a total of 49,489 products (GTIN code).

In order to account for the economic importance of each item in its market, which is essential in index number literature, Laureti and Polidoro (2017) estimated weighted CPD models using expenditure share as weights. In this way it may be possible to mitigate the effects of the large fluctuations in quantities purchased in response to price discounts which still emerge using monthly average prices. Moreover, in order to explore the effects of including information on the type of outlet and retail chain hedonic weighted CPD models were estimated.

Table 3 reports sub-national PPP results for the “Pasta products and Couscous” products. Significant differences can be observed between the results obtained from the unweighted CPD and weighted CPD models. The hedonic WCPD models confirmed large differences in price levels among regional chief towns, with higher prices observed for most (but not all) of the Northern regions than in Rome. These results appear to be coherent with the territorial characteristics of the Italian macro areas and with previous findings.

Table 3

Estimates of sub-national PPPs for regional capitals using CPD models and scanner data, ROME=100 (Pasta products and Couscous BH)

	UNWEIGHTED CPD			WEIGHTED CPD weights=expenditure share			HEDONIC WEIGHTED CPD weights=expenditure share		
	Coeff	**PPPs**	**Sig**	**Coeff**	**PPPs**	**Sig**	**Coeff**	**PPPs**	**Sig**
North									
Aosta	0.0123	101.23	***	0.0295	103.00	***	0.0256	102.59	***
Torino	-0.0208	97.94	***	-0.0170	98.31	***	-0.0074	99.26	***
Genova	0.0034	100.34	**	0.0017	100.17		0.0206	102.08	***
Milano	-0.0172	98.30	***	-0.0068	99.32	***	0.0195	101.97	***
Trento	0.0134	101.35	***	-0.0190	98.12	***	0.0078	100.78	***
Venezia	-0.0343	96.63	***	-0.0384	96.23	***	-0.0188	98.14	***
Trieste	-0.0089	99.11	***	0.0145	101.46	***	0.0424	104.33	***
Bologna	-0.0145	98.56	***	-0.0084	99.17	***	0.0143	101.44	***
Centre									
Firenze	-0.1099	89.60	***	-0.0997	90.51	***	-0.0770	92.58	***
Ancona	0.0006	100.06		-0.0001	99.99		0.0183	101.85	***
Perugia	-0.0284	97.20	***	-0.0272	97.32	***	-0.0131	98.70	***
South and Islands									
L'Aquila	0.0053	100.53	***	0.0059	100.59	***	0.0101	101.02	***
Campobasso	-0.0292	97.12	***	-0.0453	95.57	***	-0.0201	98.01	***
Napoli	-0.0169	98.32	***	-0.0178	98.24	***	-0.0121	98.80	***
Potenza	-0.0376	96.31	***	-0.0556	94.59	***	-0.0405	96.03	***
Bari	-0.0757	92.71	***	-0.0884	91.53	***	-0.0641	93.79	***
Catanzaro	-0.0428	95.81	***	-0.0605	94.13	***	-0.0390	96.18	***
Palermo	-0.0239	97.64	***	-0.0576	94.40	***	-0.0404	96.04	***
Cagliari	-0.0079	99.21	***	-0.0030	99.70	*	0.0114	101.15	***
Chain2							-0.0015		
Chain3							-0.0192		***
Chain4							0.0111		***
Chain5							0.0590		***
Chain6							-0.0020		*
Outlet (ref.hypermarket)							0.0281		***
Obs.	188,830			188,830			188,830		
Root MSE	0.10137			0.09807			0.09515		
AIC	-325923			-338393			-349836		

Source: Own elaboration.

Within the ISTAT project, Laureti *et al.* (2017) carried out interesting experiments using both scanner and CPI data. More specifically, they focused on the issue of compiling regional spatial price indexes for perishables and seasonal products such as vegetables, fruit and meat since these products are sold at price per quantity and are not pre-packaged with GTIN codes. Therefore scanner data cannot be used for these products and traditional CPI data obtained from the large-scale retail trade should be considered. The authors integrated scanner data with prices for fruit and vegetables which are traditionally collected for CPI production in modern distribution. After data quality controls and preliminary analyses of the basket, the CPI dataset used includes annual average prices for 151 vegetable products collected in the 20 regional capitals and include information on chains. Sub-national PPPs were estimated by using unweithed CPD models since CPI data do not include information on expenditure.

5. FUTURE DEVELOPMENTS REGARDING SUB-NATIONAL PPPS: NEW SOURCE OF DATA AND METHODS

The recent availability of high-frequency ‘‘scanner data’’ in addition to other sources of data enables price statisticians to deal with the sub-national PPP issue from a renewed approach. Scanner data can be defined as: «*transaction data obtained from retail chains containing data on turnover, quantities per item code based on transactions for a given period and from which unit value prices can be derived at item code level*».

