

## Notaires-INSEE indices

### Introduction

Housing price indices can meet a variety of needs: from a simple analysis of price changes to an analysis of risk for lenders, to the implementation of monetary policy and national accounting.

The simplest index is based on the average sale price observed over a given period of time, but this may be skewed since sales observed over this period are not necessarily representative of all the goods to be studied. An index based on the median price of transactions is less sensitive to any extreme values that are observed, but it is still affected by bias, as the quality of the goods sold may vary over time.

In some countries like the United States, where residential mobility by home-owners is high, the repeat sales method is used. A repeat sales index is calculated by comparing the price observed for the same dwelling for two successive sales, assuming that its quality remains the same. However, it is a difficult and costly process to check that the quality of the goods really has remained unchanged. In addition, this method is not immune to selection bias, because the goods being exchanged, especially if they are exchanged several times, may not be representative of the market overall.

The hedonic approach, the approach used in France, is based on the characteristics of the dwellings through a designated number of observed characteristics, each of which, with its own specific value, contributes to the total value. The combination of characteristics is therefore assumed to determine the value of the good in question. The model is estimated from the prices and characteristics of dwellings exchanged during what is called the “estimation” period. It is then used to measure the estimated value of a selection of dwellings over time. Thus the model can be used to estimate the value of a dwelling, even if it has not been sold during the estimation period. Such an approach solves the three problems that make it difficult to construct a price index: (1) change in the quality of the dwellings: they have many characteristics and some of these may change over time; (2) the fact that the price of any given dwelling is only observed very infrequently, and (3) the fact that the selection of dwellings to be monitored must be defined precisely, since the type of goods bought and the housing stock change over time.

This type of modelling comes with a cost since prices and sales characteristics must be observed and recorded. In countries where the law requires that sales of dwellings be registered before a notary, data are more readily available. This is the case in France, where data have been collected and centralised since 1994; a quarterly hedonic price index has been calculated since 1998. Notaries are responsible for collecting data and calculating indices, while the Scientific Board defines the scientific procedure for the hedonic method and INSEE checks its implementation and validates the results.

### The institutional framework in France

The French institutional framework consists of a network of notaries, who have a monopoly on registering property transactions, and a national statistics body, INSEE. Notaries check the property rights, draw up the deed of sale, send the property documents to the mortgage registry and receive the property transfer rights in the name of the State.

By virtue of their function, notaries have access to the transaction price and to some of the housing characteristics that have to be mentioned in the deed of sale. In addition, every notary must send information on the sale price to the authorities, since the fees collected on the sale are dependent on the price. In theory, the data cover all sales, providing real transaction prices on a regular basis and

over a long period of time. Information is available promptly as there is a time limit set for notaries to pass on all taxes, a limit that in recent years was reduced to one month.

INSEE is responsible for providing official statistics, including the national accounts and price indices. Until the end of the 1990s, INSEE published no housing price indices, except for the city of Paris which has had a series of indices for second-hand apartments since 1983.

In 1997, the Higher Council of the Notariat (Conseil Supérieur du Notariat - CSN) decided to create a price index for housing located outside Paris. It turned to INSEE and its research centre (CREST) for advice on methodology. Formal agreements were signed in 1998 and 1999 between the notaries and INSEE. Notaries collected the data and calculated the indices every quarter. In this way, many indices became available to the public free of charge.

A considerable amount of work goes into the process between drawing up a deed of sale and publishing the final index. Each of the 5,500 notarial offices was invited to send a copy of every deed of sale, either by post, or in a paperless format, as has become more common since 2010. The speed with which each office sends in the data is crucial for the quality of the index: a deed that arrives too late for the quarter being considered is not used when calculating the index. Hence the advantage of paperless systems, which avoid delays from grouping batches of deeds together and postal delays; the majority of deeds are now sent in this way.

There are two databases: the BIEN database, run by the interdepartmental chamber of Paris notaries (Chambre interdépartementale des notaires de Paris – CINP), which covers the Ile-de-France region (less than 20% of transactions), and the PERVAL database, managed for the CSN by the group ADSN, which covers the rest of France (more than 80% of transactions).

The vast majority of property transactions are housing transactions, with houses and apartments in fairly equal proportions. Some of the remaining transactions cover other types of property (garages, commercial premises, entire apartment buildings, plots, etc.), but these are not included when calculating the price indices.

Until the end of 2016, notaries sent in the data on a voluntary basis, thus the coverage rate for the database was not 100%. Since January 1, 2017, the law obliges notaries to supply the bases, therefore coverage rates are rising even though the phase of progression has not ended yet. It is around 60% (75% in Ile de France and 55% in provincial France) at the end of 2017. This coverage rate is estimated by comparing the amount of transfer taxes appearing in the database with the amount received by the Directorate General of Public Finance (Direction Générale des Finances Publiques - DGFIP).

