New or old, why would housing price indices differ? An analysis for France

Thomas Balcone* and Anne Laferrère*

Abstract – In France, housing price indices have been computed since 1996 for second-hand dwellings (*Indices Notaires-Insee*) and since 2013 for new dwellings (*Indice des prix des logements neufs*). The evolution and volatility of the two indices differ. We explore why, using notarial data and surveys on new home sales (*ECLN*) and the price of building land (*EPTB*). The computation methods and the scopes of the indices contribute relatively little to the differences. The location of new and old housing differs, but this still explains only part of the differences. Decomposing a home value between that of land and that of buildings reveals that the share of land is higher for second-hand than for new dwellings. Land and second-hand dwellings price indices evolve in a very similar way. Structures prices have a greater influence on the price index of new homes. Nevertheless, construction costs are sensitive to the trend in land prices. The counter-cyclical construction of social housing may have contributed to reduce the volatility of new homes prices.

JEL Classification: C43, C81, E31, R31
Keywords: indices, hedonic models, housing demand

* *Insee at the time this research began (thomas.balcone@enseignementsup.gouv.fr; anne.laferrere@dauphine.fr)*

We are grateful to the Scientific Board of the Indices Notaires-Insee*, to Ronan Le Saout and Benjamin Vignolles for helpful discussions. The latter also provided his computation of the distance variable in* EPTB*. We thank Cédric Cailly for his help with the* EPTB *data and Insee and* SDES (Service de la donnée et des études statistiques, of the ministère de la Transition écologique et solidaire) *for allowing access to restricted data (convention n° 2014065NF). We are also grateful to the notaries for providing the data used for the methodology comparison. Last but not least, two referees provided invaluable critics and comments. We alone remain responsible for errors.*

Received on 8 June 2017, accepted after revisions on 17 May 2018

To cite this article: Balcone, T. & Laferrère, A. (2018). New or old, why would housing price indices differ? An analysis for France. *Economie et Statistique / Economics and Statistics*, 500-501-502, 69–95. https://doi.org/10.24187/ecostat.2018.500t.1946

Housing price indices have been published quarterly in France since 1996 for second-hand dwellings[1]. The sales prices are recorded by notaries and used to compute the so-called *Indices Notaires-Insee (INI* hereafter), at the national and various local levels[2]. Since 2013, an index for new dwellings has been computed by Insee, the *Indice de prix des logements neufs (IPLN)* relying on price data from a survey on the commercialisation of new dwellings (*ECLN*[3]). It soon appeared that the evolution of *INI* and *IPLN* indices were somewhat different. Over the 2006-2015 period, the volatility of the *INI* index was more than twice that of the *IPLN* index. This article aims at exploring why.

A first section summarizes the current methods of computation of the new and second-hand housing price indices and points to the differences in evolution of these two indices. Next, two potential sources of methodological differences are examined: the effect of the computation methods and that of the exclusion of individually built single-family houses (IP) from the *IPLN* index, i.e. the different scope of the two indices. We complement the *IPLN* index with an index for those individually built single-family houses*,* using another rich data source, the *Enquête sur le prix des terrains à bâtir (EPTB)*; it allows computing a new, extended housing index. We call it *IPLN++*. Eliminating such differences in method and scope does not suppress the differences in time evolution and volatility. We then turn to some other sources of differences between the indices. The markets for new and second-hand homes differ; in particular, new homes are not located in the same areas as existing homes: they are mostly built in the periphery of cities, where land is available and cheaper. Hence the share of land in the home value would be smaller for new homes than for old ones and it might influence price evolution. We conduct two types of experiments. The first draws from the limited geographical information provided by the data used in computing the two indices, *INI* and *IPLN.* The distance to the city centre is not known, only the municipality. We compute a second-hand price index for dwellings situated in the same municipalities as the new ones. The difference between the two indices is somewhat reduced but not eliminated. Secondly, drawing from the *Enquête sur le prix des terrains à bâtir* (*EPTB*), separate indices for land and structure are computed and compared to the *INI* and to the *IPLN*. Land price evolution seems to be driving second-hand housing prices more than construction prices. The latter have more influence on new homes prices than land prices. However, the construction costs are also sensitive to the general trend in land prices. Looking for other potential explanations, two features are striking: the difference between the two indices is particularly important during the 2008-2009 crisis; and the prices for second-hand dwellings are more volatile than those of new dwellings. Looking at time series of sales and constructions suggests that the countercyclical building of social housing might have contributed to the lower volatility of new homes prices.

Current methods of computation of new and second-hand dwellings price indices in France

The approach to the second-hand (old) dwellings price index is that of hedonic imputation, while that for the new dwellings price index is that of a time dummies hedonic model.

Second-hand dwellings: the *Indice Notaires-Insee (INI)*

The second-hand dwellings price index (*INI*) is computed by a hedonic method based on the estimation of disaggregated models in homogeneous zones, separately for houses (181 zones) and flats (112 zones). Price zones are determined with an ascending hierarchical classification based on various statistics at the neighbourhood or canton level adding a criterion of geographic contiguity outside Île-de-France. The data consist of transaction prices collected by the notaries (see Gouriéroux and Laferrère (2009) and Clarenc *et al.* (2014) for details). Each quarter, the models are used to estimate the price of a fixed reference basket of dwellings in each zone. The reference basket is made up of two years of transactions and updated every two years. Hence the index is an index of transactions, not of the whole stock; but the "basket" is large enough not to be biased by short-term changes in the market[4].

1. *Second-hand in the fiscal sense, i.e. aged more than 5 years or being sold for the second time.*
2. *See: https://www.insee.fr/fr/statistiques/series/102770558. Excel files are also attached to each quarterly publication. See https://www.insee.fr/fr/statistiques?debut=0&theme=30&conjoncture=56.*
3. Enquête sur la commercialisation des logements neufs *conducted by the SDES. See Balcone (2013, 2018) for details.*
4. *1/20th of the extreme values are omitted from the computation.*

The basic model is as follows (omitting zone indices):

$$\log p_i = \log p_0 + \sum_{a=1}^{2} \mu_a Y_{a,i} + \sum_{m=1}^{12} \theta_m M_{m,i} + \sum_{k=1}^{K} \beta_k X_{k,i} + \varepsilon_i \qquad (1)$$

where:

- p_i is the price (per m² for flats) of dwelling *i*;

- $Y_{a,i}$ is a dummy for the year of sale of dwelling *i*;

- $M_{m,i}$ is a dummy for the month of sale of dwelling *i*;

- $X_{k,i}$ are *K* characteristics of dwelling *i*, including physical characteristics (size, number of rooms, of bathrooms, of floors – interacted with existence of a lift for flats –, garage), date of construction, plot size for houses, etc., and neighbourhood dummies, proxying for local amenities;

- p_0 is the price of the "reference" dwelling defined by the omitted characteristics in (1)[5].

