Journal of Official Statistics, Vol. 38, No. 1, 2022, pp. 1–28, http://dx.doi.org/10.2478/JOS-2022-xxxx

Sub-National Spatial Price Indexes for Housing: Methodological Issues and Computation for Italy

Tiziana Laureti[1], *Ilaria Benedetti*[1], *and Luigi Biggeri*[2]

It is essential to measure within-country differences in housing costs in order to evaluate costs of living, assessing and comparing poverty levels, quantifying salaries and disposable income of families and finally for designing housing policies at local level. To the authors knowledge, no studies have yet been carried out on the computation of Space Price Indexes for Housing Rents (SPIHRs). In this paper we computed preliminary estimates of sub-national SPIHRs by using hedonic regression model, which is an extension of the Country Product Dummy method, for all the Italian regions. The hedonic regression is generally used to obtain multilateral spatial indexes, thus allowing us to obtain multilateral SPIHRs for the Italian regions. The estimates have been done using 2017 data from the Real Estate Market Observatory which is a part of the Italian Agency of Revenue and Tax. This data source is the most comprehensive source of information on Italian houses price rents with a wide geographical coverage, including data for each Italian municipality. The obtained results show significant differences across the Italian regions, thus highlighting the importance of calculating SPIHR in Italy on a regular basis and the need to continue researches in this field.

Key words: Spatial rental price indexes; hedonic regression; country product dummy method; symbolic data analysis.

1. Introduction

It is essential to measure within-country differences in housing costs in order to evaluate costs of living, assessing and comparing poverty levels, quantifying salaries and disposable income of families and finally for designing housing policies at a local level (Bishop et al. 2017).

It is well known that when making comparisons, the sub-national (regional) values of the economic indicators should be adjusted for sub-national price differences in order to avoid misleading regional analyses and the consequent policy implications and outcomes, which is done by computing sub-national Spatial Price Indexes (SPIs) as illustrated by Biggeri et al. (2017) and Laureti and Rao (2018).

In this context, the computation of sub-national SPIs for housing costs is particularly important since housing costs can be a substantial financial burden to households, especially for low-income families.

These sub-national SPIs have proven to be useful for assessing and comparing poverty levels and for implementing housing policies for the poor at local level, however they can also be used as proxies of general sub-national household consumption SPIs or in

[1] University of Tuscia, Via del Paradiso 47, 01100, Viterbo (VT), Italy. Email: laureti@unitus.it and i.benedetti@unitus.it
[2] The University of Florence, Viale Morgagni, 59, 50134 Firenze, 50134, Italy. Email: Luigi.biggeri@unifi.it

combination with them, for adjusting poverty thresholds (Bishaw 2009; Bishop et al. 2017; Biggeri and Pratesi 2017).

To the authors' knowledge, official estimates of sub-national Spatial Price Indexes for Housing Rent (SPIHRs), useful for the estimation of the poverty indicators in real terms, are mainly provided in the United States (US).

During the 1970's and 1980's several researchers worked on the computation of SPIs for housing costs which, however, were based on limited data and could not be used for countrywide poverty measures.

In the early 1990's, the National Academy of Science (NAS) Panel on Poverty and Family Assistance proposed that, as "a first and partial step", poverty thresholds must be adjusted for differences in the costs of housing across geographical areas, taking into account that housing is the most expensive cost of the household budget especially for poor households, and further research is required in order to develop more refined methods and data (Citro and Michael 1995).

The U.S. Bureau of Census computed a housing cost index using data from the decennial census and the method used by the U.S. Department of Housing and Urban Development for computing Fair Market Rents using data from a specific type of houses. After this first step, the NAS panel suggested improving the method and using better data in order to gain a better understanding of poverty levels which are known as Q1 Supplementary Poverty Measures (Jolliffe 2006; Renwick 2009; Bishop et al. 2017)[1]

The turning point in the computation of spatial housing cost differences occurred in 2007 when the American Community Survey (ACS) was conducted by U.S. Census Bureau. The ACS provides detailed annual data on housing costs and a set of interarea price indexes compiled by the researchers of the U.S. Bureau of Economic Analysis (BEA) and by the U.S. Bureau of Labor Statistics (BLS).

In the beginning, interarea price indexes were used to develop an index based on 2007 Q2 gross rental costs considering all rental units (Bishaw 2009)[2]. Finally, after numerous experimental estimates had been conducted in collaboration with other US federal agencies (Martin et al. 2013), the BEA computed Regional Price Parities (RPPs) which combine the CPI price data with rental data. Both spatial indexes regarding the prices of goods and services and housing rents, were elaborated using hedonic regression models and following the methods suggested by the International Comparison Program (ICP) in Q3 order to obtain multilateral indexes[3].

Since 2014, the RPPs and the price-adjusted estimates of regional personal income have become official statistics (Aten and Figueroa 2015).

[1] The differences between the Official and Supplemental Poverty Measures are summarized in a paper published by Dalaker (2017).

[2] A similar computation was developed by Biggeri and Pratesi (2017) to estimate a spatial index for housing cost for the Italian regions, by using the median monthly expenses of the households to rent an apartment using data on the Household Budget Survey (HBS). The use of these indexes reshuffle in some way the territorial distribution of the income and poverty indicators.

[3] International spatial price comparisons for measuring differences in price levels and Purchasing Power Parities of currencies (PPP) are compiled by the ICP, administered by the World Bank and in collaboration with OECD, Eurostat and other organizations (World Bank 2013) by allowing cross-country comparisons of Gross Domestic Product (GDP) and its major aggregates.

Italian researchers have recently conducted a series of experiments for computing sub-national SPIs (Laureti and Rao 2018; Laureti and Polidoro 2018; Biggeri and Laureti 2018; Ferrante et al. 2019).

However, no complete studies have yet been carried out on the computation of SPIHRs, even if Italian households pay housing costs equal to approximately 35% of their total expenditures (Istat 2017) and significant differences in average housing costs were observed across the various Italian regions as highlighted in the report published by Istat in Q4 2010[4].

In order to fill this gap, we believe that it is important to compute preliminary estimates of sub-national SPIHRs by using hedonic regression models for all the Italian regions.

To this aim, we used data from OMI (Osservatorio del Mercato Immobiliare; Real Q5 Estate Market Observatory) which is a part of the Italian Agency of Revenue and Tax[5]. The OMI database is a comprehensive source of information on house prices and rents in Italy with a wide geographical coverage, which provides quotations on a semi-annual basis regarding minimum – maximum market prices (sale or rental) per unit area (square meters), and information on the location and the characteristics of the neighbourhood of the dwellings, which are considered the most important determinants of housing prices and rents.

The rest of the article is organized as follows: Section 2 reviews the literature on using hedonic regression models for temporal and spatial price index construction. Section 3 focuses the main characteristics of the hedonic approach for estimating spatial price indexes for housing rents. Section 4 describes the characteristics of the OMI database used. Section 5 shows the specification of the hedonic models in our study and the approaches employed to estimate regression models with symbolic interval data referring to the minimum and maximum values of the rental prices. Section 6 discusses the estimated results and the sub-national SPIHRs obtained. Finally, Section 7 presents some concluding remarks and future research suggestions.

2. Literature Review on Using Hedonic Regression Models for Price Indexes: A Focus on Spatial Price Indexes for Housing Costs and Rents

Hedonic pricing regression models have long been used for conducting empirical analyses during the 1920's; subsequently they have been brought to light through the theoretical works written by Griliches (1971) and Rosen (1974), whose conceptual foundations were derived from Lancaster (1966), and from the application in the automobile industry made by Griliches (1961) and by Ohta and Griliches (1975).

Hedonic models are a useful tool for computing quality-adjusted "temporal" price indexes, in particular for housing prices or rents and, in fact, numerous theoretical and application works have been written on this topic, including some reviews and surveys of

[4] An estimate of the territorial distribution of house prices for Italy in 2002 has been done by Cannari and Faiella (2008) by using three different data sources. The results confirmed the dualism between the Northern and Southern Italy showing lower house prices than the national average in the southern Regions.

[5] It is important to note that various researchers and in particular OMI's researchers have worked on the OMI data by using hedonic regression models in order to take into account the different characteristics of dwellings in determining housing prices and rents (see in particular: Ghirardo 2013 and Ghirardo et al. 2014).

the literature (Malpezzi 2003; Herath and Maier 2010; Hill 2013; De Haan and Diewert, 2013). This is also highlighted in the "Handbook on Residential Property Price Indices" Q6 drafted by six international organizations (Eurostat 2013)[6].

As highlighted by various authors, there are both advantages and disadvantages by using hedonic model since the results of the estimates may be sensitive to the method used (Silver 2011). However, Hill (2013) concluded that the advantages of the hedonic approach outweigh its disadvantages.

The first study on using hedonic regression methods for constructing spatial price indexes for housing costs, was published by Gillingam (1975) who made inter-area comparisons of rent levels by estimating hedonic equations for ten major cities. He combined survey microdata on individual rental units on multiple unit dwellings with information on neighbourhood quality characteristics taken from the 1960 census. The estimations of the hedonic equations for type of rental units for each of the ten cities, were computed using alternative functional forms. Two summary indexes were constructed, the first representative of all multi-unit apartments and the second limited to five room apartments.

The results revealed a substantial variation among place-to-place rent indexes, and that the specification of the group of units to be represented by the index is a crucial aspect of the index design.

After Gillingam's study, several economists at the BLS and Urban Institute focussed on producing an experimental interarea price index (Primont and Kokoski 1990; Kokoski Q7 et al. 1984) and more specifically estimates of interarea differences in the shelter cost (Follain and Malpezzi 1980; Ozanne and Thibodeau 1983; Cobb 1984; Can 1992). Moulton (1995) made several novel contributions. He used detailed rent and imputed owners' equivalent rental estimates for approximately 60,000 housing units included in the housing survey within the CPI program data collection, that matched with the neighbourhood data from the decennial Census. The hedonic bilateral price indexes were Q8 computed and then multilateral indexes were developed using GEKS[7].

