Appendix 1: Transmission of deeds of sale to the notarial databases and estimating coverage rates

Virtually all housing transactions[56] generate a deed (*acte*) certified by a notary. This document is passed on to the tax authorities (first the Mortgage Registry [*Conservation des hypothèques*], then the Land Registry [*Services du Cadastre et des Domaines*]), for registration and payment of transfer tax in most cases, and VAT if the sale concerns a new dwelling. An exhaustive source does therefore exist. However, this is a paper document and does not resemble a standard administrative form as it consists of a written text. Consequently, if this is to become a usable computer file, i.e. a database, the information has to be captured in a coded file. This is a cumbersome and costly operation given the size of the documents (usually several pages) and the volume they represent (approximately 800,000 transactions for used dwellings per year across the whole of France).

The transmission of information from the deeds and the data capture was undertaken by the notaries themselves, first by the Chambre interdépartementale des Notaires de Paris (CINP) in the late 1970s (and especially since 1990), then by the Conseil Supérieur du Notariat (CSN), which set up a limited company, Perval, specifically for this purpose in 1993. These two bodies receive information from the notarial offices and feed it into one of the databases; the first covers the Paris Region, the second covers the Provinces and the French overseas *départements*. It was this arrangement that made it possible – again on the notaries' initiative and with the collaboration of INSEE – to create the price indices described here. Note, however, that there was an index already in existence: it was launched in 1983 and tracked sales of used apartments in Paris. This index was calculated from an average of prices weighted into 72 strata according to the housing stock at the last population census. It was revamped and now applies the same methodology as the other Notaires-INSEE indices.

Until 2010 the information was transferred on paper.[57] Since this date, some of the deeds are now captured in digital format. Teletransmission is still in the process of being introduced in the notarial offices, for feeding the mortgage registry (Projet Télé@ctes) as well as the notarial databases (project to introduce teletransmission to the notarial databases). These long-term undertakings have involved modifying all the software for drawing up deeds (LRA) and changing the way the notarial offices were organised, to ensure that all the necessary information was entered in the LRA. By the end of 2012, about 50% of the deeds had been teletransmitted for the purposes of the indices.

Unfortunately, the notarial offices do not pass on all their deeds, and not all offices transmit anything at all. The coverage rate of the notarial databases varies from one *département* to another. It is important to be aware of this since it does affect the calculation applied when weighting the indices.

Coverage rate

The coverage rate of the notarial databases is obtained from the ratio of the total value of transactions recorded in the databases to the total value of all transactions.[58] The rate is estimated by comparing the value of transactions in the notarial databases with the assessment base for transfer taxes collected by the Directorate General for Public Finances (DGFiP). As the DGFiP supplies the amounts of transfer taxes collected in each *département*, the coverage rate can be calculated by *département*.

However, it is not possible to compute the rate separately for used housing, as there is no difference in the amount of transfer tax to which they are subject. Before 1999, there was a specific transfer-tax rate for used housing and the DGFiP accounts recorded the corresponding revenues. Then all that had to be done was to divide this amount by the tax rate to obtain the assessment base. Since the reforms of 1998 and 1999, used housing transactions are included in the standard regime assessment base. Today there are three tax regimes that cover property transfer tax:

- The "ordinary law" (*droit commun*) regime

[56] With the exception of certain transactions carried out by general government, or property transfers arising from corporate acquisitions.
[57] The notarial offices photocopied the deed (or more often an extract from the deed) and to this they attached a slip of paper which gave information not necessarily provided on the deed extract.
[58] Rates calculated from values are higher than those calculated from numbers of transactions, because the *départements* where the coverage rate is highest are usually the *départements* where the average transaction amount is highest, especially in the Paris Region.

This applies exclusively to transactions involving payment for used property and real estate not subject to value-added tax. Since January 1, 2011, it includes a levy of 3.8%[59] that goes to the *département* and 1.2% to the municipality, 5.0% in all. It used to be 3.6% to the *département*, 1.2% to the municipality and in the last years of the period considered, 0.2% to the State, but this is no longer collected.

- The "special rules" regime.

  This regime applies:

  - to sales of new and assimilated properties purchased with a view to a quick resale ("property developers") and real estate subject to value-added tax,
  - to certain property transactions where there is no charge (donations); these transactions represent about 40% of the assessment base for this tax regime on average in the whole of France.

  Since January 1, 2011, the special rules regime includes a levy for the *département* of 0.7% on transactions with a financial consideration and 0.6% on donations. Previously, this regime included 0.6% for the *département*, and in recent years during the period considered and for transactions with a financial consideration, a levy of 0.1% to the State, which is no longer collected.

- Transfer tax exemption

  This regime applies mainly to purchases by the State, local government and some public establishments.

Since 1999, as used housing transactions are no longer counted separately in the DGFiP accounts, we need to estimate their total value. At the end of 1999, used housing transactions represented about 80% of the ordinary law assessment base. These have been checked for coherence since then and this high proportion has proved to have varied little.[60]

If the coverage rate of the notarial databases were the same for used housing and for total property sales covered by the ordinary law regime, the problem would be solved: the proportion of used housing covered by the entire ordinary law regime would be the same whether all transactions were considered together or only sales recorded in the notarial databases. However, for years prior to 2000, when different types of properties could be differentiated within the transfer tax receipts, the coverage rate of the notarial databases was slightly higher (by about 2%) for used housing than for all transactions covered by the current ordinary law regime. This difference varied considerably from one *département* to another.

We hypothesise that the used housing coverage rate and the rate for all transactions under the ordinary law regime are similar. This may result in a large error in the coverage rate for a given *département*, but this has no effect on the value of the Notaires-INSEE index for this *département*. Any possible error in the value of very aggregated indices (France, Paris Region and Provinces) is also low, as errors in weighting for the entire country are mutually compensated. It is at supra-departmental level, where data are not very aggregated (e.g. a region made up of two *départements*) that the effect on the value of the indices is potentially highest, however, tests have been carried out which have shown that it still remains very low.

[59] At the time of writing, a project was announced to raise this rate to 4.5% (a 0.7% increase) from March 1, 2014 in those *départements* who so wish.

[60] These are checks made by the Scientific Board of the Notaires-INSEE indices when calculating the number of used housing transactions.

Table A1.1 – Coverage rate of notarial databases (transfer tax under ordinary law regime)

Year	Paris Region	Provinces	France
2000	74.0%	58.2%	62.7%
2001	78.4%	59.0%	64.5%
2002	77.4%	53.6%	60.7%
2003	80.3%	55.3%	62.8%
2004	81.7%	54.1%	62.3%
2005	81.5%	55.6%	63.3%
2006	80.3%	56.5%	63.6%
2007	83.0%	57.6%	64.8%
2008	80.4%	58.6%	64.8%
2009	76.6%	56.0%	62.2%
2010	80.0%	56.5%	63.1%
2011	70.1%	51.0%	56.8%
2012	73.4%	51.9%	58.4%

Source: DGFiP and Notaires of France - PERVAL and BIEN databases
How to read this table: coverage rates are relatively stable over time, however, they did fall considerably in 2011. This can be explained by the trend in the number of transactions, which picked up after decreasing substantially in 2008-2009. The increase was not matched by the coding measures when they were put in place.

Appendix 2: Geographic divisions

The primary used-housing price indices are calculated at a very precise geographic level, called the stratum, then aggregated in the form of totals weighted by the structure of the reference stock.[61] Thus the country is divided up into units that are as small as possible to give a network on which to base calculations. A stratum must respect two conditions for it to be of an adequate standard:

- It must include a minimum number of transactions, 110 in each quarter for the period being studied.
- Within the stratum, prices must be homogeneous.

Stratification was carried out using different methods for the Provinces and the Paris Region, given the different contexts of these real estate markets.

- In the Provinces, local markets may be very compartmentalised (with coastal resorts, ski resorts etc.). The process is based on a geographic proximity criterion with the municipality as the primary unit.
- In the Paris Region, we first define price areas: two for the apartment market (where the city of Paris makes up one area and the rest of the Paris Region the other), one for houses (because there are too few houses in Paris). Next, the municipalities are grouped together according to criteria of price level homogeneity. Municipalities assigned to the same stratum are not always contiguous.

In both cases, the aim is to have homogeneous prices within a stratum. This homogeneity is gauged according to price: sale price for houses, sale price per sq.m. for apartments. We also check afterwards that the strata differ one from another.

In the Provinces, studies have determined a higher number of strata in the new version of the indices:

- 97 strata for apartments, compared with 88 in 2002,
- 174 strata for houses in 2008 compared with 146 in 2002.

The increased number of strata stems from the improvements made in feeding the database as well as from an increase in the total number of transactions.

The exhaustive list of municipalities that make up each stratum is given in the attached Excel file.

In the Paris Region, on the other hand, studies have determined that the number of strata should be reduced:

- 5 strata for apartments in Paris, compared with 18 in 2002,
- 10 strata for apartments in the Paris Region, excluding Paris, compared with 37 in 2002,
- 7 strata for houses in the Paris Region, as in 2002.

The exhaustive list of municipalities that make up each stratum is given in the attached Excel file.

Stratification in the Provinces

The transactions analysed cover the period from the 1st quarter of 1998 to the 3rd quarter of 2007, or 39 quarters in all. Division of the territory begins at a fine level of geographic stratification, i.e. the municipality. Any municipality that exceeds the minimum threshold of 110 transactions per quarter is designated a stratum. In practice this only happens for the largest municipalities. When this is not the case, municipalities are aggregated into the same stratum if their price levels are consistent. There are also two types of municipality that are differentiated as being specific cases: ski resorts and coastal resorts.

Thus the Provinces are divided into strata by geographic level and according to different types of property market. The levels of division may be:

- the municipality: this is the smallest unit.[62] This division is the first step in the process. The number of quarterly sales is calculated for all the municipalities in France; we retain those that exceed the minimum threshold of 110 transactions per quarter.
- the urban unit: strata are combined into urban units and we remove those municipalities that have already been designated as strata. Thus there may be strata that consist of only one suburb or others

[61] See definitions in Chapter 3.
[62] There is no stratum for *arrondissements*. The *arrondissements* of Lyon and Marseille are dealt with using dummies in the regressions.

that consist of the entire agglomeration when a large city (or municipality) does not have sufficient transactions to form a stratum by itself.

- Alpine winter sports resorts: these are grouped into strata by mountain range or by *département* when this is possible. The *départements* concerned are Savoie, Haute-Savoie and Isère, given the size of their ski areas. Three strata of municipalities are created for apartments, by mountain range:
    - Stratum 01: Tarentaise
    - Stratum 02: Maurienne, Val d'Arly, Beaufortin, Les Bauges, Chablais, Giffre, Les Bornes, and Chartreuse,
    - Stratum 03: Mont-Blanc and Aravis.

For Isère, we create a specific class for municipalities that include a ski resort (i.e. stratum 04: Isère). For the other mountain ranges (southern Alps, Pyrénées, Vosges and Massif Central), it is not possible to create specific strata because the volume of sales is too small.

- coastal municipalities: the strata are obtained by grouping together all municipalities in the same *département* or region on a geographic basis. They are differentiated from nearby municipalities because their average house price is higher than in the surrounding municipalities (the divergence is set at 15 or 20% depending on the specific case).

Municipalities are grouped together on the basis of geographic proximity, and according to national territorial divisions: the municipality, the urban unit, group of coastal municipalities, group of ski resorts, up to a *département* or even a region.

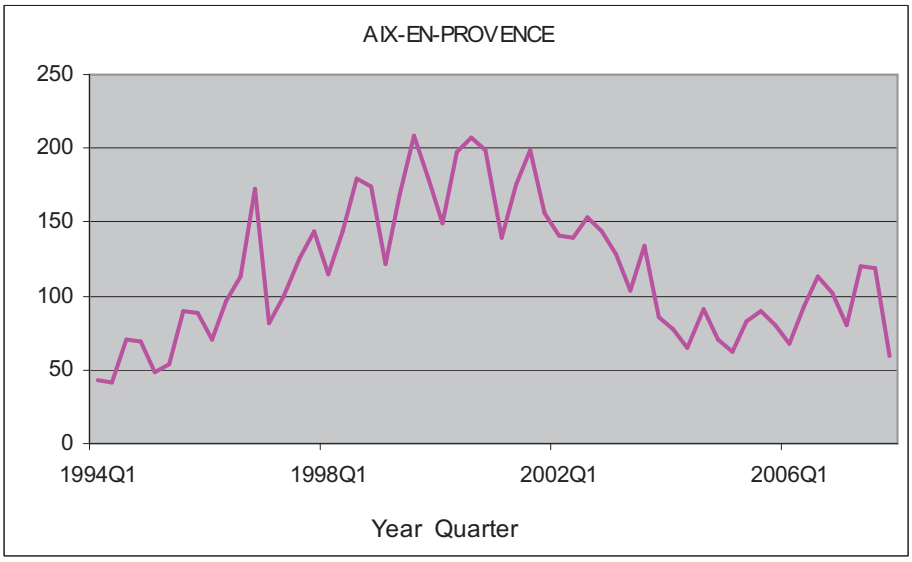
A few strata that come just below the 110 transaction cut-off for just a few quarters could be selected if the development in the number of transactions seemed favourable. Thus, for apartments, the Aix-en-Provence stratum was removed because the number of transactions fell considerably. It regularly dropped below the threshold from 2003 onwards. In Caen, on the other hand, the number of transactions increased over the period and remained above the threshold from 1998. This stratum was therefore created. (Figures A2.1 and A2.2).

Next we deal with urban units, coastal municipalities and municipalities within the mountain ranges most popular with tourists, and rural municipalities. Rural municipalities do not have enough sales per quarter to be selected, nor do they have any particular feature that could bring them up. We therefore group them into classes according to the quantiles for average income per inhabitant for 2006*,* with each class constituting a stratum. This is done using a descending process, looking first at national economic and development areas ("Zones for Study and Development" - ZEAT). Each stratum has to fulfil the following two criteria:

- a volume of transactions greater than 8,000 or 9,000 transactions between 1998 and 2007,
- a minimum of 110 transactions per quarter between 1998 and 2007.

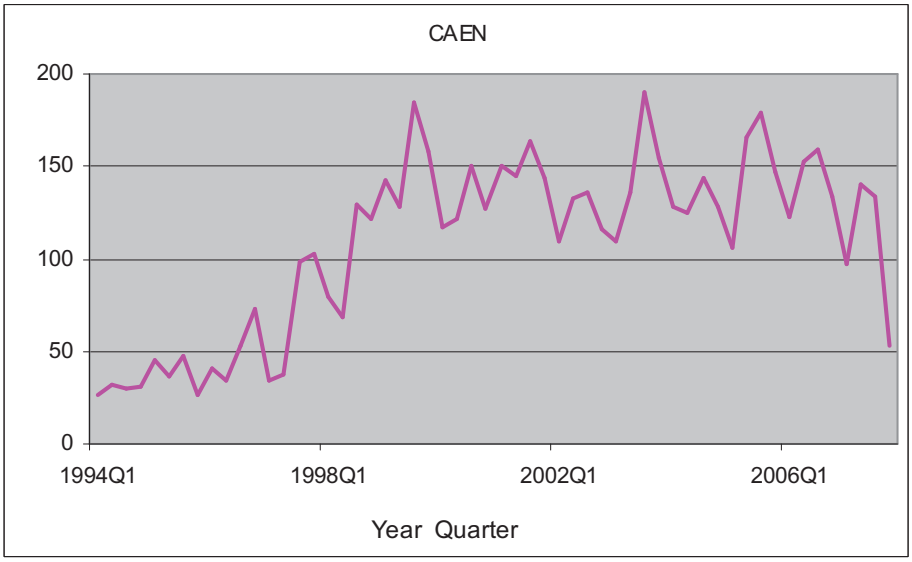
If the thresholds set for a ZEAT are widely exceeded, then we move down to regional level. The same is then done for a given region, and sometimes we have to move down to an even smaller level, the *département*. In other instances, we group together two *départements* from the same region to form one stratum (cf. Table A2.1 showing the grouping for the Ardennes and Marne *départements,* for example).

Figure A2.1 - Number of quarterly transactions in Aix-en-Provence (apartments)



Source: Perval database

Figure A2.2 - Number of quarterly transactions in Caen (apartments)



Source: Perval database

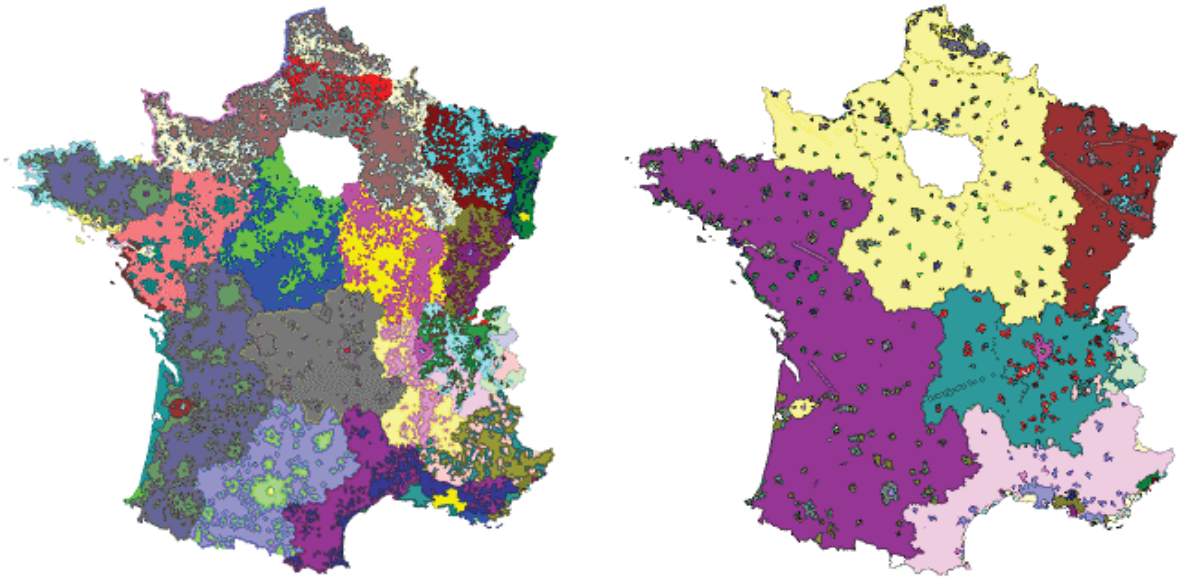
Table A2.1 - Example of breakdown of geographic areas by "income" variable (houses)

Stratum code	*Département*	Region	Lower bound of income for the stratum (€)	Upper bound of income for the stratum (€)	Total number of quarters when 110 dwellings threshold exceeded	Total sales from 1998 to 2007
99210	Grouping of Ardennes and Marne *départements*	Champagne Ardennes	7,824	14,480	39	9,564
99211			14,480	18,821	39	11,535
99212			18,821	55,946	39	10,678
99213	Grouping of Aube and Haute-Marne *départements*		8 ,321	15,106	39	8,921
99214			15,106	55,835	39	8,640
99220	Somme	Picardie	9,342	13,543	39	8,848
99221			13,543	15,638	39	9,373
99222			15,638	40,192	39	9,159
99223	Aisne		9,998	14,889	39	11,616
99224			14,889	17,456	39	11,741
99225			17,456	20,190	39	11,635
99226			20,190	72,021	39	12,147
99227	Oise		8,299	13,431	39	8,717
99228			13,431	15,505	39	8,868
99229			15,505	44,817	39	9,118

