Methodology Report for the Construction of Residential Property Price Indices

Real Estate Unit

Economic Research Department

June 2011

1. Introduction

The current report presents the methodology developed by the Central Bank of Cyprus's Real Estate Unit (REU), which is part of the Economic Research Department (ERD), for the construction of residential property price indices. The report is organised as follows: section two describes the data collection process while section three documents the cleaning and filtering procedures developed to remove outliers, incomplete and erroneous entries in the data. The methodology employed for the construction of the price indices is presented in section four.

2. Data Collection

The data collection process was carried out in two phases. During the first phase, the REU proceeded with the collection of historical valuation data from a large number of Monetary Financial Institutions (MFIs). This was deemed necessary given that until 2010 there was no existing database of residential property valuations. The data cover the period 2006-2009[1] and are referred to in this report as the Historical Valuations Database (HVD). The second phase of the data collection process began in January 2010. This involves the collection of data in a standardised format through an electronic data submission system specifically developed at the Central Bank of Cyprus (CBC) for this purpose. A total of 12 MFIs, which have partnered with the CBC through a memorandum of understanding for the purpose of constructing property price indices, are contributing data to this database. These data are referred to in this report as the Electronic Data Submission Platform Database (EDSPD).

According to the literature, price indexes based on valuation data tend to present smaller fluctuations than price indices based on transaction data (Lai and Wang, 1998). This is due to the fact that the market value of a property is typically estimated from the observed transaction prices of comparable properties in its proximity. As a result, the valuation price is likely to differ from the actual transaction price, particularly if during the appraisal period there have been a very limited number of transactions taking place in the proximity of the

---

[1] Data prior to 2006 were omitted due to the very limited number of valuations collected.

property being appraised. In addition, price indices based on valuation data tend to exhibit a lagging relation when compared to actual changes in market prices (see Case and Wachter, 2005).

In spite of these differences, price indices based on valuation data can still provide a representative and reliable representation of fluctuations in market prices. Furthermore, in the case of the valuation data collected by the CBC, the plethora of the available characteristics and the large number of observations in the data can further guarantee the quality and validity of the results.

2.1. Historical Valuations Database

Given that the historical valuation data were submitted retrospectively in the HVD by the different MFIs, these were heterogeneous both in their format and content. Specifically, these differed in the number and type of characteristics they included as well as in terms of the unit of measurement used for these characteristics. For example, different MFIs used a different measurement unit for the size of a property such as the 'skala', the 'prostathi', the square foot, the hectare, the 'decario' and the square meter. In addition, the submission of data by the different MFIs was heterogeneous in terms of the code names used to describe the variables in the data.

Consequently, the collected historical data had to be homogenised to contain a common set of characteristics prior to their inclusion in the database. These could then be processed for the purpose of constructing residential price indices.

2.2. Electronic Data Submission Platform Database

The data submitted to the EDSPD are collected on a monthly basis through a customised online data submission platform which ensures their standardisation and homogeneity. Overall, since 2010 more than 70 independent property surveyors and more than 100 evaluators submitted data to the EDSPD through the various MFIs, this contributing to the diversity, comparability and representativeness of the data used to construct the price indices as well as to the broad coverage of the different types of property and geographical areas. These data are composed of a wide range of characteristics for each type of property (residential, commercial premises, land). For each type of property a maximum of 57 characteristics are collected which refer to the location, space and other qualitative and quantitative characteristics of the property (see Table 1). Thus, the new database enables a more in-depth and detailed analysis of the data.

The quality, completeness and accuracy of the EDSPD data are guaranteed through a number of validation procedures. These procedures are

incorporated into the online data submission platform of the CBC and are performed at the time of the submission of the data. These are primarily intended to verify the quality of the data using a set of rules based on plausibility. These rules include:

- The completion of all mandatory fields.
- Ensuring that the right type of information for each attribute is sent (e.g. numerals where numbers are required).
- Ensuring the accuracy of the value of a variable depending on its type: integer, non-negative or simple numeric values, correct planning zones, sheets, plans, etc.
- Ensuring correct values for specific variables are entered based on a predetermined range of typical values (e.g. the identification number for the district is acceptable if it is an integer ranging between 1 and 6).