As previously mentioned in the introduction, BEA continued these experiments in collaboration with other US federal agencies. The Regional Price Parities were computed using the same updated data sets and various methods were used for constructing Q9 multilateral indexes[8] (Kokoski et al. 1999; Aten, 2005, 2008; Martin et al. 2013).

A substantial improvement was achieved when it was possible to use data obtained from the new American Community Survey (ACS) conducted by the U.S. Census Bureau which provided 500,000 rent price observations per year. As illustrated by Aten (2017, 2019), the estimation of the RPPs is now obtained using the CPI microdata collected by the BLS which are not publicly available, while the ACS rental data are accessible users as Public Use Microdata Samples.

[6] Various National Statistical Institutes computed and published these indexes for house or rent prices (see, for example Behrmann and Goldhammer 2017).
[7] The GEKS is a method proposed by Gini, Elteto, Koves and Szulc that is used to produce a transitive price index from a non-transitive one. For the application in the field of spatial multilateral price indexes see World Bank (2013).
[8] Only indexes for 38 metropolitan and nonmetropolitan urban areas were estimated and then the estimates of the indexes for 50 states were included, plus District of Columbia, and 383 Metropolitan Statistical Areas

Outside of the US, during a consultation project in Cambodia concerning the estimation of poverty lines (thresholds), Dalén (2006) computed the housing component of the nonfood poverty line over time and space using household data from the Cambodian Socio-Economic Surveys funding interesting results.

The literature review reveals that: (1) the results of the spatial price index for housing cost estimations depend on the model used, as well as the availability of data characteristics; (2) it is essential to have detailed territorial information especially regarding neighbourhood quality characteristics; (3) using the rental price as dependent variable is proved to be a good choice because it well represents the cost of shelter for homeowners; (4) all the researchers concluded that the log transformation of the rental prices has numerous advantages for obtaining valid SPIHRs.

3. Main Features of the Hedonic Approach for Estimating Spatial Price Indexes for Housing Rents

The main characteristic of the hedonic approach to estimate spatial price indexes for housing rents can be summarized by referring to the above-mentioned papers and in particular to those published by Malpezzi (2003) and Herath and Maier (2010).

3.1. The General Model

The hedonic pricing method for housing costs is basically a regression of the price of the house (rent or value) against known relevant determinants (characteristics of the unit) that indirectly affect the price. A classical hedonic equation is as follows:

$$R = f(S, N, L, E, t)$$

where

$R =$ rental price (or value) of the housing unit;
$S =$ structural characteristics of the unit (characteristics of the building: age, type of building as number of units in the building, number of floors, structural materials, presence of garages and gardens, etc; characteristics of the unit: size or floor area, number of rooms, bedrooms and bathrooms, type of plumbing and heating systems, and so on);
$N =$ neighbourhood (access to services as shops, schools and other essential amenities; perhaps an overall neighbourhood review, quality of schools, socio-economic characteristics of the neighbourhood);
$L =$ locational variables (distance from the city center, business district and sub-centres of employment);
$E =$ environmental characteristics (air quality, proximity to parks, beaches, landfills, an so on) and t is an indicator of time.

For the specification of the model a semi log linear form is mainly used, where most of the characteristics and their classifications are dummy variables. There are advantages and disadvantages in using this kind of model: (1) it enables to estimate coefficients which are easy to interpret, as the proportions of the price that are attributable to the respective characteristics; (2) the independent variables (characteristics) can be sub-divided into more specific variables giving flexibility to the specification of the model (for example, the structural characteristics of the building and the dwellings), (3) it often mitigates the common statistical issue known as heteroscedasticity.

For the time t, the general hedonic model can therefore be:

$$\ln R = \beta_0 + \beta_1 S + \beta_2 N + \beta_3 L + \beta_4 E + \varepsilon$$

where ε is the error term.

3.2. *The Hedonic Model for Estimating Spatial Price Indexes for House Rents (SPIHRs)*

For estimating spatial price indexes using the hedonic method, it is essential to adopt the "space dummy method", that is to include a geographical area variable (A) among the independent variables. The variable A is a set of area dummies that takes into account the number of the areas to be estimated. In this case, considering a sample of *n* independent observations of house' prices, the semi-log formulation can be expressed as follows:

$$lnr_{ij} = \sum_{j=1}^{M} \alpha_j A_j + \sum_{k=1}^{K} \beta_k C_{ik} + \varepsilon_{ij} \tag{1}$$

Where:

- r_{ij} is a vector of price of property *i* (with $i = 1,2, \ldots i, \ldots, n)$ in the different areas (regions or provinces) *j* with $j = 1,2, \ldots, j, \ldots, M;$
- A_j is a vector of geographical areas dummies;
- C_{ik} is the matrix of the house characteristics (with $k = 1, \ldots, K$);
- β_k is the vector of hedonic regression coefficients called also characteristic shadow prices;
- ε_{ij} is the error term, that satisfies the standard assumption of a multiple regression model.

C_{ik} can be subdivided into the above-mentioned sub-groups of characteristics, such as S = structural characteristics of the building, N = locality/neighborhood and E = environmental characteristics.

The hedonic approach for estimating SPIHRs expressed by model (1), refers to a relatively new strand of the stochastic approach for price index number construction. According to Rao and Hajargasht (2016), model (1) in its multiplicative form postulates that the observed rent prices, r_{ij}, can be expressed as the product of three components: the general price level in area *j* relative to reference base area (denoted by $SPIHR_r$), the price level of the *k*-th characteristic to a base characteristic (denoted by P_k) and a random disturbance term u_{ij},

$$r_{ij} = SPIHR_j \cdot P_{ik} \cdot u_{ij}$$

The additive form of the model is obtained by taking logarithms of both sides Equation (2):

$$lnr_{ij} = a_j + b_{ik} + \varepsilon_{ij}$$

This model can be expressed as a regression equation for each rent where the independent variables are dummy variables, thus obtaining model (1). In order to estimate SPIHRs by model (1), it is essential to establish the reference or base area that could be the specific area j* (region or province) or the average of areas (regions or provinces). Moreover, for each set of *h* classifications, one βhk is equal to zero so that the equation is not overidentified. The remaining parameters were estimated using $j*$ and $\alpha*$ data as reference. In this case α_j is the difference of (fixed) effects associated with the geographical areas compared with the base area j*. After having estimated the α_j parameters, the SPIHR of the area *j* with respect to the base area is given by SPIHR $j = e^{\alpha j}$.

4. The Characteristics of the OMI Data-Base and the Specific Data Used

In order to compute SPIHRs for Italy at sub-national level we used freely available user data from OMI (which is a part of the Italian Agency of Revenue and Tax) focused on Q10 residential buildings (OMI 2017). This data source is the most comprehensive source of information on Italian houses price rents with a wide geographical coverage, including data for each Italian municipality (8,046 municipalities). It provides rental value estimates on a half-yearly basis for different types of dwellings, including the maintenance and conservation status of the building, according to positional factors.

The OMI's objective is to estimate the value of possible rents for all residential building stock in each municipality, on the basis of the Cadastre of Real Estate databases. The OMI data production system stands out both for the survey design and for the rental evaluation of the dwellings which are explained in depth in a specific Manual (OMI 2018).

Q11

4.1. Survey Design

The survey design adopted for each municipality takes into account the complexity and heterogeneity that characterize the housing market and the fact that the positional factor is the factor that strongly affects property and rental value.

In order to consider the "locational" factor and the economic and socio-environmental characteristics of the areas, a hierarchical territorial classification for each Italian municipality is adopted. Firstly, each municipal territory is divided into districts, which represents a territorial area with a precise geographical location in the municipality and reflects a consolidated urban location. Subsequently, to obtain adequate estimation of the rents, each district is further sub-divided into homogeneous zones defined by referring to the housing market characteristic. These homogeneous zones are known as OMI zones in Italy.

More specifically, the homogeneous zone is identified through a price and quality evaluation process, according to homogeneous conditions in terms of economic and socio-environmental characteristics, such as presence and accessibility to public and private facilities and services, the quality of urban and suburban transport services, road connections, the presence of educational, health, sports, commercial and tertiary facilities. It was decided to establish ex-ante territorial segmentation due to the fact that the heterogeneity of dwellings is often expressed using territorial clusters that takes into account the location and neighbourhood of the dwellings. This implies that the segmentation of the housing market based on homogeneous zones can significantly reduce the variability of the rental values.

The identification of the districts, which are equal to five specified as B = central area, C = semi-centre, D = outskirts, E = suburbs and R = rural and extra-urban and the specification of the homogeneous zones are strongly affected by the size of the municipality. In Italy there are many small municipalities: approximately 2,000 municipalities have less than 1,000 inhabitants and 5,500 municipalities have less than 5,000 inhabitants. Therefore, frequently, in the small municipalities only one or two districts are identified and within the districts sometimes only one OMI zone is identified.

In fact, from the OMI report, we know that 5,667 municipalities have at maximum three of OMI zones; 1,883 municipalities have at maximum from four to six OMI zones; 358 municipalities have at maximum from seven to ten OMI zones, 106 municipalities have at maximum from 11 to 20 OMI zones, 19 municipalities have at maximum 21 to 30 OMI zones, while 13 municipalities have more than 30 OMI zones. Bearing in mind that not all the OMI zones are included in each municipality, the frequency distribution of the number of OMI zones in the 8,046 Italian municipalities produces a total of occurrences equal to 27,426.

In conclusion, referring to the classification of independent variables mentioned in the Section 3, surely the districts can be considered as a location variable, while the OMI zones capture both the neighbourhood variable and the environmental characteristics, whose effects on price rent have almost never been captured in the previous studies mentioned in the literature review. However, it is necessary to point out that OMI zones are also distinct from each other according to their location in the various districts. Therefore, the OMI zone classification contains detailed information regarding the location of the dwellings. In order to verify the validity and coherence of the definition of the zones in terms of their characteristics and homogeneity of prices for the different type
Q12 of dwellings, OMI also conducts an additional stratified sample survey on a regular basis[9].