A 100% coverage rate is nonetheless not necessary when calculating a hedonic index, as the hedonic model is estimated over a reference period and is based on a fixed “basket” of goods, defined for a sufficient number of years of transactions. This number of years (2 years in the latest versions of the Notaires–INSEE indices) is determined with two goals. The first is having enough observations so that the estimations are of good quality; the second is that the period is recent and relatively short, so that the evolutions on the valuation of the real property are correctly and quickly taken into account on the estimations.

The database used for the production is anonymised to comply with the recommendations of the French data protection authority (Commission Nationale Informatique et Liberté – CNIL). The location characteristics are the INSEE code for the municipality, and in some cases (large cities, coastal areas, mountain resorts, etc.) a district code (as in the 80 administrative districts in Paris).

Houses and apartments are treated separately. Location is very important, not only in terms of a district’s amenities and facilities, but also in terms of construction characteristics. For example, buildings constructed in 19th century Paris by Haussmann are of better quality than buildings from the same period in other areas. This is why hedonic regressions are estimated at a detailed local level and why models can include district variables.

## The hedonic method

The hedonic method is based on an econometric model linking price to the characteristics of the dwelling. The model is estimated from the prices and characteristics of the goods sold over what is called an estimation period. It is then used to adjust the quality of the goods over subsequent periods. When this is done well, quality is controlled, the value of a good can be estimated when there is no sale and the selection of goods monitored can be defined and can change over time.

Some use an alternative hedonic approach. They carry out a single regression of the sale price, with time dummies added to the model, acting as a basis for calculating an index. This method is generally applied across two adjacent periods, rolling over in time. It is more sensitive to biases due to the sampling of observed sales than the method that consists of evaluating a fixed basket of goods at each date; however, it simply picks up gradual change in the price of characteristics.

Among all the hedonic housing price indices, the unique feature of the Notaires-INSEE indexes is that it measures variation in prices from a base period ("reference period"), representing a reference stock. In addition, to take into account variations in the price of characteristics in time and space, the model is estimated in areas where prices are assumed to be similar (strata). Ideally, these strata could be small sub-markets. In practice, however, the number of strata has been limited to around 350 to ensure a sufficient number of sales within each stratum (over 440 per year).

The model links the price of dwellings (per m<sup>2</sup> for flats, overall for houses) to their characteristics and their location (a district in a stratum). The characteristics are: the district, the period of construction, the number of rooms, the number of bathrooms, the number of parking garages, the condition of the good, the energy condition, the distance to the closest city of 50 000 inhabitants (or to Paris for Ile-de-France); and for apartments the floor, the presence of a balcony (terrace or loggia), of a cellar, of a lift from the 4th floor, the surface crossed with the number of rooms; and for houses the presence of a basement, the number of levels, the surface of the ground and the habitable space. The location of the housing inside strata, summarized by the district (crossed with the number of rooms), is considered as one of its characteristics, in the same way as the number of rooms or the habitable space.

The model is estimated per stratum, differentiating between apartments and houses, and this implies that all the variables interact implicitly with the stratum. Since the version in force since 2011, we use a period of two years for the validity of the models and hence for the stability of relative prices. As a result, models are estimated every two years, so that the constraints of calculating an index of constant quality, of ensuring a good awareness of changes in the market and the requirements of regularly producing data can all be reconciled. Each quarter, the price of the reference dwelling in each stratum (i.e. the reference dwelling used in the regressions) is calculated from the coefficients of the model applied to the transactions for the quarter. Change in the value of this reference dwelling therefore gives us the change in the elementary index for each stratum. These elementary changes are not published but they are aggregated at higher levels for publication. This aggregation is made on the basis of weighted geometric means at infra-departmental geographic levels and weighted arithmetic means at departmental and more aggregated levels, with the value from the weighted coefficients reflecting the structure of the housing stock.

At a higher geographic level, the price index is a Laspeyres index, defined as the ratio of the estimated value of a reference stock of dwellings to its value at the period of the base index. For example, the index for the "Provinces" measures the change in value of the entire reference stock for the Provinces.

There are some strong seasonal variations, especially for houses. These are linked with families' residential mobility, especially in relation to the school year. INSEE calculates and publishes deseasonalised indices.

## **Conclusion : adaptable indices**

There is a strong and growing demand for reliable housing price indices.

Housing prices, and property prices in general, may evolve differently from the prices of other consumer goods and all economic actors need to be aware of these differences. The Notaires–INSEE indices can therefore provide essential input for official statistics.

To meet demand, the data collection process and the calculation methodology are revised regularly to provide analysis instruments that are always relevant. The work of the notaries is essential to ensure that the indicators are both exhaustive and representative and that the information collected is of optimal quality.

The law of 28 March 2011 confirms the public service mission that notaries fulfil in this area.

Data produced jointly by the Insee and the notaries are certified by the Authority of the Public Statistics.