These data benefit from an impressive coverage of transactions along with the availability of information on sales, prices, quantities and quality characteristics of products sold (brand, size and type of outlet) provided at barcode level or, more precisely, the GTIN code. Currently, scanner data predominantly replace price collection in supermarkets and hypermarkets, especially for food, beverages and personal and home care products.

As mentioned above, however, scanner datasets provide both opportunities and challenges for price statisticians since they must deal with huge amounts of highly detailed data on consumer purchases with high variability of products sold among cities.

Indeed, scanner data are characterized by a high attrition rate of goods and volatility of prices and quantities due to sales. This new source of data is able to capture frequent, and often large, shifts in quantities purchased in response to price changes. Moreover, using highly detailed data on consumer purchases implies that it is essential to deal with the issue of aggregation of individual items from both a theoretical and a practical perspective when computing price indexes.

Over the last decade numerous studies have analysed the impact of different aggregation methods on inflation estimates using scanner data as various NSOs are interested in using these data in their official price statistics (see for example

Ivancic *et al.*, 2011). However, to our knowledge, few studies have used scanner data and carried out experiments on aggregation issues when comparing consumer prices across space (Heravi, *et al.*, 2003; Laureti and Polidoro, 2017).

Therefore, it is essential to determine how best to use scanner data and how to combine them with other data from various sources in order to construct sub-national PPPs. (i) The first aspect to consider is that scanner data may help to overcome the issue of price data availability in the various areas involved in the comparisons that fulfill the requirements of representativeness and comparability that emerge when compiling regional spatial consumer price indices. Due to the high territorial coverage which characterizes scanner data, we are able to compare price levels among the various cities within a country. In addition, it is worth noting that GTIN codes describe the products in detail and they are generally the same for the each item at national level. In this way, we can solve the issue of comparability. (ii) Since detailed information on turnover and quantities for each item code in every city are available, it is possible to account for the economic importance of each item in its market, thus fullfilling the representativeness requirement. (iii) Moreover, as different modern retail chains can sell products of different quality and offer additional services information on the type of outlet and retail chain can be included in order to account for these "quality" characteristics that may influence the price of a product. (iv) Finally, using scanner data to carry out spatial comparisons will increase cost efficiency since price data collection may be limited to traditional outlets thus lowering data collection costs for the NSIs.

From a methodological point of view, when we use scanner data, we can refer to a wide range of methods for calculating spatial price indexes due to the availability of quantities and expenditure shares (Heravi *et al.*, 2003). In this way, we can improve the quality of sub-national PPPs, since we can also use the Geary-Khamis (GK) method and weighted CPD methods at BH level. In addition, scanner data can provide a basis for further statistical developments in the estimation of sub-national PPPs. Indeed, detailed price and quantity data are available on a monthly basis thus allowing us to add the time dimension to multilateral price comparisons.

A relatively new area of research within sub-national PPPs is how to address the issue of spatial dependence which is inherent in consumer price levels (Aten, 1996; Rao, 2004). Several researchers have found that consumer prices are more similar in geographically proximate locations, thus observing a significant positive correlation between the law of one price deviations and distance. This spatial effect may reflect transport costs as well as local distribution costs, which are likely to be similar in nearby locations if the distribution of goods is labour intensive and labour markets are geographically integrated. However, in spite of its theoretical attraction, as yet few studies have been carried out on spatial dependence in spatial price index constructions and the empirical evidence of spatial correlation has only been analysed at aggregate national-level by Aten (1996) and Rao (2009). Focusing on methodological issues, Rao (2004) drew attention to the effects of spatial autocorrelation and the adjustments required for the estimates. Biggeri *et al.* (2017) were the first authors to explore this issue at sub-national level by estimating a CPD model with a spatial first-order autoregressive process for the error terms, in which the spatial weights matrix **W** was based on an economic distance.

The development of spatial econometric models for estimating sub-national PPPs may be encouraged by the availability of scanner data due to their high territorial coverage and the detailed information on outlet location they provide.