As a consequence, from an econometric point of view these two variables are collinear and cannot be included in the same model specification. As we will illustrate in Section 5, we decided to use these two different variables in order to better understand their influence on rental prices.

4.2. *The Process for Estimation of Housing Rents*

The rental valuation should be carried out for each dwelling within each OMI zone. However, this is not a practical method as the *average rent* for the different building characteristics of the dwellings is estimated by taking into account the conservation status of the building.

The following typologies of dwellings are defined according to their homogeneous distribution, organizational and functional characteristics, as follows[10]: well-finished

Q13

[9] The OMI also conducts a regular stratified sample survey on the real estate market in the municipalities and zones where numerous sales and lease agreements are stipulated.
For this reason, the survey is only conducted in approximately 1,500 municipalities in order to follow the evolution of the market and to support the computation of temporal House Price Index (Istat 2020). The sample survey also aims to support the estimations of the appraisal process and to obtain more detailed information on the characteristics of the residential buildings and dwellings which would increase the number of quality characteristics to be taken into account when applying the hedonic model. However, this kind of data is not publicly available, but it is only accessible after signing a research protocol with the OMI Agency. The coherence between the confidence interval of the estimates of the average rental values and the interval min-max quotations estimated by the appraisal process are verified.

[10] These typologies are a little bit different from the cadastral typologies, also if are based on them.

apartments; economic apartments, detached houses, luxury dwellings, houses and dwellings typical of the area. Not all dwellings types are located in each OMI zone and the most commonly found dwelling are detached houses, well-finished apartments and economic apartments.

Detached houses generally have gardens while well-finished apartments are dwellings made with expensive materials and finishing (floors, walls, fixtures, plant accessories) of high architectural value. These dwellings have average sized-rooms with at least two bathrooms. Economic apartments, that also includes cheap dwellings, are made with low-cost buildings materials and finishing. The surface area of the building is exploited maximally, with small sized-rooms, only one bathroom and are located in working-class areas.

It is worth noting that social housings are included in the economic apartments. For this dwelling type and for dwellings located in rural districts a corrective coefficient is applied in order to take into account the specific characteristics of the property, both intrinsic and extrinsic, which determine the different value of these properties with respect to properties with average characteristics of the area.

The conservation status of the building, which include the quality of the building materials and workmanship and the general atmosphere of the property, is divided into three categories: excellent, normal and poor quality.

As regards the survey design, there are two methods for estimating rents: the indirect or appraisal method (based on the appraisal expertise of the local housing market experts) and the direct method (based on the estimations of rents through a sample survey of the housing market rents).

In order to estimate the average rents for all municipalities and zones, an appraisal process is carried out by local technicians and verified by specific technical committees at provincial and central level, which use both data included in the cadastre data base and specific rental data obtained from rental agencies and from specific surveys.

The valuation process is based on the suggestions and methods specified in the "International Valuation Standards: A guide to the Evaluation of Real Property Assets", established by a specific international Council (Parker 2016).

The average rental values are expressed per square meter, that is they are standardized according to the dimension of the dwelling. However, it is difficult to carried out a point estimation of the rental value and therefore, the OMI committees set a *minimum* and *maximum* values for each typology of dwelling and establish that the deviation between the two values must not be higher than 50%.

It is worth noting that constructing representative samples of rental values is statistically challenging due to the heterogeneity of the market and to the fact that transaction data are often incomplete (BIS 2019). Indeed, the estimation may prove to be unreliable due to a reduced number of observations (Lopez and Hewings 2018). However, the OMI carries out quality control of sample survey data following the international standards established in the Italian Code of Official Statistics.

In conclusion, by taking into account the general survey design and the process used for the estimation of rental values, we believe that the OMI data are a useful tool for calculation of the SPIHRs.

4.3. Available and Used Data

As already underlined, the OMI data base contains two quotations (minimum–maximum) for each type of dwelling (detached houses, well-finished apartments and economic apartments) and their concerning conservation status (excellent, normal, poor quality) for each homogenous zone in each municipality for a total of 54,214 observations. In order to enhance the quality of price comparisons, we modified the OMI zone classification by aggregating specific zones for various municipalities characterized by similar socio-economic features and rental price distributions, thus obtaining 35 different OMI zones with a reliable overlap of the rental prices among geographical areas. Regarding the type of dwellings, it is worth noting that for the year 2017, 28% of the 54,214 residential buildings are well-finished apartments, 30% are economic apartments and 42% are detached dwellings.

Economic apartments are mainly present in the Southern Italian Regions (Basilicata, Calabria, Molise and Sicilia), while well-finished apartments are mainly found in Trentino-Alto Adige, Valle d'Aosta and Lazio. Two observations can be made from the available quotations. Firstly, very few quotations consider poorly conserved dwellings (only 131). Secondly, in some regions there are few quotations for excellently conserved apartments, especially regarding economic apartments in Trentino Alto Adige. Table 1 and 2 show the descriptive statistics of the rents per m^2 (euro per month) based on the quotations for all of the Italian municipalities and on the average quotations of the for the 20 Italian regions according to type of dwellings and conservation status. Quotations for poorly conserved dwellings are very few to compute adequate descriptive statistics.

Table 1 shows the variability of the rental quotations observed for the 8,045 municipalities and across the Italian regions and how type of dwelling and conservation status affect the distribution of the minimum and maximum rental prices. As illustrated in Table 1, the highest rental value is observed for well-finished apartments, for which the rental price ranges from 3.99 to 5.37 EUR/m^2. In particular, the mean rental price ranges from 4.90 to 6.37 EUR/m^2 for excellently conserved apartments (the maximum rental price shows as well the highest level of variability). Contrastingly, the lowest rental value is observed for economic apartments, for which the rental prices range from 2.73 to 3.72 EUR/m^2.

From the second part of the Table 1, we can see that also the variability of the average rental values across regions is very high and this surely justify the importance to compute the SPIHRs for the Italian Regions.

5. The Specification of the Hedonic Models and the Estimation Procedure

Q14 As regards the estimation strategy[11], it is worth noting that, as already mentioned in Section 4, the OMI database provides minimum–maximum range of rental prices according to OMI zones, type of dwelling and conservation status. Therefore, for each observation two rental prices are specified: the minimum and the maximum rental value.

[11] In the OMI data there are not specific weights for each municipality. In our analysis, we considered an implicit system of weight given by the number of homogeneous zone in each municipality. This is an issue that could be useful to deal with in future works.

Table 1. Descriptive statistics based on the quotations available for all the municipalities. Italy, 2017.

	Conservation status					
	Total		Excellent		Normal	
Type of dwelling	Min	Max	Min	Max	Min	Max
Detached						
Mean	3.60	4.86	4.72	6.12	3.36	4.58
Standard deviation	2.04	2.78	2.42	3.39	1.86	2.55
Coefficent of variation	0.57	0.57	0.51	0.55	0.55	0.56
Skewness	2.41	2.65	2.50	2.85	2.36	2.45
Well-finished						
Mean	3.99	5.37	4.90	6.37	3.83	5.19
Standard deviation	2.09	2.94	2.54	3.76	1.96	2.73
Coefficent of variation	0.52	0.55	0.52	0.59	0.51	0.53
Skewness	2.39	2.94	2.89	3.89	2.13	2.32
Economic						
Mean	2.74	3.73	3.50	4.38	2.68	3.68
Standard deviation	1.59	2.13	1.97	2.53	1.54	2.08
Coefficent of variation	0.58	0.57	0.56	0.58	0.58	0.57
Skewness	2.41	2.25	3.43	3.12	2.19	2.09
Total						
Mean	3.46	4.67	4.58	5.91	3.27	4.47
Standard deviation	2.00	2.73	2.44	3.46	1.85	2.53
Coefficent of variation	0.58	0.59	0.53	0.59	0.57	0.57
Skewness	2.36	2.68	2.66	3.26	2.20	2.31

As a consequence our dependent variable in model (1) is an interval-valued variable which is defined for a given set of N entities $S = \{s_1, \ldots, s_N\}$, by an application $Y\colon S \longrightarrow B$ such that $s_i \longmapsto Y(s_i)$, $Y(s_i) = [l_i,\ u_i]$, B is the set of intervals of an underlying set $O \subseteq \mathbb{R}$, In other words, the value of the interval-valued variable Y for each $s_i \in S$ is usually defined by the lower and upper bounds l_i and u_i of $I_i = Y(s_i)$, In order to estimate the hedonic regression model [1] with $r_{ij} = [l_{ij},\ u_{ij}]$, we used the Symbolic Data Analysis approach (SDA) which provides a framework for the representation and analysis of data that comprehends inherent variability (Brito 2014).

Several approaches have been introduced for estimating regression models with symbolic interval data, Billard and Diday (2000) proposed a first approach which consisted in fitting a regression line to the centres of the intervals or midpoints defined as $c_{ij} = (l_{ij} + u_{ij})/2$. Therefore, model (1) can be expressed as:

$$lnc_{ij} = \sum_{j=1}^{M} \alpha_j A_{jC} + \sum_{k=1}^{K} \beta_k C_{ikC} + \varepsilon_{ijC} \tag{2}$$

Table 2. Descriptive statistics based on averages computed for the 20 Italian regions. Italy, 2017.