Source: PERVAL database

Figure A2.3 – Strata for apartments in the Provinces

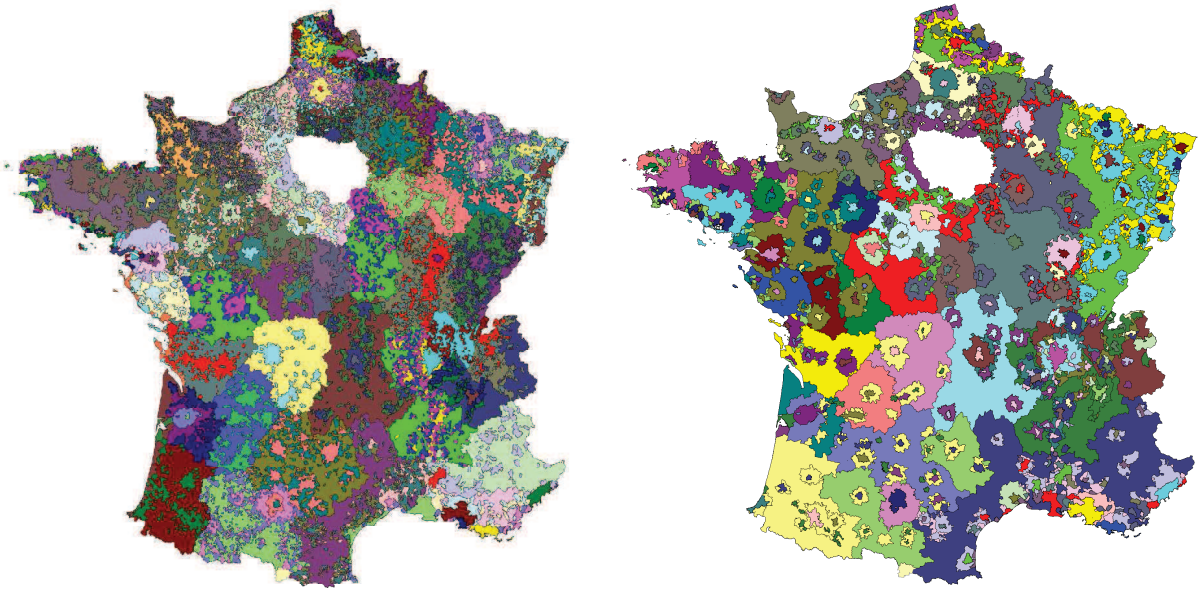
New database (indices version 3) *Old database (indices version 2)*



Source: PERVAL database
How to read this chart: strata are distinguished by their colour. For example, in version 2, the large yellow expanse around the Paris Region corresponds to the strata of rural municipalities in North-West France. As there are a large number of strata, some are shown on the maps in the same colour. They can be differentiated one from another by the fact that they are not contiguous.

Figure A2.4 – Strata for houses in the Provinces

New database (indices version 3) *Old database (indices version 2)*



Source: PERVAL database
How to read this chart: see above.

Stratification in the Paris Region

In the Paris Region, the transactions analysed are those from 2003 to 2007, or 20 quarters in all.

The first step consists of combining the information using factorial correspondence analysis (CA). The stratification per se is carried out in a second step using a hierarchical ascendant classification (HAC).

- First step: combine the information

As the HAC can accept only qualitative variables, the continuous variables are discretised. We look for the best possible summary of the information by reducing the number of axes for analysis. Each variable modality can be seen as an axis for analysis. We check that the contribution of each variable is not too dominant compared with the average contribution on the first axes. To establish the cut-off points, the method relies on statistical indicators, such as centiles, mean and standard deviations. Next, the little used modalities are combined. If the number of instances for a modality represents less than 5% of the total, it is combined with another.

- Second step: compile the strata

Before starting the stratification, we compile statistics at neighbourhood level (for Paris) or canton level (for the rest of the Paris Region), to cross the following variables: living space in the dwelling (bracketed values), plot size for houses (also by bracketed values), number of bathrooms, number of rooms, construction period, presence of lift for apartments, floor of the building, presence of cellars, number of garages and presence of balconies or terraces. Paris, where average prices are higher, is dealt with separately. The price per sq.m. for apartments and sale price for houses are incorporated as supplementary variables in our analysis. They can also be used to define the resulting classes.

The classification is done step by step, aggregating observations that are “most similar”. To do this, we use the distance between two observations or two classes to group together those that are closest. As the aggregation progresses, the classes include more and more observations.

- Example: the classification for apartments in Paris

First the number of classes is determined. Using the *Tree analysis* (*Figure A2.5*) we can determine visually the optimal number of classes. This graph shows the order in which classes are formed and also measures the value of aggregation distance between classes. A large “jump” on the tree corresponds to a large increase in the value for this distance. In practice, we select a cut-off in the upper part of the diagram, which is where we have classes that are still constructed with fairly short aggregation distances (hence classes that are fairly homogeneous), and just before a large jump (making sure not to go too

high up in the tree where the classes are too varied). In our case, the large jump corresponds to a cut-off below the split into 5 classes.

A second diagram (Figure A2.6) shows for each class, the modalities of variables and the individual variables (neighbourhoods grouped by *arrondissement*) that contribute most to defining the class. The diagram shows, in terms of weight in the total for transactions in Paris, the five classes of modalities of variables that are well-represented in the class and the geographic distribution of dwellings. Crossing these three criteria gives the profile for each class.

Profile of class 1

This class represents 20% of transactions in Paris. The best represented modalities of variables are:

- high price per sq.m.,
- older construction period,
- apartment without cellar,
- apartment without lift,
- apartment without garage.

The geographic distribution of dwellings in this class is as follows:

- 100% of dwellings in the 1st, 2nd,3rd, 4th and 6th *arrondissements*,
- 100% of dwellings in the 9th *arrondissement*,
- 72% of dwellings in the 10th *arrondissement*,
- 70% of dwellings in the 5th *arrondissement*,
- 50% of dwellings in the 7th *arrondissement*.
- 20% of dwellings in the 12th *arrondissement.*

Profile of class 2

This class represents 34% of transactions in Paris. The best represented modalities of variables are:

- medium price per sq.m.,
- older construction period,
- small living space between 20 and 40 sq.m.,
- apartment without lift,
- apartment without bathroom,
- apartment without garage.

The geographic distribution of dwellings in this class is as follows:

- 100% of dwellings in the 11th and 18th *arrondissements*,
- 63% of dwellings in the 17th *arrondissement*,
- 50% of dwellings in the 20th *arrondissement*,
- 39% of dwellings in the 14th *arrondissement*
- 30% of dwellings in the 5th *arrondissement*,
- 28% of dwellings in the 10th and 13th *arrondissements*.

Profile of class 3

This class covers 24% of transactions in Paris. The best represented modalities of variables in this class are:

- fairly high price per sq. m.,
- fairly recent construction period,
- average living space,
- apartment with garage,

- apartment with cellar
- apartment with lift.

The geographic distribution of dwellings in this class is as follows:

- 100% of dwellings in the 15th *arrondissement*,
- 80% of dwellings in the 12th *arrondissement*,
- 47% of dwellings in the 14th *arrondissement*,
- 30% of dwellings in the 19th *arrondissement*,
- 27% of dwellings in the 20th *arrondissement*.
- 12% of dwellings in the 17th *arrondissement*.

Profile of class 4

This class represents 10% of transactions in Paris. The best represented modalities of variables are:

- low price per sq. m.,
- recent construction period,
- apartment with garage,
- apartment with cellar.

The geographic distribution of dwellings in this class is as follows:

- 72% of dwellings in the 13th *arrondissement*,
- 70% of dwellings in the 19th *arrondissement*,
- 23% of dwellings in the 20th *arrondissement*,
- 15% of dwellings in the 14th *arrondissement*.

Profile of class 5

This class represents 12% of transactions in Paris. The best represented modalities of variables are:

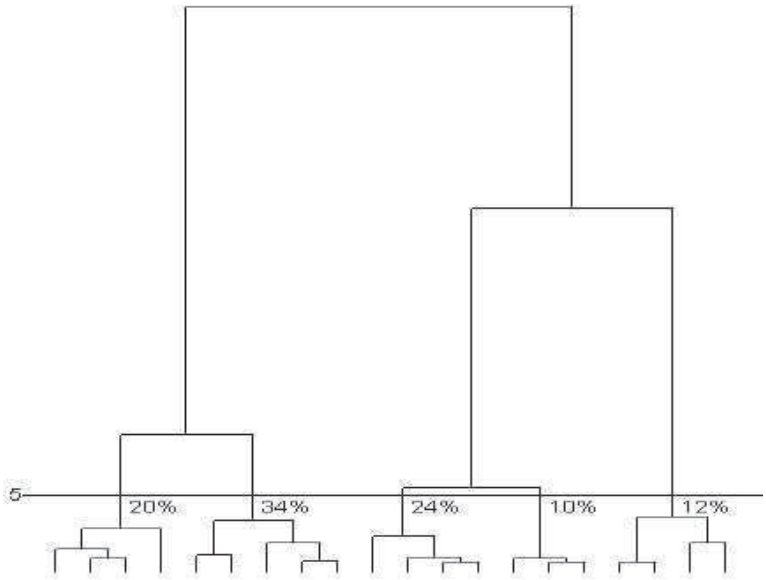
- high price per sq. m.,
- construction period between 1950 and 1970,
- apartment with large living space,
- apartment with 2 or more bathrooms,
- apartment with lift.

The geographic distribution of dwellings in this class is as follows:

- 100% of dwellings in the 8th and 16th *arrondissements*,
- 50% of dwellings in the 7th *arrondissement*,
- 15% of dwellings in the 17th *arrondissement*.

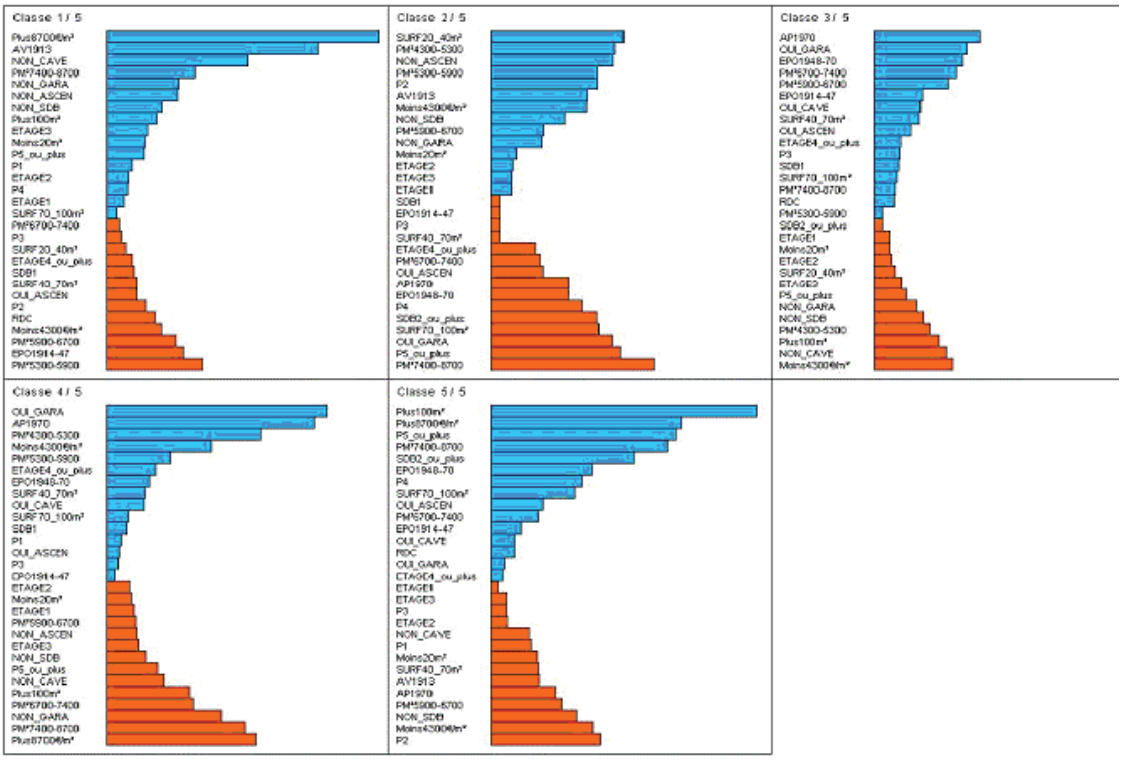
Figure A2.7 gives a visualisation of the spatial distribution of the five classes for apartments in Paris.

Figure A2.5 – Tree analysis for apartments in Paris



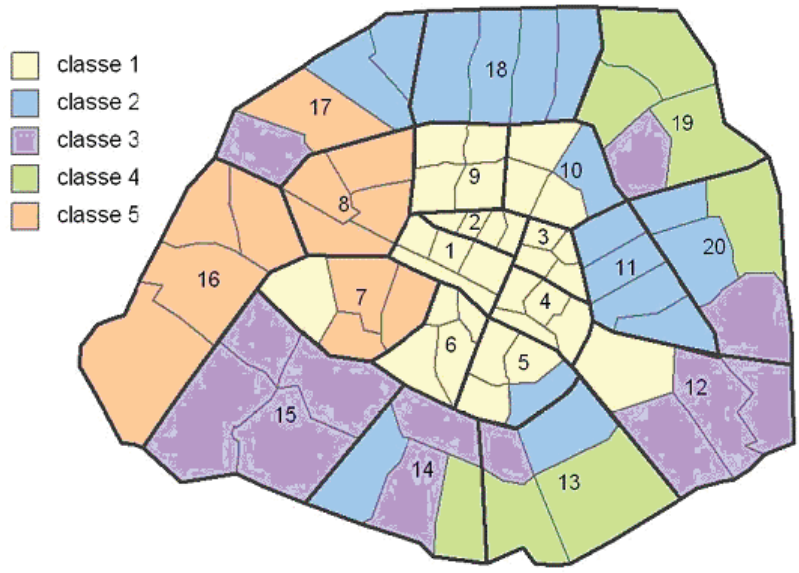
Source: BIEN database
How to read this chart: the straight line defining the division into 5 classes corresponds to an optimal choice in terms of strategy.

Figure A2.6 – Looking for over-represented variables for each class



Source: BIEN database
How to read this chart: The blue bars correspond to over-represented variables in the class.

Figure A2.7 - Division into classes for apartments in Paris



Source: BIEN database

Testing for strata homogeneity

Once the division into strata has been completed, we check the homogeneity of prices within a stratum, and also heterogeneity between strata. We propose to take as an example the studies carried out by the PNS for apartments in Paris. The studies by Min.not are similar.

Table A2.2 provides a first visual examination of average prices by strata for apartments in Paris. From 2003 - 2007, these averages varied from €4,514 to €6,549 per sq.m. Confidence intervals are very often unconnected, which seems to indicate that prices are different from one stratum to another.

Table A2.2 – Statistics for prices per sq.m. in apartments in Paris strata

Stratum	Number of transactions	Average	Standard deviation	Standard error	95% confidence interval for mean		Minimum	Maximum
					Lower bound	Upper bound		
7501	24,724	6,190.61	2,037.04	12.96	6,165.21	6,216.00	322.58	2,4851.37
7502	40,714	4,919.52	1,323.92	6.56	4,906.66	4,932.38	148.24	1,8152.17
7503	28,771	5,453.07	1,361.13	8.02	5,437.34	5,468.80	435.15	1,9402.71
7504	12,488	4,514.04	1,208.99	10.82	4,492.83	4,535.24	431.97	1,7784.15
7505	16,339	6,549.16	1,874.81	14.67	6,520.41	6,577.90	877.37	2,2000.00
Total	123,036	5,474.97	1,707.11	4.87	5,465.43	5,484.51	148.24	2,4851.37

Source: PNS

The Levene method is used to ensure that n samples extracted from the same population are of equal variance (Table A2.3). Finally, we carry out a Fisher's test (Table A2.4). One-factor analysis of variance also enables us to see whether the geographic area (factor) has an influence on the price of the property. The homogeneity hypothesis in terms of prices of strata can be rejected when significance level $\alpha = 0.05$. There are therefore different variances and different mean effects. But we do not know if the five strata are different or if it is only some of the strata that are not homogeneous one with another. To solve this problem, a post-hoc test is subsequently carried out (cf. Table A2.5) to see if all the strata are different when taken two by two. In our example, this test compares the five strata in Paris. For each stratum, the significance level is below the threshold $\alpha = 0.05$. Price differences are therefore observed between the two strata for apartments in Paris. Similar results are found for the other strata in the Paris Region and in the Provinces.

Table A2.3 - Test for equality of variances

Levene's test statistic W	Degrees of freedom 1	Degrees of freedom 2	P value of test
1,276.53	4	123,031	0.000

Source: BIEN database
H0: no difference in variance between strata
H1: at least one different variance in one stratum
If the hypothesis H0 is not confirmed, this indicates that the result does not derive from a simple sampling from one and the same population. In this case, variances of the groups are no longer equal amongst themselves at significance level α.
When significance level α is equal to 0.05, hypothesis H0 is rejected; the variances are not equal.

Table A2.4 – Analysis of variance for strata of apartments in Paris

	Sum of squares	Degrees of freedom	Mean square	Fisher's statistic	Significance level of Fisher's test F (Pvalue)
Inter-stratum variance	55,621,427,170.7	4	13,905,356,792.5	5,647.5	>0.000
Intra-stratum variance	302,929,525,100.8	123,031	2,462,221.1		
Total	358,550,952,270.9	123,035			

Source: BIEN database
H0: there is no difference in price between means for the strata
H1: there is a difference in price between means for the strata

Table A2.5 - *Post-hoc* test for strata of apartments in Paris

(I) Reference stratum	(J) Compared with strata	Difference in means (I-J)	Standard error	Significance	95% confidence interval	
					Upper bound	Lower bound
7501	7502	1,271.08	14.52	0.03	1,227.42	1,314.75
	7503	737.54	15.24	0.03	690.34 1	784.74 1
	7504	1,676.57	16.88	0.02	626.24	726.90
	7505	-358.55	19.57	0.02	-417.15	-299.95
7502	7501	-1,271.08	14.52	0.03	-1,314.75	-1,227.42
	7503	-533.55	10.37	0.03	-566.18	-500.92
	7504	405.49	12.65	0.00	370.30	440.67
	7505	-1,629.63	16.07	0.00	-1,674.40	-1,584.86
7503	7501	-737.54	15.24	0.03	-784.74	-690.34
	7502	533.55	10.37	0.03	500.92 901.05	566.18
	7504	939.04	13.47	0.00	-1,143.17	977.02
	7505	-1,096.08	16.72	0.00		-1 049.00
7504	7501	-1,676.57	16.88	0.02	-1,726.90	-1,626.24
	7502	-405.49	12.65	0.00	-440.67	-370.30
	7503	-939.04	13.47	0.00	-977.02	-901.05
	7505	-2,035.12	18.23	0.01	-2,086.88	-1,983.36
7505	7501	358.55	19.57	0.02	299.95	4,17.15
	7502	1,629.63	16.07	0.00	1,584.86	1,674.40
	7503	1,096.08	16.72	0.00	1,049.00	1,143.17
	7504	2,035.12	18.23	0.01	1,983.36	2,086.88

Source: BIEN database

Description of strata

Strata are geographic areas within which prices are homogeneous (cf. Chapter 3). Details of their breakdown into municipalities are given in an Excel file that can be obtained on request from notaries.