Similar models can be estimated at each date *t*, allowing estimating the price of the reference basket at each date. In practice the models are only revised every two years[6]. The ratio of the estimated values of the reference basket between *t* and *t* – 1, provides the index. Indices are then chained from period to period. The sub-indices by zones and type of dwellings are aggregated at higher geographic levels for publication[7]. Aggregation uses geometric means when the geographic level is infra-*département*, that is a small enough zone where the consumer is assumed to make her residential choices, and arithmetic means at higher geographic levels, with weights reflecting transaction values[8].

New dwellings: the *Indice de prix des logements neufs (IPLN)*

The data source is the *ECLN*. The survey covers all building permits of 5 units or more: single-family units that are part of a development (called "*individuel groupé*", IG), and all multi-family units (in apartment buildings, called "*collectif*", Coll). Individually built single-family units, i.e. units built one by one, called "*individuel pur*" (IP), are left out of the index because, except when they are built by a developer and sold "*clefs en main*", no proper sale price is recorded. This is consistent with the scope of the European new dwellings index[9]. The *ECLN* survey only covers dwellings built for the private market. Subsidized construction for the social sector is left out. Table 1 compares the scope of housing price indices to that of all new home constructions and sales during the period 2006-2012. Social housing represented about 14% of new construction, and 18% when adding the increasing part of social housing built by private developers. Social housing is left out of the index computation because no costs or prices are recorded. The *IPLN* index covers 61% of all new housing units aimed at the private market and 98% of those for which a price is recorded.

For one newly built private dwelling sold, nearly four second-hand dwellings are sold. This is why price indices for second-hand dwellings can be computed at various geographical levels, and separately for houses and flats, while the *IPLN,* the official price index for new dwellings, has been computed only at the national level. Besides, because of the difference in the number of available recorded prices, the indices are computed with different methodologies: hedonic imputation for the *INI* and adjacent two-period time dummies for the *IPLN*.

Each quarter, the model is estimated on two successive quarters of data. The data are available only at the level of a construction program, not at the dwelling level. The following information is provided by type of construction ("*individuel groupé*", IG or "*collectif*", Coll[10]) for each class of number of rooms (from 1 to 6 or more): the total number of sold dwellings, the average size in m² and the average price of the sold dwellings[11].

5. *For instance the reference house is 100 m² on a 610 m² plot, has 4 rooms, 2 levels, a garage, one bathroom, of unknown construction date and is sold in December of year 2 of the reference period. R^2 ranges from 0.25 to 0.85.*
6. *The model estimated over the period 2009-2010, was used to compute the indices for the period 2012-2013.*
7. *Only the sub-indices in cities or* départements *with enough transactions get the* Notaires-Insee *label.*
8. *By construction such models only allow getting different price evolutions by zones, separately for houses and apartments. They assume that the price evolution of a given basket of homes is the same within a zone, or whatever the number of rooms, or the date of construction. Details of the method can be found in Clarenc* et al.*, 2014 at https://www.insee.fr/fr/information/2569926.*
9. *See Owner-Occupied Housing regulation (EU Commission regulation N° 93/2013).*
10. *And whether the sold dwellings are private ordinary dwellings or part of residences offering specific collective services.*
11. *Hence, only the mean characteristics of the sold dwellings of a program are available. For similar houses, it is not a big issue. For apartments, some important characteristics, such as the level in the building, are unknown. It is of no consequences if the relative price of the omitted variable is constant and the frequency of the characteristics is also constant over time.*

Table 1
Newly built or sold second-hand dwellings and prices indices coverage

	Number of dwellings	Share of all dwellings (in %)	Type of dwellings	Share of new dwellings by type (in %)	Indices
New dwellings built for:	372,866	33	-	-	-
the private market	304,580	27	"Pure" single-family units (IP)	39	*IPLN*
			"Grouped" single-family units (IG)	12	
			Flats ("collective") + Residences	49	
the non- private market	68,286	6	"Pure" single-family units	1	
			"Grouped" single-family units	25	
			Flats	74	
Sales of second-hand dwellings	740,571	67			*INI*
Total	1,113,438	100			

Note: Annual average over 2006-2012. Units built with permits of 2 to 4 units are in principle included in "Grouped" single-family units or in "collective" dwellings, but they are excluded of the *ECLN* survey and consequently from the scope of the *IPLN* (they represent only 2% of new dwellings for the private market). Homes transmitted by bequest or gift are excluded.
Sources: *Sit@del2* for total new constructions. The number of constructions for the non-private market is estimated from the number of homes built by the social sector including homes built in "VEFA" (dwellings sold before they are built - *vente en l'état futur d'achèvement*) by the private sector for the public sector (see CDC, 2015). Sales of second-hand homes are estimated from CGEDD, from DGFiP (*MEDOC*) and notaries database (http://www.cgedd.developpement-durable.gouv.fr/nombre-et-montant-des-ventes-immobilieres-a1003.html). See Friggit (2014) for the method.

The hedonic model is the following:

$$\forall t = T-1, T \text{ and } \forall i = 1,\ldots, nb\_obs(t),$$
$$\ln(\overline{p}_{i,t}) = \alpha + \beta_S \ln(S_{i,t}) + \sum_{k=1}^{K} \beta_k I_{i,t,k} + \delta_T D_{i,t,T} + \varepsilon_{i,t} \qquad (2)$$

where:

- $\overline{p}_{i,t}$ is the average price of dwellings of program[12] *i* sold in quarter *t*;

- $S_{i,t}$ is the average size in m² of the dwellings of program *i* sold in quarter *t*;

- $I_{i,t,k}$ is a vector of characteristics of the dwellings: type of construction (IG or Coll), number of rooms, dummies for "standing", presence of a swimming pool, air conditioning, balcony for flats, etc. Location is taken into account through 14 dummies for areas that are homogeneous in terms of price per square meter[13], and dummies for some characteristics of the municipality (e.g. sea[14], ski or hiking resorts);

- $D_{i,t,T}$ is a time dummy for quarter *T*.

The quarterly change in the index is obtained by the exponential of the coefficient of the quarter dummy. Contrary to second-hand dwellings, there is only one model for the whole of France[15]. Moreover, houses and flats are not separated because of the small number of quarterly observations in the *ECLN*[16]. Balcone (2013, 2018) describes the computation method in details. The main differences between the new and second-hand dwellings price indices are summed up in Table 2.