	Conservation status					
	Total		Excellent		Normal	
Detached	Min	Max	Min	Max	Min	Max
Min	2.51	3.44	3.00	3.00	2.50	3.42
Max	6.04	8.14	10.09	13.00	6.60	8.97
Standard Deviation	2.04	2.78	2.42	3.39	1.86	2.55
Well-finished						
Min	1.86	2.43	2.46	3.04	1.68	1.23
Max	5.59	7.46	9.41	12.58	5.24	7.00
Standard Deviation	2.09	2.94	2.54	3.76	1.96	2.73
Economic						
Min	1.25	1.67	1.62	2.18	1.23	1.66
Max	7.50	11.65	6.30	7.97	6.30	7.97
Standard Deviation	1.59	2.13	1.97	2.53	1.54	2.08
Total						
Min	1.64	2.16	2.51	3.10	2.51	3.10
Max	5.29	7.22	9.18	12.28	9.18	12.28
Standard Deviation	2.00	2.73	2.44	3.46	1.85	2.53

The classical least squares estimator may be used and standard statistical inference will apply under the standard assumptions regarding the error term of the regression. Although this method can provide information on the average centrality of the intervals, it does not Q15 take into account the range of the intervals[12].

Billard and Diday (2002) also proposed an alternative method which consisted of estimating two different regression lines, one for the minimum and another for the maximum value of the intervals with no restrictions across lines as follows:

$$\begin{aligned} lnr_{ijL} &= \sum_{j=1}^{M} \alpha_j A_{jL} + \sum_{k=1}^{K} \beta_k C_{ikL} + \varepsilon_{ijL} \\ lnr_{ijU} &= \sum_{j=1}^{M} \alpha_j A_{jU} + \sum_{k=1}^{K} \beta_j C_{ikU} + \varepsilon_{ijU} \end{aligned} \tag{3}$$

The OLS estimator may be used for minimizing: $\min \sum_i (\varepsilon_{ijL}^2 + \varepsilon_{ijU}^2)$ which entails performing two separate minimizations. However, the OLS estimator would not be the

[12] There are several proposals aimed to incorporate the length of the interval into the analysis (Neto and De Carvalho, 2008; 2010). However, these methods will produce the same estimates obtained by using the midpoint to method.

most efficient estimator because it is likely that the two error terms may be correlated given that $r_{ijL} < r_{ijU}$.

In this paper we propose expressing models (3) as a system of seemingly unrelated regression Equations (SURE) which take into account the properties of the error terms (Zellner 1963).

Indeed, this set of equations has a contemporaneous cross-equation error correlation (i.e., the error terms in the regression equations are correlated). At first glance, the equations seem unrelated, but the equations are related through the correlation in the errors which is tested using the Breush-Pagan χ-statistics (Cameron and Trivedi 2009) The generalized least-squares algorithm, which produces asymptotically efficient and feasible estimates (Greene 2012) was used.

In order to better understand the influence of the locational factors and the neighbourhood and environmental characteristics on rental prices, we estimated two different model specifications: the first model considers a broader classification for the locational characteristics as specified by the "district variable"; the second model considers mainly the neighbourhood and environmental characteristics as specified by the "OMI zones variable". Since the OMI zone classification is embedded into the district classification and include also information on the location, substantially this variable can be considered as a more detailed definition of the locational characteristics that contains the specific information above mentioned. It is obvious that, for this reason, from an econometric point of view these two variables are collinear and cannot be included in the same model specification.

6. Results: Presentation and Some Comments

The SPIHRs were estimated using both models (2) and (3) by considering first all the dwelling types included in the OMI database and then only the subgroup of the economic apartments. In this way useful information may be provided to economic and social policy Q16 makers. The results of the estimates are reported in Tables 3A and 3B respectively[13].

No evidence of functional form misspecification was observed from the Ramsey tes (1969). The estimation results of models (2) and (3) at regional level for all the dwelling types are reported in Table 3A by considering two model specifications. The first three columns refer to districts, while the last three columns refer to homogeneous territorial zones.

As regards the estimation results, the best fitting-models are obtained when homogeneous zones are included. Moreover, model (3) is preferable to model (2) as it allows for the correlation between ε_{ijL} and ε_{ijU} and the estimation of the full variance–covariance matrix of the coefficients. As shown in the last rows of Table 3, the correlation of the residuals in the minimum and maximum rent equations is equal to 0.975. We can reject the hypothesis that this correlation is zero on the basis of the Breush-Pagan test of independence results. The correlation is strong, so the efficiency gains with the SUR estimation is great (Cameron and Trivedi 2009

[13] The mid-point included in model estimation (2) is an arithmetic mean of the min and max rental values. Estimation results of model (2) computed with the geometric mean are not significantly different from the estimated results obtained using the arithmetic mean.

Table 3A shows different distributional patterns as rental prices vary greatly across the Italian regions, thus confirming territorial heterogeneity. The lowest rental prices were observed for Basilicata and Molise while the highest rental prices were observed in Trentino Alto-Adige and Liguria (both for midpoint and minimum- maximum models).

As expected, the localization factor plays an important role in determining rental prices. However, it is interesting to note that the rents of houses located in central areas of the municipalities cost 11.4% – 11.8% less than the rents of dwellings on the outskirts as shown in Table 3A while houses located in rural areas cost 25.5% less than houses located in central areas of the municipalities. This effect may be due to the heterogeneity of houses included in the district classification which does not take into account the socio-economic characteristics of the territorial area. More specifically, the district variable captures an average effect of the locational characteristics since it is obtained by considering rental prices related to all the municipalities included in each region. Yet, Italy is characterized by the presence of numerous small municipalities in which the rent values are smaller than those observed in the big municipality and the more expensive houses (well-finished dwelling and detached houses) are located in semi-central and outskirts districts. Contrastingly, we have found that Italian regional capitals, such as Florence (Tuscany), Rome (Lazio) or Milan (Lombardy), show a different behaviour of housing rental prices with more expensive houses located in the central districts.

When considering a detailed specification for the localization factors (OMI zones), the effects of type-of-dwelling and property status do not vary: detached houses cost 13.6% and 12.6% more than rented well-finished apartments, while excellently conserved houses cost 22.8% and 17.2% more than those with normal conservation status.

Table 3B reports hedonic regional model estimation results for economic apartments by using models (2) and (3) specifications. As regards the estimation results, the best fitting-models are obtained when homogeneous zones are included, and model (3) is preferable to model (2). The number of observations is equal to 16,040 and estimated results confirm high territorial heterogeneity across the Italian regions, highlighting disparities between the Northern and Southern regions. As regards the localization factors, the coefficients associated to the districts dummy continue to assume the same signs but are lower than the coefficients estimated in Table 3A for all the dwelling type thus confirming the influences of houses rental prices in small municipalities.

Table 3A. Hedonic regional models for all dwelling types with different territorial classifications, first semester (january–june), year 2017.

	Hedonic model with districts			Hedonic model with homogeneous zones		
	Model [2]	Model [3]		Model [2]	Model [3]	
	Midpoint	Min	Max	Midpoint	Min	Max
Regions (*ref. Italy*)						
North						
Piemonte	−0.079	−0.082	−0.075	−0.065	−0.070	−0.061
	(0.005)	*(0.005)*	*(0.005)*	*(0.005)*	*(0.005)*	*(0.005)*
Valle d'Aosta	0.216	0.240	0.197	0.227	0.251	0.209

Table 3A. Continued

	Hedonic model with districts			Hedonic model with homogeneous zones		
	Model [2]	Model [3]		Model [2]	Model [3]	
	(0.022)	*(0.022)*	*0.023*	*(0.022)*	*(0.022)*	*(0.022)*
Lombardia	0.173	0.191	0.159	0.174	0.192	0.161
	(0.005)	*(0.004)*	*(0.005)*	*(0.004)*	*(0.004)*	*(0.004)*
Liguria	0.426	0.436	0.419	0.423	0.433	0.416
	(0.008)	*(0.008)*	*(0.008)*	*0.007*	*0.007*	*(0.008)*
Trentino-Alto Adige	0.412	0.422	0.403	0.415	0.425	0.406
	(0.008)	*(0.008)*	*(0.008)*	*(0.008)*	*(0.008)*	*(0.008)*
Veneto	0.102	−0.176	−0.174	−0.176	−0.176	−0.174
	(0.009)	*(0.009)*	*0.010*	*(0.009)*	*(0.009)*	*(0.009)*
Friuli-Venezia Giulia	0.020	−0.022	0.048	0.028	−0.014	0.056
	(0.009)	*(0.009)*	*(0.009)*	*(0.009)*	*(0.009)*	*(0.009)*
Emilia-Romagna	0.222	0.235	0.213	0.218	0.231	0.209
	(0.007)	*(0.007)*	*(0.007)*	*(0.007)*	*(0.007)*	*(0.007)*
Centre						
Toscana	0.435	0.436	0.435	0.433	0.434	0.432
	(0.007)	*(0.007)*	*(0.007)*	*(0.007)*	*(0.007)*	*(0.007)*
Umbria	−0.071	−0.074	−0.068	−0.080	−0.083	−0.078
	(0.013)	*(0.013)*	*(0.013)*	*(0.013)*	*(0.013)*	*(0.013)*
Marche	0.024	0.021	0.026	0.024	0.022	0.027
	(0.008)	*(0.008)*	*(0.008)*	*(0.008)*	*(0.008)*	*(0.008)*
Lazio	0.387	0.388	0.386	0.386	0.388	0.385
	(0.008)	*(0.008)*	*(0.008)*	*(0.008)*	*(0.008)*	*(0.008)*
Abruzzo	−0.179	−0.210	−0.157	−0.182	−0.212	−0.159
	(0.008)	*(0.008)*	*(0.008)*	*(0.008)*	*(0.008)*	*(0.008)*
South and Islands						
Molise	−0.308	−0.351	−0.276	−0.284	−0.328	−0.251
	(0.023)	*(0.023)*	*(0.023)*	*(0.023)*	*(0.023)*	*(0.023)*
Campania	−0.174	−0.168	−0.179	−0.195	−0.188	−0.201
	(0.005)	*(0.005)*	*(0.005)*	*(0.005)*	*(0.005)*	*(0.005)*
Puglia	−0.122	−0.104	−0.134	−0.150	−0.131	−0.163
	(0.008)	*(0.008)*	*(0.008)*	*(0.008)*	*(0.008)*	*(0.008)*
Basilicata	−0.639	−0.623	−0.651	−0.626	−0.611	−0.637
	(0.012)	*(0.012)*	*(0.012)*	*(0.012)*	*(0.012)*	*(0.012)*
Calabria	−0.362	−0.359	−0.364	−0.372	−0.369	−0.374
	(0.008)	*(0.008)*	*(0.008)*	*(0.007)*	*(0.007)*	*(0.008)*
Sicilia	−0.333	−0.346	−0.323	−0.336	−0.349	−0.326
	(0.007)	*(0.007)*	*(0.007)*	*(0.007)*	*(0.007)*	*(0.007)*
Sardegna	−0.149	−0.147	−0.151	−0.135	−0.132	−0.136
	(0.010)	*(0.010)*	*(0.010)*	*(0.010)*	*(0.010)*	*(0.010)*
Type of dwelling (*ref. well-finished apartment*)						
Economic apartment	−0.175	−0.176	−0.174	−0.175	−0.176	−0.174
	(0.004)	*(0.004)*	*(0.004)*	*(0.004)*	*(0.004)*	*(0.004)*
Detached	0.125	0.130	0.121	0.130	0.136	0.126
	(0.004)	*(0.004)*	*(0.004)*	*(0.004)*	*(0.004)*	*(0.004)*