Data is recorded in the statistical database only after it has successfully passed all the validation procedures. Apart from the individual validation procedures performed separately for each characteristic at the time of submission, in-house validation and cross-validation techniques are also performed, which further ensures the consistency and reliability of the data (e.g. cross-examination to guarantee that the municipality or the registered communal authority belongs to the municipalities and communal authorities of the recorded district). Apart from ensuring the validity and accuracy of the data, these validation and cross-validation procedures help avoid any discrepancies in the results due to inconsistent and incorrect entries in the data.

3. Cleaning and Filtering Procedures

Prior to the econometric analysis of the data, further cleaning and filtering procedures were implemented to remove any erroneous recordings or outliers in the data. This was deemed necessary, particularly in the case of the historical data, which were not collected through a customised data submission system. Consequently, for this dataset no validation or cross-validation steps were implemented to ascertain the validity and quality of the submitted data as was the case for the data submitted through the CBC online data submission platform.

The various cleaning and filtering procedures were developed based on the observed empirical distribution of the individual variables corresponding to a specific type of property. Subsequently, these are implemented independently of each of the variables and for each type of property. For example, observations with a recorded zero market value for a property are removed as are observations with a recorded age greater than 100 years for a house, observations with incomplete information on the collected variables, etc. Overall, the cleaning and filtering of each variable is based either on widely used

statistical techniques (e.g. the cleaning of outliers in the market value per square meter variable is carried out using the Stock and Watson, 2004 technique) or on objectively acceptable criteria.

4. Index Construction Methodology

In constructing the various valuation price indices and sub-indices, the hedonic methodology was used. Hedonic techniques are among the most computationally advanced methods in the literature for the construction of price indices and the most widespread. Additionally, these are advocated in the literature as being the most reliable methods (Gourieroux and Laferrere, 2006, Hoffman and Lorenz, 2006, Wen, Jia and Guo, 2005 and Maurer, Pitzer, and Sebastian, 2004). This class of techniques has already been employed for the construction of residential property price indices for Cyprus by Platis and Nerouppo (2005) and Pashardes and Savva (2009).

The basic premise of this class of index construction techniques is the determination of the price of a property relative to the set of attributes it possesses, which themselves do not encompass observable market prices. Even though the price of a property depends on the value that a buyer places on the set of the characteristics of the property, both quantitative and qualitative (location, internal covered area, proximity to the sea, etc.), these values are not directly observable.

Using multivariate regressions, the hedonic methodology disaggregates statistically the price into its constituent characteristics (Thwaites and Wood, 2003). Hence it is able to disentangle that part of the price variation which is: a) due to changes in the mix of characteristics of the property; and b) caused by inflationary factors and market conditions.

Consequently, the results obtained from hedonic techniques, unlike those obtained from simple average-based techniques, are comparable and quality constant. This key advantage stems from the fact that this class of techniques takes into account the heterogeneous nature of the physical, structural and locational characteristics of the property and consequently allows for the influence of these characteristics on the valuation price.

One of the first basic steps in the creation of a residential property price index using the hedonic methodology is the identification of the most important characteristics of the property that explain statistically and to the greatest extent possible its market value. In the literature there are no clear guidelines as to how one can choose the optimal relationship between the market value and the characteristics of the property, as this depends on the type of market under consideration (i.e. its cultural and demographic characteristics, living standards of the country, etc.).

To determine the statistically most important characteristics of a typical residential property in Cyprus, an extensive procedure of analysing the data using correlation coefficients and statistical regression methods were implemented by the REU. Variables which were classified as having a multicollinear relationship between them were excluded from the analysis. The characteristics which were found to explain about 70% of the market value of a residential property in the CBC databases are the following:

- Characteristics of size, which explain most of the variation in property prices: the internal covered area, the area of covered and uncovered verandas.
- Quality characteristics: the age of the dwelling, its condition and its level of luxury.
- Location characteristics: the district in which the property is registered, the planning zone of the property and therefore its building coefficient, whether it is located in a tourist area and whether it has a sea view.
- The existence of central heating and central air conditioning units.
- For house properties only: the distinction between a detached house, semi-detached house and maisonette, whether the house is built in a plot or field, the land area occupied by the house, whether it is built in a rural or in an urban area.