Comparing price indices of new and second-hand dwellings

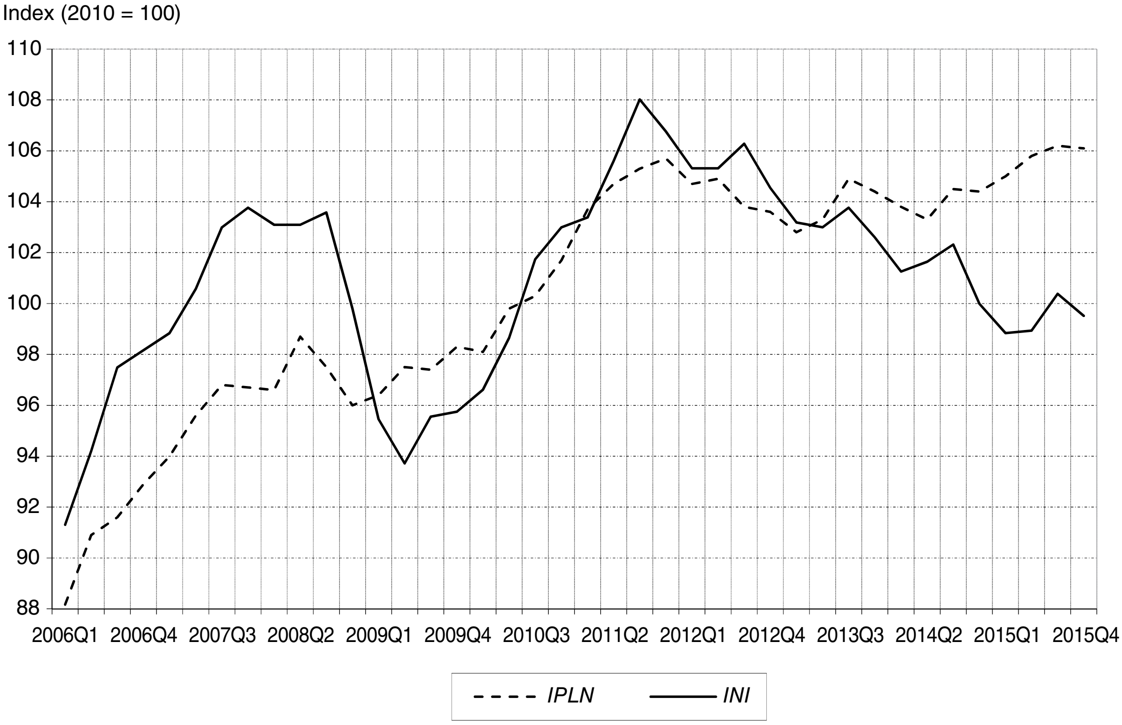
Over the 2006-2015 period, the new and existing second-hand dwellings indices evolved differently (Figures I to III): the average difference in absolute value between the quarterly growth rates of the two indices is not negligible (1.2 percentage points). The same is true for the annual growth rates (2.4 percentage points). Actually, both indices follow the same trend except over the crisis 2008_Q4-2010_Q1 period and again in 2014_Q4 (Figure III). In 2008_Q4, the *INI* falls by 4 points, it falls by another 4 points in 2009_Q1, and by 2 points in

12. *More exactly, the price corresponds to a program and a class of number of rooms.*
13. *The 14 zones have been computed by a ascending hierarchical classification based on 8 large regions (ZEAT) and 9 urban unit sizes, from dwelling prices and sizes. An urban unit is a municipality or group of municipalities with a continuously built zone (i.e. less than 200 meters between two constructions) with more than 2,000 inhabitants.*
14. *The law* Littoral *n° 86-2 (1986) defined the classification. A municipality is coastline* "littorale (or maritime)" *if on the seaside, near ocean or salty marshes;* "arrière-pays littoral" *is a non-coastline municipality within a coastline* canton *(a group of municipalities with at least one coastline municipality).*
15. *Metropolitan France excluding Corsica.*
16. *Over the period 2006_Q1-2012_Q3, the average quarterly number of observations is 8 194 "programs x number of rooms", corresponding to 26 105 new dwellings. In the regression each observation is weighted by the corresponding number of sold dwellings.*

Table 2
Main differences between the new and second-hand dwellings indices

	IPLN	*INI*
Data	Survey on the commercialisation of new dwellings: *ECLN*	Transactions registered in the Notaries databases: *BIEN* (Base d'informations économiques notariales) database for Île-de-France and *Perval* (Min.not ADSN) database for the Province.
Method	Adjacent two-period time dummy hedonic model	Hedonic imputation
Geographical effect	14 dummies for zones + municipality characteristics, in a single model	One hedonic model for each of the 293 zones + neighbourhood dummies

Figure I
New dwellings (*IPLN*) and second-hand dwellings (*INI*) price indices, 2006-2015



Sources: *ECLN*, *BIEN* and *Perval* data bases (see Table 2); authors' computation.

2009_Q2. The *IPLN* only drops by 2 points in 2008_Q4. In 2014_Q4 the *INI* falls by 2 points, while the *IPLN* does not change. The index for second-hand dwellings appears twice more volatile than the index for new dwellings (*IPLN*): the standard deviation of its annual evolution rate over this period was 4.59%, when it was 2.31% for the *IPLN*.

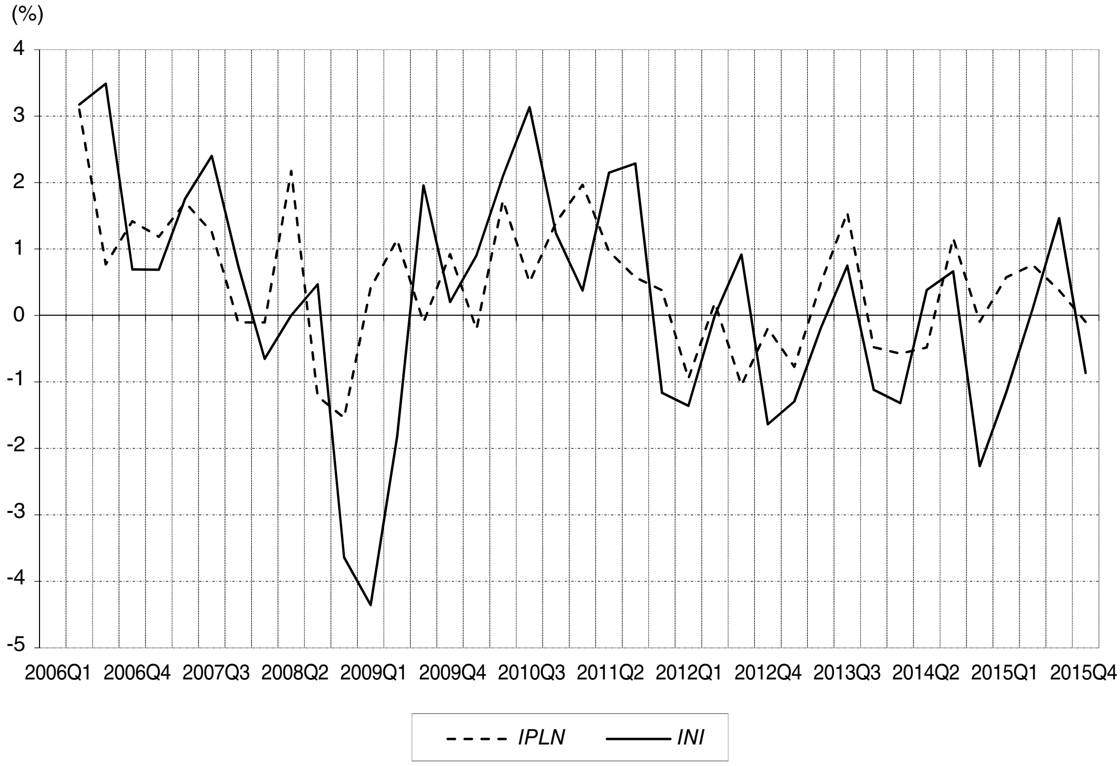
Potential methodological sources of differences between the *IPLN* and *INI* indices