Table 3A. Continued

	Hedonic model with districts			Hedonic model with homogeneous zones		
	Model [2]	Model [3]		Model [2]	Model [3]	
Status of the dwelling (*ref. Normal*)						
Excellent	0.194	0.226	0.170	0.196	0.228	0.172
	(0.005)	*(0.005)*	*(0.005)*	*(0.005)*	*(0.005)*	*(0.005)*
District (*ref. Central*)				No	No	No
Semi-central	0.326	0.316	0.333			
	(0.007)	*(0.007)*	*(0.007)*			
Outskirts	0.116	0.114	0.118			
	(0.005)	*(0.005)*	*(0.005)*			
Suburban	0.031	0.033	0.030			
	(0.005)	*(0.005)*	*(0.005)*			
Rural and extra-urban	−0.255	−0.255	−0.255			
	(0.005)	*(0.005)*	*(0.005)*			
Homogeneous zones	No	No	No	Yes	Yes	Yes
Intercept	1.058	1.058	1.374	1.189	1.020	1.332
	(0.004)	*(0.004)*	*(0.004)*	*(0.004)*	*(0.004)*	*(0.004)*
Correlation		0.975			0.974	
Breusch-Pagan test		51565			51467	
Adjusted R2	0.463	0.478	0.478	0.485	0.471	0.471
RMSE	0.375	0.374	0.381	0.368	0.367	0.372
AIC	47.719	−67.210	−67.210	45.462	−69.667	−69.667
BIC	47.977	−66.644	−66.644	45.969	−68.564	−68.564
N	54,214	54,214	54,214	54,214	54,214	54,214

Source: our elaboration from OMI data 2017.
Note: standard error in brackets; all the coefficients are significant at $p < 0.01$.

Table 4 reports the SPHIRs obtained when homogeneous territorial zones are included in the models (2) and (3) if we consider the 20 Italian regions and compare every region to the overall mean (Italy = 100). The SPHIRs are calculated both for all dwelling type and the economic apartments.

On observing regional models reported in Table 4, our results highlight high territorial heterogeneity and confirming significant rental price differences across the Italian regions thus demonstrating that rents are higher in the Northern-Central regions than in the South in respect to the whole of Italy. Basilicata proved to be the less expensive region for rental properties while Tuscany, Liguria while the highest rents were found in Trentino-Alto Adige. Significant differences were observed in the estimated SPHIRs, when all of the dwelling type were included with the exception of the estimation for economic apartments.

Table 3. Hedonic regional models for economic apartments with different territorial classifications. First semester (January–june), year 2017.

	Hedonic model with districts			Hedonic model with homogeneous zones		
	Model [2]	Model [3]	Model [3]	Model [2]	Model [3]	Model [3]
	Midpoint	Min	Max	Midpoint	Min	Max
Regions (*ref. Italy*)						
North						
Piemonte	−0.044	−0.064	−0.028	−0.033	−0.054	−0.016
	(0.013)	(0.013)	(0.013)	(0.013)	(0.013)	(0.013)
Valle d'Aosta	0.146	0.174	0.125	0.161	0.189	0.141
	(0.038)	(0.038)	(0.038)	(0.038)	(0.038)	(0.038)
Lombardia	0.196	0.213	0.183	0.190	0.207	0.178
	(0.013)	(0.013)	(0.013)	(0.013)	(0.013)	(0.013)
Liguria	0.452	0.445	0.457	0.450	0.443	0.455
	(0.016)	(0.016)	(0.016)	(0.016)	(0.016)	(0.016)
Trentino Alto Adige	0.673	0.691	0.659	0.641	0.666	0.623
	(0.177)	(0.179)	(0.178)	(0.175)	(0.175)	(0.175)
Veneto	0.238	0.246	0.153	0.237	0.250	0.158
	(0.026)	(0.026)	(0.026)	(0.026)	(0.026)	(0.026)
Friuli Venezia Giulia	0.020	−0.027	0.052	0.031	−0.017	0.064
	(0.019)	(0.019)	(0.019)	(0.019)	(0.019)	(0.019)
Centre						
Emilia-Romagna	0.213	0.223	0.205	0.211	0.221	0.203
	(0.016)	(0.016)	(0.016)	(0.016)	(0.016)	(0.016)
Toscana	0.390	0.399	0.384	0.389	0.397	0.382
	(0.017)	(0.017)	(0.017)	(0.017)	(0.017)	(0.017)
Umbria	−0.070	−0.065	−0.074	−0.076	−0.071	−0.079
	(0.027)	(0.027)	(0.027)	(0.027)	(0.027)	(0.027)
Marche	0.014	0.012	0.016	0.016	0.013	0.018
	(0.016)	(0.016)	(0.016)	(0.016)	(0.016)	(0.016)
Lazio	0.433	0.427	0.438	0.436	0.430	0.441
	(0.018)	(0.018)	(0.018)	(0.018)	(0.018)	(0.018)
Abruzzo	−0.217	−0.249	−0.193	−0.217	−0.249	−0.193
	(0.016)	(0.016)	(0.016)	(0.016)	(0.016)	(0.016)
South and Islands						
Molise	−0.360	−0.402	−0.329	−0.339	−0.382	−0.307
	(0.035)	(0.035)	(0.035)	(0.035)	(0.035)	(0.035)
Campania	−0.212	−0.214	−0.211	−0.230	−0.231	−0.230
	(0.013)	(0.013)	(0.013)	(0.013)	(0.013)	(0.013)
Puglia	−0.140	−0.117	−0.156	−0.158	−0.135	−0.175
	(0.016)	(0.016)	(0.016)	(0.016)	(0.016)	(0.016)
Basilicata	−0.734	−0.720	−0.743	−0.724	−0.712	−0.733
	(0.020)	(0.020)	(0.020)	(0.020)	(0.020)	(0.020)
Calabria	−0.368	−0.363	−0.371	−0.373	−0.368	−0.375
	(0.015)	(0.015)	(0.015)	(0.015)	(0.015)	(0.015)
Sicilia	−0.415	−0.427	−0.405	−0.419	−0.431	−0.409
	(0.014)	(0.014)	(0.014)	(0.014)	(0.014)	(0.014)
Sardegna	−0.215	−0.212	−0.218	−0.193	−0.190	−0.197

Table 3B. Continued

	Hedonic model with districts			Hedonic model with homogeneous zones		
	Model [2]	Model [3]	Model [3]	Model [2]	Model [3]	Model [3]
	(0.020)	(0.020)	(0.020)	(0.020)	(0.020)	(0.020)
Status of the dwelling (*ref. Normal*)						
Excellent	0.193	0.246	0.153	0.198	0.250	0.158
	(0.012)	(0.012)	(0.012)	(0.012)	(0.012)	(0.012)
District (*ref. Central*)						
Semi-central	0.312	0.305	0.317			
	(0.012)	(0.012)	(0.012)			
Outskirts	0.117	0.116	0.117			
	(0.008)	(0.008)	(0.008)			
Suburban	0.009	0.010	0.008			
	(0.009)	(0.009)	(0.009)			
Rural and extra-urban	−0.294	−0.295	−0.293			
	(0.008)	(0.008)	(0.008)			
Homogeneous zones	No	No	No	yes	yes	yes
Intercept	1.078	0.906	1.222	1.046	0.876	1.189
	(0.011)	(0.011)	(0.011)	(0.011)	(0.011)	(0.011)
Correlation		0.978	0.978		0.977	0.977
Breusch-Pagan test		15,345	15,345		15,322	15,322
Adjusted R2	0.457	0.463	0.449	0.477	0.482	0.470
RMSE	0.373	0.376	0.374	0.365	0.369	0.366
AIC	13.873	−22.242	−22.242	13.275	−22.888	−22.888
BIC	14.080	−21.781	−21.781	13.698	−21.958	−21.958
N	16,040	16,040	16,040	16,040	16,040	16,040

Source: our elaboration from OMI data 2017.
Note: standard error in brackets; all the coefficients are significant at $p < 0.01$.

As regards economic apartments, the SPHIRs estimated for Trentino-Alto Adige are not significant due to the reduced number of observations. In this case, the SPHIRs estimated for Veneto and Lazio are higher than the SPHIRs estimated for all types of dwelling, while the SPHIRs estimated for Toscana, Abruzzo, Molise, Sicilia and Sardegna are lower than the SPHIRs estimated for all dwelling types. These findings may be influenced by two contrasting effects which emerge when estimating SPHIRs at regional level: the inclusion of tourist destinations characterized by higher rents and small municipalities with lower rents. Taking into account the various housing characteristics provided by the OMI data

Table 4. Regional SPIHRs (Italy = 100), year 2017.

