

REPEAT ASSESSED VALUES MODEL FOR HOUSING PRICE INDEX

Manuela Carini, ass.prof., PhD

University of Calabria, 87036 Rende (CS), Italy
e-mail: manuela.carini@unical.it

Marina Ciuna, ass. prof., PhD

University of Palermo, 90100 Palermo, Italy
e-mail: marina.ciuna@unipa.it

Manuela De Ruggiero, PhD

University of Calabria, 87036 Rende (CS), Italy
e-mail: manueladeruggiero@gmail.com

Francesca Salvo, ass. prof., PhD

University of Calabria, 87036 Rende (CS), Italy
e-mail: francesca.salvo@unical.it

Marco Simonotti, prof

University of Palermo, 90100 Palermo, Italy
e-mail: marco.simonotti@unipa.it

Abstract

This study proposes an innovative methodology, named Repeat Appraised Price Model (RAV), useful for determining the price index numbers for real estate markets and the corresponding index numbers of hedonic prices of main real estate characteristics in the case of a lack of data. The methodological approach proposed in this paper aims to appraise the time series of price index numbers. It integrates the principles of the method of repeat sales with the peculiarities of the Hedonic Price Method, overcoming the problem of an almost total absence of repeat sales for the same property in a given time range; on the other hand, the technique aims to overcome the limitation of the repeat sales technique concerning the inability to take into account the characteristics of individual properties.

Key words: *real estate market, price index numbers, appraisal system approach, hedonic prices.*

JEL Classification: R30.

Citation: Carini M., Ciuna M., De Ruggiero M., Salvo F., Simonotti M., 2017, *Repeat Assessed Values Model For Housing Price Index*, Real Estate Management and Valuation, vol. 25, no. 4, pp. 25-39.

DOI: 10.1515/remav-2017-0027

1. Introduction

For investors, buyers and sellers, it is important to be aware of price trend levels in a given geographical area to know and predict the dynamics of market prices and the phase of the real estate cycle. These trends can be explained through information arising from knowledge of real estate prices, and the analysis of these values (recorded over time) is useful in creating time series of index numbers, which are numerical indicators helpful for understanding the profitability of land phenomena in different market segments (CASE and QUIGLEY 1991).

In their general meaning, the index numbers are useful indicators for making predictions, taking decisions, and studying price movement trends in various sectors of the economy. These are frequently used in the analysis of time series and, in particular, the historical study of long-term trends, seasonal variations, and cyclical developments (FREUD and WILSON 1997). In real estate, the construction of index numbers of real estate prices in different market segments provides important information about real estate trends, investment profitability, and capital appreciation/depreciation. For real estate, the index number time series is exhaustive if expressed on a fixed basis. These real estate numbers express the percentage change of a variable over time compared to a fixed period called the base period.

The present research allows for building the index numbers of unit prices of a building area in complex real estate contexts and proposes a methodology, named Repeat Assessed Values model (RAV), that is useful in determining prices and corresponding hedonic price indexes of real estate market characteristics. The RAV procedure refers conceptually to the repeat price model proposed by Bailey, Muth and Nourse (BMN) (BAILEY et. al., 1963; SHILLER 1993), considering, at the same time, the incidence of real estate hedonic prices, as indicated by the hedonic price method (HPM) and Appraisal System Approach (ASA) (SIMONOTTI 2006) from an operational point of view. In this work, index numbers are obtained from the time series of market prices by following the proposed methodological approach, with the aim of overcoming the operational difficulties in the application of the repeat sales technique and the need for a large amount of data required to implement the HPM. In some real estate markets, such as the Italian one, the data required in order to apply these techniques is not available, due mainly to:

- 1) the lack of transparency in price formation,
- 2) insufficient competitiveness,
- 3) general viscosity of market prices.

Although the new Real Estate Taxation System introduced by the 2005 Finance Act has improved the resilience of the residential market, the lack of transparency still survives in properties that are not residential.

In light of the above, the proposed methodology (in affinity with the method of repeat sales) evaluates prices by considering a sample of housing units which had been purchased and resold during the specified time frame. Moreover, hedonic prices of specific property characteristics are employed in order to derive the time series of hedonic price index numbers (without large data samples), as it is required by the HPM method. The calculated index numbers are transactions based on the systematic and continuous detection of market prices. By knowing time series, it is possible to describe the evolution of the market segment of a building area by a specific analysis (SALVO 2008; CIUNA, SALVO, D'AMATO 2013; CIUNA, DE RUGGIERO, SALVO 2014). Furthermore, the study focuses on the evaluation of relative rates of appreciation/depreciation, and index numbers of the hedonic prices of real estate properties are also clarified (CIUNA, SALVO 2008).

In order to validate the proposed methodology, the RAV model has been tested in the city of Cosenza, in southern Italy, considering the time period between 1995 and 2013. It must be noted that no national or local statistics about the housing market are available in the reference area of this study. However, official quotes about the trend of the housing market are provided by agencies (such as the Land Registry), which offer ranges of values generally not supported by an adequate appraisal and a deep statistical analysis.

We have to underline that the proposed methodology can be used for real estate types other than residential, upon properly considering the corresponding features and their hedonic prices.

2. Literature review

The issue of limited data quantities in calculating price indexes is well known in real estate literature, thus increasing sample selection bias and heterogeneity.

In particular, Geltner and Pollakowski (GELTNER, POLLAKOWSKI 2007) addressed this issue by applying a ridge regression technique to the repeat-sale index. Although this technique was designed for commercial real estate transactions, it could overcome the data scarcity of residential markets, as this methodology requires only information about prices and sales data.

The Bailey, Muth, and Nourse method (BMN) (BAILEY et al. 1963) employed a linear regression to compute price index values by using log price differences between sales pairs of houses. The log price difference was thought to be equal to the difference in the respective log indexes in addition to a

homoscedastic error term. Therefore, only houses, which had been sold twice were used to calculate the index, while the remaining observations were omitted.

Case and Shiller (CASE, SHILLER 1987, 1989) expanded the BMN setup by assuming that the error terms are heteroscedastic. They reasoned that a time range among sales should increase the variance of the log price differences between sale pairs. To compute the house price index they followed the BMN procedure, but observations were weighted depending on the gap time between sales when constructing estimates.