A first possible source of differences between the evolutions of the two indices comes to mind: the bias due to different methods, adjacent two-period time dummy method on a single model versus hedonic imputation method at a disaggregated geographical level. This has been explored in details by Balcone (2013, 2018). The next section summarizes his results. The period of observation is restricted to 2006-2010 because of the availability of the Notaries data for this simulation.

The difference in computation methods has little effect

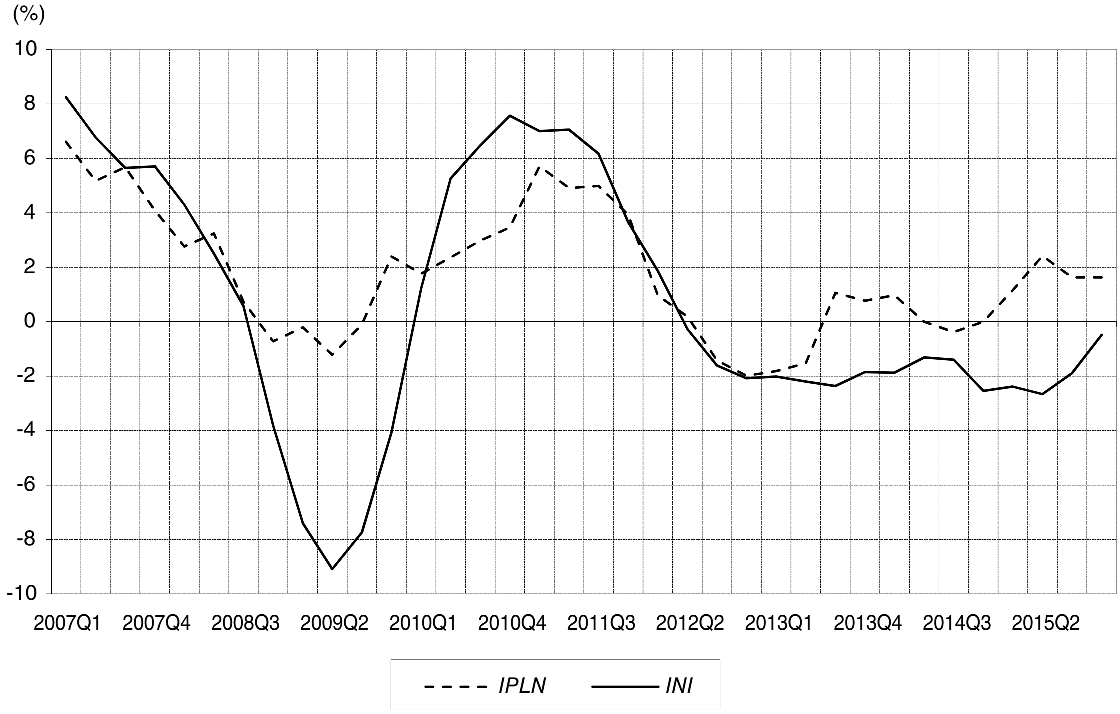
Taking the second-hand dwellings sales data used for computing the *INI*, Balcone (2013) used the same adjacent two-period time dummy hedonic model as the one used for the

Figure II
Quarterly growth rate of the new dwellings (*IPLN*) and the second-hand dwellings (*INI*) price indices, 2006-2015



Sources: *ECLN, BIEN* and *Perval* data bases (cf. Table 2); authors' computation.

Figure III
Annual growth rate of the new dwellings (*IPLN*) and the second-hand dwellings (*INI*) indices, 2006-2015



Note: Annual growth is the year-on-year evolution of the quarterly index.
Sources: *ECLN*, *BIEN* and *Perval* data bases (cf. Table 2); authors' computation.

new dwellings index. Two indices were computed: one for houses and one for flats. To get closer to the smaller number of explanatory variables in *ECLN*, the number of explanatory variables was reduced compared to model (1), the geographical dummies were the same as in model (2) and a single model was used for the whole country.

Such experimental time dummy indices for second-hand houses and flats are compared to the *INI* indices. Since the only source of difference between these two sets of indices is the computation method, the gap between the indices is used to assess the potential bias due to different methods**.** For houses, the absolute value of the difference between the two indices is not higher than 2.6 index points (Figure IV-A), even if seasonal changes in the "time dummy" index are more pronounced than in the *INI* (Figure IV-B). The annual profiles are similar, and differences are most often no more than 1 percentage point (not shown). The difference is larger (2 percentage points) in 2009_Q3 in the middle of the crisis. The difference in the computation methods clearly cannot account for the difference between second-hand and new price indices for houses. For second-hand flats, the absolute value of the difference between the two indices is not higher than 2.6 index points (Figure V-A). Moreover, the evolution rates of both indices are also close (Figure V-B).

The results, both for second-hand houses and flats, lead to conclude that even if the computation method implemented for the *INI* is very different from a simplified adjacent two-period time dummy method, the final difference at the national level is surprisingly small. The time dummy method would lead to a less volatile index, as was the case for the *IPLN* compared to the *INI*. For instance for houses over the period 2006-2010, the standard deviation of the annual evolution rate is 5.90% for the *INI* and 5.58% for the "time dummy" index. However, the method does not explain the gap between the *INI* and the *IPLN* that was observed during the 2008_Q4-2010_Q1 crisis period.