	All dwelling types			Economic apartments		
	Model [2]	Model [3]		Model [2]	Model [3]	
Region	Midpoint	Min	Max	Midpoint	Min	Max
North						
Piemonte	93.72	93.28	94.09	96.76	94.70	98.37
Valle d'Aosta	125.48	128.58	123.26	117.52	120.78	115.18
Lombardia	119.00	121.12	117.45	120.98	122.99	119.48
Liguria	152.73	154.18	151.58	156.87	155.81	157.66
Trentino-Alto Adige	151.39	152.95	150.06			
Veneto	110.09	83.82	84.02	126.69	128.45	117.14
Friuli-Venezia Giulia	102.83	98.61	105.81	103.15	98.28	106.59
Emilia-Romagna	124.37	126.00	123.23	123.43	124.76	122.52
Centre						
Toscana	154.13	154.31	154.08	147.48	148.78	146.58
Umbria	92.29	92.01	92.53	92.66	93.14	92.37
Marche	102.45	102.19	102.70	101.59	101.35	101.82
Lazio	147.11	147.34	147.01	154.65	153.66	155.47
Abruzzo	83.36	80.89	85.26	80.49	77.94	82.45
South						
Molise	75.31	72.05	77.80	71.27	68.28	73.55
Campania	82.29	82.90	81.80	79.43	79.36	79.46
Puglia	86.09	87.72	84.94	85.38	87.33	83.98
Basilicata	53.47	54.30	52.89	48.48	49.09	48.06
Calabria	68.96	69.16	68.83	68.90	69.20	68.70
Sicilia	71.44	70.53	72.15	65.79	64.96	66.45
Sardegna	87.41	87.62	87.26	82.41	82.73	82.15

Source: our elaboration from OMI data 2017.

set, it is clear that the role played by tourism is stronger than the effect of the prevalence of low-cost houses with economic apartments.

In order to explore the territorial heterogeneity of rents within each Italian region, Figures 1–4 show sub-national SPHIRs estimate with model (3) if we consider homogeneous territorial zones for the Italian regions (Italy = 100).

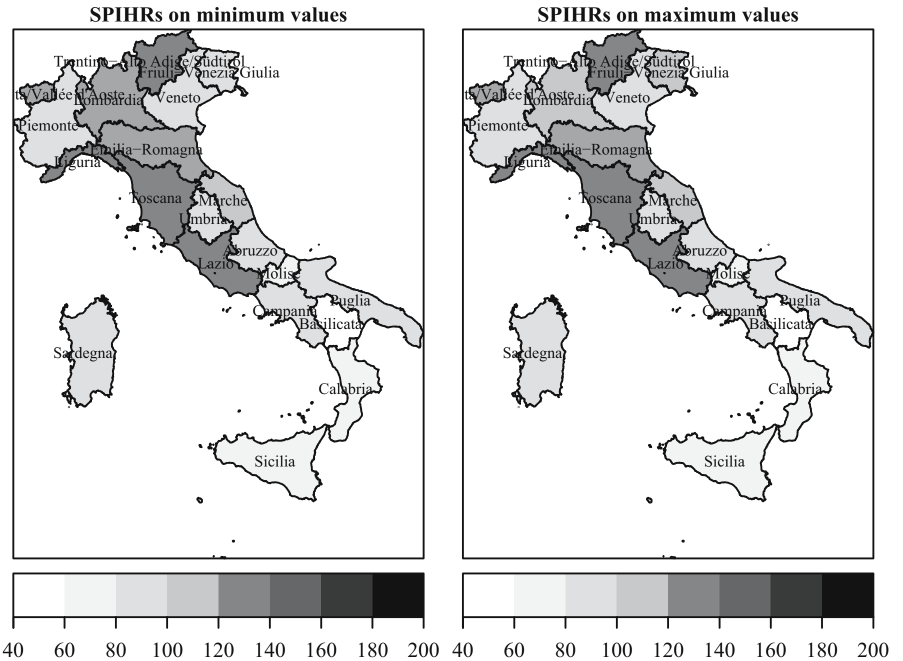
Figures 3 and 4 show sub-national SPHIRs with the corresponding intervals at 95%. Although most regions have small confidence intervals, few regions have overlapping confidence intervals. Overlapping confidence intervals were observed between Lombardia and Liguria SPIHRs and Calabria and Sicilia SPIHRs for the model estimated for all dwelling types and for the economic apartments[14].

Q17

7. Concluding Remarks

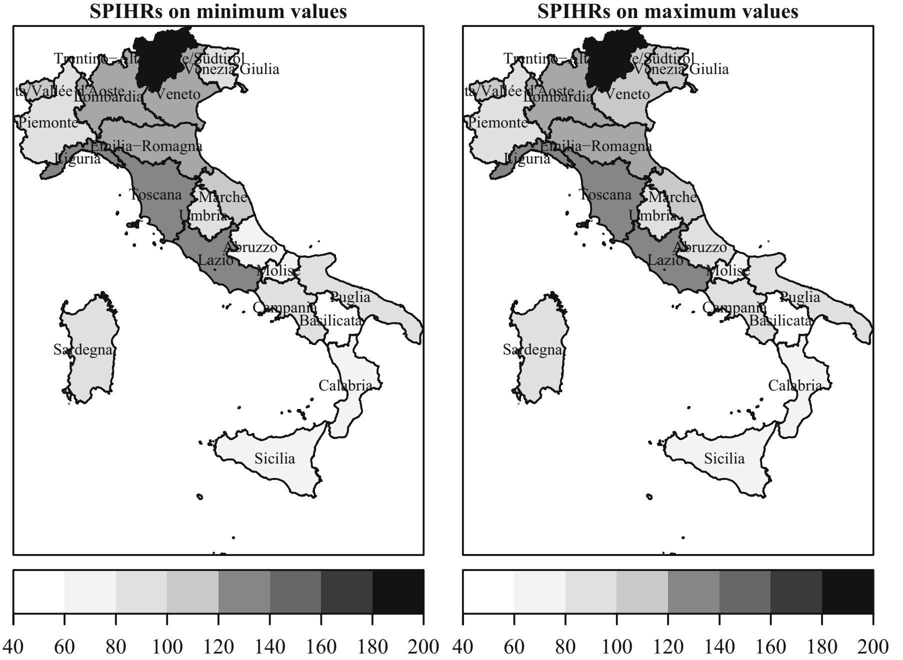
The main objective of this paper was to explore the feasibility of using the OMI data set for compiling SPIHRs for the Italian regions on a regular basis. Spatial price indexes measuring

[14] With the aim of making multiple comparisons based on interval confidence, it would be advisable to use simulation methods as suggested by Marshall and Spiegelhalter (1998) and Leckie and Goldstein (2011) and applied by Biggeri et al. (2017) in the field of price statistics.



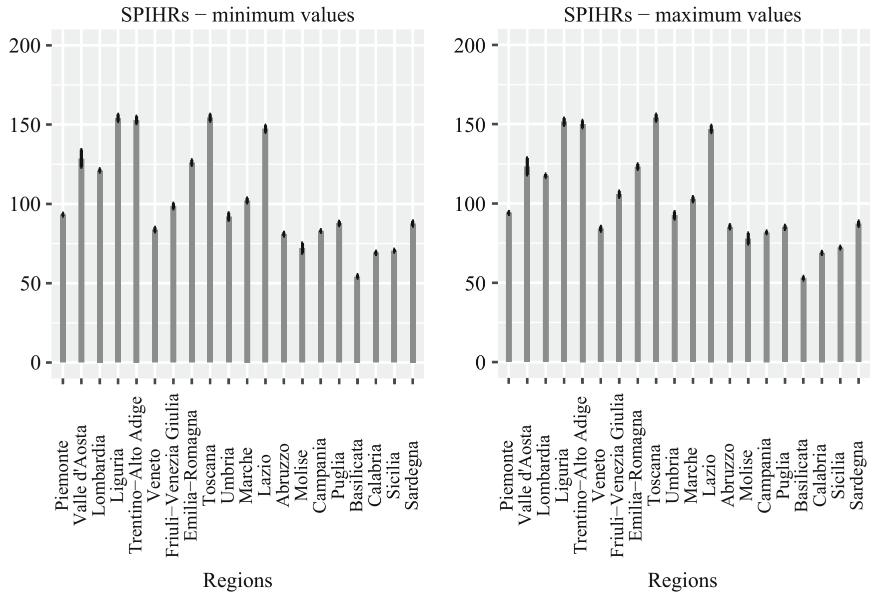
Source: our elaboration from OMI data 2017

Fig. 1. SPIHRs for Italian regions for all dwellings types, year 2017.



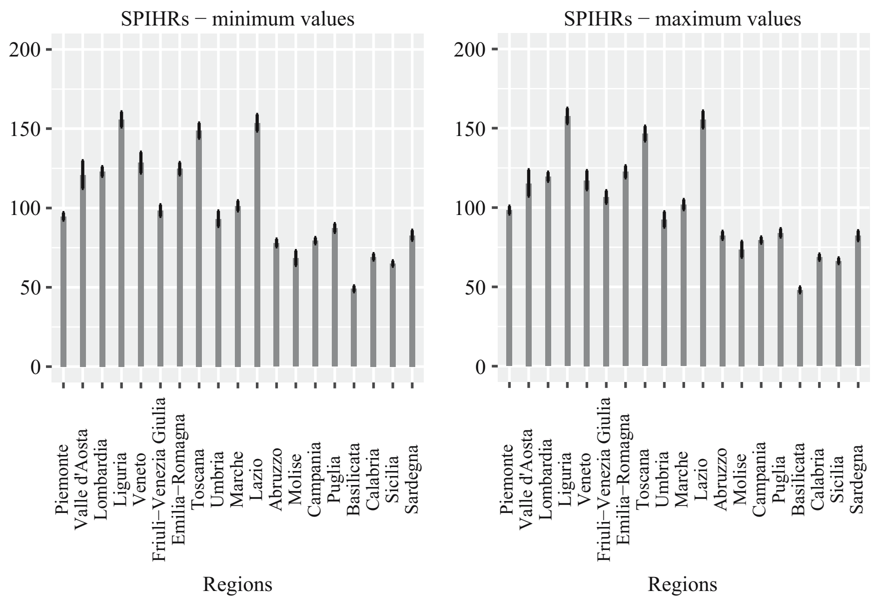
Source: our elaboration from OMI data 2017

Fig. 2. SPIHRs for Italian regions for economic apartments, year 2017.