Strata in the Provinces

For the Provinces, the construction of the strata is based on a geographic proximity criterion, starting with the municipality. Table A2.6 shows the distribution of transactions in the reference stock of 2007-2008 in the Provinces, by part of urban unit 1999.

Table A2.6 - Provinces: transactions (apartments and houses) by part of urban unit and municipality grouping (%)

Part of urban unit 1999	Rural	UU < 5,000 inhab.*	5,000-10,000 inhab.	10,000-20,000 inhab.	20,000-50,000 inhab.	50,000-100,000 inhab.	100,000-200,000 inhab.	Over 200,000 inhab.	Total
Apartments									
Rural	6%	0%	0%	0%	0%	0%	0%	0%	6%
Central city	0%	2%	3%	5%	6%	9%	8%	33%	67%
Suburb	0%	0%	0%	1%	1%	2%	3%	20%	27%
Total	6%	2%	3%	6%	8%	11%	11%	53%	100%
Houses									
Rural	39%	0%	0%	0%	0%	0%	0%	0%	39%
Central city	0%	8%	6%	5%	5%	4%	2%	5%	35%
Suburb	0%	1%	1%	2%	2%	3%	3%	14%	26%
Total	39%	8%	7%	7%	7%	8%	5%	20%	100%

* Inhab. = inhabitants
2007-2008 Reference stock

Strata for apartment indices in the Provinces

The entire country was divided into 97 strata for used apartments: 33 strata consist of a single city (Table A2.7), thus an index is calculated for each of these 33 cities. For the other strata (Tables A2.8 et seq.), 12 are in the suburbs of these cities, 5 correspond to an entire urban unit, 14 consist of atypical municipalities (coastal towns and winter sports resorts); the remaining 33 are groupings of towns made according to the region and the average income per inhabitant (see definitions below).

A certain number of series of indices disseminated by the notaries do not have the Notaires-INSEE designation. These are mainly regional and departmental series and indices for the 33 municipalities listed in Table A2.7.

These indices are disseminated on the websites immoprix.com and perval.fr, also via press conferences organised by the Chambers of Notaries on property prices and via survey reports produced by the notaries.

The list of indices that may be disseminated is checked approximately every two years. An index series can be disseminated when the number of transactions per quarter is greater than or equal to 110 for several consecutive quarters.

Table A2.7 – List of 33 cities in the Provinces for which an apartment index is calculated

Municipality	Population*	Number of dwellings*	Number of transactions in reference stock**
Antibes	76,994	27,135	1,871
Agde	22,487	3,862	1,719
Annecy	50,115	20,372	1,020
Besançon	117,599	36,396	1,974
Bordeaux	235,891	78,724	2,539
Brest	142,097	40,430	2,411
Caen	109,899	33,206	942
Cannes	72,939	26,181	2,836
Le Cannet	40,940	15,169	1,205
Clermont-Ferrand	139,006	42,322	1,349
Dijon	151,576	52,136	2,628
Grenoble	156,659	61,561	2,212
Le Havre	178,769	32,486	1,304
Lille	225,784	61,103	2,328
Lyon	474,946	188,776	6,808
Marseille	851,420	244,867	6,160
Menton	28,833	9,775	1,133
Metz	122,838	35,076	1,566
Montpellier	252,998	84,784	2,399
Nancy	106,361	40,915	1,756
Nantes	283,288	82,438	4,123
Nice	439,553	132,175	8,541
Orléans	113,257	28,652	1,593
Pau	84,036	27,637	1,209
Reims	181,468	33,262	1,913
Rennes	206,655	65,019	2,972
Rouen	109,425	38,595	1,448
Saint-Étienne	172,696	53,944	1,743
Strasbourg	272,116	84,070	2,929
Toulon	166,733	50,808	1,906
Toulouse	439,553	152,830	4,674
Tours	135,480	37,007	1,380
Villeurbanne	141,106	47,334	1,963

* Source: population census 2008

** 2007-2008 stock

With the exception of Lyon and Marseille, the associated indices do not have the Notaires-INSEE designation.

Table A2.8 - List of 17 strata (suburb or urban unit in the Provinces) for which an apartment index is calculated

Stratum	Number of transactions in reference stock*
Suburb of Lille (French part)	1,903
Suburb of Annecy	1,111
Suburb of Bordeaux	1,861
Suburb of Grenoble	2,340
Suburb of Nancy	1,458
Suburb of Nantes	1,099
Suburb of Rouen	936
Suburb of Strasbourg (French part)	1,547
Suburb of Nice	4,941
Suburb of Toulon	3,807
Suburb of Lyon	4,689
Suburb of Marseille-Aix-en-Provence	2,108
Urban unit of Chambéry	1,148
Urban unit of Mulhouse	1,602
Urban unit of Saint-Nazaire	848
Urban unit of Genève-Annemasse (French part)	1,944
Urban unit of Bayonne	1,953

* 2007-2008 stock

Twelve strata are composed of the suburbs of an urban unit and five are composed of the entire urban unit. The associated indices do not have the Notaires-INSEE designation and are not published.

Table A2.9 – List of 14 strata of winter sports and coastal resorts for which an apartment index is calculated

Stratum	Number of transactions in reference stock*
Coastal towns in 'Grand Nord', Nord - Pas-de-Calais and Picardie regions	2,013
Coastal towns in Haute and Basse Normandie regions	2,228
Coastal towns in Bretagne *départements* Ille-et-Vilaine and Morbihan	2,062
Coastal towns in Bretagne *départements* Finistère and Côte d'Armor	2,262
Coastal towns in Pays de la Loire region	1,140
Coastal towns in South-West Poitou-Charentes and Aquitaine regions	2,584
Coastal towns in Languedoc-Roussillon South	2,457
Coastal towns in Languedoc-Roussillon North	2,038
Coastal towns in PACA West	1,551
Coastal towns in PACA East	3,058
Ski resorts in Tarentaise	1,871
Ski resorts in Maurienne, Val d'Arly, Beaufortin, les Bauges, Chablais, Giffre, Les Bornes and Chartreuse	1,775
Ski resorts in Mont-Blanc and Aravis	1,136
Ski resorts in Isère	1,198

* 2007-2008 stock

Fourteen strata are composed of "atypical" municipalities (e.g. coastal municipalities, winter sports resorts), which had to be isolated in order to obtain coherent strata in the regressions. The associated indices do not have the Notaires-INSEE designation and are not published.

Table A2.10 - Other strata for which apartment indices are calculated, by average income 2006

Stratum	Number of transactions in reference stock by income bracket*		
	Low income	Medium income	High income
Other municipalities in Alsace	2,052		1,594
Other municipalities in Bourgogne	1,703		1,444
Other municipalities in Franche-Comté	1,657		1,512
Other municipalities in Lorraine	2,669		2,505
Other municipalities in the West except Pays de la Loire	2,580		1,812
Other municipalities in PACA North	1,775	1,667	1,454
Other municipalities in Picardie	1,867		1,419
Other municipalities in the Centre	2,526		1,956
Other municipalities in Limousin and Auvergne	1,763		1,503
Other municipalities in Midi-Pyrénées	2,076		1,596
Other municipalities in the North except Picardie	1,908	2,118	1,518
Other municipalities in Pays de la Loire	1,936		1,689
Other municipalities in Rhône-Alpes-East	2,051	1,900	2,498
Other municipalities in Rhône-Alpes-West	2,404		2,198
Other municipalities in the South-East (Languedoc-Roussillon and PACA)	2,016		3,236

* 2007-2008 stock

Municipalities in the 33 strata listed above have insufficient sales per quarter, and do not have any particular features to distinguish them. They have therefore been classified according to the average income per inhabitant for 2006.
The associated indices do not have the Notaires-INSEE designation and are not published.

Strata for house indices in the Provinces

There are 37 strata that concern towns, suburbs or an entire urban unit (Table A2.11); the other 137 strata are groupings of towns by their region and their average income per inhabitant in 2006 (Table A2.12).

Table A2.11 – List of 37 towns, suburbs and urban units for which a house index is calculated

Municipality	Number of transactions in reference stock**
Amiens	970
Bordeaux	1,192
Le Havre	868
Le Mans	1,394
Lille	931
Marseille	826
Nantes	1,146
Toulouse	744
Tourcoing	1,060
Suburb of Bordeaux	3,631
Suburb of Lille (French part)	6,772
Suburb of Marseille-Aix-en-Provence	1,990
Suburb of Nantes	2,640
Suburb of Toulouse	2,222
Urban unit of Avignon	949
Urban unit of Béthune	1,417
Urban unit of Brest	1,227
Urban unit of Calais	1,129
Urban unit of Clermont-Ferrand	929
Urban unit of Dijon	907
Urban unit of Douai-Lens	3,298
Urban unit of Dunkerque	970
Urban unit of Grenoble	800
Urban unit of Lyon	3,811
Urban unit of Maubeuge (French part)	903
Urban unit of Metz	1,547
Urban unit of Montpellier	815
Urban unit of Nancy	2,258
Urban unit of Nice	3,241
Urban unit of Orléans	1,966
Urban unit of Reims	906
Urban unit of Rennes	1,047
Urban unit of Rouen	2,708
Urban unit of Saint-Nazaire	770
Urban unit of Toulon	2,089
Urban unit of Tours	1,709
Urban unit of Valenciennes (French part)	1,991

* 2007-2008 stock

With the exception of the Lille index, these indices do not have the Notaires-INSEE designation and are not published.

Table A2.12 - Other strata for houses in the Provinces

Stratum	Number of transactions from the reference stock**			
	Income			
	Very low	Low	Medium	High
Ski resorts in Isère, Savoie and Haute Savoie				1,597
Coastal towns in Gironde, Landes and Pyrénées-Atlantiques		1,573		1,127
Coastal towns in Loire-Atlantique, Vendée and Charente-Maritime		2,186		1,762
Coastal towns in Ille-et-Vilaine, and Morbihan	1,128	985	958	1,044
Coastal towns in Finistère and Côte d'Armor	1,434	1,263	1,139	1,186
Coastal towns in Nord - Pas-de-Calais and the two Normandies	1,244	1,385	1,359	1,148
Coastal towns in the Mediterranean		1,789	1,490	1,583
Other municipalities in Alsace		1,567	1,685	1,686
Other municipalities in Aquitaine North-East		2,020		1,852
Other municipalities in Aquitaine South-West		2,179		2,301
Other municipalities in Auvergne		1,409	1,779	1,985
Other municipalities in Basse-Normandie	2,320	2,220	2,219	2,134
Other municipalities in Bourgogne East (21,71)		2,287	2,254	2,184
Other municipalities in Bourgogne West (89,58)		1,453	1,620	1,967
Other municipalities in Bretagne North		2,640	2,455	2,309
Other municipalities in Bretagne South		2,371		2,451
Other municipalities in Champagne-Ardenne North		1,951	1,996	1,940
Other municipalities in Champagne-Ardenne South		1,836		1,703
Other municipalities in Franche-Comté		2,308	2,148	2,134
Other municipalities in Gironde		1,908		2,003
Other municipalities in Somme		1,611	1,633	1,511
Other municipalities in Ain		2,117		2,236
Other municipalities in Aisne	1,760	1,976	1,723	1,934
Other municipalities in Eure		1,999	1,853	1,870
Other municipalities in Isère		1,889		1,973
Other municipalities in Loire-Atlantique		2,354		2,141
Other municipalities in Oise		1,468	1,455	1,450
Other municipalities in Lorraine East (54,57)		2,429	2,408	2,548
Other municipalities in Lorraine West (55,88)		2,279		2,126
Other municipalities in Midi-Pyrénées North		1,272	1,144	1,537
Other municipalities in Midi-Pyrénées South		1,500	1,479	1,624
Other municipalities in PACA North		1,708		1,717
Other municipalities in PACA South		1,398	1,570	1,187
Other municipalities in Poitou-Charentes North		1,703	1,792	1,621
Other municipalities in Poitou-Charentes South		3,159		3,103
Other municipalities in Savoie and Haute-Savoie		1,532	1,525	1,425
Other municipalities in Seine-Maritime		1,670	1,917	1,844
Other municipalities in Vendée		1,703		1,609
Other municipalities in Eure-et-Loir		1,733	1,713	1,870
Other municipalities in Indre and Indre-et-Loire		1,314	1,442	1,536
Other municipalities in Languedoc-Roussillon	1,569	1,611	1,917	1,763
Other municipalities in Limousin		1,953		2,034
Other municipalities in Loir-et-Cher		1,879		1,785
Other municipalities in Loiret		1,493	1,403	1,574
Other municipalities in Cher		1,362		1,283
Other municipalities in Nord	1,608	1,604	1,580	1,542
Other municipalities in Pas-de-Calais		1,625	1,660	1,768
Other municipalities in Pays de la Loire North	2,904	2,731	3,012	2,942
Other municipalities in Rhône-Alpes West	1,883	2,031	1,936	2,017

* 2007-2008 stock

The associated indices do not have the Notaires-INSEE certification and are not published.

Strata in the Paris Region

In the Paris Region, strata were determined by principal component analysis of the average property price by municipality. The result was that the strata were not constructed according to the geographical proximity criterion and could not be identified by name as was the case in the Provinces.

In the table below, we show the distribution by *département* of the number of transactions by stratum (transactions in reference stock 2007 -2008) and by dwelling type.

Table A2.13 – Paris Region: number of strata by *département* (excl. Paris) and corresponding numbers in the reference stock*; apartments

Apartments*								
Département	77	78	91	92	93	94	95	Total
Stratum								
1	2,089	434	2,086	-	-	126	986	5,721
2	2,229	158	73	-	-	162	314	2,936
3	1,217	475	1,119	-	1,204	-	1,546	5,561
4	523	1,532	1,728	-	2,348	2,863	1,895	10,889
5	1,612	3,628	3,927	-	846	831	2,021	12,865
6	-	1,596	-	5,308	4,319	6,102	2,082	19,407
7	-	2,172	-	1,527	-	-	-	3,699
8	271	3,636	1,781	3,147	-	31	326	9,192
9	-	-	-	8,528	-	3,238	-	11,766
10	-	-	149	7,294	6,150	4,649	-	18,242
Total	7,941	13,631	10,863	25,804	14,867	18,002	9,170	100,278

* 2007-2008 stock

Table A2.14 – Paris Region: number of strata by *département* (incl. Paris) and corresponding numbers in the reference stock*; houses

Houses*									
Stratum	75	77	78	91	92	93	94	95	Total
1	133	234	1,321	259	2,248	-	1,163	1,081	6,439
2	-	-	-	-	690	709	379	-	1,778
3	-	-	984	1,082	109	4,988	2,040	2,623	11,826
4	-	1,603	466	1,959	-	770	829	-	5,627
5	-	2,304	956	828	-	-	484	3,193	7,765
6	-	2,867	3,918	3,009	49	-	332	154	10,329
7	-	6,801	2,110	1,817	-	-	-	1,300	12,028
Total	133	13,809	9,755	8,954	3,096	6,467	5,227	8,351	55,792

How to read this chart: For houses, stratum 1 includes Paris.
* 2007-2008 stock

Strata for apartments in the Paris Region

Table A2.15 - List of 5 strata for apartment indices in Paris

Stratum	Neighbourhood	Number of transactions in reference stock*
7501	1 to 17, 19 to 24, 28, 33 to 39 and 48	9,346
7502	18, 40 to 44, 49, 56, 67 to 72, 77, 79	15,139
7503	45 to 47, 52, 53, 55, 57 to 60, 65, 76, 80	10,689
7504	50, 51, 54, 73 to 75, 78	4,635
7505	25 to 27, 29 to 32, 61 to 64, 66	6,002
Total	80 neighbourhoods	45,811 transactions

* 2007-2008 reference stock
Five strata were constructed from the 80 neighbourhoods in Paris.

Table A2.16 – List of 10 strata for apartment indices in Paris Region, excl. Paris

Stratum	Number of municipalities	Number of transactions in reference stock*
1	790	5,721
2	87	2,936
3	23	5,561
4	41	10,889
5	184	12,865
6	51	19,407
7	3	3,699
8	61	9,192
9	13	11,766
10	27	18,242
Total	1,280 municipalities	100,278 transactions

* 2007-2008 reference stock

Ten strata were constructed for apartments from the 1,280 municipalities in the Paris Region, excl. Paris. These indices do not have the Notaires-INSEE designation and are not published.

Strata for houses in the Paris Region

Table A2.17 – List of 7 strata for house indices in Paris Region

Stratum	Number of municipalities	Number of transactions in reference stock*
1	64	6,439
2	23	1,778
3	81	11,826
4	56	5,627
5	98	7,765
6	286	10,329
7	673	12,028
Total	1,281 municipalities	55,792 transactions

* 2007-2008 reference stock

Seven strata were constructed for houses in the Paris Region, from the 1,281 municipalities. Paris is included in stratum 1 because the volume of house sales is low. These indices do not have the Notaires-INSEE designation and are not published.

Appendix 3: Updating the model specifications

The specifications for the models are reviewed each time the database is revised. Most of the changes in version 3 compared with version 2 relate to modalities for the plot size and living space variables in houses; we also wanted to harmonise as completely as possible the models for the Paris Region and the Provinces.

Changes to processing the surface area in houses

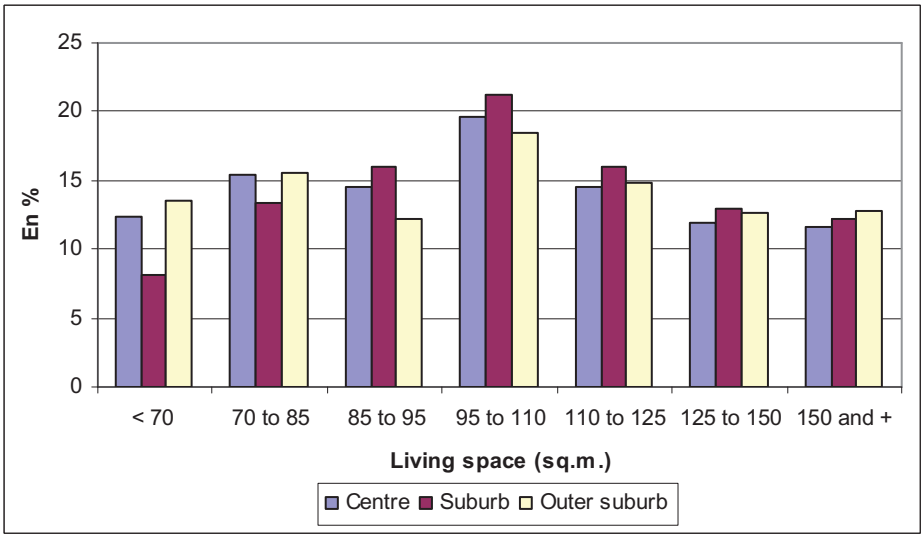
Analysis of surface area variables

The prices that we observe are used to estimate the value of a reference property, which in turn is used to estimate the value of the reference stock at time *t*. In the earlier version of the indices, the reference value for continuous variables was not fixed, and thus had a zero default setting. The estimation of relative prices has not changed. However, it is more correct to choose as a reference property one with a plot size and living space that are not zero.