Adding individually built single-family homes to the *IPLN*

As shown in Table 1, the effective scope of new and second-hand dwellings indices differs: *IPLN* covers only the "*individuel groupé*" (IG) and "*collectif*" (Coll) dwellings whereas *INI* also includes "*individuel pur*" (IP). To get a comprehensive new dwellings index, we use another rich data source, the *EPTB* survey[17]. It covers the building permits of "individually built" single-family houses (IP) and provides the prices and features of land plots (size in m², whether it was purchased or not, purchase date, servicing works done or not, etc.). It also provides the expected price of the construction, and some of the house features: floor space, nature of the main coordinator of the works, type of heating system. We keep 344,847 observations for which land was purchased between 2006 and 2012, was located in metropolitan France[18], and for which the size of the purchased plot is equal to that registered in the building permit[19]. From these data, a quality-adjusted price index for new IP single-family houses is computed using the adjacent two-period time dummy method.

An "individually built" single-family houses price index

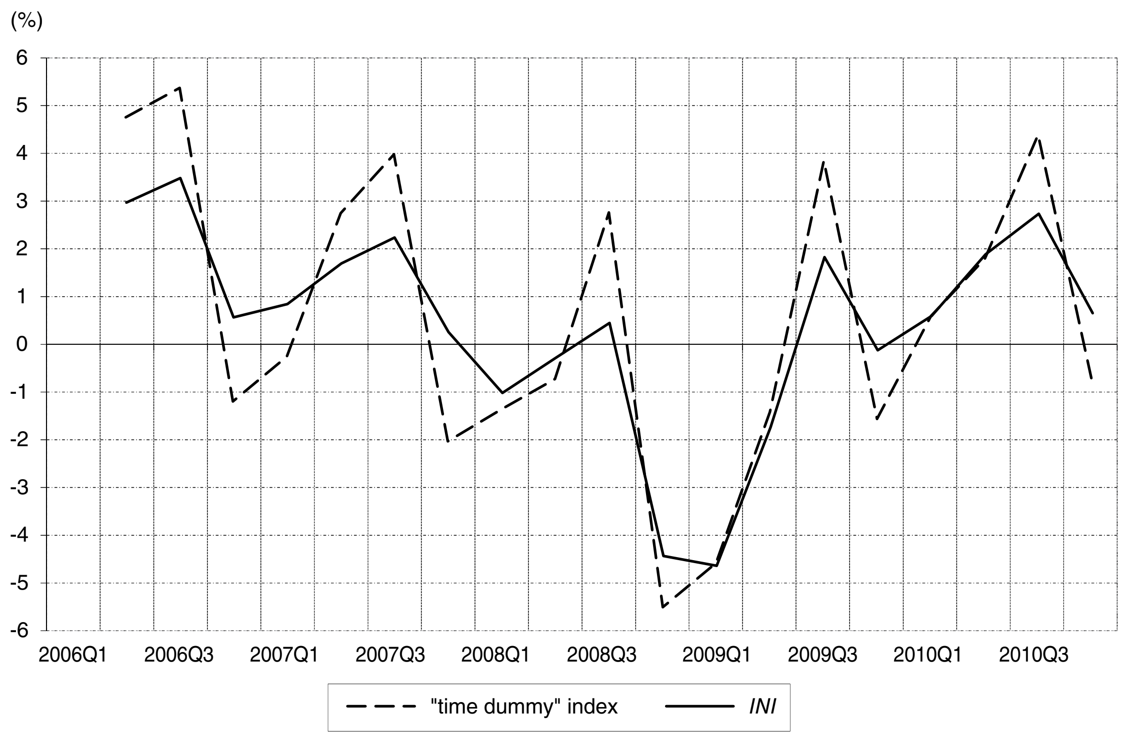
Since the *EPTB* data are richer than the *ECLN*, the location effect is taken into account by estimating one hedonic model per region. Dummies for municipal amenities are included: coastline, estuary, touristic "*arrière-pays littoral"* or ski or alpine resort. We also control for the type of urban unit: if the municipality is a single urban unit, it is a "*ville isolée"*; if it belongs to an urban unit made of several municipalities, it is then either a city centre or a suburb[20]; municipalities outside urban units are called rural. We finally add the straight-line distance in kilometers between the municipality where the land is bought and the closest urban centre[21]. Moreover, differences of quality between houses are taken into account by adding dummies for the construction characteristics. The hedonic model (3) used for each of the 21 regions r is the following (omitting the region index):

17. Enquête sur le prix des terrains à bâtir *conducted annually by the* ministère de la Transition écologique et solidaire.
18. *Excluding Corsica, as for the new dwellings index.*
19. *The registered plot size is the total underlying land (ground floor + gardens and outhouses). We exclude the 4.6% cases where the individual buys just an extension of a plot which was his already, and the 5.2% cases where he buys a large plot, then divides it and uses only part of it for the building. Then the mean price per square meter may be lower because only part of the purchased land may have building permission, the remaining corresponding to farmland for instance.*
20. *Definitions can be found at : http://www.insee.fr/en/methodes/default.asp?page=definitions/ville-centre-et-banlieue.htm.*
21. *The urban center (*pôle*) is an urban unit offering at least 10,000 jobs and not located in the crown of another urban centre. The crown from an urban cluster covers all the municipalities in the urban area to the exclusion of its urban centre.*