Source: our elaboration from OMI data 2017

Fig. 3. SPIHRs for Italian regions with confidence intervals for all dwellings types, year 2017.



Source: our elaboration from OMI data 2017

Fig. 4. SPIHRs for Italian regions with confidence intervals for economic apartments, year 2017.

differences in price levels across regions within a country are essential for comparing real income, standards of living and consumer expenditure patterns. Housing is the most expensive cost of the household budget, and several economic and social players have highlighted the importance of providing spatial rental price indexes for countries like Italy.

With the aim of comparing rental prices across the Italian territory using hedonic regression models, various classifications of "positional factors" are available on OMI database.

The results of our computations are well-suited for the hedonic models used and for the significant differences observed across the various Italian regions. The SPIHRs for economic apartments are particularly interesting for policy makers and confirm the usefulness of calculating SPIHRs on a regular basis, every two years. However, it is important to carry out further analysis by computing the SPIHRs for other years from 2010 up to now considering also spatial dependencies inherent to rental prices will be considered by estimating spatial models.

Second, if the access to additional data is granted, further analyses will be carried out for the year 2019 by using detailed information on all of the dwelling and building characteristics for a sample of approximately 1,500 municipalities collected by OMI in order to compare the results.

Finally, it is worth noting that the strong differences found in the regional estimates of the SPIHRs surely reshuffle in some way the territorial distribution of the income and poverty indicators.

However, to assess to what extent our findings impact the regional differences in the standard of living and poverty levels in the Italian context, other sources of data on income and living conditions should be used and analysed by considering different poverty indicators computed using adjusted poverty thresholds. This important topic will constitute a further line of research.

8. References

Q18 Aten, B. 2005. *Report on Interarea Price Levels, 2003*, working paper 2005–11, U.S. Bureau of Economic Analysis. Available at: https://www.bea.gov/system/files/papers/WP2005-11.pdf.

Q19 Aten, B. 2008. *Estimates of State and Metropolitan Areas Price Parities for Consumption Goods and services in the United States, 2005*', U.S. Bureau of Economic Analysis. Available at: https://www.bea.gov/system/files/papers/P2008-1.pdf.

Aten, B.H. 2017. "Regional price parities and real regional income for the United States", *Social Indicators Research*, 131(1): 123–143 DOI: https://doi.org/10.1007/s11205-015-1216-y.

Aten, B.H. 2019. "Case Study of Subnational PPPs: United States", paper presented at the 4th Task force meeting the ICP, World Bank, in May, Paris. Available at: https://pubdocs.worldbank.org/en/148641591895276884/pdf/ICP-Annual-Report-2019.pdf.

Aten, B.H. and E. Figueroa. 2015. "Real Personal Income and Regional Price Parities for States and Metropolitan Areas 2009-2013", *Survey of Current Business*, 95(7): 1–4.

Q20 Behrmann, T., and B. Goldhammer. 2017. "New developments in the field of house and rental price indices in German price statistics in the light of the hedonic method." In: Deutshe Bundesbank-Ottawa Group (2017), 15th Meeting of the Ottawa Group on Price Indices, Hosted by the Deutshe Bundesbank, 10–12 May, Eltville am Rhein, Germany.

Available at: https://www.bundesbank.de/resource/blob/635276/865b648d8fb3d6a23-d327e59c2f48040/mL/2017-05-10-ottawa-group-02-2-abstract-data.pdf.

Biggeri, L., T. Laureti, and F. Polidoro. 2017. "Computing sub-national PPPs with CPI data: an empirical analysis on Italian data using country product dummy models", *Social Indicators Research* 131(1): 93–121. DOI: https://doi.org/10.1007/s11205-015-1217-x.

Q21 Biggeri, L., and M. Pratesi. 2017. "Monetary poverty indicators at local level: definitions, methods of estimations and comparisons in real terms." In Proceedings of the 60th World Statistics Congress of the International Statistical Institute, 2017. Available at: http://www.centrodagum.it/wp-content/uploads/2018/07/Biggeri-Pratesi-ISI2017.pdf.

Q22 Biggeri, L., and T. Laureti. 2018. "Publications, experiments and projects on the computation of spatial price level differences in Italy", paper presented at the 3rd Task force meeting the ICP, World Bank, 27th September, Country case studies: Italy. Available at: www.worldbank.org/en/programs/icp/brief/tf03-doc03.

Billard, L, and E. Diday. 2000. *Regression analysis for interval-valued data, In Data Analysis, Classification, and Related Methods*. Springer: 369–374. DOI: https://doi.org/10.1007/978-3-642-59789-3_58.

Q23 Billard, L., E. Diday. 2002. "Symbolic regression analysis, In: Classification, Clustering and Data Analysis", In Proceedings of the Eighteenth Conference of the International Federation of Classification Societies (IFCS'02), Springer: 281–288. DOI: https://doi.org/10.1007/978-3-642-56181-8_31.

Q24 BIS. 2019. *Mind the data gap: commercial property prices for policy*, Irving Fisher Committee on Central Bank Statistics, Bank for International Settlements. Available at: https://www.bis.org/ifc/publ/ifc_report_cb_trade_rep_deriv_data.pdf.

Q25 Bishaw A. 2009. "Adjusting Poverty Thresholds Based on Differences in Housing Cost: Application of American Community Survey", Poster Presentation Prepared for the Population Association of America April 30–May 2, 2009.

Q26 Bishop J.A., J. Lee, and L.A. Zeager. 2017. "Incorporating spatial price adjustments in US public policy analysis", *ECINEQ WP*, 438. Available at. www.ecinq.org.

Brito, p. 2014. "Symbolic Data Analysis: another look at the interaction of Data Mining and Statistics", *Data Mining and Knowledge Discovery* 4(4): 281–295. DOI: https://doi.org/10.1002/widm.1133.

Can, A. 1992. "Specification and estimation of hedonic housing price models", *Regional Science and Urban Economics* 22: 453–474. DOI: https://doi.org/10.1016/0166-0462(92)90039-4.

Q27 Citro C.F., T.T. Michael. 1995. *Measuring Poverty: A New Approach*, National Academy Press, Washington, D.C. Available at: https://library.econ.upd.edu.ph/ir/handle/123456789/6535.

Cobb, S.A. 1984. "The impact of site characteristics on housing cost estimates", *Journal of Urban Economics*, 15: 26–45. DOI: https://doi.org/10.1016/0094-1190(84)90021-4.

Q28 Dalaker, J. 2017. P*overty in the United States in 2016*: In Brief, U.S. Congressional Research Service, 7–5700.

Q29 Dalén, J. 2006. "Spatial price comparisons in poverty measurement, An example from Cambodia". Ninth Meeting of the Ottawa Group-International Working Group on Price Indices, London. Available at: https://www.ottawagroup.org/Ottawa/

ottawagroup.nsf/home/Meeting + 9/$file/2006%209th%20Meeting%20-%20Jorgen%20Dalen%20-%20Spatial%20price%20comparisons%20in%20poverty%20measurement.pdf.

Q30 De Haan, J. and E. Diewert. 2013. "Hedonic Regression Methods", In OECD et al. *Handbook on Residential Property Price indices*, Eurostat, Luxembourg: chap. 5. Available at: https://ec.europa.eu/eurostat/documents/3859598/5925925/KS-RA-12-022-EN.PDF.

Q31 Eurostat. 2013. *Detailed Technical Manual on Owner-Occupied Housing for Harmonised Index of Consumer Prices*, Eurostat, Luxembourg. Available at: https://ec.europa.eu/eurostat/documents/7590317/0/Technical-Manual-OOH-HPI-2017/.

Q32 Ferrante C., T. Laureti, and F. Polidoro. 2019. "Combining data coming from scanner, traditional CPI data collection and other sources to compile sub-national PPPs in Italy". poster presented at the Sixtheenth Meeting of the Ottawa Group, 8–10 May, Rio de Janeiro.

Q33 Follain, J.R., and S. Malpezzi. 1980. "Dissecting housing Value and rent: Estimates of hedonic indexes for thirty-nine large SMSAs", *Urban Institute Press*, Washington D.C.

Q34 Gillingam, R. 1975. "Place to Place Rent Comparisons", *Annals of Economic and Social Measurement*" 4(1): 153–173. N.B.E.R. Available at: http://www.nber.org/books/aesm75-1.

Q35 Ghirardo, E. 2013. *Aggregazione di coefficient di un modello edonico, un approccio di regressione geografica ponderata*, Quaderni OMI, Roma.

Q36 Ghirardo, E., R. Ciula, and M. Festa. 2014. *Effetti di non linearità nei modelli edonici: un esercizio di regressione quantilica*, Quaderni OMI, Roma.

Greene, W.H. 2012. "Econometric Analysis", 7th edition, Upper Saddle River, NJ: Prentice Hall.

Griliches, Z. 1961. "Hedonic price indexes for automobiles: an econometric analysis of quality changes", In *The Price Statistics of Federal Government*, Washington D.C: Government Printing Office. DOI: https://doi.org/10.4159/harvard.9780674592582.c4.