The question then had to be asked as to whether different reference vales should be used for central cities, suburbs and rural municipalities. We therefore compared the respective distributions according to municipality type. In the end, we used the same values for the Paris Region and the Provinces.

For the Provinces, the modal class of living space in a dwelling is the same (Figure A3.1) whatever the municipality type [class 95 to 110 sq.m.]. There is not a great deal of difference in the three series, although we note that municipalities in the suburbs are less well represented in the classes for small living space (less than 85 sq.m.).

Figure A3.1 - Distribution of transactions by living space, houses in the Provinces

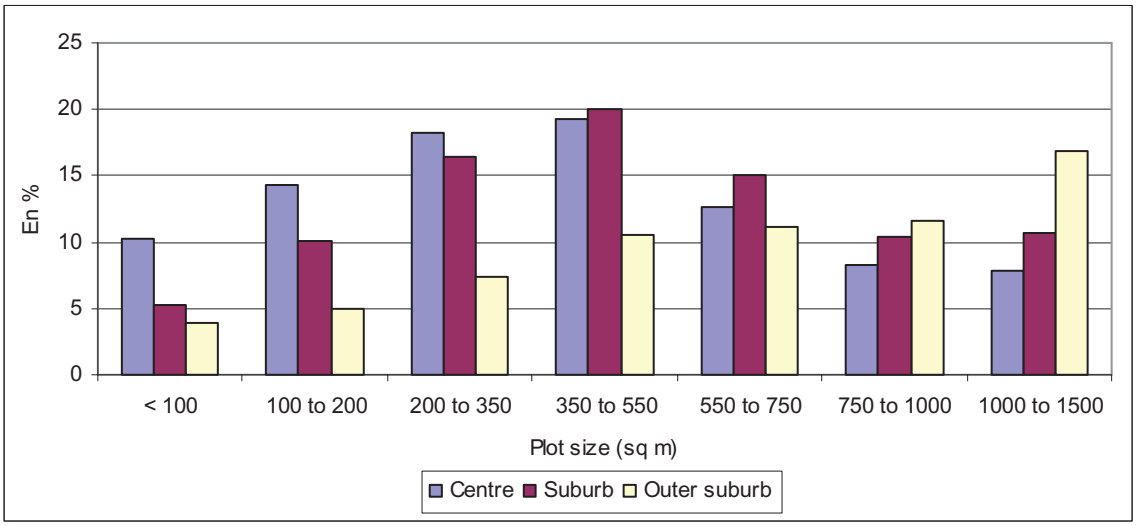


Source: PERVAL database, transactions observed between 2004 and 2008
How to read this chart: Houses with 95 to 110 sq.m. represent more than 20% of all houses located in suburbs.

Differences are more pronounced when considering plot size (Figure A3.2). In outer-suburban municipalities, plots are larger: 50% of houses in outer-suburban municipalities have plots of over 1,000 sq.m. compared with only 23% of suburban municipalities and 17% of central city municipalities. This is reflected in the value of the modal classes for the series: for outer-suburban municipalities the interval is 1000 to 1500 sq.m. compared with a smaller interval class of 350 to 550 sq.m. for the other two types of municipality.

However, it is not possible in all cases to take different reference modalities according to municipality type, as some strata contain both central city municipalities and outer-suburban municipalities.

Figure A3.2 - Distribution of transactions by plot size, houses in the Provinces

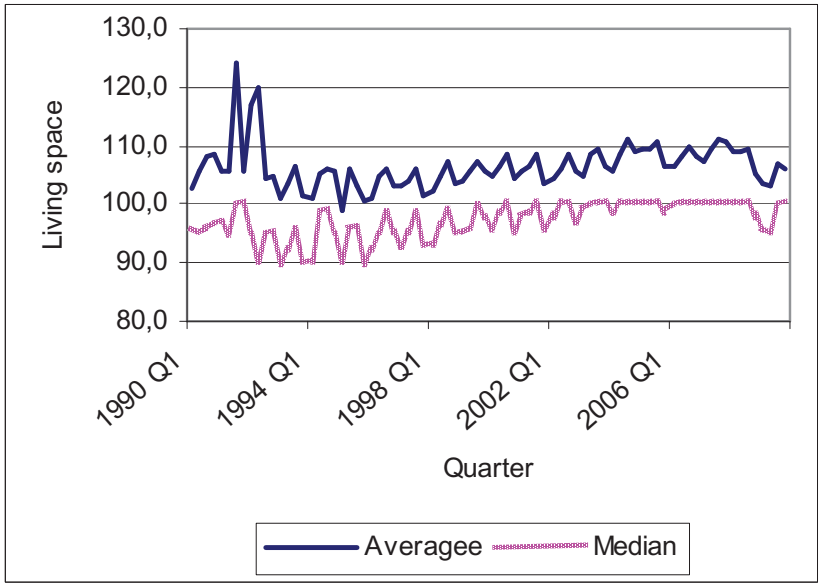


Source: PERVAL database, transactions observed between 2004 and 2008
How to read this chart: Houses with a plot size of 350 to 550 sq.m. represent almost 20% of all houses located in suburbs.

With regards to changes over time, if we take the example of the Paris region, the quarterly living space average has an irregular profile in the period from 1990 Q1 to 2009 Q4. The median value, which is less sensitive to extreme values, is more stable, especially from 2003 onwards. Considerable seasonal variation can be seen (Figure A3.3).

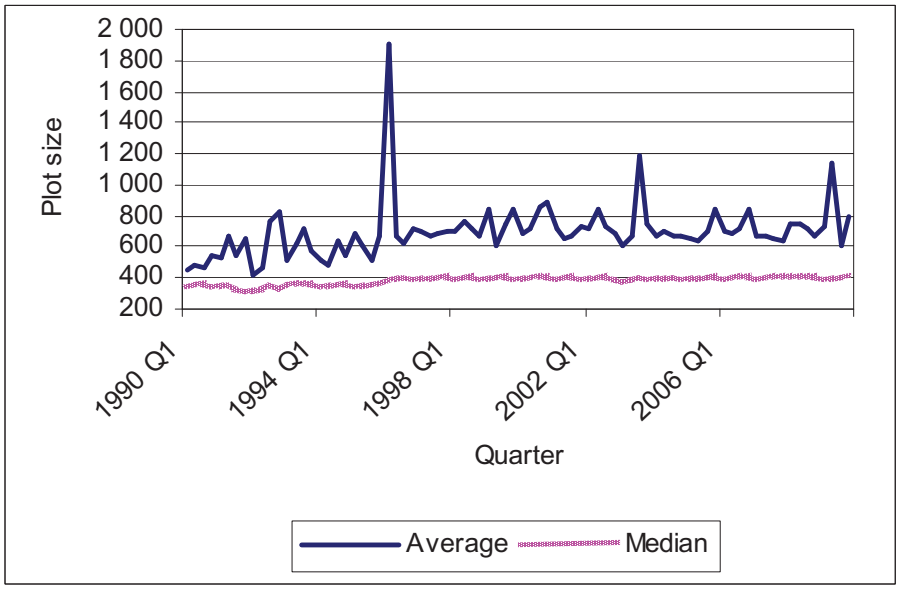
With regards to plot size, the median fluctuates little, while the average is much more volatile (Figure A3.4).

Figure A3.3 - Living space, median and average values, houses in Paris Region



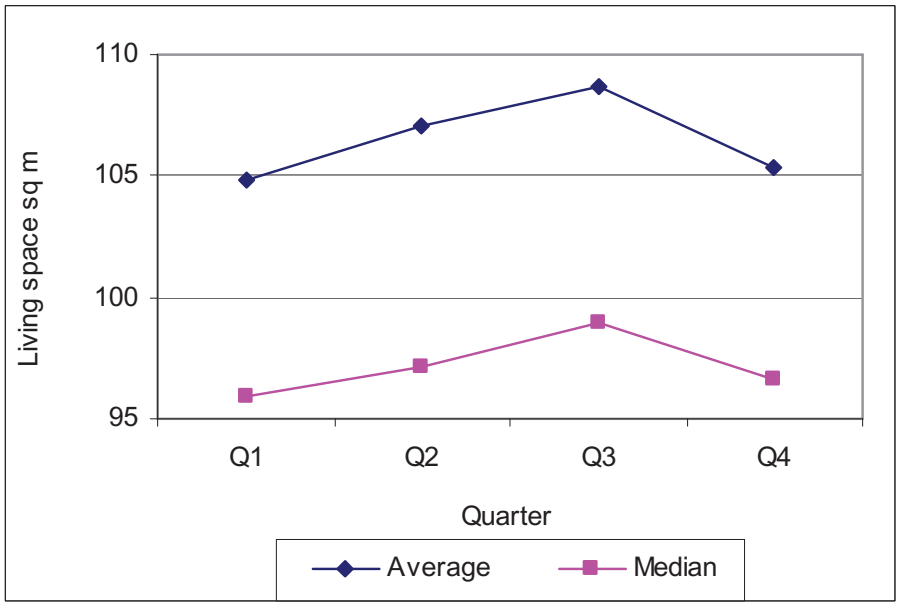
Source: BIEN database, transactions observed between 1990 Q1 and 2009 Q4

Figure A3.4 - Plot size, median and average values, houses in Paris Region



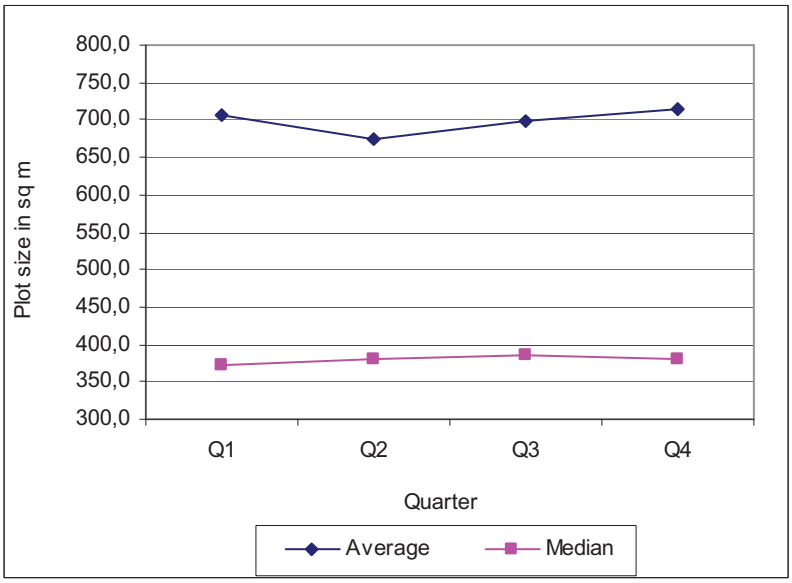
Source: BIEN database, transactions observed between 1990 Q1 and 2009 Q4
Comment: the change in the series is affected by the addition of the area of the outer Paris suburbs from 1996

Figure A3.5 - Living space, median and average values, houses, Paris Region by quarter



Source: BIEN database, transactions observed by quarter (between 1990 Q1 and 2009 Q4)

Figure A3.6 - Plot size, median and average values, houses in Paris Region by quarter



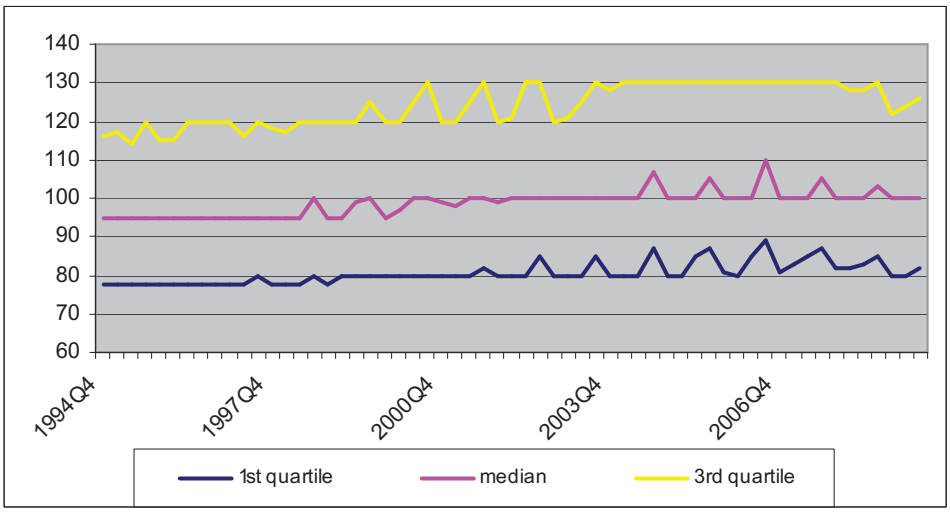
Source: BIEN database, transactions observed by quarter (between 1990 Q1 and 2009 Q4)

Choice of reference values

The volatility of plot size is less than that of living space. There is little seasonal variation in the median value.

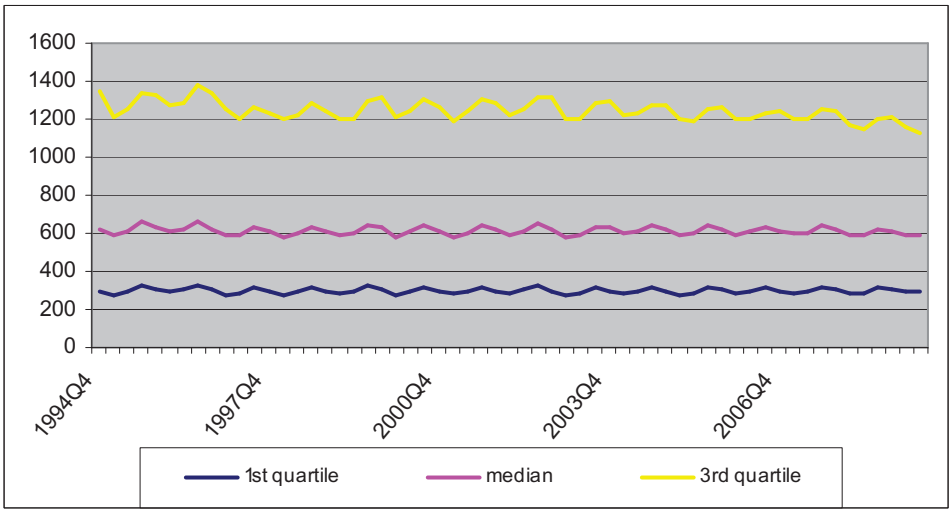
Figures A3.7 and A3.8 show the change in the median and the two extreme deciles for the two variables for the Provinces.

Figure A3.7 - Living space in houses in the Provinces for three quantiles



Source: PERVAL database, transactions observed between 1994 Q4 and 2009 Q2
How to read this chart: the median of living space in houses was relatively stable over the last ten years.

Figure A3.8 - Plot size in houses in the Provinces for three quantiles



Source: PERVAL database, transactions observed between 1994 Q4 and 2009 Q2

Taking living space into account in the models

The reference house selected has living space measuring 100 sq.m., and plot size 610 sq.m. These values are the same for the Paris Region and the Provinces.

We also need to know if we must introduce a transformation of the living space variables, for example by using their logarithm or breaking them down into segments. The choice depends on our idea of the link between the price of the property and living space (or plot size). If we think that the link is linear, then we opt for logarithms. The hypothesis that we then make is that price-elasticity is constant. A log-log model has an advantage in that we are able to read elasticities directly. If, on the other hand, we think that the link between price and plot size or living space is not linear, then a solution is to divide the variables into classes. In addition, this discretisation makes it easier to deal with discontinuity in the living space. Living space in houses is often rounded up or down (to the nearest 5 or 10 square metres).[63]

In the previous version of the indices, plot size and living space were expressed as raw data and not as logarithms (we hypothesised that house prices were an exponential function of $\beta_{k,\tau}$ coefficients of the hedonic model, instead of being a linear function). When determining house prices, we therefore tended to assign rather too little influence to these two variables compared with the other variables. This resulted in an adjustment coefficient close to zero, which in turn increased the seasonal variation in the series of price indices artificially.

In order to simplify this revision, the choice was made to use logarithms. Bracketing categories requires many steps, which can be carried out during the next revision: finding the optimal number of categories (5 to 7 normally[64]), determining their values; differentiating them, where necessary, according to strata; testing the hypothesis of price elasticity stability for each stratum in relation to living space.

Impact of change in the method on the series of indices

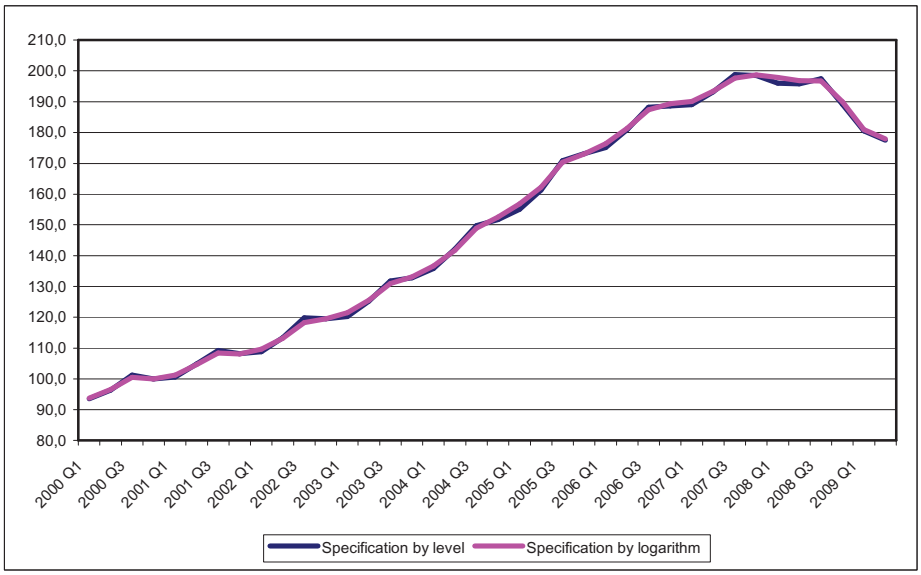
We also studied the impact of the new method on determining coefficients of the characteristics vector and on the price index series. Here, we present results for the Provinces.