6. CONCLUSIONS

The main objective of the paper is to provide an assessment of the current status of spatial price comparisons across regions within a country. Relative to cross-country comparisons of prices and real incomes within the International Comparison Program, work on compilation of sub-national PPPs is in a state of infancy. Though a lot of progress has been made in compiling regional price parities within the USA using data drawn from price data collected for CPI, work on sub-national price comparisons is gathering some momentum in other countries. The paper presented an overview of work on sub-national PPPs in China, India, Italy and other countries. Detailed descriptions of the index number methods used in spatial price comparisons are provided for the benefit of the readers. Empirical results from recent work in Italy on sub-national PPPs are reported. The challenging and exciting prospects offered through the availability of high-frequency scanner data with price and quantity data on each transaction at the point-of-sale are discussed. A case is made for the use of stochastic approach that allows for explicit modelling of spatial structures in prices. The authors hope that this paper serves as a useful review on spatial price comparisons across regions within countries.

BIBLIOGRAPHY REFERENCES

ATEN, B. H. (1996). "Evidence of Spatial Autocorrelation in International Prices". *Review of Income and Wealth*, June, 42(2), pp. 149-63.

ATEN, B. H. (1999). "Cities in Brazil: An inter-area price comparison". In International and inter-area comparisons of income, output, and prices (pp. 211-229). Chicago: University of Chicago Press.

ATEN, B. H. (2005). "Report on inter-area price levels". *Working paper no*. 2005-11. Bureau of Economic Analysis, May.

ATEN, B. H. (2006). “Inter-area price levels: An experimental methodology”. *Monthly Labor Review*, 129(9), pp.47-61.

ATEN, B. H. (2008). “Estimates of state and metropolitan price parities for consumption goods and services in the United States, 2005”. *Bureau of Economic Analysis*, April. http://bea.gov/papers.

ATEN, B. H., and FIGUEROA, E. B., (2015). “Real personal income and regional price parities for state and metropolitan areas, 2009–2013”. *Survey of Current Business*, 95(7), pp. 1-4.

ATEN, B. and T. MENEZES (2002), “Poverty Price Levels: An Application to Brazilian Metropolitan Areas”- Paper presented at CONFERENCE ON THE INTERNATIONAL COMPARISON PROGRAM WASHINGTON, D.C., MARCH 11 - 15, 2002

BALK, Bert M. (2008). “Price and Quantity Index Numbers." Cambridge Books.

BIGGERI, L., FERRARI, G., and ZHAO, Y. (2017). “Estimating cross province and municipal city price level differences in China: Some experiments and results”. *Social Indicators Research*, 131(1), pp.169-187.

BIGGERI, L., LAURETI, T., and POLIDORO, F., (2017). “Computing sub-national PPPs with CPI data: an empirical analysis on Italian data using country product dummy models”. *Social Indicators Research*, 131(1), pp. 93-121.

BRANDT, L., and HOLZ, C. A. (2006). “Spatial price differences in China: Estimates and implications”. *Economic Development and Cultural Change*, 55(1), pp. 43-86.

COONDOO, D., MAJUMDER, A., and RAY, R. (2004). “A method of calculating regional consumer price differentials with illustrative evidence from India”. *Review of Income and Wealth*, 50(1), pp. 51-68.

COONDOO, D., A. MAJUMDER, and S. CHATTOPADHYAY, (2011). “Estimating Spatial Consumer Price Indices Through Engel Curve Analysis”. *Review of Income and Wealth*, 57, 138-155.

DEATON, A., and DUPRIEZ, O. (2011). “Spatial price differences within large countries”. *Manuscript*, Princeton University, July.

DEATON, A. and TAROZZI, A. (2005). “Prices and poverty in India”. Chapter 16 in Angus Deaton and Valerie Kozel, (eds.): *The great Indian poverty debate* (pp. 381-411). New Delhi: Macmillan.

DIEWERT, W. E. (1976). “Exact and superlative index numbers”. *Journal of econometrics*, 4(2), pp.115-145.

DIEWERT, W.E. (2013). “Methods of Aggregation above the Basic Heading Level within Regions”. Chapter 5 in World Bank (ed.): *Measuring the Real Size of the World Economy* (pp. 121-168). Washington DC: World Bank.

DIKHANOV, Y., PALANYANDY, C., and CAPILIT, E. (2011). “Sub-national purchasing power parities toward integration of international comparison program and the consumer price index: The case of Philippines”. *ADB Economics Working Paper Series*, No. 290, Mandaluyong City: Asian Development Bank.

ECE-ILO (2010). *Consumer price index manual: Theory and practice.* Geneva.

ELTETÖ, O. and P. KÖVES, (1964). “On a Problem of Index Number Computation Relating to International Comparison”. *Statisztikai Szemle*, 42, pp.507-18.

FEENSTRA, R. C., HONG MA, J. PETER NEARY AND D.S. PRASADA RAO, (2013). “Who Shrunk China? Puzzles in Measurement of Real GDP”. *Economic Journal*, Vol. 123, Issue 573, pp. 1100-1129.