The Office of Federal Housing Enterprise Oversight (OFHEO) releases a repeat sales index called the House Price Index (HPI), which is based on the BMN method. Like the Case-Shiller method, a heteroscedastic error term is incorporated, but the form of the error term is different in two aspects: i) the fixed component of the variance is not considered in order to avoid problems in the weighting step of the estimation procedure, ii) error terms are not independent across multiple sales of the same house.

In 1991, Standard and Poor's published the S&P/Case-Shiller (S&P/C-S) Home Price Index which is based on the arithmetic index proposed by Shiller (SHILLER 1991). This index uses sale prices instead of log prices and is not based strictly on price differences as is the case for the BMN method. This change is due to an easier interpretation of the index together with a different house appraisal according to their initial sale price.

An alternative method, combining single sales and repeat sales, was proposed by Englund et al. (ENGLUND et al. 1998). It is an additional research about hybrid indexes developed in Quigley (QUIGLEY 1995) and Hill et al. (HILL et al. 1997). This method manages information on all sales, as well as all available data on housing attributes in order to estimate prices trends.

It must be highlighted that a distinctive feature of the repeat sales method is that it completely considers housing quality requiring less data in comparison to hedonic or hybrid methods, under appropriate assumptions. Today, it is common for real estate transactions to be geocoded in order to associate Geographic Information Systems (GIS) latitude and longitude coordinates to each sale. This makes it possible to develop models with finely grained spatial details. For example, spatial models can use distances between each pair of transactions and the GIS can associate Census tracts with geocoded sales. Thus, neighborhood variables (income and other socioeconomic characteristics) can be used as part of local price index estimation.

In light of the abovementioned facts, another interesting work was presented by Pace (PACE 1997) and Pace et al. (PACE et al. 1998) with the development of semi-parametric models. In particular, spatial autoregressive models (explanatory variables include neighboring prices) were built and combined with a temporal autoregressive process (explanatory variables include lagged prices) to produce a spatial/temporal model. In addition, local housing price estimation was also proposed by Goetzmann and Spiegel (GOETZMANN, SPIEGEL 1997) using geocoded data through the distance-weighted repeat sales method.

Nagaraja, Brown, and Zhao (NAGARAJA et al. 2011) proposed an autoregressive index (AR). This index is computed using all sales; however, as described below, repeat sales are considered more important during the index calculation, because it is possible to know more information about a house after it has been sold many times. Conceptually, the autoregressive index is a repeat sales index even though it is not based on BMN methodology. This model is made up of three components: an index, the effect of home location in a particular ZIP code area and autoregressive time series AR (1,0,0), which automatically take into account the time gap between sales.

In many cases, it is necessary to introduce a weight when the exclusion of the outliers from the sample produces either an impressive decrease of data in the sample or a reduction of information. The idea to maintain all the estate data in a sample had already been proposed in the sales price appraisal ratio (SPAR) method (BOURASSA et al., 2006). This method combines information from appraisals and sale prices. On the other hand, Salvo, Ciuna and De Ruggiero (CIUNA, DE RUGGIERO, SALVO 2014) proposed the sales price weighted ratio (SPWR) approach, where the weighting system is based only on sale prices, eliminating the probabilistic component that characterizes the appraised values. The main differences among these methods are:

1. The BMN is the only method which assumes that errors are homoscedastic; in the house price setting, this means that the errors do not depend on the gap time between sales.

2. Case-Shiller (C-S), S&P/C-S and OFHEO indexes weight observations differently depending on assumptions about the variance of the error terms.
3. The S&P/C-S method manages prices instead of differences in log prices between sales of the same house and uses instrumental variables in the analysis.
4. The AR method uses an underlying autoregressive time series approach, including single sales and location information (ZIP code).
5. The SPAR method is an alternative to hedonic or repeat sales indexes. The method is based on matched pairs, even if it uses all price data available for the period under observation.
6. The SPWR is based on the mechanisms of stock index formation and investigates analogies between stock and property information, according to the peculiarities of property trade. Furthermore, it leads to a methodological approach (derived from the Simple Price Index Method) which is able to consider possible anomalies in the collected sample of sales prices using weighting coefficients. These are based on reliability coefficients of property sale prices.

However, the currently available methodologies have some drawbacks seeing as how they are statistical methods (with all their limitations) as opposed to proper appraisal tools. Although the need to perform significance tests is evident, it is important to underline that a statistically correct result may be not always be significant from the point of view of an appraisal. Therefore, alternative indexing models may be more suitable seeing as how they are based on linear equation systems.

Many studies analyze the possibility of using assessed values in the building of index price numbers. CLAPP and GIACCOTTO (1999) appraise price trends by using the sales price, assessed value and date of sale for every residential property transaction between independent parties. This assessed value (AV) methodology is compared to the repeat sales (RS) method. The work develops a simple method of correcting the effect of the measurement errors associated with the assessed value.

Several other authors (BLUNDELL and WARD 1987; GELTNER 1989, 1991, 1993a; QUAN and QUIGLEY 1989, 1991) (CIUNA, SALVO, D'AMATO 2013) have developed models to extract an underlying market value series from the appraisal series. Although the theory (QUAN and QUIGLEY 1989) allows for a time-varying smoothing parameter, it is normal in empirical studies to estimate a constant smoothing parameter. Smoothing might reasonably be assumed to be related to liquidity, where "liquidity" is defined as the volume of trading in the asset market. As liquidity increases, more signal information is available on current transactions, so appraisers place more weight on this relative to past valuations, thus resulting in less smoothing. In literature on the securities market, "liquidity" refers to two related concepts: the speed of conversion to cash and market "depth" (the inverse of price elasticity with respect to trading volume – that is the ability to trade without influencing price) (HOWELLS and BAIN 1990). In the securities market, liquidity is commonly measured by the bid-ask spread or by measures related to market turnover (KLUGER and STEPHAN 1997), but only the latter type of information (on turnover or asset market trading volume) is meaningful and readily available in the private real estate market. MATYSIAK and WANG (1995) apply a rigorous econometric approach, using the results of more extensive testing than has been undertaken to date. Their study supports the claim that valuers may not be using all of the information available at the time of valuation. The results imply that performance measurement and league table rankings may need to be qualified and that the routine application of fixed parameter unsmoothing procedures to valuation-based property indices may be invalid

3. Data and Methods

3.1. Overview

Real estate data represent the essential input for creating real estate indexes; indeed, several methodologies which are based on the direct processing of real estate data are available in appraisal literature. These techniques can be used to create past index price series (Eurostat 2013). Generally, index numbers are indicators of price trend movements in many economic sectors. The construction of index numbers for property prices in different market segments provides important information about real estate trends, direct profitability and the revaluation/devaluation of capital.