The old method gives a more volatile series of results than the new (Figure A3.9). Seasonal variation in the indices is reduced using the new method. Divergences between the indices are fairly low, less than about 2% in absolute value.

[63] Plot sizes are measured more carefully, however, because they correspond to the cadastral surface areas provided in the notarial deeds.

[64] There should not be too many groupings, otherwise we tend to revert to a single continuous case. This also increases the number of coefficients to be estimated.

Figure A3.9 - House price indexes for the Provinces



Source: PERVAL database, transactions observed between 2000 Q1 and 2009 Q3
How to read this chart: the series with surface area variables given by level shows a more pronounced seasonal variation in its profile than the series with the surface area expressed as logarithm.

Figure A3.10 - Divergences in indexes between the old method and the new method

Source: PERVAL database, transactions observed between 2000 Q1 and 2009 Q3
How to read this chart: the divergence corresponds to the "old method – surface area variables by level" index minus the "new method - surface area variables as logarithm" index.

Specification of the new models

The dependent variables have not been modified: for houses, the total price is used whereas for apartments, we take the price per sq.m. For the Provinces, we keep the same explanatory variables as in version 2. There is one change: in version 2, the number of rooms and the neighbourhood[65] were introduced separately, whereas in version 3, they have instead been cross-tabulated.

[65] In the general sense of an area included in a stratum: it can be the neighborhoods in the cities but also larger groupings outside the cities.

The models for the Paris Region and the Provinces cannot be exactly alike, because some explanatory variables are not found in both notarial databases (e.g. the variable to codify the general condition of the property is only present in the PERVAL database).

See the tables below for the list of variables introduced into the models. All qualitative variables have been discretised.

Table A3.1 – List of explanatory variables in models for apartments

Explanatory variables	Present in Paris Region model	Present in Provinces model
Year of transaction	Yes	Yes
Month of transaction	Yes	Yes
Construction period	Yes	Yes
Number of bathrooms	Yes	Yes
Cellar	Yes	Yes
Number of garages	Yes	Yes
Floor*	Yes	Yes
Lift*	Yes	Yes
Cross-tabulation lift and number of floors*	Yes	Yes
Cross-tabulation number of rooms and neighbourhood	Yes	Yes
Floor space per room**	Yes	Yes
Condition of property	No	Yes
Terrace or balcony	No	Yes

* For apartments located on upper floors (4th floor and higher), we crossed the variables "number of floors" and "presence of lift". Below this, we believe that the variable "lift" has no effect on prices and we used only the variable "floor".
** Logarithm

Table A3.2 – List of explanatory variables in models for houses

Explanatory variables	Present in Paris Region model	Present in Provinces model
Living space*	Yes	Yes
Plot size*	Yes	Yes
Year of transaction	Yes	Yes
Month of transaction	Yes	Yes
Construction period	Yes	Yes
Number of bathrooms	Yes	Yes
Basement	No	Yes
Number of garages	Yes	Yes
Number of storeys	Yes	Yes
Condition of property	No	Yes
Cellar	Yes	No
Cross-tabulation number of rooms and neighbourhood	Yes	Yes

* Logarithm

Appendix 4: Stability tests and duration of reference period

The relative prices of property characteristics, estimated from the hedonic models, will be used to calculate price indices with constant characteristics. Hedonic models are estimated from transactions over a given period, called the "estimation period".[66] As transactions are relatively infrequent but relative prices change over time, the estimation period selected must be long enough to provide sufficient observations but not so long that the model coefficients can be considered as constants. Stability test are carried out to determine the optimal duration that should be selected (in number of years). We must ensure that in statistical terms the $\hat{\beta}_k$ coefficients of the characteristics can be considered as constant over the estimation period. The time effect is captured by the time dummies in the econometric equation.

Stability tests in version 2

In version 2 the estimation period was sixteen quarters (1998 to 2001). We tested the stability of the $\hat{\beta}_k$ coefficients across the estimation period by verifying that the gap between the value of properties with the characteristic X_k estimated by the model, and their actual selling price (residual ε_i) meets the model's stochastic hypotheses, and in particular that it does not contain any overlooked deterministic trend.

Let $\overline{\varepsilon}_{t,1}$ be the mean of the residuals measured for all dwellings in quarter t that have characteristic X_1 . We build the sequence of mean residuals $\overline{\varepsilon}_{1,1}$, $\overline{\varepsilon}_{2,1}$,...., $\overline{\varepsilon}_{16,1}$ for the sixteen quarters in the estimation period. Bearing in mind that the residuals have a zero mathematical mean, and diverge from 0, we were able to conclude that overall the coefficients were stable across the period 1998 to 2001.

Stability tests in version 3

The sharp rise in prices observed in the 2000s which was then interrupted in 2008 mean that it is particularly crucial to consider this question of model stability. A stability test in the form of a Chow test was applied to each econometric equation. In addition, to ensure that conditions were valid for the Chow test, in a second step we estimated the models using the quasi-generalised least squares estimator, to guard against any distribution problems that might arise due to the existence of heteroscedasticity.

The Chow test is carried out on the sum of the squared residuals from the second step in the estimation. This is a traditional analysis of variance test, in the form of a Fisher's test. The scope of the analysis is observations from the estimation stock. Observations are divided into two groups of years and two sub-models: the years 2004-2005, the "years at the beginning of the period" group and years 2006-2007, the "years at the end of the period" group. Tests showed that the coefficients estimated across the two periods were significantly different. An estimation was applied with variable selection but did not improve the results.

Table A4.1 shows an example of the analysis results for houses and apartments in the Provinces. It gives the number and the weight of the strata where the coefficient stability hypothesis was rejected. We could have gone on to differentiate the duration of the reference stock by strata according to the test results: for example, four years for the models where the stability hypothesis was accepted and two years for the rest. However, for practical reasons –in particular, different methods were needed for chain-linking the price indices– we decided to retain a single estimation duration in the model for all strata. The proportion of strata where the zero hypothesis was rejected increased naturally with the type 1 error value. For the most frequently uses error level (α=0.05), this proportion was high for houses (46.0%). It was even more so for apartments (for α=0.01, this share reached 80.4% and rose to 85.6% for α=0.05). Table A4.2 shows that correcting the heteroscedasticity does not alter the results: the model stability hypothesis cannot be sustained, either for houses or apartments, across the period 2004-2007. The relative prices of the characteristics are not constant over time. The hypothesis of hedonic model stability over a four-year period cannot be sustained. The same result was obtained for the Paris Region.

[66] See definitions in Chapter 3.

Table A4.1 – Number of strata where stability hypothesis is rejected at threshold α,[67] OLS estimator

Significance level	Houses		Apartments	
Level of risk	Number of strata	%	Number of strata	%
α=0.001	27	15.5	60	61.9
α=0.01	52	29.9	78	80.4
α=0.05	80	46.0	83	85.6

Source: PERVAL database
Number of strata for houses, 174; for apartments, 97
H0: the model is stable over time
H1: the model is not stable (at least one coefficient is significantly modified)

Table A4.2 – Number of strata where stability hypothesis is rejected at threshold α; QGLS estimator

	Houses		Apartments	
Significance level	Number of strata	%	Number of strata	%
α=0.001	20	11.5	58	59.8
α=0.01	44	25.3	72	74.2
α=0.05	81	46.6	86	88.7

Source: PERVAL database
Number of strata for houses, 174; for apartments, 97
With the OLS estimator a majority of hedonic models do not satisfy the stability hypothesis. However, the reliability of the tests may need to be treated with caution when heteroscedasticity is present, in which case the OLS estimator is not effective and the usual tests no longer work. Graphs were produced linking the squared residuals with the predicted values for the most influential variables but were not able to determine which variables were associated with the heteroscedasticity. We therefore calculated the quasi-generalised least squares (QGLS) estimator, by hypothesising a different variance across each period. This method was used when the shape of the variance-covariance matrix was unknown. We applied this adjustment when estimating the hedonic function in the second step, i.e. after removing the atypical points discovered during the first step. The aim was to find a good estimator of the variances-covariances of the risks.

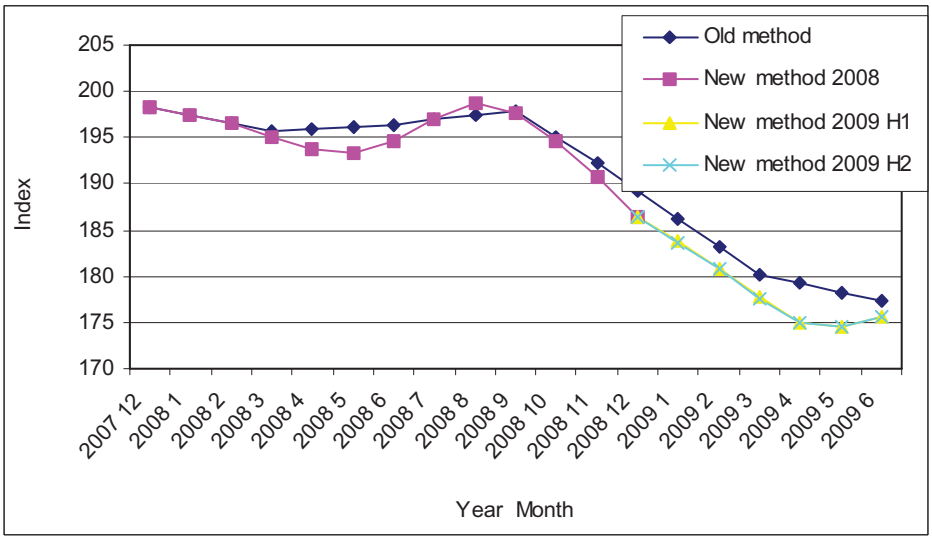
Choice of reference periods and updating the vector of relative prices of dwelling characteristics

Thus in version 3, the estimation period was reduced to two years, whereas it had been four years in version 2. This reduction was also applied to the reference stock. The definition of the strata and the specification for the econometric equations will remain the same for the lifetime of version 3 but the β_k coefficients will be updated every two years. Note that we could have decided to update them every year. These decisions were based on a comparison for 2008 and 2009 of quarterly price indices calculated using different methods relating to the length of the estimation period and the frequency with which coefficients are updated (Figure A4.1).

All three graphs have almost the same pattern of change. However, the shape of the series calculated using a four-year estimation stock ("old method") is less uneven than the other two series, which in fact are combined for 2008 as they were calculated from the same estimation stock (that of 2006-2007). Divergences between the series calculated with the four-year estimation stock and those based on the two-year stock are relatively large, ranging from 1.5 points to -4 points. The graph showing the "old method" series is always higher than the other two except for the 2nd and 3rd quarters of 2008. For the series calculated using the two-year estimation stock, the calculation method made little difference: divergences were zero in 2008 and ranged from -0.0012 to 0.1317 percentage points in 2009.

[67] A type 1 error (error of the first kind) is the probability that the hypothesis of coefficient stability be incorrectly rejected.

Figure A4.1 - Index of used houses for the Provinces

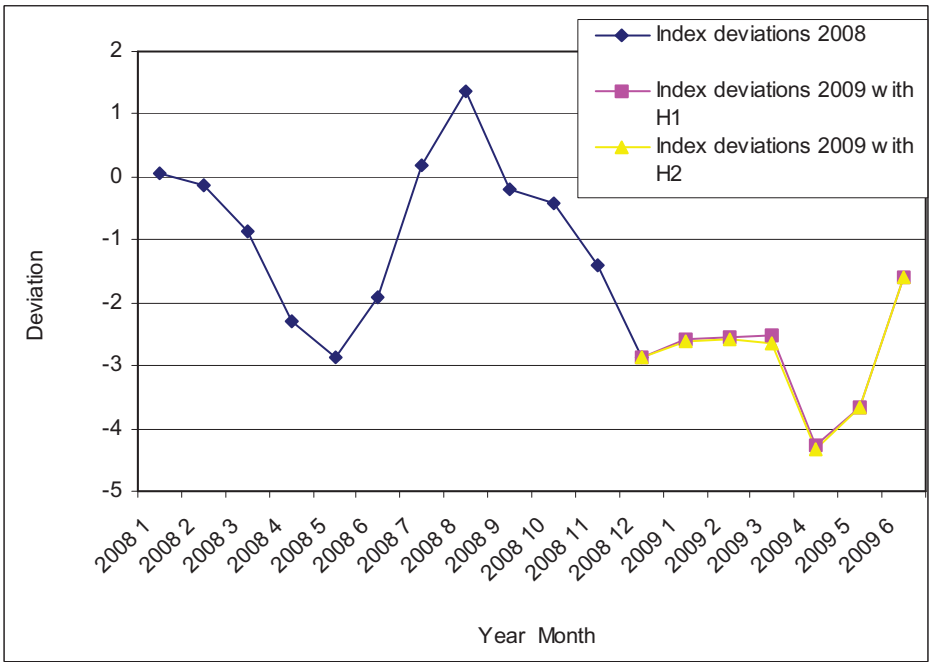


Source: PERVAL database

How to read this chart:

- old method = 4-year estimation stock;
- new method H1: 2-year estimation stock without updating β_k across the period); the β_k coefficients for 2008 and 2009 are estimated from 2006-2007;
- new method H2: 2-year estimation stock, with annual update of β_k.; the β_k coefficients for 2008 are estimated from 2006-2007 and those for 2009, from 2007-2008;

Figure A4.2 - Index of used houses for the Provinces (index deviations)



Source: PERVAL database

How to read this chart: the "new methods" series are identical in 2008 (we use the same set of coefficients) and different in 2009 (we have two different sets).

Table A4.3 - Examples of Chow test for strata for houses in the Provinces

	Total model		Model 1: 2004 and 2005		Model 2: 2006 and 2007			
Stratum code	degree of freedom	sum of squared residuals	degree of freedom	sum of squared residuals	degree of freedom	sum of squared residuals	Chow test statistic	Pvalue
751	1,915	114.8	874	54.2	1,007	56.9	1.82	0.0028
752	2,690	142.4	1,325	76.3	1,326	61.7	2.23	0.0000
754	3,136	211.2	1,477	105.3	1,620	102.3	1.38	0.0581
755	7,182	412.1	3,456	204.1	3,677	202.2	2.09	0.0000
756	4,656	288.1	2,289	142.6	2,313	140.0	1.66	0.0018
757	8,938	428.0	4,574	225.8	4,320	198.4	1.80	0.0009
758	4,055	211.7	1,820	100.4	2,186	108.6	1.07	0.3401
6701	7,477	906.8	3,601	427.4	3,817	464.5	2.11	0.0000
13055	1,676	123.9	824	66.0	818	53.7	1.67	0.0092
21701	2,418	75.3	1,262	41.7	1,122	32.4	1.06	0.3742
29701	2,825	122.2	1,430	63.2	1,351	56.4	1.36	0.0573
31555	1,556	83.5	774	42.8	743	38.3	1.18	0.2137
31701	5,032	150.5	2,614	82.4	2,379	66.4	1.50	0.0245
33063	2,742	174.6	1,346	80.3	1,352	87.4	2.53	0.0000
33701	8,568	382.9	4,379	205.9	4,125	172.8	1.48	0.0080
34701	1,734	81.3	780	39.4	920	39.8	1.32	0.1061
35701	2,257	93.6	1,154	46.7	1,069	43.7	2.32	0.0000
37701	3,801	146.3	1,756	70.9	2,001	73.5	1.12	0.2656
38701	1,920	94.0	968	49.8	918	42.0	1.33	0.0977
44109	2,617	111.9	1,245	54.9	1,333	55.7	0.76	0.8619
44601	2,324	211.3	1,264	120.0	1,026	85.7	1.85	0.0021
44701	5,774	178.3	2,772	89.9	2,943	86.2	1.25	0.0982
45701	4,853	134.6	2,501	72.2	2,303	60.8	1.17	0.1984
51701	2,245	80.7	1,115	41.8	1,096	36.0	2.34	0.0000
54701	4,770	229.5	2,190	105.6	2,536	121.4	1.19	0.1790
59350	1,874	95.1	941	48.9	889	42.1	1.90	0.0004
59502	2,049	184.8	1,020	90.1	995	91.1	1.17	0.2332
59599	2,198	91.0	1,083	47.9	1,091	42.0	1.07	0.3675
59601	2,133	90.0	1,115	49.6	979	38.5	1.16	0.2291
59701	3,622	321.8	1,894	176.3	1,684	140.9	1.17	0.2019
59702	15,671	652.9	7,985	348.2	7,582	294.6	2.35	0.0000
62601	2,537	130.2	1,317	72.0	1,191	55.6	1.76	0.0072
63701	2,105	126.3	1,070	66.5	1,001	55.9	1.91	0.0013

How to read this chart: In the Pvalue column, shaded boxes indicate strata where the stability hypothesis H0: $\beta^1 = \beta^2 = \beta$, is rejected for a significance level $\alpha=0.05$.
Source: PERVAL database

Table A4.4 - Examples of Chow test for strata after adjusting for heteroscedasticity for houses in the Provinces

	Total model		Model 1: 2004 and 2005		Model 2: 2006 and 2007			
Stratum code	degree of freedom	sum of squared residuals	degree of freedom	sum of squared residuals	degree of freedom	sum of squared residuals	Chow test statistic	Pvalue
751	1,921	462.5	865	209.0	1,014	238.3	1.52	0.0186
752	2,687	608.3	1,316	310.9	1,324	278.4	1.81	0.0007
754	3,150	812.7	1,471	393.5	1,632	402.7	1.37	0.0478
755	7,180	1,699.6	3,452	826.1	3,671	846.4	2.03	0.0000
756	4,653	1,143.1	2,281	560.9	2,310	557.9	1.61	0.0018
757	9,047	1,944.6	4,567	997.2	4,428	926.5	1.88	0.0001
758	4,064	919.9	1,814	415.3	2,193	490.9	1.06	0.3476
6701	7,474	2,607.6	3,588	1 235.4	3,819	1,340.0	1.38	0.0216
13055	1,700	465.3	817	229.1	841	218.7	1.55	0.0145
21701	2,408	422.3	1,252	227.8	1,114	185.3	1.25	0.1307
29701	2,882	593.9	1,423	293.7	1,407	284.9	1.45	0.0205
31555	1,634	419.7	810	212.2	792	198.1	1.15	0.2649
31701	5,028	865.1	2,607	461.2	2,374	392.3	1.44	0.0269
33063	2,734	693.5	1,339	332.5	1,343	335.6	1.95	0.0001
33701	8,561	1,782.7	4,367	926.2	4,122	834.3	1.49	0.0049
34701	1,734	384.0	774	176.4	918	197.7	1.07	0.3569
35701	2,270	462.2	1,151	231.6	1,077	215.3	1.81	0.0012
37701	3,798	737.0	1,750	347.8	1,996	379.3	0.97	0.5372
38701	1,914	431.9	960	222.5	912	198.4	1.16	0.2183
44109	2,611	538.1	1,239	262.0	1,325	267.4	0.90	0.6630
44601	2,326	698.5	1,257	383.6	1,027	294.5	1.64	0.0061
44701	5,775	1,006.5	2,771	498.0	2,937	495.4	1.12	0.2280
45701	4,850	800.1	2,497	422.9	2,296	366.6	1.13	0.2318
51701	2,238	418.5	1,106	213.5	1,090	192.0	1.68	0.0042
54701	4,789	1,044.2	2,185	477.8	2,552	553.2	1.17	0.1855
59350	1,868	417.4	930	208.7	886	189.9	1.65	0.0027
59502	2,039	610.0	1,014	302.4	983	293.6	1.11	0.2883
59599	2,190	458.6	1,074	228.2	1,084	223.7	1.00	0.4673
59601	2,126	437.4	1,107	232.5	972	193.7	1.16	0.2139
59701	3,616	1,074.0	1,885	571.3	1,679	487.5	0.98	0.5089
59702	15,683	3,148.4	7,976	1 634.3	7,595	1,467.6	2.08	0.0000
62601	2,541	576.5	1,311	304.4	1,193	259.6	1.51	0.0259
63701	2,097	505.9	1,061	260.8	994	230.9	1.41	0.0429

How to read this chart: In the Pvalue column, the shaded boxes indicate strata where the stability hypothesis is rejected H0: $\beta^1 = \beta^2 = \beta$, for a significance level $\alpha=0.05$.
Source: PERVAL database

Appendix 5: Alternative calculation methods

We have described in this volume the method for constructing a hedonic price index and implementing it to produce the Notaires-INSEE indices. There are several steps in this method:

- definition of strata, where price changes are assumed to be homogeneous;
- introduction of adjustments for quality effects, ar stratum by stratum;
- estimation of adjustment effects from an estimation stock;
- calculation of price changes by stratum from total transactions;
- calculation of the index by observing the change in value of a reference stock;
- regular publication of indices and sub-indices.