FEENSTRA, R. C., MA, H., and RAO, D. S. PRASADA. (2009). “Consistent comparisons of real incomes across time and space”. *Macroeconomic Dynamics*, 13(S2), pp.169-193.

FENWICK, D., and O’DONOGHUE, J. (2003). “Developing estimates of relative regional consumer price levels”. *Economic Trends*, 599, pp.72-83.

FERRARI, G., G. GOZZI AND M. RIANI (1996). "Comparing GEKS and EPD Approaches for Calculating PPPs at the Basic Heading level". In *Improving the Quality of Price Indices: CPI and PPP*, EUROSTAT.

GEARY, Roy C., (1958). “A Note on the Comparison of Exchange Rates and Purchasing Power between Countries”. *Journal of the Royal Statistical Society, Series A (General),* 121(1), pp.97-99.

GINI, C. (1931). “On the circular test of index numbers”. *Metron*, 9(9), pp. 3-24.

HAJARGASHT, G. and RAO D.S. PRASADA, (2010). “Stochastic Approach to Index Numbers for Multilateral Price Comparisons and their Standard Errors”. *Review of Income and Wealth*, Series 56, *Special Issue 1*, June, 2010, pp. S32-S58.

HAJARGASHT, G., HILL, R. SHANKAR S. AND RAO D.S. PRASADA, (2017). “Spatial Chaining in International Comparisons of Prices and Real Incomes”. Mimeographed.

HERAVI, S., HESTON, A., and SILVER, M. (2003). “Using scanner data to estimate country price parities: A hedonic regression approach”. *Review of Income and Wealth*, 49(1), pp.1-21.

ISTAT (2008). “Le differenze nel livello dei prezzi tra i capoluoghi delle regioni italiane per alcune tipologie di beni Anno 2006”. Roma.

ISTAT (2010). “La differenza nel livello dei prezzi al consumo tra i capoluoghi delle regioni italiane, Anno 2009”. Roma.

IVANCIC, L., DIEWERT, W. E., and FOX, K. J. (2011). “Scanner data, time aggregation and the construction of price indexes”. *Journal of Econometrics*, 161(1), pp.24-35.

KHAMIS, SALEM H., (1972). “A New System of Index Numbers for National and International Purposes”. *Journal of the Royal Statistical Society, Series A (General)*, 135(1), pp.96-121.

KOCOUREK, A., ŠIMANOVÁ, J., and ŠMÍDA, J. (2016). “Estimation of Regional Price Levels in the Districts of the Czech Republic”. *Statistika: Statistics and Economy Journal*, 96(4), pp. 56-70.

KOKOSKI M. (1991). “New Research on Interarea Consumer Price Differences”. *Monthly Labor Review*, Vol. 114, 31.

KOKOSKI, M., CARDIFF, P., and MOULTON, B. (1994). “Interarea price indices for consumer goods and services: an hedonic approach using CPI data.”. *US Department of Labor, Bureau of Labor Statistics, Working Paper,* 256.

KOKOSKI, M. MOULTON B, and ZIESCHANG K. (1999). “Interarea Price Comparisons for Heterogenous Goods and Several Levels of Commodity Aggregation”. In *International and Interarea Comparisons of Income, Output and Prices*, ed. by Alan Heston and Robert Lipsey, 123-66. University of Chicago Press.

LAURETI T. and POLIDORO F. (2017). “Testing the use of scanner data for computing sub-national Purchasing Power Parities in Italy”. *Proceeding of 61st ISI World Statistics Congress,* Marrakech.

LAURETI T, FERRANTE C. and DRAMIS B. (2017). “Using scanner and CPI data to estimate Italian sub-national PPPs”. *Proceeding of 49th Scientific Meeting of the Italian Statistical Society,* Florence, pp.581-588.

LI, X., ZHANG, L., and DU, Y. (2005). “Study on the Method of Regional Purchasing Power Parity in China”. *China-USA Business Review*, 4(7), pp.41-48.

MAJUMDER, A., R. RAY, AND K. SINHA, (2012). “The Calculation of Rural Urban Food Price Differentials from Unit Values in Household Expenditure Surveys: A New Procedure and Comparison with Existing Methods”. *American Journal of Agricultural Economics*, 94, pp. 1218-1235.

MAJUMDER, A., RAY, R., and SINHA, K. (2015a). “Estimating purchasing power parities from household expenditure data using complete demand systems with application to living standards comparison: India and Vietnam”. *Review of income and wealth*, 61(2), pp. 302-328.