The proposed methodology is conceptually based on the hedonic price method, the repeat sales method and the SPAR, because it simulates the mentioned techniques by using the fundamental characteristics of each one, as follows:

- it provides values for the hedonic prices of the real estate features while trying to overcome the difficulties related to data quantities (according to the hedonic price method),
- it conceptually refers to the same property purchased and sold in different periods (according to the repeat sales method) while overcoming the loss of information related to the hedonic price (also making it possible to calculate the time series of the hedonic price index) and the objective to detect properties suitable for the method,
- it simulates the same conceptual approach of the sale price appraisal ratio method with a different operational framework.

The technique proposed in this paper, while based on BMN, differs by being based on appraised values rather than on sales prices. In fact, given that in some markets there aren't repeated sales for the same property, the RAV technique simulates ideal repeated sales over the years of a virtual property for which values are appraised; these values (named appraised prices) take the place of sales prices for calculating index numbers.

From an operational point of view, the RAV is based on the ASA, from which it takes procedures and algorithms (see sub-section 3.3).

3.2. Appraisal System Approach (ASA)

The appraisal system is based on equations related to comparisons between the property being appraised and every single real estate property of a comparable sample (SIMONOTTI 2006). The elementary equation affirms that the price difference between two properties is a function of differences in their characteristics; therefore, it can be stated that equal properties are priced the same.

The method is a formalization of the market comparison approach, intended to appraise the value of the subject and hedonic prices when data are sufficient for solving the linear equation system but not sufficient for implementing regression models.

It is useful in a typical real estate market, such as the Italian one, and allows for appraising every kind of property for which comparison data are sufficient.

The appraisal system implements a set of pairwise comparisons from which an equation system is derived. It can be solved in terms of the subject market value and the hedonic prices of the considered real estate characteristics. The implicit hedonic price of a characteristic is the change in the total price of the property corresponding to unitary variation in the value of a characteristic. Hedonic price values are expressed per unit of each characteristic and are most commonly estimated using statistical analysis. In the real estate context, appraisal sales adjustment grids are special cases of hedonic price application in terms of "adjusted dollars" obtained by multiplying the unit price by the difference between the analyzed properties and comparable properties. Statistical models are independent of the building which is to be evaluated, in contrast with sales adjustment grids.

The system can be solved in terms of both the appraisal value and the hedonic prices of the real estate characteristics. Data collection concerns the following: i) x_{ji} values (amount of each characteristic) with $j = 1, 2, \dots, m$ and $i = 1, 2, \dots, n$, that represent real estate properties and characteristics, respectively; ii) the known market prices - P_j ; iii) the value of the characteristics of the property to be evaluated - x_{0i} . The appraisal method is based on a generic comparison function (1) for which the difference in price between two properties is a function of the differences in the characteristics of the properties. This relation may be written as follows:

$$P_j - V = \sum_{i=1}^n (x_{ji} - x_{0i}) \cdot p_i \quad (1)$$

$$\begin{cases} P_1 = V + \sum_{i=1}^n (x_{1i} - x_{0i}) \cdot p_i \\ P_2 = V + \sum_{i=1}^n (x_{2i} - x_{0i}) \cdot p_i \\ \dots = \dots \\ \dots = \dots \\ P_m = V + \sum_{i=1}^n (x_{mi} - x_{0i}) \cdot p_i \end{cases} \quad (2)$$

In Formula (2), V is the appraised market value of the subject. The resolution of the system is performed by applying the least squares method when the system is determined, otherwise it is performed by applying the Moore-Penrose technique.

3.3. Repeat Assessed Values (RAV) Model

For the purpose of real estate, the time series of index numbers is essentially expressed in fixed base. This type of numbers expresses the percentage changes in a variable period compared to a fixed one (that is considered as reference). The index numbers are based on the detection of market prices and are related to a specific type of property. These monetary indexes are analyzed on an annual basis and according to the systematic detection of market prices (transaction-based). Their calculation, moreover, refers to simple price index methods, which determine the position indexes (average) of the unit market prices in the time range between one index and the next. In this paper, the ASA has been implemented by simulating the BNM technique in order to produce the time series of index numbers of properties prices.

The RAV technique is used to overcome the critical points of the two cited methodologies and, at the same time, to address the appraisal protocol for the construction of index numbers in order to ensure fast and straightforward application.

3.4. Additional considerations regarding BMN methodology

The BMN method considers a sample of repeat sales of properties that were bought and sold in a pre-determined period of time. It is widely applied in the determination of the efficiency of real estate markets, mainly in chattel and artwork markets. This method represents a good tool for the monitoring of prices on the housing market and is able to quantify the percentage differences in prices considering heterogeneous characteristics (ABRAHAM and SCHAUMAN 1991; SHILLER 1991). This approach allows for the characteristics of real estate to be taken into account more explicitly because it is based only on changes in the price of the property between two transactions, under the assumption that the characteristics of the property are unchanged over time. In this way, this index takes into account the heterogeneity related to the structural characteristics and the housing location. The necessary information concerns only the price and timing of transaction, and there is no need for detailed information about the characteristics of houses, which are essential for the appraisal of indexes using the hedonic approach (DREIMAN et al. 2004). Index construction is based on the regression of the logarithmic difference of each unit price between two transactions and on a set of dummy variables indicative of the times at which the transactions took place. In addition to the problems related to the efficiency of the appraisal method, the main limitation of repeat sales models is the considerable loss of information due to the reduction in size of the sample used for the evaluation, as the analysis discards all properties that have not been subjected to at least two transactions in the reference period. This implies a considerable loss of efficiency in the evaluation of the price index.

Moreover, it is hard to assume that the subset of real estate units which had been sold at least twice is representative of the underlying population. Therefore, it is plausible to assume the presence of sample selection bias in the construction of a price index with the repeat sales model (DOMBROW et al. 1997).