Figure IV
Second-hand houses: "time dummy" and *INI* indices, 2006Q1-2010Q4

A – Indices

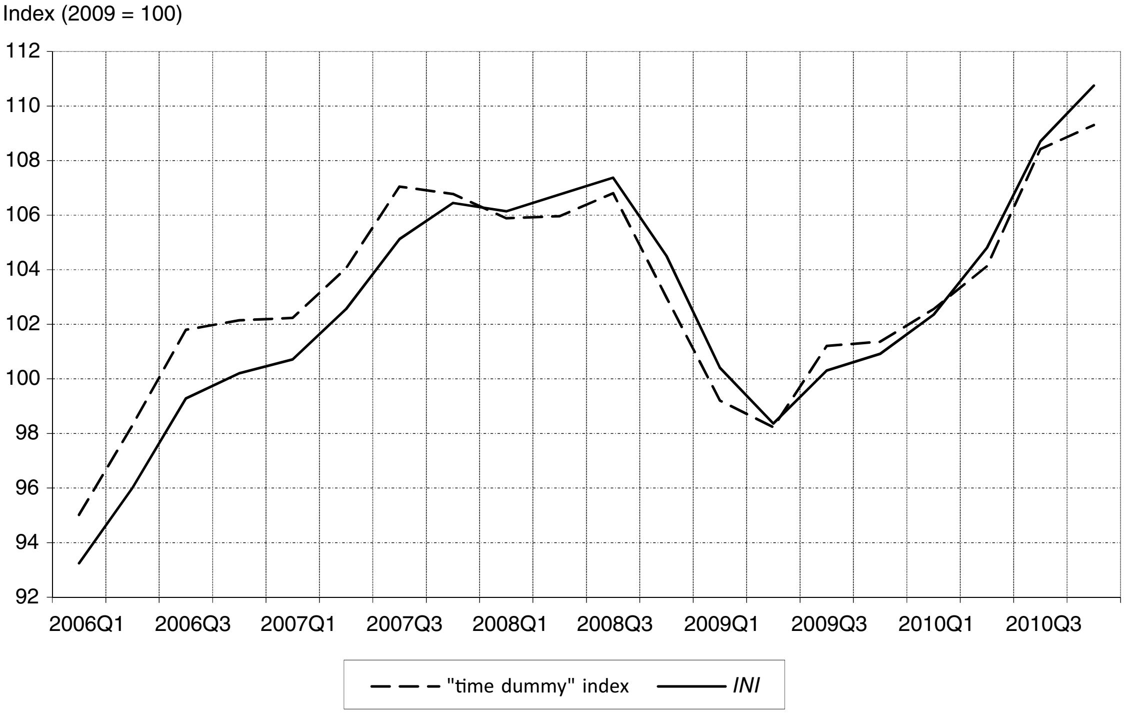
B – Quarterly growth rates



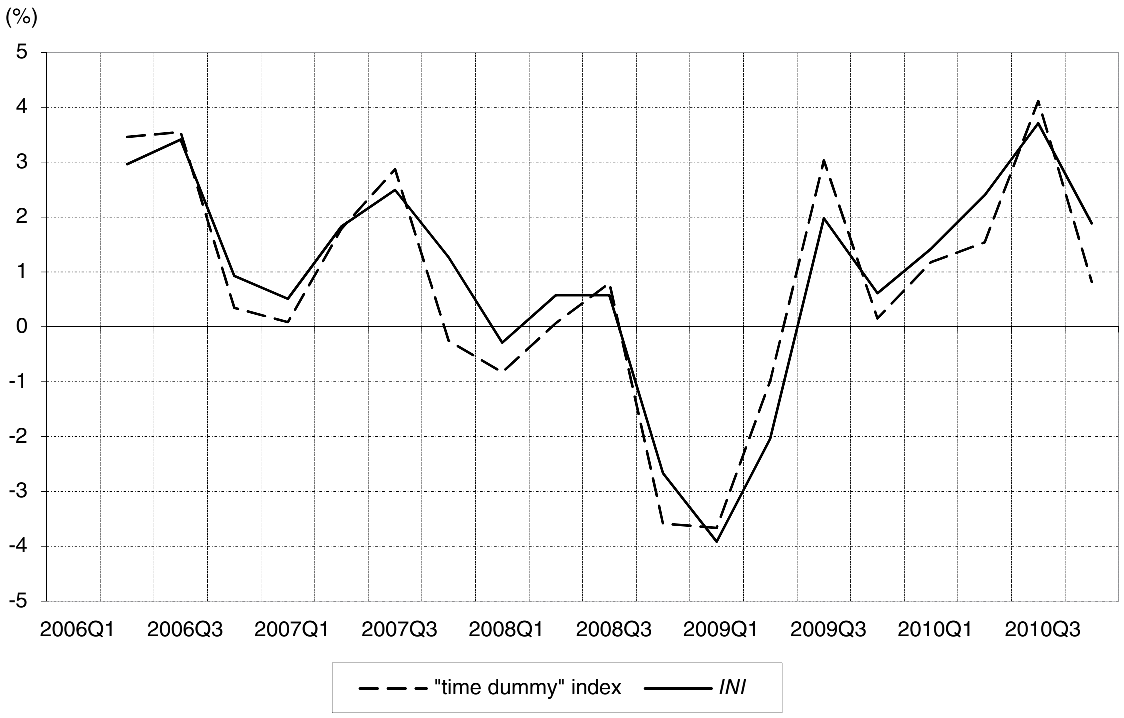
Sources: *ECLN*, *BIEN* and *Perval* data bases (cf. Table 2); authors' computation.

Figure V
Second-hand flats: "time dummy" and *INI* indices, 2006Q1-2010Q4

A – Indices



B – Quarterly growth rates



Sources: *ECLN*, *BIEN* and *Perval* data bases (cf. Table 2); authors' computation.

$$\forall A = 2007,...,2012, \forall a = A-1, A, \forall i,$$

$$\ln\left(p\_V_{i,a}\right) = \alpha_{A,r} + \beta^{A}_{s\_L}\ln\left(s\_L_{i,a}\right) + \beta^{A}_{shon}\ln\left(shon_{i,a}\right) + \beta^{A}_{dist}dist_{i,a} + \sum_{k=1}^{K}\beta^{A}_{k}I_{i,a,k} + \delta_{A,r}D_{i,a,A} + \varepsilon_{i,a} \quad (3)$$

Where, for plot i in year a:

- $p\_V_{i,a}$ is the house value, the sum of the land price and the construction price;

- $s\_L_{i,a}$ is the plot size in m²;

- $shon_{i,a}$ is the size in m² of the house[22];

- $dist_{i,a}$ is the distance to the nearest urban centre;

- $I_{i,a,1},\ldots, I_{i,a,k},\ldots, I_{i,a,K}$ is a vector of K dummies for the characteristics of the land and of the structure : the geographical dummies described above to which we add whether the plot was serviced ("*viabilisé*"), or was bought through an intermediary or not and its type, degree of finish of the structure (totally fitted, ready to decorate, only "*clos et couvert*"), heating mode (gas, electricity, renewables, etc.), type of builder (architect, developer, artisan, self-building, other);

- D_A is year A dummy.

From the models, we compute, for each of the 21 regions r an annual constant quality price index for year A, 100 = A – 1, $I\_V_{r,A/A-1}$:

$$I\_V_{r,A/A-1} = \exp\left(\hat{\delta}_{A,r}\right)*100,$$

where $\hat{\delta}_{A,r}$ is the OLS estimator of $\delta_{A,r}$.