A procedure like this must be systematic and integrated. In this chapter, we examine alternative methods used in France and in other countries too, to determine whether they are theoretically better, or if they are preferred for practical reasons (budgetary constraints, data availability, different publication requirements), or because of differences in household behaviour (greater mobility leading to repeat sales). In the last section, we describe procedures used by other bodies in France, and in several other countries, with an emphasis on data-collection issues and the bodies responsible for calculating the indices.

If we consider the range of methods used internationally to compile housing price indices, we find some that are relatively simple and others that are more sophisticated. For practical reasons related to cost or for legal reasons, many housing price statistics use average or median prices to track changes in the housing market. We have explained why these are biased (Chapter 2, paragraph 2.1.2).

The most sophisticated methods usually use the hedonic approach or the repeat-sales approach. The former is an econometric approach; it uses regressions incorporating quality effects and a time effect, and assimilates this time effect to a pure constant-quality price effect. The second approach aims to eliminate quality effects by retaining only data on successive sales, or repeat-sales. We shall now look briefly at their underlying principles and hypotheses and discuss some of their variants.

Econometric (hedonic) approach: interpreting time coefficients in a regression as a pure constant-quality price effect

Let us assume that we have several instantaneous cross-sections of samples of real-estate transactions, which give us a measurement for the price and dwelling/building/neighbourhood characteristics and which we can use to estimate a hedonic model. The model defines the price level as a function of the dwelling characteristics. The coefficients of these characteristics are assumed to be stable over time, and a time effect v_t is introduced into the constant term.

The model is written:

$$\log p_{i,t} = a + \sum_{k=1}^{n} \beta_k X_{k,i,t} + v_t + e_{i,t}$$

where the errors $e_{i,t}$ are assumed to be independent, identically distributed and zero-mean. Note that in order to simplify the discussion, the stratum effect is not introduced. To make the time effect identifiable, we assume $v_{t_0} = 0$ for a given date, which then becomes the base period. The change in the index between t_0 and t is equated with the value v_t and the change in the index between *t-1* and *t* is equated with $v_t - v_{t-1}$.

The hypotheses underlying this specification are as follows:

- The variables selected to characterise a dwelling are incorporated in additive form (recall that this is after a transformation of initial characteristics, if needed);
- The relative prices β_k of the characteristics X_k are time-independent, which means that these variables are assumed to have no cross-effects over time;

- After adjustment for effects of characteristics, price variabilities are constant (absence of heteroscedasticity);
- Transactions introduced as observations are representative of the total set of properties for which we want to calculate the index, after quality-effect adjustment.

Some of these aspects, such as the time-dependence of the parameters β_k , are classic subjects for study in the literature. While the long-term constancy of these coefficients is usually rejected, which explains the need for regularly revising the index, it is generally accepted for shorter periods of around 4 to 5 years. There is a simple way of verifying this stability over time. If, for example, the coefficient of X_1 depends on date *t*, the initial model will be seen as missppecified, as the cross variables $1_\tau(t)X_{1,i,t}$, with τ varying, have been omitted. These variables are orthogonal to one another. We need only record the empirical correlation between the omitted variable $1_\tau(t)X_{1,i,t}$ and the estimation residual $\hat{e}_{i,t}$ as a function of date τ . If these correlations are close to zero, the β_1 coefficient is regarded as stable over time. Otherwise, the form of the change in this correlation as a function of τ gives information on the type of change in the coefficient.

One example of this method is the method over "adjacent periods". We consider a model with time dummies based on two consecutive periods. At each period *t* the model is reestimated. The time dummy measures price change with constant characteristics between dates *t-1* and *t*. the index between two dates is obtained by chaining one to another the indices of the periods between these two dates.

The representativeness of the transaction samples at each date *t* relative to what is theoretically desirable for an index is an issue that is less often discussed. For example, if the sample at date *t* includes transactions of which the prices are systematically higher than the theoretical price for the overall population (i.e. the price of a fixed reference stock and not of transactions that vary from one date to another), the time coefficient $\hat{v}_t$ estimated from data will reflect both the theoretical price level and the bias due to the non-representativeness of the sample, although it will not be possible to distinguish these effects easily (problem of identifiability). However, there are two approaches that we can use to detect any non-representativeness.

- The first was suggested by Griliches (1971, p.7-8) in a different setting (price index for cars). It consists of distinguishing repeat sales in samples for dates *t-1* and *t* for example. We can then calculate the mean residuals for these repeat sales on date *t-1* and date *t* and see if they are close to zero. However, the aim of this approach is to verify the representativeness of the repeat-sales sub-sample rather than that of the complete samples. It is also difficult to implement in our context, where the number of repeat-sales is small.
- Another approach relies on the assumed change in the index. Given a set of observations $\hat{v}_1, \ldots, \hat{v}_T$ of the index between 1 and *T*, we can construct a dynamic model which will supply a forecasting interval:

$$\left[\tilde{v}_{T+1}, \tilde{\tilde{v}}_{T+1}\right]$$

for the future value v_{T+1} . If the estimate for date $T+1$ does not lie within this interval, we can either assume that the sample for date $T+1$ is unrepresentative, or we can look for a structural cause for this sudden price change.

Repeat-sales method

In 1943, Gaston Duon, who worked for the French National Statistics Service, the forerunner of INSEE, applied what was called the repeat-sales method to compile price indices for dwellings in Paris between 1790 and 1944 (Duon, 1943 and 1946). In 1956, Léo Grebler, who was apparently unaware of the work being done by Gaston Duon, then applied this method to compile a housing price index in 22 cities of the United States.[68] Both Duon and Grebler were well aware that the quality of a given property changes over time as a result of the effects of two conflicting phenomena, on the one hand the general obsolescence of buildings and on the other, the improvements made over the years. They both compiled two series of indices, before and after adjustment for

[68] Historically, housing price indices usually use the repeat-sales method. This was the method used by Piet Eicholtz to construct the Herengracht index. The exceptions are the indices by d'Avenel covering the period 1200-1800 and which are thus, apparently, the property indices that go back farthest in time, but which may contain some uncertain data.

these effects. The deviation in growth for these two series was around 1% per year both for Duon and for Grebler.[69]

The repeat-sales method was then applied, with more powerful computing resources, by Bailey, Muth and Nourse (1963). The authors do not directly pose the question of the fixed basket of goods, but they do make two comments. First, "because of the variation in the quality of sold dwellings from one period to another, the mean price varies more than the price of each property taken individually"; second, a gradual change in the quality of traded dwellings over time biases the change in average prices.

It was difficult to specify a hedonic model with a time dummy, given the frequent lack of data on the characteristics of dwellings. To overcome this difficulty, they suggested using the fact that certain dwellings are sold several times in succession. These repeat-sales data would obviate the need for detailed information on the characteristics of the properties.[70]

The lack of data on dwelling characteristics and the technical difficulty of applying the econometric approach are still today the main reasons given in favour of a repeat-sales approach. This approach is easy to explain when repeat-sales occur on dates *t* -1 and *t*. In this case, we equate the price change to the mean change observed for these repeat sales, to give a model of this type:

$$\log p_{i,t} = \log p_{i,t-1} + b_t + u_{i,t-1,t}, i \in I_{t-1,t}$$

at a given date *t*. $I_{t-1,t}$ denotes total dwellings traded both at *t* -1 and *t*, and b_t is the change that we are looking for.

However, even with high mobility, few repeat sales are so very close together in time. The approach is therefore extended to take account of repeat sales taking place at two dates t_1 and t_2 , $t_1 < t_2$, which can be further apart. The underlying model now becomes:

$$\log p_{i,t_2} = \log p_{i,t_1} + \sum_{t=t_1+1}^{t_2} b_t + u_{i,t_1,t_2}, i \in I_{t_1,t_2}$$

with t_1 and t_2 varying.

It can be rewritten by including the explanatory variables for dates $Z_{i,t} = 1$, if date *t* lies in the period between the trading dates, if not it is $Z_{i,t} = 0$. The model now becomes:

$$\log \frac{p_{i,t_2}}{p_{i,t_1}} = \sum_{t=T_1}^{T_2} b_t Z_{i,t} + u_{i,t_1,t_2}, i \in I_{t_1,t_2}$$

where $[T_1, T_2]$ gives the time interval that is the union of all the intervals $[t_1, t_2]$. In this form, it is a linear model in the parameters of interest b_t, with *t* varying, which gives price changes at the various dates. This model is generally estimated using ordinary least squares.

However, if the time interval between repeat sales covers more than two periods, then overlaps will occur for periods concerning the various dwellings, and hence there may be correlations between error terms, which should be taken into account in the estimation method. To illustrate this potential problem and understand why the quality effects have disappeared, it is worth returning to the hedonic model described in the previous section:

$$\log p_{i,t} = a + \sum \beta_k X_{k,i,t} + v_t + e_{i,t}$$

with $e_{i,t} = u_{i,t} + \eta_i$,

[69] On secular housing price indices, and in particular a comparison of the work of Duon and Grebler, cf. Comparing Four Secular Home Price Indices, J. Friggit, June 2008, http://www.cgedd.developpement-durable.gouv.fr/IMG/doc/house-price-index-Paris-and-others-secular_cle7fed11.doc.

where the error terms $u_{i,t}$, η_i are assumed to be independent of each other, and zero-mean. If all repeat sales taking place in t_1, t_2 are representative of sales in the total set of dwellings (after adjustment for the characteristic effect), we have:

$$\log \frac{p_{i,t_2}}{p_{i,t_1}} = v_{t_2} - v_{t_1} + u_{i,t_2} - u_{i,t_1}, i \in I_{t_1,t_2}$$

We confirm that the repeat sales approach with a least-squares estimation is consistent with $b_t = v_t - v_{t-1}$ and $\sum_{t_1+1}^{t_2} b_t = v_{t_2} - v_{t_1}$, since the error terms $u_{i,t_1,t_2} = u_{i,t_2} - u_{i,t_1}$ are indeed independent, with identical distribution.

Finally, we can make the model more complex and refine it by incorporating the natural depreciation of dwellings (net of improvements made), or other quality variables that may change between two sales (e.g. a comfort feature added to the dwelling).

To sum up, the repeat-sales method is based on the following assumptions:

- relative prices of characteristics are constant over time. This is the same hypothesis as in the standard applications of the hedonic method. But it is undoubtedly less likely to be fulfilled over the relatively long period between two successive sales of the same dwelling;
- there is no selection bias. However, dwellings sold frequently are probably not representative either of total transactions, or of the housing stock. These may be, for example, small dwellings (first homes for young couples), whose prices may move differently from those of larger dwellings; or they may be dwellings that are resold very quickly and which may have unobservable characteristics, related to the vendor, which may account for large capital gains. Clapp *et al*. (1991) find a difference in the short-term change in repeat-sales indices relative to hedonic indices, but this difference disappears in the long term (3 years), which they see as logical: if the market is working, there can be no long-term disequilibrium in relative prices.[71] Given the distortions for periods of less than three years, they recommend using hedonic methods. Case *et a*l. (1997) suggest adjusting for this effect by incorporating information on the link between a dwelling's appreciation rate and transaction frequency;
- the dwelling is effectively the same. Of course renovations or even more substantial alterations (extensions, improvements) are common, with the result that the dwelling is no longer the same. These last two problems can be addressed by combining the hedonic model with repeat-sales data; this will also correct a flaw in the repeat-sales method: the fact that it uses few observations relative to total transactions. The reason is that dwellings change hands fairly infrequently, at least this is the case in France;
- the error terms $u_{i,t}$ are indeed independent, zero-mean and with identical variance. However, this hypothesis is presumably not satisfied. These errors concern prices, which, on average, increase over time, hence generating heteroscedasticity. We also expect a stronger correlation between prices for dates that are close than for dates that are further apart. It therefore seems important to introduce a time correlation between errors $e_{i,t}$, which modifies the method for estimating rates of change b_t.

Note that for France, results from the repeat-sales method,[72] for overlapping periods and areas, do not differ very much from those of the Notaires-INSEE indices.[73]

[71] Mark and Goldberg (1984) for their part, find a persistent long-term difference.

[72] Applied to notarial data over a long period (price comparison, without econometric approach). See Friggit J., "Comparing Four Secular Home Price Indices", CGEDD, June 2008, http://www.cgedd.developpement-durable.gouv.fr/prix-immobilier-evolution-1200-a1048.html.

[73] The repeat-sales method gives poor results for recent years, because of the small number of observations and the selection bias mentioned above for short periods of ownership (capital gains on properties resold quickly exceed the capital gains on the index). The comparison was therefore made after incorporating an adjustment coefficient benchmarked to Paris.

Laspeyres index or chain index

A large amount of classical index theory is devoted to the choice of weightings. Three options are available in general: they can be set once and for all, defined as equal to a set of quantities traded at an initial date (Laspeyres index), or at a terminal date (Paasche index), or else they can be modified at each index calculation date (chain index).

The quantities exchanged are related to the composition of the housing stock by stratum. The repeat-sales method leads naturally to the perspective of a chain index, where the weights assigned to the thinnest stratum level vary over time, depending on the degree of aggregation used for the calculation. Different degrees of detail may become necessary in the aggregation, according to the number of useable sales there are in the calculations. Since the structure of repeat sales shifts over time, the way the chaining procedure is applied cannot be specified *a priori*. This problem is not addressed by the users or the advocates of the method.

By contrast, the hedonic approach can be used to calculate indices of the Laspeyres, Paasche or chain type, since it comes to the prix of a reference property. However, in order to facilitate comparison with other types of investments, eg financial ones, it seems preferable not to choose chaining from one period to another, but to maintain a reference stock over a certain period of time. Indeed, for financial investment the usual practice is to track a fixed ("crystallised") portfolio so as not to mix up price effects with the effects of updating the portfolio (investment strategy).[74]

This seems all the more justified for housing price indices since housing is a durable asset that carries high transaction costs and for which investor preferences between different types of goods evolve slowly. As a result, the index of housing prices is different from that of consumer prices. The argument for chaining consumer price indices (or those for the annual weighting revision) is based on the fact that we want to track changes in consumer preference, that are revealed by the share of the budget allocated to each major category of assets. In the housing sphere, adaptations are slow in all likelihood. At the detailed level, the indices are aggregated multiplicatively to take into account a possible substitutability of properties. At the higher level of aggregation, they are aggregated additively.

[74] The housing price indices are chained Laspeyres indices. Compared to version 2, the new calculation method corresponds to a succession of calculations of Laspeyres closer in time– every two years instead of every five years (cf. Chapter 3).

Appendix 6: Advance price indices

Usually, before a property is sold, a pre-contract is signed.[75] There are normally about three months between the signing of the pre-contract and the signing of the actual sale contract. The introduction of advanced price indices calculated from the pre-contracts will therefore enable us to produce indices that are more reactive to the state of the market.

Collecting pre-contracts

Notaries are legally obliged to pass on information about pre-sale contracts to the housing databases, according to the Law of March 2011. However, the notariat has collected pre-contract data from the notarial offices since mid-2009. There have been two different systems. First, a portal was created for the notarial offices in the provinces to enter basic pre-contract information. With this system a large number of pre-contracts were collected (Table A6.1), however, it did not contain all the variables needed for calculating the indices.

Table A6.1: Number of sales and pre-contract sales of used dwellings collected in 2012

	Sales	Pre-sales
Paris Region	95,927	15,526
Provinces	308,650	111,216

The decision was therefore taken to develop a specific module for the teletransmission of pre-contracts, incorporated into the software used in the notarial offices to draw up deeds of sale. This was along the same lines as the teletransmission module for sales, which gradually replaced the need for the notaries to send paper copies of deeds.

The pre-contract teletransmission tool was introduced gradually into the notarial offices from 2010 (using IT systems and the deed of sale software). It has grown since then depending on the demands of database managers and developments in notarial IT services.

Based on the new version of the software, a procedure to pair up the pre-contract and the sale was put in place during 2013. It is now possible to carry out a number of calculations that had not been done previously due to a lack of information: in particular, it is now possible to study the time lag between pre-contract and sale, or the proportion of pre-contracts that do not ultimately become a sale.

Methodology and dissemination

Given this background and the volume of pre-contracts transmitted, it is not yet possible to produce a price index based solely on this data, however, this could be an ultimate goal. A provisional calculation method has been developed based on current indices.

The working hypothesis is that the econometric models developed for sales should also be valid for pre-sale contracts. Thus by applying these models to the pre-contract data it should be possible to calculate standardised prices ("reference-property equivalent" prices[76]) for the pre-contracts and compare them with the standardised prices for sales for the preceding quarters. We will therefore be able to calculate advance indicators, provided that the volume of pre-contracts received is sufficient.

Advance indicators for price changes in Paris and in the Hauts-de-Seine *département* are disseminated by the Chamber of Notaries of Paris through their economic outlook reports and press conferences. In the Provinces, a general indicator is given in the quarterly economic outlook reports and indications about price changes in some large provincial cities are also given at the annual national conference of the *Conseil Supérieur du Notariat*.

[75] All pre-contracts or "promises of sale" (*compromis* or *promesse de vente*) are collected by the notarial databases.
[76] Cf. Chapter 3.