MAJUMDER, A., RAY, R., and SINHA, K. (2015b). “Spatial comparisons of prices and expenditure in a heterogeneous country: Methodology with application to India”. *Macroeconomic Dynamics*, 19(5), pp. 931-989.

MAJUMDER, A., MAJUMDER, A., RAY, R., RAY, R., SINHA, K., and SINHA, K. (2016). “A unified framework for the estimation of intra and inter country food purchasing power parities: India, Indonesia and Vietnam”. *Indian Growth and Development Review*, 9(1), pp. 2-31.

MELSER, D., and HILL, R. (2005). “Developing a methodology for constructing spatial cost of living indexes”. Sidney, Australia: Economic Department, University of New South Wales.

MENGGEN, C, Y. WANG and D.S. PRASADA RAO (2017). “Measuring the Spatial Price Difference in China: Regional Price Parities 2000-2014”. Mimeographed, Beijing Normal University, Beijing.

MISHRA, A., & RAY, R. (2014). “Spatial variation in prices and expenditure inequalities in Australia”. *Economic Record*, 90(289), pp. 137-159.

MOULTON, B. R. (1995). “Interarea indexes of the cost of shelter using hedonic quality adjustment techniques”. *Journal of Econometrics*, 68(1), pp. 181-204.

NEARY, J. P., (2004). “Rationalizing the Penn World Table: True Multilateral Indices for International Comparisons of Real Income”. *American Economic Review*, 94(5), pp. 1411-28.

ONS. (2011). “UK relative regional consumer price levels for goods and services for 2010”, UK, pp. 1-15.

OECD (2012). *EUROSTAT-OECD Methodological Manual on Purchasing Power Parities*. OECD Publishing.

PRIMONT, D. F., and KOKOSKI, M. F. (1990). “Comparing prices across cities: A hedonic approach.” *US Bureau of Labor Statistics*.

RAO, D. S. PRASADA. (2004). “The Country-Product-Dummy Method: A Stochastic Approach to the Computation of Purchasing Power Parities”. In the ICP, Paper presented at the SSHRC Conference on Index Numbers and Productivity Measurement, Vancouver, June 30-July 3, 2004.

RAO D.S. PRASADA, (1976). "Existence and Uniqueness of a System of Consistent Index Numbers". *Economic Review*, 27, pp. 212-218.

RAO D.S. PRASADA, (2009). "Generalised Elteto-koves-Szulc (EKS) and Country-Product-Dummy (CPD) Methods for International Comparisons". In Prasada Rao (ed.): *Purchasing Power Parities: Recent Advances In Methods And Applications* (pp. 86-120). Edward Elgar Publishing Company.

RAO, D.S. PRASADA, (2013a). "The Framework for the International Comparison Program (ICP)". Chapter 1 in World Bank (ed): *Measuring the Real Size of the World Economy* (pp.13-45). Washington DC: World Bank.

RAO, D.S. PRASADA, (2013b). "Computation of Basic Heading Purchasing Power Parities (PPPs) for Comparisons within and between Regions". Chapter 4 in World Bank (ed.): *Measuring the Real Size of the World Economy* (pp. 93-119). Washington DC: World Bank.

RAO, D.S. PRASADA, and HAJARGASHT, G. (2016). "Stochastic approach to computation of purchasing power parities in the International Comparison Program (ICP)". *Journal of Econometrics*, 191(2), pp. 414-425.

ROOS, M. (2006). "Regional price levels in Germany". *Applied Economics*, 38(13), pp.1553-1566.

SHERWOOD, M. K. (1975). "Family budgets and geographic differences in price levels". *Monthly Lab. Rev.*, 98, pp.8-15.

SZULC, B. J., (1964). "Indices for Multiregional Comparisons". *Przeglad Statystyczny 3, Statistical Review,* 3, pp. 239-54.

SUMMERS, R. (1973). "International Price Comparisons Based upon Incomplete Data". *Review of Income and Wealth*, 19(1), pp.1-16.

WASCHKA, A., MILNE, W., KHOO, J., QUIREY, T., and ZHAO, S. (2003). "Comparing living costs in Australian capital cities. A progress report on developing experimental spatial price indexes for Australia". Australian Bureau of Statistics.

WINGFIELD, D., FENWICK, D., and SMITH, K. (2005). "Relative regional consumer price levels in 2004". *Economic Trends (Office for National Statistics, UK)*, 615, pp. 36-46.

WORLD BANK (2013). "Measuring the Real Size of the World Economy: the Framework, Methodology and Results of the International Comparison Program-ICP". Washington.

WORLD BANK (2015). "Purchasing Power Parities and the Real Size of the World Economies - A Comprehensive Report of the 2011 International Comparison Program". Washington, DC: World Bank.