The national index for year A (100 = A – 1), $I\_V_{A/A-1}$ is the weighted average of the 21 regional indices:

$$I\_V_{A/A-1} = \sum_{r=1}^{21} w\_V_{r,A-1} * I\_V_{r,A/A-1},$$

where $w\_V_{r,A-1}$ is the share of the expenses for single-family units in region r, in year A-1 (see Appendix 1). This index is chained to get a national annual price index for new "individually built" single-family houses for year A, 100 = 2006, ∀ A = 2007, …, 2012,

$$I\_V_{A/2006} = \left(\prod_{i=0}^{A-2007}\frac{I\_V_{A-i/A-i-1}}{100}\right)*100$$

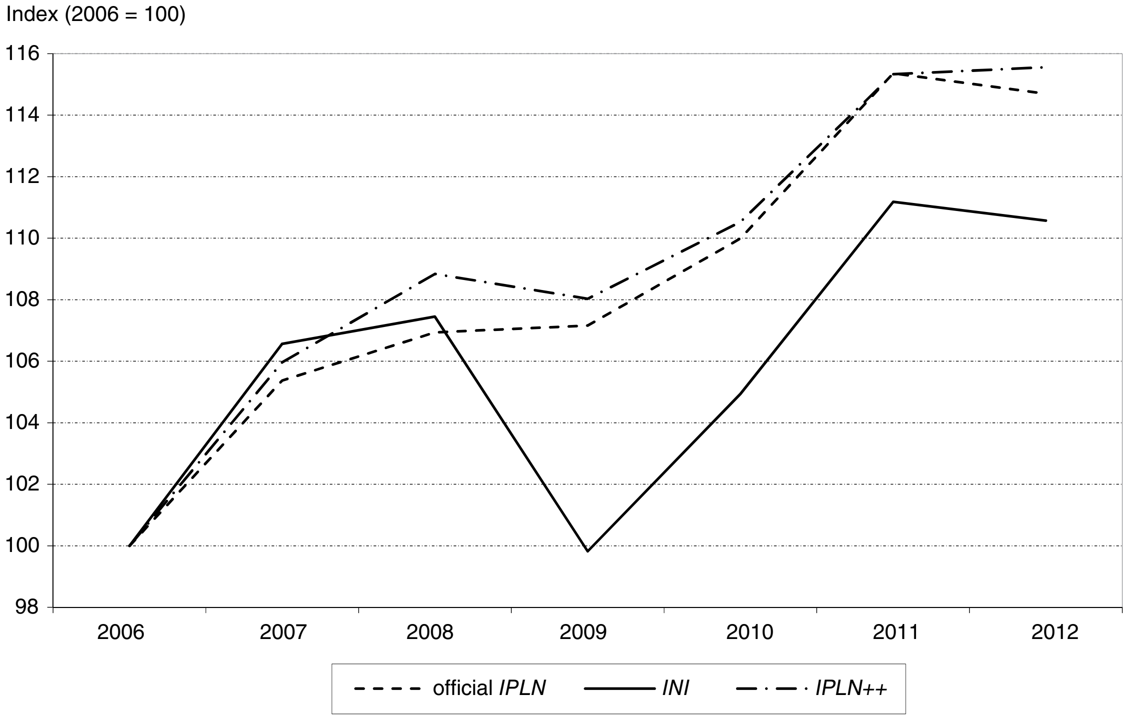
A comprehensive price index for new homes

The index for newly built single-family houses is then aggregated to the current official *IPLN* index to get a comprehensive price index for new homes, adding "*individuel pur", "individuel groupé"*, and "*collectif",* that we call *IPLN++*. This is done for each year. Over the period where the comparison is possible the *IPLN* index and the comprehensive price index *IPLN++* are extremely close (Figures VI and VII). Extending the scope of the price index of new homes to include single-family units does not modify it significantly. The comprehensive index *IPLN++* does not get closer to the *INI*. The gap even slightly increases: the average absolute difference in the annual growth rates of the two indices is now 2.0%, whereas it was 1.8% with the current official *IPLN*.

Since the effect of the methodology has been ruled out, we hypothesize that the differences in evolution of the prices of new and second-hand dwellings come from deeper differences between the two markets. New homes are not located in the same areas as old ones; they are mostly built in the periphery of cities, where land is available and cheaper. This leads to reflect on the decomposition of the price of a home into that of the structure and that of the land, and introduce the notion of land leverage, the share of land in the home value. Then two types of experiment are conducted. The first one draws from the limited geographical information provided by the data used to compute the two indices, *INI* and *IPLN*. The distance to the city centre is not known, only the municipality. We compute the evolution of a price index for second-hand dwellings situated in the same municipalities as those newly constructed. The difference in the evolution of the two indices is somewhat reduced but not eliminated. Secondly, drawing from the *EPTB* survey*,* separate indices for land and structure are computed. Land prices seem to be driving second-hand housing prices more than construction prices. Conversely, structure prices have more influence on changes in new homes' prices. Still, the evolution of the construction cost index seems also sensitive to the general trend in land prices.

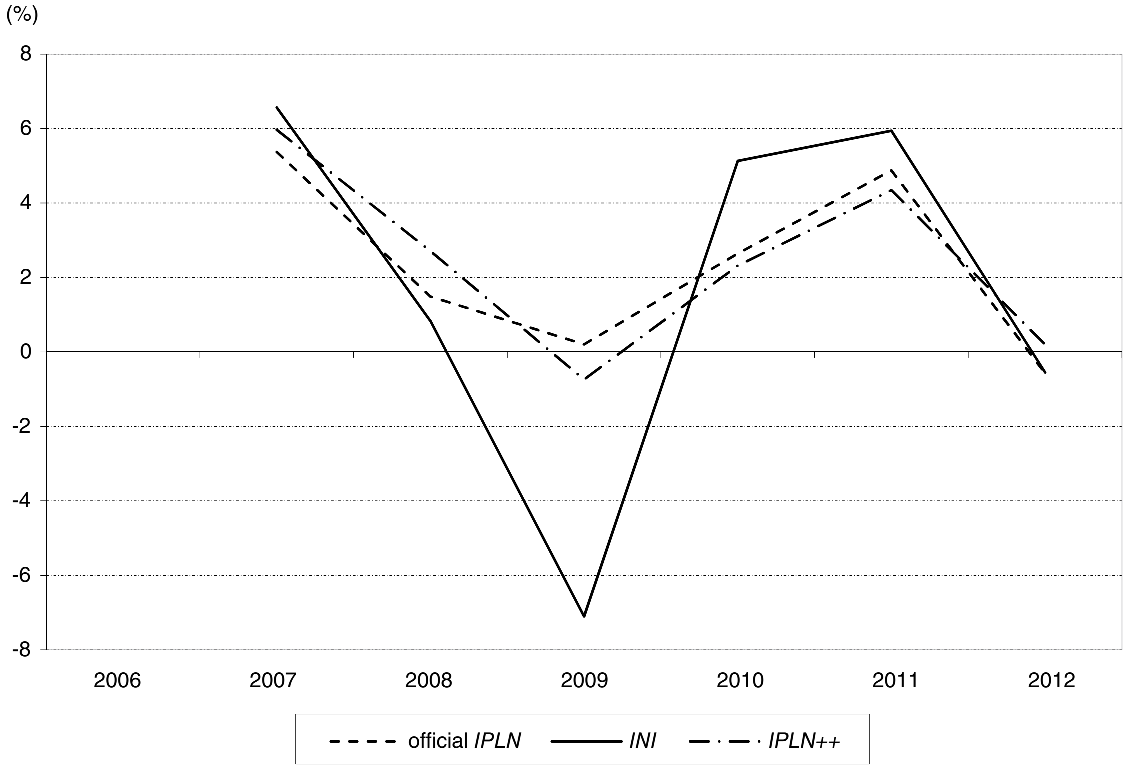
22. *Floor area (*GFA*) replaced the net ground area (*SHON*) on 1st March 2012.*

Figure VI
The *INI*, the official *IPLN* and the *IPLN++* comprehensive price index



Sources: *ECLN*, *BIEN* and *Perval* data bases (see Table 2); authors' computation.