Appendix 7: Agreement with PNS

Agreement regarding the continuation and expansion of the partnership between INSEE and PNS in the field of housing price indices in the Paris Region.

between:

- **CHAMBRE INTERDEPARTEMENTALE DES NOTAIRES DE PARIS**, represented by Mr. Benasse, notary, its chairman,

- CHAMBRE DES NOTAIRES DE SEINE ET MARNE, represented by Mr. Hautebas, notary, its chairman,

- CHAMBRE INTERDEPARTEMENTALE DES NOTAIRES DE VERSAILLES, represented by Mr. Savoure, notary, its chairman,

- **CHAMBRE DES NOTAIRES DE L'ESSONNE,** represented by Mr. Lemoine, notary, its chairman,

- CHAMBRE DES NOTAIRES DES HAUTS DE SEINE, represented by Mr. Herrnberger, notary, its chairman,

- **ASSOCIATION DES NOTAIRES DU CHATELET** (**P**ARIS **N**OTAIRES **S**ERVICES), a not-for-profit organisation administered under the 1901 Act, represented by Mr. Cauro, notary, its chairman;

Acting jointly and hereafter referred to as PNS,

the first party,

and:

INSTITUT NATIONAL DE LA STATISTIQUE ET DES ÉTUDES ÉCONOMIQUES, hereafter referred to as INSEE and represented by its Director-General Mr. Jean-Luc Tavernier,

the second party,

it has been agreed and decided as follows:

Preamble

PNS and INSEE signed an initial agreement on December 6, 1990, establishing a partnership for the calculation and publication of the price index for vacant used apartments sold in Paris.

A second agreement, dated May 16, 2000 (agreement 2000 00094), and an addendum dated November 26, 2002 have revised the method for calculating this index and extended the scope of application of this partnership to other price indices for vacant used housing sold in all or part of the Paris Region.

A third agreement, dated December 8, 2005 (agreement 2005 00353), and an addendum dated November 26, 2009 widened the scope of application of the indices even further and put in place combined dissemination and publication of the said indices.

The agreements signed in 1990, 2000 and 2005 have now reached their expiry dates. They are replaced by the present agreement, the purpose of which is to continue and strengthen the partnership between PNS and INSEE in the field of housing price indices. The present agreement takes into account recent progress in this partnership.

Article 1: Purpose of agreement

The present agreement defines the conditions for PNS and INSEE to participate in the calculation, validation, publication and dissemination of price indices for vacant used housing. These are calculated from property transaction data in the BIEN database, belonging to and managed by PNS, and fed with data by the notaries of the Paris Region.

PNS informs INSEE that the computerised system for processing information on named individuals in the BIEN database has been declared to the National Commission on Information Technology and Civil Liberties (*Commission Nationale de l'Iinformatique et des Libertés – CNIL*), in accordance with the provisions of Article 16 of the Act of January 6, 1978 Act on Information Technology, Data Files and Civil Liberties, amended by law 2004-801 of 6 August 2004.

The indices covered by the present agreement are the following, it being understood that the quarters referred to are calendar quarters:

- quarterly housing price indices (apartments and houses) for the entire Paris Region,
- quarterly housing price indices (apartments and houses) for each of the seven *départements* in the Paris Region, excluding Paris, for the three *départements* in the inner suburbs taken together, for the four *départements* in the outer suburbs taken together and for all of the Paris Region, excluding Paris,
- quarterly price indices for apartments for the entire Paris region,
- quarterly price indices for apartments in each of the eight *départements* in the Paris Region, for the three *départements* in the inner suburbs taken together, for the four *départements* in the outer suburbs taken together and for all of the Paris Region, excluding Paris,
- quarterly price indices for houses for the entire Paris Region,
- quarterly price indices for houses in each of the seven *départements* in the Paris Region excluding Paris, for the three *départements* in the inner suburbs taken together, for the four *départements* in the outer suburbs taken together and for all of the Paris Region, excluding Paris.

Each index is to be produced as a provisional index, calculated about one and a half months after the end of the quarter to which it relates, and a definitive index, calculated about four and a half months after the end of the quarter to which it relates.

In addition, each of these indices will be produced without adjustment for seasonal variations and also with adjustment for seasonal variations.

Hereafter, the indices covered by the present agreement will be referred to generically as “the indices”.

PNS informs INSEE that monthly used housing price indices are also calculated quarter-on-quarter by PNS for quarters that do not correspond to calendar quarters. These monthly quarter-on-quarter indices are calculated using the same method as the indices covered by the present agreement and cover the same time period.

Article 2: The “Indice Notaires-INSEE” designation and approval by the *Autorité de la statistique publique* (National public statistics authority)

Subject to the provisions of Article 5 of the present agreement, the indices will be allowed to use the designation “Indice Notaires-INSEE”. In this respect, INSEE authorises PNS to use this designation in external communications and in relations with third parties.

PNS is informed that similar indices calculated for the Provinces by the company Min.not and the Conseil Supérieur du Notariat in partnership with INSEE also use the “Indice Notaires-INSEE” designation, as do the indices for metropolitan France calculated by Min.not from the “Paris Region” index by PNS and the “Province” index by Min.not.

In addition, the Paris Region indices have received approval from the *Autorité de la statistique publique* (National public statistics authority) to use the designation. The Notaires-INSEE indices for the Paris Region have also received this approval.

Article 3: Scientific Board

The role of the Scientific Board for the Notaires-INSEE indices will be one of deliberation and advice for the parties of the present agreement. Its work will concern mainly (but not exclusively) the way in which the indices are calculated and the way the notarial databases are fed with data.

For example, it defines coverage rates below which the indices are not validated.

The Scientific Board is composed of:

i) a chairperson, chosen jointly by INSEE, PNS, Min.not and the Conseil Supérieur du Notariat;

ii) two representatives from INSEE;

iii) two representatives from PNS;

iv) one representative from Min.not;

v) one representative from Conseil Supérieur du Notariat.

The Scientific Board may call on the assistance of any qualified persons that it may judge necessary for carrying out its duties, for the duration of the present agreement or only occasionally.

The Board shall meet, at the request of the chairperson, at least once per quarter. Meetings shall be organised by INSEE and written minutes shall be kept of the proceedings. Each party may put before the chairperson a request for a meeting, as required.

The Scientific Board is the same for both agreements concerning the calculation, validation and dissemination of indices with the "Indice Notaires-INSEE" designation:

- the agreement linking PNS and INSEE;
- the agreement linking Min.not, the Conseil Supérieur du Notariat and INSEE.

Article 4: Amendments to the present agreement

Any change made to the present agreement shall be set out in an addendum.

Article 5: INSEE commitments

INSEE shall:

(i) provide PNS with its statistical expertise and advice for calculating the indices and shall supply data in its possession that can ensure the quality of the indices;

(ii) provide PNS in August every year with the correction coefficients for seasonal adjustments needed to calculate the seasonally adjusted indices in their index series;

(iii) include the indices covered by the present agreement in its own publications where dissemination seems justified;

INSEE shall put in place a method for monitoring the quality of the indices. At least one week before publication, the indices shall be submitted to INSEE for approval. INSEE shall respond within no more than two working days. If there is no response after two working days, this shall be the equivalent of approval.

Achieving minimal coverage rates is now a necessary condition for validation of the indices. The methodology for the indices is described in detail in a special issue of the *INSEE-Méthodes* series which is regularly updated by INSEE.

Article 6: PNS commitments

PNS shall:

(i) comply with the index calculation procedures recommended by the Notaires-INSEE indices Scientific Board,

(ii) allocate the human and physical resources (especially IT resources) needed to maintain and update these procedures and in particular shall adapt these resources to the increase in workload that will result when the obligation for notaries to supply data to the databases used to calculate the indices comes into force; PNS shall put in place a system to assess these human and financial needs on a regular basis and will evaluate the resources actually used over the last few quarters and anticipated use in the next quarters. These assessments will be put before the Scientific Board of the Notaires-INSEE indices;

(iii) preserve and improve the current system for collecting information on real estate sales in the Paris Region: in particular the collection rate and data incorporation time shall be monitored via dashboards (number of notarised transactions and number of transactions added per month, stock awaiting codification); these will be submitted to INSEE every quarter. PNS agrees to inform INSEE before any change is made to the process of creating the notarial databases (definition of fields, changes that could impact the collection and incorporation rates, data-completion rates for some fields, index values);

(iv) provide INSEE with the information defined as agreed with them so that they can approve the indices. In particular this will include detailed indices, price levels and the transaction amounts and volumes in the course of the quarter;

(v) publish the indices quarterly;

(vi) in accordance with the aim of the agreement, provide INSEE with all information extracted from PNS databases that could be of use for INSEE's internal statistical studies, and any price index series calculated by PNS for its own requirements and which would be useful to the Institute for study purposes;

(vii) announce in its publications any change made to the method used to calculate the monthly quarter-on-quarter indices not covered by the present agreement.

Article 7: Joint publication and dissemination of the indices

Publication and dissemination of the indices, by INSEE and PNS, are subject to embargo.

The date and time when the embargo is lifted, i.e. the date and time when INSEE and PNS are authorised to publish and disseminate the indices, are determined as follows:

- The dates for publication by INSEE and for the Chambre interdépartementale des notaires de Paris to hold its press conference are proposed by the Notaires-INSEE indices Scientific Board and approved by INSEE at least one quarter in advance. Basically, the principle behind the schedule is to have the dates in the last week of February, May, August and November. However, the date for the end of August will systematically be moved back by about two weeks.
- The time that the embargo is lifted is 8.45am.

Information about the indices is disseminated and published as follows:

i) for the attention of the media, by PNS:

By means of a press conference, which will take place systematically on the morning of the day the embargo is lifted, at 8.45am, barring exceptions. Once the press conference is over, information can be disseminated to all the media by any appropriate means;

In accordance with the recommendations of the Autorité de la Statistique Publique, the press file handed to journalists and the documents presented at the conference must clearly differentiate those indices which are designated "Indice Notaires-INSEE" from any other information.

ii) for the attention of the media, by INSEE:

Information is relayed to the press agencies at 8.30am and to the rest of the media at 8.45am, on the morning the embargo is lifted;

iii) for the attention of the general public, by PNS and / or by INSEE:

Information is communicated from 8.45am onwards, by any appropriate means.

The cut-off dates for the BIEN database when calculating the indices are determined by PNS in such a way as to ensure that the indices can be disseminated on the date the embargo is lifted, taking into account the time needed to carry out the necessary calculations, to check them and validate them, and also to prepare publications.

Coefficients and details of the zoning used in the regressions may not be made public except with the express agreement of INSEE and PNS.

Indices that are not approved by INSEE may not be published with the "Indice Notaires INSEE" designation.

Article 8: Financial conditions

In consideration of the fact that the preparation, monitoring and regular publishing of the indices is a task that falls within the respective remits of all partners, and that the cooperation required under the terms of the present agreement entails an evenly balanced workload for each party, the present agreement shall be entered into without financial considerations.

Article 9: Duration of contract

The present agreement will come into force as soon as it is signed by all the partners. It is concluded for a period of five years.

Article 10: Termination of the agreement

The present agreement may be cancelled by any one of the parties by registered letter with acknowledgement of receipt. Cancellation shall take effect six months after receipt of this letter.

In the event of cancellation, PNS and INSEE waive the right to future use of the "Indice Notaires-INSEE" designation.

However, the parties agree in advance to consult with each other in order to allow the continuation of financial operations undertaken by financial institutions or others, under any licensing agreements entered into by PNS.

Article 11: INSEE and PNS representatives

The following persons shall execute the terms of the present agreement:

i) on behalf of INSEE, the head of the Housing division in the Department of Consumer Prices, Household Income and Living Conditions;

ii) on behalf of PNS, the current chairman of the ASSOCIATION DES NOTAIRES DU CHATELET (PARIS NOTAIRES SERVICES).

Article 12: Enforcement clauses

The present agreement is exempt from stamp duty and formal registration requirements.

The implementation of the present agreement shall have as an immediate consequence the abrogation of agreement 2005 00353 and its addendum dated 2009.

Signed in Paris in seven original copies, 13 November 2013,

For PNS and the CHAMBRES DES NOTAIRES D'ILE-DE-FRANCE

Signed: Maître Bénasse

Signed: Maître Hautebas

Signed: Maître Savoure

Signed: Maître Lemoine

Signed: Maître Herrnberger

Signed: Maître Cauro

For INSEE, the Director General

Signed: Monsieur Jean-Luc Tavernier

Appendix 8: Agreement with the Conseil supérieur du Notariat *and Min.not*

Agreement regarding the continuation and expansion of the partnership between INSEE, the *Conseil supérieur du Notariat* and Min.not in the field of housing price indices in the provinces.

between:

- The **CONSEIL SUPERIEUR DU NOTARIAT,** represented by Me Jean Tarrade, notary, its chairman,

- The Min.not Company - **Real Estate Market Notaries**, SASU capital of € 2,450,000, registered with the RCS Aix en Provence under number 381 000 611, represented by M. Daniel BOUCHON, its chairman,

the first party,

and:

INSTITUT NATIONAL DE LA STATISTIQUE ET DES ÉTUDES ÉCONOMIQUES, hereafter referred to as INSEE and represented by its Director-General Mr. Jean-Luc Tavernier,

the second party,

it has been agreed and decided as follows:

Preamble

The CNS, Perval and INSEE signed an initial agreement on June 15, 1998, establishing a partnership for the calculation and publication of the price index for vacant used apartments sold in the central cities of agglomerations of 10,000 or more inhabitants, located in France outside Ile-de-France.

A second agreement, dated October 7, 1999 has extended the scope of application of this partnership to the calculation and dissemination of a price index of used houses in all municipalities of the Provinces, as well as an index of prices of used apartments in all municipalities of province towns of over 10,000 inhabitants (instead of only central cities).

The CNS, Perval and INSEE signed an second agreement in 2005 to continue and strengthen the partnership between the CSN, Perval and INSEE in calculating indices in housing prices, by extending the scope to indices calculated in the administrative regions and large cities, since statistical quality of such evidence would be considered satisfactory by the parties. This second agreement was the subject of an amendment in 2010 on the one hand to describe the nature of the information transmitted by the Scientific Board for the Notaries-INSEE indices, for the quarterly validation of indices, and on the other hand to specify the conditions for the dissemination of data, including advancing the dates of publication of indices to reflect the improvement of the information collected by notarial databases.

The agreements signed in 1998 and 2005 have now reached their expiry dates. They are replaced by the present agreement, the purpose of which is to continue and strengthen the partnership between the CNS, Perval and INSEE in the field of housing price indices. The present agreement takes into account recent progress in this partnership.

Article 1: Purpose of agreement

The present agreement defines the conditions for the CNS, Perval and INSEE to participate in the calculation, validation, publication and dissemination of price indices for vacant used housing. These are calculated from property transaction data fed by the notaries in the Provinces on the one hand and by the price indices of used housing in Paris région calculated par Paris Notaires Service (PNS) on the other hand.

Min.not informs INSEE that the computerised system for processing information on named individuals from the housing transaction by the notaries in the Provinces and managed by ADSN is referenced in the list maintained by the Correspondent Infromatique et liberté, designated by Min.not in accordance with Act no. 78-17 "Informatique et Libertés" of January 6, 1978.

Each index covered by the present agreement is to be produced as an advanced provisional index calculated about one and a half months after the end of the quarter to which it relates, a provisional index, calculated about three months after the end of the quarter to which it relates, a semi-definitive index calculated about four and a half months after the end of the quarter to which it relates, and a definitive index, calculated about six months after the end of the quarter to which it relates.

In addition, each of these indices will be produced without adjustment for seasonal variations and also with adjustment for seasonal variations.

The indices covered by the present agreement are quarterly used housing price indices, without adjustment for seasonal variations and also with adjustment for seasonal variations, for:

- apartments and houses taken together,
- apartments,
- houses;

in the following geographic fields:

- metropolitan France,
- Provinces,
- Nord-Pas-de-Calais Region (provisional, semi-final and final indices only),
- Provence-Alpes-Côte-d'Azur Region (provisional, semi-final and final indices only),
- Rhône-Alpes Region (provisional, semi-final and final indices only) ;

In addition, the following indices will be produced without adjustment for seasonal variations and also with adjustment for seasonal variations:

- houses in the agglomeration of Lille (provisional, semi-final and final indices only),
- apartments:
    - in the agglomerations of 10,000 or more inhabitants in the Provinces taken together (provisional, semi-final and final indices only),
    - in the central towns of the agglomerations of 10,000 or more inhabitants in the Provinces taken together (provisional, semi-final and final indices only),
    - in the suburbs of the agglomerations of 10,000 or more inhabitants in the Provinces taken together (provisional, semi-final and final indices only),
    - in the rural municipalities and municipalities of the agglomerations with less than 10,000 inhabitants in the Provinces taken together (provisional, semi-final and final indices only),
    - in the municipality of Lyon (provisional, semi-final and final indices only),
    - in the municipality of Marseille (provisional, semi-final and final indices only).

Min.not informs INSEE that monthly used housing price indices are also calculated quarter-on-quarter by Min.not for quarters that do not correspond to calendar quarters. These monthly quarter-on-quarter indices are calculated using the same method as the indices covered by the present agreement and cover the same time period.

Hereafter, the indices covered by the present agreement will be referred to generically as "the indices".

Article 2: The "Indice Notaires-INSEE" designation

Subject to the provisions of Article 5 of the present agreement, the indices will be allowed to use the designation "Indice Notaires-INSEE" and optionally a custom region as "Index Notaires-INSEE Rhône-Alpes" appellation. In this respect, each party, the CSN, Min.not and INSEE, is authorized to use this designation in external communications and in relations with third parties.

The CSN and Min.not are informed that similar indices calculated for The Paris region by Paris Notaires Services 'PNS) in partnership with INSEE also use the "Indice Notaires-INSEE" designation.

Article 3: Scientific Board

The role of the Scientific Board for the Notaires-INSEE indices will be one of deliberation and advice for the parties of the present agreement. Its work will concern mainly (but not exclusively) the way in which the indices are calculated and the way the notarial databases are fed with data.

For example, it defines coverage rates below which the indices are not validated.

The Scientific Board is composed of:

i) a chairperson, chosen jointly by INSEE, PNS, Min.not and the Conseil Supérieur du Notariat;

ii) two representatives from INSEE;

iii) two representatives from PNS;

iv) one representative from Min.not;

v) one representative from Conseil Supérieur du Notariat.

The Scientific Board may call on the assistance of any qualified persons that it may judge necessary for carrying out its duties, for the duration of the present agreement or only occasionally.

The Board shall meet, at the request of the chairperson, at least once per quarter. Each party may put before the chairperson a request for a meeting, as required. Meetings shall be organised by INSEE and written minutes shall be kept of the proceedings.

The Scientific Board is the same for both agreements concerning the calculation, validation and dissemination of indices with the "Indice Notaires-INSEE" designation:

- the agreement linking PNS and INSEE;
- the agreement linking Min.not, the Conseil Supérieur du Notariat and INSEE.

Article 4: Amendments to the present agreement

Any change made to the present agreement shall be set out in an addendum.

Article 5: INSEE commitments

INSEE shall:

(i) provide Min.not with its statistical expertise and advice for calculating the indices and shall supply data in its possession that can ensure the quality of the indices;

(ii) provide Min.not in August every year with the correction coefficients for seasonal adjustments needed to calculate the seasonally adjusted indices in their index series;

(iii) include the indices covered by the present agreement in its own publications, where dissemination meets dissemination criteria defined by the Scientific Board.