The results from the application of the hedonic methodology using the set of characteristics selected from the above analysis are then used to construct the various residential indices and sub-indices. The CBC residential property price indices employ the Fischer index which is the geometric average of the Laspeyres and Paasche indices. Fisher type indices are preferred in the literature as they avoid the various shortcomings and weaknesses of the Paasche and Laspeyres indices. For example, the Laspeyres index tends to give greater weight to properties whose prices have risen while the Paasche index tends to give greater weight to properties whose prices have fallen (Diewert, 1976 and 2004). The aggregated residential property price index is then calculated as the weighted sum of the two by type indices, i.e. houses and flats, using their respective quarterly number of observations as weights.

The statistical analysis of the results obtained from the application of the hedonic regressions and of the produced indices and sub-indices confirms their reliability, robustness and accuracy. These are characterised by a high coefficient of determination (R^2 above 0.70), highly statistically significant coefficients (at significance level 0.1%) and small confidence intervals for the estimated coefficients.

5. Conclusion

As mentioned above, the construction of residential price indices using the hedonic methodology is based on property valuation data recorded by the partner banks, which in turn receive the relevant information from property appraisal agencies for the purpose of granting loans. These data exhibit a high correlation with residential property transaction data and they constitute an alternative data type for the construction of residential property price indices.

The construction of residential property price indices constitutes an elaborate process both from a practical and from an analytical perspective. Consequently, it is very unlikely that different price indices will record the same numerical results, particularly in the case were the methodologies and/or the database used differ, even though one should expect that these follow the same trend. Nevertheless, the availability of a range of alternative residential property price indices is of importance given that no price index alone can encompass all possible advantages or exclude all possible disadvantages (Fenwick, 2005 and 2006).

Bibliography

Case, B and S. Wachter (2005) "Residential real estate price indices as financial soundness indicators: methodological issues", in *Real estate indicators and financial stability*, BIS Papers 21, April.

Diewert, W.E. (1976) "Exact and superlative index numbers" *Journal of Econometrics*, 4(2): 115–146.

Diewert, W.E. (2004) "A new axiomatic approach to index number theory", Discussion Paper 05, Department of Economics, University of British Columbia.

Fenwick, D. (2005), "Statistics on Real Estate Prices: The Need for a Strategic Approach", BIS Papers No 21, Bank for International Settlements, Washington DC: The International Monetary Fund pp. 368-372.

Fenwick, D. (2006), "Real Estate Prices: The Need for a Strategic Approach to the Development of Statistics to Meet User Needs", paper presented at the OECD-IMF Workshop on Real Estate Price Indexes held in Paris, November 6-7, 2006.

Fleming, M. and J. Nellis (1984) "The Halifax house price index: technical details", Halifax Building Society.

Lai T-Y. and K. Wang (1998), "Appraisal Smoothing: The Other Side of the Story", Journal of Real Estate Economics, Vol. 26 No. 3, pp. 511-535.

Nationwide House Price Index (2010) "The Nationwide house price index methodology", Nationwide.

Gourieroux, C. and A. Laferrere (2006) "Managing hedonic housing price indexes: the French experience", paper presented at the OECD-IMF Workshop on Real Estate Price Indexes, Paris, 6-7 November.

Hoffman, J. and A. Lorenz (2006) "Real estate price indices for Germany: past, present and future", paper presented at the OECD-IMF Workshop on Real Estate Price Indexes, Paris, 6-7 November.

Maurer, R., M. Pitzer and S. Sebastian (2004) "Construction of a transaction based real estate index for the Paris housing market", *Journal of the German Statistical Society (Allgemeines Statistisches Archiv)*, 88: 303-326.

McAllister, P. and F. Fuerst (2010) "Constructing real estate indices for Cyprus: an evaluation of the options", Royal Institution of Chartered Surveyors (RICS), Cyprus.