In order to produce a series of index numbers, the hedonic price method uses regression models considering the date of sales as an independent variable designed to quantify the rate of real estate revaluation. The repeat sales method is impossible to implement in practice, especially in countries lacking data.

In light of the above, the methodological approach proposed in this paper, which is intended to construct the time series of sale prices for which the index numbers are built, is designed to overcome the operational difficulties connected with using the repeat sales technique. At the same time, it is useful for avoiding the necessity of a considerable number of transaction data to implement the HPM.

3.5. Formulation

The ASA was modified in order to build a model (RAV) with the purpose of defining the price of a virtual property whose characteristics are constant for all of the analyzed years. The RAV model compares the mean detected house prices (comparable) and the unknown average price of a virtual property to determine the desired mean price and hedonic prices of real estate features, in relation to which the series of both price index numbers and hedonic price index numbers can be built.

The application of the RAV to real estate data (referring to individual years) allows for the hedonic price of each characteristic and the unit price of a virtual property with the same real estate features to be valued. In this way, considering that the market price of a single property under investigation is

known for each year, it is possible to draw up time series of index numbers by applying the repeat sales technique. Hedonic prices of the real estate characteristics were considered and processed in order to obtain the time series of hedonic price index numbers.

An overview of real estate market trends is provided by the index numbers of the nominal unit prices and the hedonic price index numbers. The analysis is carried out by organizing a detected sample in a number of sub-samples t , each one including all the transactions recorded in each investigated year. It must be underlined that the methodology defines a virtual real estate reference property whose characteristics are an expression of the average values of the whole sample. From the sample data, the following values are calculated: i) index numbers of the recorded prices; ii) index numbers of the appraised prices; iii) index numbers of the hedonic prices. Index numbers of the recorded prices are calculated starting from the arithmetic mean of total prices recorded for each year and the mean of the corresponding surfaces (see Table 2, Column 3). The index numbers of the appraised prices and appraised hedonic index numbers are obtained by the RAV method.

For each year t , the RAV is presented as a system of linear equations, where each one is a comparison function that is the difference between the average price of a generic detected property \bar{p}_{tj} ($j=1..m$) and the unit price of a virtual property \bar{p}_t . In other words, Equation (3) compares the features of a generic comparable property and those of a virtual one:

$$\bar{p}_{tj} - \bar{p}_t = \sum_{i=1}^n (x_{ji} - x_{0i}) \cdot p_{ti} \quad (3)$$

where m is the number of sales data available in the database, n is the number of features considered in the analysis and p_{ti} are the hedonic prices.

System 2 may be written as follows:

$$\begin{cases} \bar{p}_{t1} = \bar{p}_t + \sum_{i=1}^n (x_{1i} - x_{0i}) \cdot p_{ti} \\ \bar{p}_{t2} = \bar{p}_t + \sum_{i=1}^n (x_{2i} - x_{0i}) \cdot p_{ti} \\ \dots = \dots \\ \dots = \dots \\ \bar{p}_{tm} = \bar{p}_t + \sum_{i=1}^n (x_{mi} - x_{0i}) \cdot p_{ti} \end{cases} \quad (4)$$

where:

$\bar{p}_{t1}, \bar{p}_{t2}, \dots, \bar{p}_{tm}$ are the average prices detected in sub-sample t for the year under consideration; \bar{p}_t is the price (appraised) of a virtual ordinary property with known and constant characteristics for each investigated year; x_{ji} , with indexes $j=1,2,\dots,m$ and $i=1,2,\dots,n$, are the numbers of considered real estate features for the examined properties which are a part of the sub-sample for the given year; x_{0i} represents a characteristic of the virtual property which is the same for each year. Further, each virtual characteristic value is obtained as the arithmetic mean (or frequency in the case ordinal/binary values) of the corresponding quantities collected in the full sample (considering the entire investigated period); p_{ti} is the hedonic price of each real estate characteristic.

In Formula (4), the subject unit price \bar{p}_t and hedonic prices p_{ti} are unknown variables, whereas \bar{p}_{tj} are constant variables. Unknown variables are represented by the s vector:

$$s = \begin{bmatrix} \bar{p}_t \\ p_{t1} \\ p_{t2} \\ \dots \\ p_{tn} \end{bmatrix}, \quad (5)$$

The vector p includes all the constant terms \bar{p}_{tj} :

$$p = \begin{bmatrix} \bar{p}_{t1} \\ \bar{p}_{t2} \\ \dots \\ \bar{p}_{tm} \end{bmatrix} \quad (6)$$

The coefficients are contained in matrix D :

$$D = \begin{bmatrix} 1 & x_{11} - x_{01} & x_{12} - x_{02} & \dots & x_{1n} - x_{0n} \\ 1 & x_{21} - x_{01} & x_{22} - x_{02} & \dots & x_{2n} - x_{0n} \\ \dots & \dots & \dots & \dots & \dots \\ 1 & x_{m1} - x_{01} & x_{m2} - x_{02} & \dots & x_{mn} - x_{0n} \end{bmatrix}. \quad (7)$$

Formula (4) can also be written in a compact form, as follows:

$$D \times s = p \quad (8)$$

Solutions from (8) are found according to the instructions in 3.2.

Once the price \bar{p}_t and hedonic prices p_{it} of the real estate characteristics are obtained for each year by solving the system (4), it is possible to calculate index numbers of property prices and hedonic prices of property characteristics.

As previously specified, index numbers are calculated on a fixed base and express the percentage variations of the real estate phenomenon during a given time range with respect to a fixed reference period.

According to the RAV method, the corresponding series of simple index numbers on a fixed base k is thus defined by the relationship between the sequence of prices \bar{p}_t and \bar{p}_k during the period k . The period is an integer that corresponds to the last year of data collection and represents the reference year of the time series of index numbers. Therefore, the price index of a generic t -th period with respect to the base period is:

$$I_{RVM,t}^{\bar{p}} = \frac{\bar{p}_t}{\bar{p}_k}, \quad (9)$$

The hedonic price index $I_{RVM,t}^p$ of a general i -th characteristic is, on the other hand, obtained as follows:

$$I_{RVM,t}^p = \frac{p_{it}}{p_{ik}}, \quad (10)$$

where: p_{it} is the hedonic price of a given characteristic i at time t , and p_{ik} is the hedonic price of the i characteristic during the reference period k .