INSEE shall put in place a method for monitoring the quality of the indices. At least one week before publication, the indices shall be submitted to INSEE for approval. INSEE shall respond within no more than two working days. If there is no response after two working days, this shall be the equivalent of approval.

Achieving minimal coverage rates is now a necessary condition for validation of the indices. The methodology for the indices is described in detail in a special issue of the *INSEE-Méthodes* series which is regularly updated by INSEE.

Article 6: CSN and Min.not commitments

Min.not shall:

(i) comply with the index calculation procedures recommended by the Notaires-INSEE indices Scientific Board,

(ii) allocate the human and physical resources (especially IT resources) needed to maintain and update these procedures and in particular shall adapt these resources to the increase in workload that will result when the obligation for notaries to supply data to the databases used to calculate the indices comes into force;

(iii) monitor the collection rate and data incorporation time via dashboards (number of notarised transactions and number of transactions added per month, stock awaiting codification); these will be submitted to INSEE every quarter. Min.not agrees to inform INSEE before any change is made to the process of creating the notarial databases (definition of fields, changes that could impact the collection and incorporation rates, data-completion rates for some fields, index values);

(iv) provide INSEE with the information defined as agreed with them so that they can approve the indices. In particular this will include detailed indices, price levels and the transaction amounts and volumes in the course of the quarter;

(v) publish the indices quarterly;

(vi) in accordance with the aim of the agreement, provide INSEE with any price index series calculated by Min.not for its own requirements and which would be useful to the Institute for study purposes;

The CSN shall :

(i) preserve and improve the current system for collecting information on real estate sales in the provinces: in particular the collection rate and data incorporation time.

Article 6 bis: Protection of personal data

The CSN, Min.not and INSEE, each commit to complying with the regulations on the protection of personal data, according to Act no. 78-17 "Informatique et Libertés" of January 6, 1978 and its decrees.

Thus, they shall take all necessary precautions to maintain the security, confidentiality of personal data which they are aware in the framework of this agreement and in particular to prevent it from being distorted, damaged or communicated to persons not expressly authorized.

CSN, Min.not and INSEE commit to making any administrative process under their responsbility according to regulations related to the protection of personal data.

In this respect, it is reported that the CSN and Min.not, have each appointed a correspondant « Informatique et Libertés » responsible for maintaining and updating the register of treatment and makes it available to the Commission Nationale Informatique et Libertés.

The CSN Min.not and INSEE vouch, as defined in Article 1120 of the Civil Code, for the respect by their employees and/or possible subcontractors duly authorized, of this section. It was agreed that the present Agreement may be the object of a communication to the Commission Nationale Informatique et Libertés.

Article 6 ter: – confidentiality

The CSN Min.not and INSEE are conventionally required to absolute secrecy about all information from databases, documents or any element, in particular technical, financial or organizational they would access, under this Agreement and / or previous conventions. Each party undertakes to respect this confidentiality obligation by all of its staff, any officer, any contractor and any subcontractor.

Breach of that duty of confidentiality commits the responsibility of the CSN, or Min.Notaires-Insee or INSEE and constitutes a case of early termination which is described in section 10 "Terms of denunciation of the agreement."

Article 7: Joint publication and dissemination of the indices

Publication and dissemination of the indices, by INSEE, the CSN and Min.not, are subject to embargo.

The date and time when the embargo is lifted, i.e. the date and time when INSEE, The CSN and Min.not are authorised to publish and disseminate the indices, are determined as follows:

Dates INSEE publications are proposed by the Scientific Board for the Notaires-INSEE and validated by the INSEE at least one quarter in advance.

The time that the embargo is lifted is 8.45am.

Information about the indices is disseminated and published as follows:

i) for the attention of the media, by the CSN and Min.not:

By means of a press conference, which will take place systematically on the morning of the day the embargo is lifted, at 8.45am, barring exceptions. Once the press conference is over, information can be disseminated to all the media by any appropriate means;

In accordance with the recommendations of the Autorité de la Statistique Publique, the press file handed to journalists and the documents presented at the conference must clearly differentiate those indices which are designated "Indice Notaires-INSEE" from any other information.

ii) for the attention of the media, by INSEE:

Information is relayed to the press agencies at 8.30am and to the rest of the media at 8.45am, on the morning the embargo is lifted;

iii) for the attention of the general public, the CSN, Min.not and / or by INSEE:

Information is communicated from 8.45am onwards, by any appropriate means.

The cut-off dates for the BIEN database when calculating the indices are determined by Min.not in such a way as to ensure that the indices can be disseminated on the date the embargo is lifted, taking into account the time needed to carry out the necessary calculations, to check them and validate them, and also to prepare publications.

Coefficients and details of the zoning used in the regressions may not be made public except with the express agreement of INSEE, the CSN and Min.not.

Indices that are not approved by INSEE may not be published with the "Indice Notaires INSEE" designation.

Article 8: Financial conditions

In consideration of the fact that the preparation, monitoring and regular publishing of the indices is a task that falls within the respective remits of all partners, and that the cooperation required under the terms of the present agreement entails an evenly balanced workload for each party, the present agreement shall be entered into without financial considerations.

Article 9: Duration of contract

The present agreement will come into force as soon as it is signed by all the partners. It is concluded for a period of five years.

Article 10: Termination of the agreement

The present agreement may be cancelled by any one of the parties by registered letter with acknowledgement of receipt. Cancellation shall take effect six months after receipt of this letter.

In the event of cancellation, The CSN, Min.not and INSEE waive the right to future use of the "Indice Notaires-INSEE" designation. However, the parties agree in advance to consult with each other in order to allow the continuation of financial operations undertaken by financial institutions or others, under any licensing agreements entered into by Min.not and the CSN.

Article 11: INSEE, CSN and Min.not representatives

The following persons shall execute the terms of the present agreement:

i) on behalf of INSEE, the head of the Housing division in the Department of Consumer Prices, Household Income and Living Conditions;

ii) on behalf of the CSN, its chairman,

ii) on behalf of Min.not, its chairman.

Article 12: Enforcement clauses

The present agreement is exempt from stamp duty and formal registration requirements.

The implementation of the present agreement shall have as an immediate consequence the abrogation of agreement 2005 00329 and its addendum dated 2010.

Signed in Paris in three original copies, 25 june 2014

A Paris, le 21 mai 2014	A Paris, le 21 mai 2014	A Paris, le 11 juin 2014	A Venelles, le 20 juin 2014
Pour le ministre des Finances et des Comptes publics et par délégation le directeur général de l'Insee	Pour le ministre de l'Économie, du Redressement productif et du Numérique et par délégation le directeur général de l'Insee	Pour le CSN	Pour Min.not
Signé : Jean-Luc TAVERNIER	Signé : Jean-Luc TAVERNIER	Signé : Me Jean TARRADE	Signé : Christine REY DU BOISSIEU

Appendix 9: Definitions

Used housing

We use a tax definition for used (second-hand) housing. A transaction is deemed to concern a used dwelling if it is the first sale more than five years after the construction completion date, or, if it is a second sale, then irrespective of the construction completion date. Thus the transaction may consist of the first sale of a property that is already used (more than five years old), or the second sale of a nearly-new property. The tax distinction is reflected in different tax rates: 0.6% for new dwellings, 4.8% for existing dwellings.

Suburb

Suburbs are suburban municipalities in urban units of 10,000+ inhabitants: all units not classified as "central city" are suburbs.

Reference property

Property whose price is tracked to calculate the indices. The characteristics of this property are the reference modalities for the explanatory variables of the transaction price model (e.g. 4-room house, on two floors, with a garage and bathroom). The reference property is described in Chapter 3 (Table 3.1).

Apartment buildings

Properties in apartment buildings are categorised as apartments (studio, apartment, duplex, triplex). We exclude rooms, attics, lofts, workshops and superintendents' lodgings.

Occupancy status and purpose

The dwellings tracked in the indices are vacant at the time of sale; they are intended for residential use only, and are acquired with full property rights. We do not exclude apartments leased by the purchaser before the sale. We do remove apartments occupied by a third party or by the seller when loss of use exceeds six months and when there is a right of use and residence or a right of usufruct.

Single-family dwellings

Properties are called single-family dwellings if they are houses, whether detached or not. They have a separate, direct private entrance from outside. To maintain consistency, we exclude property types such as large estates, manors, luxury townhouses, towers, and converted watermills/windmills. We therefore include farms, townhouses in cities and villages, detached houses and villas. We also include properties of which the type is not specified.

Transaction type

The only transactions considered are private sales conducted directly between sellers and buyers or through a real estate professional. Sales by voluntary auction on the Notaires' Real Estate Market (*Marché Immobilier des Notaires*) are therefore excluded.

Seller and buyer categories

The seller may be an individual, a real estate professional, or a company. Only dwellings purchased by individuals or by real estate companies (*Société Civile Immobilière (SCI)*) are tracked to calculate the indices. Dwellings purchased by real estate professionals are therefore outside the scope of the indices. Non-responses, which are much more numerous, are included in the scope of the indices.

Table A9.1 – Number of dwellings sold by buyer type: Provinces

	Apartments						Houses	
	Central cities of UU >10,000 inhab.		Suburbs of UU >10,000 inhab.		Rural and UU<10,000 inhab.			
Buyer type	Number	%	Number	%	Number	%	Number	%
Individuals	78,763	89%	34,883	90%	14,789	88%	193,190	91%
Real-estate companies (SCI)	4,602	5%	1,714	4%	808	5%	8,377	4%
Real-estate dealers	546	1%	210	1%	53	0%	1,288	1%
Other (non-professionals)*	949	1%	480	1%	210	1%	2,912	1%
Non-responses	3,830	4%	1,632	4%	959	6%	7,310	3%
Total	88,690	100%	38,919	100%	16,819	100%	213,077	100%

* Administrations, businesses, etc.

Scope: used housing transactions in notarial databases, 2010.

Table A9.2 - Number of dwellings sold by buyer type: Paris Region (apartments)

Buyer type	Apartments							
	Paris		Inner suburbs		Outer suburbs		Total	
	Number	%	Number	%	Number	%	Number	%
Individuals	25,376	85%	33,125	92%	24,092	95%	82,593	90%
Real-estate companies (SCI)	2,140	7%	1,471	4%	661	3%	4,272	5%
Real-estate dealers	231	1%	92	0%	28	0%	351	0%
Other	691	2%	599	2%	187	1%	1,477	2%
Non- responses	1,435	5%	859	2%	449	2%	2,743	3%
Total	29,873	100%	36,146	100%	25,417	100%	91,436	100%

Scope: used housing transactions in notarial databases, 2010.

Table A9.3 - Number of dwellings sold by buyer type: Paris Region (houses)

Buyer type	Houses					
	Paris and inner suburbs		Outer suburbs		Total	
	Number	%	Number	%	Number	%
Individuals	9,303	90%	26,476	94%	35,779	93%
Real-estate companies (SCI)	468	5%	779	3%	1,247	3%
Real-estate dealers	82	1%	147	1%	229	1%
Other	229	2%	268	1%	497	1%
Non- responses	257	2%	438	2%	695	2%
Total	10,339	100%	28,108	100%	38,447	100%

Scope: used housing transactions in notarial databases, 2010.

Reference stock

Stock of which the change in value constitutes the index (cf. Chapter 3). For a given neighbourhood or city, it consists of all the transactions in the reference period that fall within the scope of the index, except sales where the price is judged to be an outlier ($1/20^{th}$ at each end of distribution).[77] The reference stock forms the basket or portfolio of properties for the index.

Estimation stock

Housing stock of which the values are used to estimate the relative prices of property characteristics. It consists of all dwellings sold in the reference period that fall within the scope of the index. We exclude transactions considered to be aberrant, i.e. those whose residuals are greater than two standard deviations, and which therefore lie outside the interval $\left[\overline{x} - 2\sigma; \overline{x} + 2\sigma\right]$. The estimation stock and the reference stock are therefore sub-sets of the set of transactions for the estimation period that fall within the scope of the index (*Table A9.4*). Non-responses and the processing of missing observations are dealt with in Chapter 4 (*Table 4.7*).

[77] In version 1 of the model, one sixth of transactions with extreme values were removed as a precaution and because the quality of these data was not known (David *et al*., 2002). Later tests showed that the indices were in fact robust and could withstand a more parsimonious elimination of outliers.

For example: for the "Lille suburbs, used houses" stratum, we have 10,337 transactions in 2007 and 2008. Of these, 2,831 had some variables missing or were not properly documented, and in 340 cases the price was considered to be higher or lower than the observed price by more than two standard deviations. This left 7,166 transactions in the estimation stock. The reference stock comprises 6,772 transactions: the original 7,506 transactions of the estimation stock (7,166 + 340), minus the outliers.

Table A9.4 – Size of reference stock and estimation stock of Notaires-INSEE indices

	Number of dwellings in reference stock*	Number of transactions In estimation stock*
Metropolitan France	723,691	764,084
Paris Region		
Apartments	146,089	154,318
Paris	45,811	48,269
Inner suburbs	58,673	62,126
Outer suburbs	41,605	43,923
Houses (total)	55,792	58,893
Inner suburbs (+Paris)	14,923	15,733
Outer suburbs	40,869	43,160
Provinces		
Apartments	209,824	222,026
UUs > 10,000 centre	128,992	137,162
UUs > 10,000 suburbs	56,502	59,827
Rural and UUs < 10,000	24,330	25,037
Houses	311,986	328,847

* Stock in 2007-2008

Table A9.5 - Paris: comparison of structures in total stock, reference stock and annual transactions, %

	Number of dwellings*	Number of reference stock transactions**	Number of transactions in 2010	Number of transactions in 2011
Stratum				
1	20.0	20.4	19.7	20.2
2	30.0	33.0	32.4	31.8
3	26.0	23.3	23.6	24.1
4	108	10.1	10.5	10.2
5	13.2	13.1	13.8	13.6
Total	100.0	100.0	100.0	100.0
Construction period				
Before 1992	95.0	86.8	79.8	75.6
Since 1992	5.0	2.2	2.6	2.5
Unknown	0.0	11.0	17.6	21.7
Total	100.0	100	100.0	100.0
Size				
1 or 2 rooms	55.8	57.0	55.7	57.9
3+ rooms	44.2	41.2	42.5	40.6
Unknown	0.0	1.8	1.8	1.5
Total	100.0	100.0	100.0	100.0

* Source: 2008 Population census
** Stock 2007-2008

Table A9.5 compares the structure of the index reference stock (2007 - 2008) with that of the housing stock in the 2008 population census, and transactions from 2010 and 2011 for Paris. For a considerable proportion of the variables, information was not supplied for the databases and in addition, the number of transactions does not include new housing stock. Comparisons are therefore approximate. However, overall, these structures show that transactions give a good representation of the total housing stock.

A detailed breakdown of the estimation stocks for apartments and houses in the Paris Region and the Provinces is given at the end of the appendices.

Base period, index base quarter

1st quarter 2010.

Estimation period or reference period

Period for which we estimate the basic models, i.e. the relative prices of property characteristics. In version 3 of the indices, this period is two years.

Property price

Seller's net price (after deduction of agency commission if this is recorded in the deed of sale), therefore excluding taxes and legal fees.

Stratum, neighbourhood

Strata are geographic areas within which prices are homogeneous (cf. Chapter 3). For the method used to determine the strata, cf. Appendix 2. For the stratification of the municipalities, cf. Excel file, obtained on request from notaries.

Thus a stratum corresponds to the scope of application of a distinct hedonic model. A neighbourhood is a finer division, whose influence is demonstrated by suitable dummies within a given hedonic model. Neighbourhoods do not necessarily correspond to administrative boundaries.

Urban unit

An urban unit is defined as a set of dwellings arranged so that no two units are more than 200 metres apart, and containing at least 2,000 inhabitants. Municipalities (*communes*) that meet these criteria make up urban units (UU), the others are said to be "rural". The concept of the urban unit, based on the continuity of the built-up area and the threshold of 2,000 "agglomerated" inhabitants, is thus a mainly visual notion, based on demographics and settlement patterns.

Central city

A municipality (*commune*) is defined as a central city if it contains more than half of the population of the agglomeration. Otherwise, all municipalities with a population exceeding half the population of the largest municipality in the agglomeration are classified as central cities, as is the largest municipality.

If the annual number of transactions is less than 110, the central city does not have its own specific index.

Appendix 10: Bibliography

For a more comprehensive bibliography, see Beauvois et al. (2005), Insee Méthodes no. 111, available online on this website.

Bailey, M., Muth, R., Nourse H.: "A Regression Method for Real Estate Price Index Construction", *Journal of the American Statistical Association*, 58, 933-942, 1963.

Beauvois, M., David A., Dubujet F., Friggit, J., Gouriéroux C., Laferrère A., Massonnet, S., Vrancken, E. : "Les indices Notaires-Insee de prix des logements anciens, Version 2 des modèles hédoniques", Insee Méthodes, no. 111, 2005.

Case, B., Pollakowski, H., Wachter S.: "Frequency of Transaction and House Price Modelling", *JREFE*, 14, 173-188, 1997.

Clapp, J., Giacotto, C.: "House Price Indices Based on All Transactions Compared to Repeat Subsamples", *AREUEA*, 19, 270-285, 1991.

Court A.: "Hedonic Price Indices with Automotive Examples", p. 99-117, in The Dynamics of Automobile Demand, New-York, General Motors Company, 1939.

David A., Dubujet F., Gouriéroux C. and Laferrère A.: "Les indices de prix des logements anciens", Insee méthodes, no. 98, 2002.

Duon, G.: "Documents sur le problème du logement", Ministère de l'Économie, Service National de la Statistique, *Études Économiques, n°1,* 1946.

Eichholtz Piet M.A., "A Long Run House Price Index : the Herengracht Index, 1629-1973", REE, 25 (2),175-192, Amsterdam, August 1996.

Grebler, Blank and Winnick, Annexe C of "Capital Formation in Residential Real Estate: Trends and Prospects", *National Bureau of Economic Research,* Princeton University Press, Princeton, 1956.

Griliches Z.: "Price Indices and Quality Change", Cambridge, Harvard University Press, 1971.

Mark, J., Goldberg M.: "Alternative Housing Price Indices: An Evaluation", *AREUEA*, 12, 30-49, 1984.

Moreau, A.: "Methodology of the Price Index for Microcomputers and Printers in France", in *Industry Productivity: International Comparison and Measurement Issues*, OECD Proceedings, 99-118, 1996.

Triplett, Jack E.: "Concept of Quality in Input and Output Price Measures: A Resolution of the User-Value Resource Cost Debate", in *The US Income and Product Accounts: Selected Topics*, ed M.F. Foss, NBER Studies in Income and Wealth, University of Chicago Press, Chicago, 1983.

Triplett Jack E. : "Handbook on Hedonic Indexes and Quality Adjustments in Prices Indexes: Special Application to Information Technology Products", *STI Working papers 2004/9, OECD*, 2004.

Warsh, D.: "Economics principals, Roots and Wings", Globe Columnist, 14 March 1999.