Griliches Z. 1971. "Introduction: Hedonic Price Indexes Revisited", In *Price Indexes and Quality Change*, edited by Z. Griliches: 3–15, Cambridge MA: Harvard University Press.

Q37 Herath, S. and G. Maier. 2010. "The hedonic price method in real estate and housing market research: a review of the literature". Presented at a Conference and published as Faculty of Business paper of the University of Wollongong, Australia. Available at: http://ro.uow.edu.au/buspapers/971.

Hill, R.J. 2013. "Hedonic price indexes for residential housing: A survey, evaluation and taxonomy", *Journal of economic surveys*, 27(5): 879–914. DOI: https://doi.org/10.1111/j.1467-6419.2012.00731.x.

Q38 Istat. 2010. *La differenza nel livello dei prezzi al consumo tra i capoluoghi delle regioni italiane, Anno 2009*, Roma.

Q39 Istat. 2017. *Spese per consumi delle famiglie, Anno 2017*, Statistiche report 2018.

Jolliffe, D. 2006. "Poverty, prices and place: How sensitive is the spatial distribution of poverty to cost of living adjustments?" *Economic Inquiry*, 44(2): 296–310. DOI: https://doi.org/10.1093/ei/cbj016.

Kokoski, M.F., P. Cardiff, and B.R. Moulton. 1994. *Interarea price indices for consumer goods and services: An hedonic approach using CPI data*, BLS working paper 256, U.S. Bureu of Labor Statistics, Washington D.C. DOI: https://doi.org/10.1016/0304-4076(94)01648-J.

Q40

Kokoski, M.F., Moulton, Moulton, B.R. 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*, edited,by A. Heston and R. Lipsey: 123–166, University of Chicago Press.

Lancaster K.J. 1966. "A New Approach to Consumer Theory", *Journal of Political Economy* 74-13257. DOI: https://doi.org/10.1086/259131-.

Laureti, T., and D.P. Rao. 2018. "Measuring Spatial Price Level Differences within a Country: Current Status and Future Developments", *Estudios de Economía Aplicada*, 36(1): 119–148. DOI: http://dx.doi.org/10.25115/eea.v36i1.2520.

Q41 Laureti, T., and F. Polidoro. 2018. "Big data and spatial comparisons of consumer prices Testing the use of scanner data for computing sub-national Purchasing Power Parities in Italy", In proceeding of 49th Scientific Meeting of the Italian Statistical Society, Palermo.

Lopez, E., and G.J.D. Hewings. 2018. "Housing price indices for small spatial units", *Regional Science and Urban Economics*. DOI: 10.1016/j.regsciurbeco.2018.02.005.

Malpezzi S. 2003. "Hedonic Pricing Models: A Selective and Applied Review." In *Housing Economics: Essays in Honor of Duncan Maclennan*, edited by A. O'Sullivan and K. Gibb: 67–89, Blackwell: Malder MA. DOI: https://doi.org/10.1002/9780470690680.ch5.

Marshall, E.C., and D.J. Spiegelhalter. 1998. "Reliability of league tables in in vitro fertilisation clinics: Retrospective analysis of live birth rates". *British Medical Journal*, 316: 1701–1704. DOI: https://doi.org/10.1136/bmj.316.7146.1701.

Martin, T., B. Aten, and E. Figueroa. 2013. "Estimating the Price of Rents in Regional Price Parities", BiblioGov. Available at: https://citeseerx.ist.psu.edu/viewdoc/download? doi = 10.1.1.440.3550&rep = rep1&type = pdf.

Moulton, B.R. 1995. "Interarea Indexes of the cost of shelter using hedonic quality adjustment Techniques", *Journal of Econometrics* 68: 181–204. DOI: https://doi.org/10.1016/0304-4076(94)01648-J.

Neto, E.D.A.L., and F.D.A. de Carvalho. 2008. "Centre and range method for fitting a linear regression model to symbolic interval data", *Computational Statistics and Data Analysis*, 52(3): 1500–1515. DOI: https://doi.org/10.1016/j.csda.2007.04.014.

Neto, E.D.A.L., and F.D.A. de Carvalho. 2010. "Constrained linear regression models for symbolic interval-valued variables", *Computational Statistics & Data Analysis*, 54(2): 333–347. DOI: https://doi.org/10.1016/j.csda.2009.08.010.

Q42 Ohta, M., and Z. Griliches. 1975. "Automobile prices revisited: Extensions of the hedonic price hypothesis" In *Household Production and Consumption, Studies in Income and Wealth*, edited by N.E. Terleckyi, 40. University of Chicago Press, for NBER, Chicago. Available at: http://www.nber.org/books/terl76-1.

Ozanne, L., and T. Thibodeau. 1983. "Explaining metropolitan housing price differences", *Journal of Urban Economics* 13: 51–66. DOI: https://doi.org/10.1016/0094-1190(83)90045-1.

Parker, D. 2016. "International Valuation Standards. A Guide to the evaluation of Real Property Assets", Wiley Blackwell, London.

Q43 Primont, D.F., and M.F. Kokoski. 1990. *Comparing Prices across cities: A hedonic approach*, BLS working paper 204, U.S. Bureau of Labor Statistics, Washington, D.C. Available at: https://www.bls.gov/osmr/research-papers/1992/ec921210.htm.

Rao, P.D., and G. Hajargasht. 2016. "Stochastic approach to computation of purchasing power parities in the International Comparison Program (ICP)". *Journal of Econometrics*, 191(2): 414–425. DOI: https://doi.org/10.1016/j.jeconom.2015.12.012.

Q44 Renwick, T. 2009. *Alternative Geographical Adjustments of U.S. Poverty Thresholds: impact on State Poverty Rates*, U.S. Census Bureau. Available at: http://citeseerx.ist.psu.edu/viewdoc/download?doi = 10.1.1.227.4438&rep = rep1&type = pdf.

Rosen S. 1974. "Hedonic Prices and Implicit Markets: Product Differentiation in Pure Competition", *Journal of Political Economy* 82(1): 34–55. DOI: https://doi.org/10.1086/260169.

Q45 Silver, M. 2011. "House Price Indices: Does Measurement Matter?" *World Economics* 12(3): 69–86. Available at: https://www.researchgate.net/profile/Mick-Silver/publication/227641359_House_Price_Indices_Does_Measurement_Matter/links/53eb9d620cf24f 241f129611/House-Price-Indices-Does-Measurement-Matter.pdf.

World Bank. 2013. *Measuring the Real Size of the World Economy*, World Bank, Washington D.C. DOI: https://doi.org/10.1596/978-0-8213-9728-2.

Zellner, A. 1963. "Estimators for seemingly unrelated regression equations: Some exact finite sample results", *Journal of the American Statistical Association* 58: 977–992. DOI: 10.1080/01621459.1963.10480681.

Received August 2020
Revised December 2020
Accepted April 2021

Author Queries

JOB NUMBER: Laureti

JOURNAL: JOS

Q1 Author: NOTE No footnotes. Either remove or embed within body text /JOS Office

Q2 Author: NOTE No footnotes. Either remove or embed within body text /JOS Office

Q3 Author: NOTE No footnotes. Either remove or embed within body text /JOS Office

Q4 Author: NOTE No footnotes. Either remove or embed within body text /JOS Office

Q5 Author: NOTE No footnotes. Either remove or embed within body text /JOS Office

Q6 Author: NOTE No footnotes. Either remove or embed within body text /JOS Office

Q7 Author This reference " Kokoski et al. 1984 " is missing in the reference list /JOS

Q8 Author: NOTE No footnotes. Either remove or embed within body text /JOS Office

Q9 Author: NOTE No footnotes. Either remove or embed within body text /JOS Office

Q10 Author: This reference "OMI 2017" is missing in the reference list /JOS Office

Q11 Author: This reference "OMI 2017" is missing in the reference list /JOS Office

Q12 Author: NOTE No footnotes. Either remove or embed within body text.The reference "Istat 2020" is missing in the reference lst /JOS Office

Q13 Author: NOTE No footnotes. Either remove or embed within body text /JOS Office

Q14 Author: NOTE No footnotes. Either remove or embed within body text /JOS Office

Q15 Author: NOTE No footnotes. Either remove or embed within body text /JOS Office

Q16 Author: NOTE No footnotes. Either remove or embed within body text /JOS Office

Q17 Author: NOTE No footnotes. Either remove or embed within body text. The reference "Leckie and Goldstein (2011)" is missing in the reference list/JOS Office

Q18 Author Add accessed date /JOS Office

Q19 Author Add accessed date /JOS Office

Q20 Author Please add acessed date /JOS Office

Q21 Author Please add date and place for this congress. Add also accessed date /JOS Office

Q22 Author Please add place and accessed date /JOS

Q23 Author Please add date and place /JOS

Q24 Author Please add accessed date/JOS

Q25 Author. Please add date place, and DOI or URL /JOS Office

Q26 Author Please check the link and add accessed date /JOS

Q27 Author Add accessed date /JOS

Q28 Author. Please provide DOI or URL /JOS

Q29 Author Add date for the meeting and accessed date /JOS

Q30 Author Add accessed date /JOS

Q31 Author Add accessed date /JOS

Q32 Author Please add DOI or URL /JOS Office

Q33 Author Add DOI or URL/JOS

Q34 Author: Add accessed date /JOS

Q35 Author Add DOI or URL /JOS Office

Q36 Author Add DOI or URL /JOS Office

Q37 Author please add date and place for the conference. Add also accessed date /JOS

Q38 Author Povide DOI or URL /JOS Office

Q39 Author Povide DOI or URL /JOS Office

Q40 Author. This reference “Koloski et al. 1994” is missing in the reference list /JOS Office

Q41 Author Please provide date and DOI or URL /JOS Office

Q42 Author Add accessed date /JOS

Q43 Author: Please add accessed date /JOS

Q44 Author Add accessed date /JOS

Q45 Author Add accessed date /JOS