4. Case study

The case study was carried out in the city of Cosenza, in Southern Italy. A sample of 1049 sales data on flats in condominiums was considered. Property data was collected from a market survey of real estate agencies, with the aim of detecting properties of the same building types that are located in the same area. In the analyzed sample, all data regard single sales with no repetitions. Transactions refer to the period from January 1995 to June 2013. For each sampled unit, the main characteristics and the price (total price) were noted.

The following characteristics were considered:

- surface area (SUR) in square meters,
- the number of restrooms (RES),
- the floor level (LEV), with an ordinal scale, indicating the level on which the flat is situated (1st, 2nd, 3rd, ...),
- the state of maintenance (MAI), measured on a scale of 1-5 (0 = very poor maintenance, 1 = poor maintenance, 2 = sufficient maintenance, 3 = reasonable state, 4 = good maintenance, 5 = excellent),
- the number of views (VIW), that indicates the number of fronts with openings (windows and balconies) (Table 1).

Table 1

Sample statistics, 1049 Real Estate Data					
Statistics	Minimum	Maximum	Average	Frequency	Standard deviazione
Total Price PRC (€)	30,987.00	371,849.00	123,516.59	-	54,647.50
Average Price PRA	245.87	2,246.56	961.12	-	312.89

(€/sqm)					
Surface (SUR) (sqm)	43.50	345.80	128.93	-	37.15
Restrooms (RES) n° (1-2-3-4)	1	4	-	429-489-124-7	-
Level (LEV) (1-2-3-4- 5-6-7)	1	7	-	229- 216-192- 160-117-106-29	-
Maintenance (MAI) (0-1-2-3-4-5)	0	5	-	70- 90-43-108- 616-122	-
View (VIW) (1-2-3-4- 5)	1	5	-	96-529-387-22-15	-

Source: Own study.

5. Empirical results

For the application of the proposed methodology, the tested sample is divided into nineteen subsamples related to each year (Table 2).

Table 2

Number of data and characteristics SUR, PRA for the 19 subsamples
(SUR = Surface; PRA = Average detected prices)

Year	Data number	SUR sqm	PRA €/sqm
2013	30	154.40	743.31
2012	38	150.00	666.98
2011	40	150.00	703.46
2010	60	123.84	729.57
2009	80	116.18	777.16
2008	64	121.52	803.09
2007	47	122.11	877.22
2006	36	142.26	908.25
2005	41	138.28	940.41
2004	65	127.20	1,004.09
2003	87	143.13	810.79
2002	90	136.02	1,062.37
2001	93	141.81	1,313.49
2000	52	111.81	1,401.22
1999	51	123.04	1,319.53
1998	68	130.30	1,350.77
1997	44	124.38	1,302.60
1996	35	102.22	1,443.88
1995	28	104.61	1,412.64

Source: Own study.

The real estate data are collected through special documents that summarize all the peculiarities of these types of property. The characteristics represented by: the surface, contract date, number of restrooms, floor levels, state of maintenance and number of views are the key characteristics considered in the descriptive statistics for the entire sample. From these data, it is possible to calculate the index numbers of the mean prices and the index numbers of the hedonic prices of each real estate characteristic (using 1995 as the base year).

For each year t , with $t = 1995, \dots, 2013$, the RAV method is presented as follows:

$$\begin{cases} PRA_{t1} = \bar{p}_t + (RES_1 - RES_0) \cdot p_{RES} + (LEV_1 - LEV_0) \cdot p_{LEV} + (MAI_1 - MAI_0) \cdot p_{MAI} + (VIW_1 - VIW_0) \cdot p_{VIW} \\ PRA_{t2} = \bar{p}_t + (RES_2 - RES_0) \cdot p_{RES} + (LEV_2 - LEV_0) \cdot p_{LEV} + (MAI_2 - MAI_0) \cdot p_{MAI} + (VIW_2 - VIW_0) \cdot p_{VIW} \\ \vdots \\ PRA_{tm} = \bar{p}_t + (RES_m - RES_0) \cdot p_{RES} + (LEV_m - LEV_0) \cdot p_{LEV} + (MAI_m - MAI_0) \cdot p_{MAI} + (VIW_m - VIW_0) \cdot p_{VIW} \end{cases}, \quad (11)$$

where: $PRA_{t1,t2,\dots,tm}$ [€/sqm] are the unit prices for the detected sales data of sub-sample t in the year under investigation;

\bar{p}_t [€/sqm] is the unknown unit price of the virtual reference property, which has known and fixed characteristics for each year under investigation (referred to the year t);

$RES_{1,\dots,m}$, $LEV_{1,\dots,m}$, $MAI_{1,\dots,m}$, $VIW_{1,\dots,m}$ represent the following characteristics: number of restrooms, floor level, state of maintenance and number of views of each real estate property of the sub-sample for the considered year, respectively;

RES_0 , LEV_0 , MAI_0 , VIW_0 represent the same characteristics in the previous list for the analyzed property (the virtual one), which are constant for each year. Moreover, each one is obtained as the arithmetic mean (or frequency) of the corresponding quantities referred to the full sample;

p_{RES} , p_{LEV} , p_{MAI} , p_{VIW} are the hedonic prices of the same characteristics expressed in €/sqm-no., €/sqm-level, €/sqm-point, €/sqm-nr., respectively.

For each year, the appraisal model provides the price of the reference property and the hedonic prices of characteristics. Once unitary prices of a reference property and hedonic prices of real estate characteristics are known (for each year), it is possible to build the time series of index numbers of mean appraised prices (Table 3, Column \bar{p}) and the time series of hedonic price index numbers (Table 4, columns RES , LEV , MAI , $VIEW$). Index prices calculated based on sales unit prices detected in the sample are reported in Table 4 (Column PRA).