Pashardes, P. and C. S. Savva (2009) "House prices in Cyprus", University of Cyprus Economic Research Centre Policy Paper 01.

Stock, J.H. and M. W. Watson (2004) "Combination forecasts of output growth in a seven-country data set", *Journal of Forecasting*, 23(6): 405-430.

Platis, S. and M. Nerouppos (2005) "Asking price and transaction-based indices for the Cyprus housing market (rebased)", BuySell Cyprus Real Estate, November.

Thwaites, G and R. Wood (2003) "The measurement of house prices", *Bank of England Quarterly Bulletin*, Spring.

Wen, H., S. Jia and X. Guo (2005) "Hedonic price analysis of urban housing: an empirical research of Hangzhou, China", *Journal of Zhejiang University SCIENCE*, 6A(8): 907-914.

Table 1: List of characteristics collected through the CBC Online Data Submission Platform

	Time Variables:		**Property Services Characteristics:**
1	Valuation date	52	Air conditioning: split units/ central AC/No AC
	Valuation Information:	53	Ventilation
2	Share of property to be valued	54	Central heating
	Valuation Variables:	55	Fire safety system
3	Market value of property	56	Burglar - Alarm system
4	Estimated Restricted Realization Price	57	Lift

5	Market value of Land		**Property Legislative Characteristics:**
6	Estimated Restricted Realisation Price of Land	58	Property with title deed
	Property Type Variables:	59	Planning and building permission
7	Property type		**Geographical Characteristics:**
8	Type of house: detached/semi-detached/terraced	60	District
9	Under construction/development	61	Municipality
	Property Internal Characteristics:	62	Communal authority
10	Number of bedrooms	63	Quarter
11	Number of WCs	64	Sheet
12	Kitchen	65	Plan
13	Locker/changing rooms	66	New code
	Property External Characteristics:	67	Block
14	Window sides	68	Registration number
15	Number of covered parking places	69	Plot number
16	Number of uncovered parking places	70	Planning zone
17	Swimming pool: private/communal/none	71	Proportion of zone to land is 100%
	Property Size Characteristics:		**Property Position Characteristics:**
18	Land area	72	Number of houses in the plot/field
19	Building internal covered area	73	Number of flats in the building
20	Building basement area	74	View: uninterrupted/limited/sea/ mountain/green/ other
21	Building ground floor area		
22	Building mezzanine area	75	Road Relation: road level/ below road level/ above road level
23	Building other floors area		
24	Office space area	76	Position of plot: corner/standard
25	Showroom area	77	Next to pedestrian walkway
26	Area of covered verandas	78	On a cul de sac
27	Area of uncovered verandas	79	Floor: semi-basement/ ground floor/ mezzanine/penthouse/other/
28	Building storage area		
29	Total consolidated area per property type		**Area Characteristics:**
30	Frontage	80	Availability of services (water and electricity)
31	Depth of shop	81	Accessibility: inaccessible/public road (asphalt/dirt)/ right of way/ road under registration/pathway
	Property Building Characteristics:		
32	Structured cabling	82	Development potential
33	Raised floors	83	Next to green area
34	False ceilings	84	Near high voltage electricity supply
35	Steel structure and walls made of blocks	85	Nearby an electricity sub-station
36	Steel structure and walls made of sandwich stone fibre panels		**Usage Of Property:**
		86	Under Rent Control Act
37	Steel structure and walls made of bricks and plaster	87	Possession
38	Roof by metallic crossings -zinc	88	Monthly rent paid
39	Roof by metallic crossings with sandwich panels	89	Subject to life interest
40	Roof by metallic crossings with zinc and thermal insulated materials	90	Groves/cultivated land
		91	Redistribution area
41	Roof by metallic crossings with plates of asbestos	92	Under division process
42	Smooth concrete floor		
43	Smooth concrete floor and hepoxy paint		

44	Ceramic tiles		
45	Undulating ground		
	Property Quality Characteristics:		
46	Age		
47	Type of office		
48	Quality of residence: very luxurious/luxurious/ ordinary/below ordinary/poor		
49	Maintenance: very good/good/average/poor		
50	Listed building		
51	Shape		