Table 3

Appraised values for each year
(*pt. average appraised price for the reference property; RES restrooms hedonic price; LEV floor level hedonic price; MAI State of Maintenance hedonic price; VIW Number of views hedonic price*)

Year	p_t (€/sqm)	RES (€/sqm-n°)	LEV (€/sqm-level)	MAI (€/sqm-point)	VIW (€/sqm-n°)
1995	463.87	21.52	25.83	19.03	56.75
1996	524.55	56.97	22.43	41.18	77.90
1997	692.50	66.79	29.03	23.39	25.43
1998	652.15	94.47	13.53	33.30	50.01
1999	749.70	29.69	16.04	20.50	16.10
2000	828.12	14.47	11.91	21.50	34.17
2001	800.26	6.68	19.07	39.64	18.08
2002	850.04	19.17	13.83	36.27	46.08
2003	777.82	24.00	36.19	66.19	28.72
2004	924.58	13.74	17.63	31.62	34.60
2005	1,151.13	26.40	56.32	38.12	38.50
2006	1,065.16	55.84	51.25	97.62	14.18
2007	1,150.83	59.51	14.89	76.27	11.01
2008	1,426.23	33.48	18.09	72.24	10.66
2009	1,220.72	60.67	51.40	12.22	28.79
2010	990.62	85.48	12.80	132.26	53.87
2011	1,529.15	64.20	13.18	117.77	54.00
2012	1,262.81	48.86	41.22	89.48	35.00
2013	1,276.20	37.33	30.13	57.89	40.00

Source: Own study.

Table 4

Index numbers for each year
(*RES restrooms hedonic price; LEV floor level hedonic price; MAI State of Maintenance hedonic price; VIW Number of views hedonic price; pt index prices based on appraised prices; PRA index prices based on detected prices*)

Year	RES	LEV	MAI	VIW	p_t	PRA
1995	100.00	100.00	100.00	100.00	100.00	100.00

1996	264.72	86.83	216.43	137.27	113.08	89.73
1997	310.34	11236	122.92	44.81	149.29	94.64
1998	438.98	52.37	175.04	88.12	140.59	98.15
1999	137.95	62.07	107.77	28.36	161.62	104.55
2000	67.22	46.10	113.01	60.21	178.52	108.04
2001	31.04	73.81	208.34	31.86	172.52	118.02
2002	89.07	53.52	190.61	81.19	183.25	122.19
2003	111.51	14.10	347.90	50.61	167.68	126.52
2004	63.86	68.26	166.16	60.97	199.32	135.08
2005	122.65	218.00	200.34	67.84	248.16	109.08
2006	259.46	198.38	513.08	24.98	229.62	142.92
2007	276.53	57.62	400.85	19.40	248.09	176.71
2008	155.55	70.02	379.66	18.78	307.46	188.51
2009	281.91	198.99	673.87	50.74	263.16	177.52
2010	397.20	49.55	695.10	94.92	213.56	181.72
2011	298.31	51.04	618.98	95.16	329.65	175.24
2012	227.04	159.58	470.27	61.68	272.23	194.25
2013	173.46	116.63	304.26	70.49	275.12	190.05

Source: Own study.

The presented index numbers are direct, annual and fixed based (year 1995). During the investigated time period (1995-2013), the real index numbers of unit prices show an upward trend, as do, for the most part, the index numbers of appraised unit prices.

The appraised index numbers are lower than those based on real mean prices (the whole trend appears smoother) and are characterized by a reliability quite superior to that of the latter, since the conceptual and methodological relevance of estimated price construction is considered. It is evident that the appraised price index numbers and those referring to real prices show a different trend, which is smooth for the first and not regular for the latter. This observed behavior can be related to the following statement: whereas index numbers of appraised prices are suitably constructed based on unit prices and refer to the same property, real prices refer to a set of different properties in a given time range, showing a formal contradiction. Once the problem of the lack of data on repeat sales is solved, index numbers relating to the appraised prices are accurate from a methodological point of view, as they are related to changes in the price of the same property over the years as this work aims to demonstrate. Therefore, these index numbers are a valuable reference in several situations in which it is necessary to carry out a property revaluation (Figure. 1).

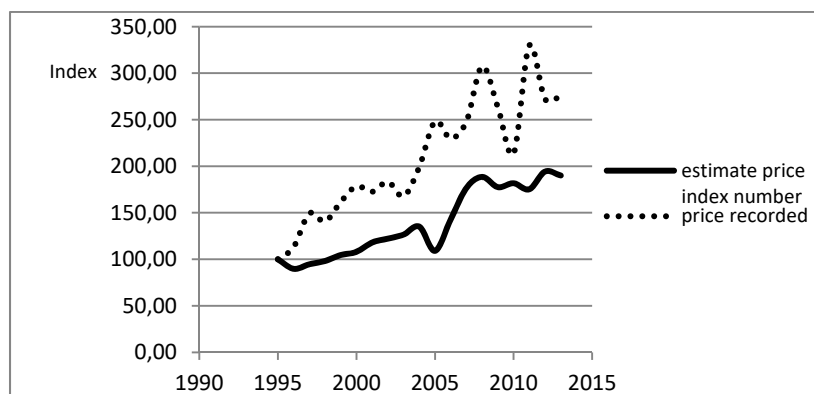


Fig. 1. Comparison between the prices of index numbers based on recorded prices and estimated values. Source: own study.

For the restrooms characteristic, the following trend was observed (see Fig. 2): first, a steady increase until 1998, followed by a substantial decrease until 2001 and, once again, a gradual increase until 2010. In 2011, a constant decline was detected; after that, a period of growth was registered (Figure 2).

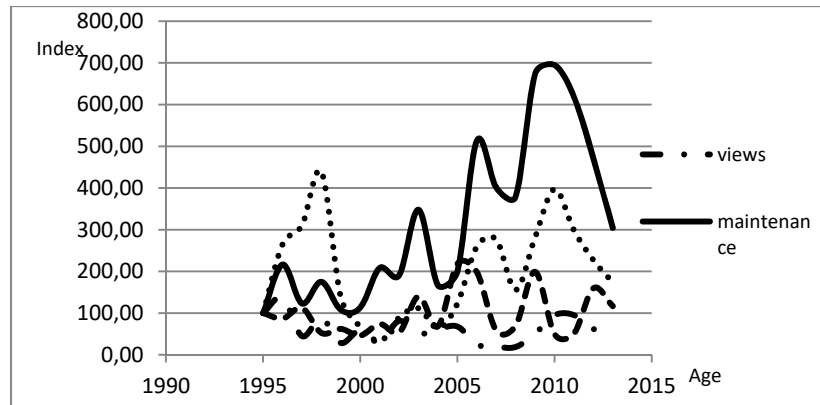


Fig. 2. Hedonic price index numbers. Source: own study.

Moreover, the following analysis ought to be highlighted from index numbers trend evolution:

- for the floor level characteristic, an alternating trend was observed, thus confirming the poor interest of buyers in this feature,
- for maintenance, an upward trend until 2010 was registered, followed by a decreasing trend,
- for the characteristic related to views, a decreasing trend until 2007 was detected, followed by an increasing trend in the subsequent period.

The index numbers calculated based on appraised prices have a lower variability level when compared to those obtained from surveyed prices, although they maintain significant fluctuations. These are not due to defects in the approach but to the database data quality, which belongs to a viscous, opaque and not very competitive market (such as the Italian one).

Figure 2 shows that the maintenance status of the investigated flat is the most appreciated feature of a property, probably because the purchase of a property with a better state of maintenance does not require renovation with consequent additional costs to the sale price. Also the convenience of an additional restroom is a very desirable feature, whereas the floor level and the number of views are not particularly valued in an urban environment, such as the city of Cosenza. Here, in fact, buildings are of moderate height (up to 7 floors) and characterized by a rather uniform number of views.

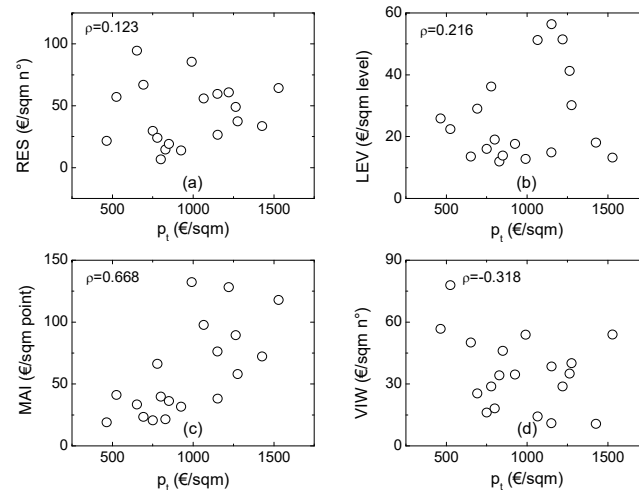


Fig. 3. Correlation plots between the reference price and the hedonic price index numbers. Only the hedonic index associated with the maintenance status is significantly correlated with the reference price. Source: own study.

In addition, we performed a correlation analysis between the reference price and the hedonic index numbers as a tool for understanding the purchaser profile. The results have been reported in Figure 3. Only the hedonic index of the maintenance status exhibits a significant correlation (correlation coefficient $\rho = 0.668$) with the reference price. This correlation suggests that in the years characterized by a higher reference price, the purchaser gives more consideration to the maintenance status of a building. On the other hand, the low correlation observed between the reference price and the hedonic

index associated with the number of restrooms ($\rho = 0.123$), floor level ($\rho = 0.216$), and number of views ($\rho = -0.318$) simply suggests that the weight given to these features by the purchaser when choosing a property is rather independent of the current status of the reference price.

6. Discussion and conclusions

Prices index numbers are a key measure for monitoring real estate profitability. Unfortunately, despite the importance of these indicators, it is widely recognized that, nowadays, only index numbers calculated based on real estate quotations may be available in some countries.

This paper proposed an operational methodology designed to overcome the problem of the lack of data, specifically of those relating to repeat sales, indicating a procedure for the construction of prices index numbers for the real estate market. Their construction is possible with the availability of a sufficient number of real estate transactions recorded during the years.

The methodology was applied in the city of Cosenza (in Southern Italy) using a sample of data on 1,049 flat sales, with transactions carried out in the period between January 1995 and June 2013. The sample was divided into nineteen subsamples relating to each year. The following real estate characteristics were considered in the descriptive statistics for the entire sample: contract date, number of restrooms, floor level, state of maintenance and number of views.

From the sample data, it was possible to calculate index numbers of recorded and appraised mean prices and index numbers of the hedonic prices of the real estate characteristics. The price index numbers were calculated on the basis of the arithmetic mean of the total prices recorded for each year and the arithmetic mean of the corresponding flat surfaces. The index numbers of appraised prices and the hedonic index numbers were determined with the use of appraisal methodology known as the appraisal system approach.

The appraisal model, used in order to build the historical series of index numbers, provided the estimated price of a virtual property considered as a subject with fixed characteristics (for each year) and hedonic prices of real estate characteristics, simulating the repeat sales technique.

The appraisal test employed in the time series analysis of real estate data demonstrated the suitability of the model for expressing the effects of the characteristics and their variation over time, as measured by the index numbers. The model presents some advantages when compared to direct hedonic models due to the absence of temporal variables. The proposed model overcame the problems of the indirect hedonic models in terms of regression instability, as it is constructed in different periods of time, and allowed for the analysis to be implemented also in the case of a lack of repeated data.

The method integrated the principles of the repeat sales method with the peculiarities of the hedonic price method, overcoming the problem of the almost total absence of repeat sales for the same property in some real estate markets, such as the Italian one. On the other hand, the proposed technique avoids the intrinsic limitations of the repeat sales technique, which are related to the inability to take into account the characteristics of individual properties.

We have also proposed a correlation analysis between the reference price and the hedonic index numbers associated with the important features of the building as a tool for understanding the purchaser profile. The obtained results indicate that the only building parameter which is significantly correlated with the reference price is the state of maintenance, thus suggesting that, when the reference price is higher, the purchaser gives more attention to the maintenance status of a property.

Finally, the method conceived in the case study characterized by flats in condominiums in the designated area of Cosenza may be configured as standardized appraisal methodology, which can be applied to any geographical context and property type, if real estate data are systematically and continuously recorded. In the absence of high data quality, no appraisal methodology can be applied, although it is formally correct.

7. References

- ABRAHAM J.M. SCHAUMAN W., 1991, *New Evidence on House Prices from Fannie Mac Repeat Sales*, AREUEA Journal, No. 19.
- BAYLEY, M.J., MUTH, R.F., NOURSE H.O., 1963, *A Regression Model for Real Estate Price Index Construction*, Journal of American Statistical Association, No. 58, pp. 933-942.

- BLUNDELL, G., WARD, C., 1987, *Property Portfolio Allocation: a Multi-Factor Model*, Land Development Studies, Vol. 4, pp. 145-56.
- BOURASSA, S.C., HOESLI M., SUN J., 2006, *A Simple Alternative House Price Index Method*, Journal of Housing Economics, 15, 80-97.
- CASE B., QUIGLEY J., 1991, *The Dynamics of Real Estate Prices*, Review of Economics and Statistics, No. 73.
- CASE K. E., SHILLER R. J., 1987, *Prices of Single-Family Homes Since 1970: New Indexes for Four Cities*, New England Economics Review, 45-56.
- CASE K. E., SHILLER R. J., 1989, *The Efficiency of the Market for Single Family Homes*, American Economics Review, 79, 125-137.
- CIUNA M., SALVO F., D'AMATO M., 2013, *The Appraisal Smoothing in the Real Estate Indexes*, in Dynamics of Land Values and Agricultural Policies, MediaMond International Proceedings, Bologna, 63-71.
- CIUNA M., DE RUGGIERO, M. SALVO, F., 2014, *Property Prices Index Numbers and Derived Indices*, Property Management, Vol. 32 (2), pp. 139-153.
- CLAPP J. M., GIACCOTTO C., 1992, *Estimating Price Trends for Residential Property: A Comparison of Repeat Sale and Assessed Value Methods*, Journal of the American Statistical Association, Vol. 87, issue 418, pp. 300-306
- CIUNA M., SALVO F., 2008, *Numeri indici dei prezzi immobiliari*, Estimo e Territorio, No. 12, pp. 30-47.
- DOMBROW J., KNIGHT J.R., SIRMANS C.F., 1997, *Aggregation Bias in Repeat-Sales Indices*, Journal of Real Estate Finance and Economics, No. 14.
- DOMBROW J., KNIGHT J., SIRMANS, C.F., 1997, *Aggregation Bias in Repeat-Sales Indices*, The Journal of Real Estate Finance and Economics, Vol. 14, issue 1-2, pages 75-88
- DREIMAN M. H., PENNINGTON-CROSS, A., 2004, *Alternative Methods of Increasing the precision of Weighted Repeated Sales House Price Indices*, Journal of Real Estate Finance and Economics, No. 28(4).
- ENGLUND P., QUIGLEY J.M., REDFEARN C., 1998, *Improved Price Indexes for Real Estate: Measuring the Course of Swedish Housing Prices*, Journal of Urban Economics, 44, 171-196.
- FREUD, R.J. and WILSON W.J., 1997, *Metodi Statistici*. Padova, Italy.
- GELTNER, D., 1989, *Estimating Real Estate's Systematic Risk from Aggregate Level Appraisal-Based Returns*, Journal of the American Real Estate and Urban Economics Association, Vol. 17, pp. 463-81.
- GELTNER, D., 1991, *Smoothing in Appraisal-Based Returns*, Journal of Real Estate Finance and Economics, Vol. 4, pp. 327-45.
- GELTNER, D., 1993a, *Temporal Aggregation in Real Estate Return Indices*, Journal of the American Real Estate and Urban Economics Association, Vol. 21, pp. 141-66.
- GELTNER, D., POLLAKOWSKI H., 2007, *A Set of Indexes for Trading Commercial Real Estate Based on the Real Capital Analytics Transaction Prices Database*, MIT Center for Real Estate.
- GOETZMANN W.N., SPIEGEL, M., 1997, *A Spatial Model of Housing Returns and Neighborhoods Substitutability*, Journal of Real Estate Finance and Economics 14, 11-31.
- HILL R., CARTER, KNIGHT J. R., SIRMANS C.F., 1997, *Estimating Capital Asset Prices*, Review of Economics and Statistics 79, 226-233.
- HOWELLS P.G.A., BAIN, K., 1990, *Financial Markets and Institutions*. London: Longman.
- KLUGER, B. D., STEPHAN, J., 1997, *Alternative Liquidity Measures and Stock Returns*, Review of Quantitative Finance and Accounting 8(1), 19-36.
- MATYSIAK G., WANG P., 1995, *Commercial Property Market Prices and Valuations: Analysing the Correspondence*, Journal of Property Research 12:181.
- NAGARAJA C. H., BROWN L. D., ZHAO L. H., 2011, *An Autoregressive Approach to House Price Modeling*, The Annals of Applied Statistics 5.1, pp. 124-149.
- PACE R. K., 1997, *Performing Large Scale Spatial Autoregressions*, Economics Letters 54, 283-291.
- PACE, R. K., BERRY R., CLAPP J., RODRIGUEZ M., 1998, *Spatiotemporal Autoregressive Models of Neighborhood Effects*, Journal of Real Estate Finance and Economics 17:1, 15-33.
- QUAN D.C., QUIGLEY J.M., 1991, *Price Formation and the Appraisal Function in Real Estate Markets*, Journal of Real Estate Finance and Economics, Vol. 4, pp. 127-46.
- QUAN D.C. AND QUIGLEY J.M., 1989, *Inferring an Investment Return series for Real Estate from Observations on Sales*, Journal of the American Real Estate and Urban Economics Association, Vol. 17, pp. 218-30
- QUIGLEY J. M., 1995, *A Simple Hybrid Model for Estimating Real Estate Price Indexes*, Journal of Housing Economics, Vol. 4, No. 1, pp. 1-12.

- SALVO F., 2008, *Principi metodologici per la costruzione di indici dei prezzi nel mercato immobiliare*, In DEL GIUDICE, V., D'AMATO, editor, *L'elaborazione dell'informazione immobiliare: segmentazione immobiliare del mercato e indici dei prezzi immobiliari*, Maggioli, Rimini, pp. 114-133.
- SALVO F., CIUNA M., D'AMATO M., 2013, *Appraising Building Area's Index Numbers Using Repeat Value Model. A Case study in Paternò (CT)*, in *Dynamics of Land Values and Agricultural Policies*, MediaMond International Proceedings, Bologna, 63-71.
- SHILLER R.J., 1991, *Arithmetic Repeat Sales Price Estimators*, *Journal of Housing Economics*, No. 1.
- SHILLER R.J., 1993, *Measuring Asset Values for Cash Settlement in Derivative Markets: Hedonic Repeated Measures Indices and Perpetual Futures*, *Journal of Finance*, No. 48 (3).
- SIMONOTTI M., 2006, *Metodi di stima immobiliare*, Dario Flaccovio Editore, Palermo