Figure VII
Growth rates of the *INI*, the official *IPLN* and the *IPLN++* comprehensive price index



Sources: *ECLN*, *BIEN* and *Perval* data bases (cf. Table 2); authors' computation.

Other potential explanations: building on land

As new buildings become old, structures quality declines while land quality might stay constant, improve (if, say, new public services, industries or transportation are emerging or developing in the area) or decline (if new sources of pollution, congestion or noise appear, or if industries disappear, or as a result of climate change). Depending on the depreciation rates of the structure and changes in land quality, the value of the "existing" dwellings will decline or increase. Supply and demand play a role, so do improvement and rehabilitation of the structure.

Separating structure from land

We start from the idea[23] that at time t the total value of a new house, V, can be separated into the land value, L, and the value of the building, the structure S:

$$V_t = L_t + S_t \tag{4}$$

Let g_L, g_S, and g_V, denote the percentage change (say between t and $t + 1$) in the land, structure, and overall property values, respectively. With these appreciation rates, the value of the same property at date $t + 1$ can be expressed in two ways:

$$V_{t+1} = V_t(1 + g_v)$$

and

$$V_{t+1} = L_t(1 + g_L) + S_t(1 + g_S),$$

with $g_S<0$ if the structure depreciates over time and the sign of g_L depending on the evolution of land value over time.

$$L_t(1 + g_v) + S_t(1 + g_v) = L_t(1 + g_L) + S_t(1 + g_S)$$

$$L_t(g_v - g_L) + S_t(g_v - g_S) = 0$$

$$g_v(L_t + S_t) = g_S S_t + g_L L_t + g_s L_t - g_s L_t$$

$$g_v(L_t + S_t) = g_S(L_t + S_t) + L_t(g_L - g_s)$$

$$g_v = g_S + (g_L - g_S)(L_t / L_t + S_t)$$

If we define the land share (or land leverage)

$$\alpha_t = L_t / L_t + S_t (0 < \alpha_t < 1),$$

we can write:

$$g_v = \alpha_t g_L + (1 - \alpha_t) g_S \tag{5}$$

From such mechanical decomposition we draw two conjectures H1 and H2.

The land share, i.e. the contribution of land in a home value, increases (decreases) over time as soon as $g_L > g_S$ $(g_S > g_L)$. It is easy to show that:

$$\alpha_{t+1} = L_{t+1} / (L_{t+1} + S_{t+1})$$
$$= L_t(1 + g_L) / [L_t(1 + g_L) + S_t(1 + g_S)]$$
$$\alpha_{t+1} > L_t / (L_t + S_t) \text{ if } g_L > g_S$$

The land share in a new house value increases over time if the structure depreciates. Hence our first conjecture.

H1: In the same location, the land share in the value of second-hand homes will be higher than in the value of new homes if structures depreciate more, or appreciate less than land.

New homes are not located in the same areas as old ones: they are mostly built in the periphery of cities, where land is cheaper. This would predict that:

H1bis: In general the land leverage will be lower for new homes than for existing homes.

For a given net depreciation rate of structure g_S and for the same g_L the land leverage will influence the evolution of the home value.

Equation (5) shows that the change in the house value between t and $t + 1$ is the weighted mean of the evolution of land and that of structure values, where the weights are function of the share of land in the total value of the house in period t. If we differentiate, $dg_v / d\alpha_t = g_L - g_S$, the difference is positive (resp. negative) when $g_L > g_S$ (resp. $g_L < g_S$).

H2: Price indices for second-hand dwellings, with a higher land leverage, would be more volatile than for new dwellings[24]. This seems to be the case at the turning point in 2009 when we compare the *INI* and the *IPLN* indices. Then the shock in demand seems to affect land values more than the value of structures.

Recalling that, for owner-occupiers, the home is both a consumption good and an investment suggests an economic interpretation.

23. *Inspired among others by Bostic* et al. *(2007), Diewert (2011), Davis and Heathcote (2007), Davis and Palumbo (2008), Diewert* et al. *(2015).*
24. *Davis and Palumbo (2008) write: "Volatility is likely an increasing function of land's share in home value."*

Residential land value reflects more the investment dimension of the property, while the value of the structure would reflect more its consumption dimension. In a boom period, and even more in case of a price bubble, land prices would change faster than structure prices. In a recession, the consumption dimension does not change, the number of transactions drops, and the investment dimension, reflected in land prices, declines more as land prices absorb more of the shocks than the structure.

To test these ideas we conduct two types of experiment. The first relies on the different location of new and old homes. The data do not provide the distance to the city centre, only the municipality, but the sheer number of municipalities in France makes it more informative than in a country with smaller numbers of larger units. The first experiment consists in computing a second-hand price index for dwellings situated in the same municipalities as the new ones. Our second experiment will compute separate indices for land and for structure from the *EPTB* data.

The different location of new construction explains only part of the difference

To correct as much as possible for the different locations of new versus second-hand dwellings, when the precise locations (for instance the distance to the city center) are not known, the second-hand transactions observed were re-sampled to mimic the locations of new dwellings. More precisely "geographical clones" of the new dwellings were created by drawing a sub-sample of second-hand transactions in the same municipalities as new constructions. The only difference between such "clones" and the original database is then the municipality distribution of dwellings. Thus, indices calculations carried out with the same methodology (the adjacent two-period time dummy hedonic model) on these two samples would allow assessing if the difference in the municipality location of new and second-hand dwellings can explain the differences between their price evolutions.

For this exercise, houses and flats are separated, hence the *IPLN* was recomputed separating houses and flats. The method is detailed in Appendix 2. We focus here on the annual growth rates of three indices. Two are computed for second-hand houses (or flats): the "clone" and the "time dummy" indices; the third is the *IPLN*, also separating houses and flats (Table 3).

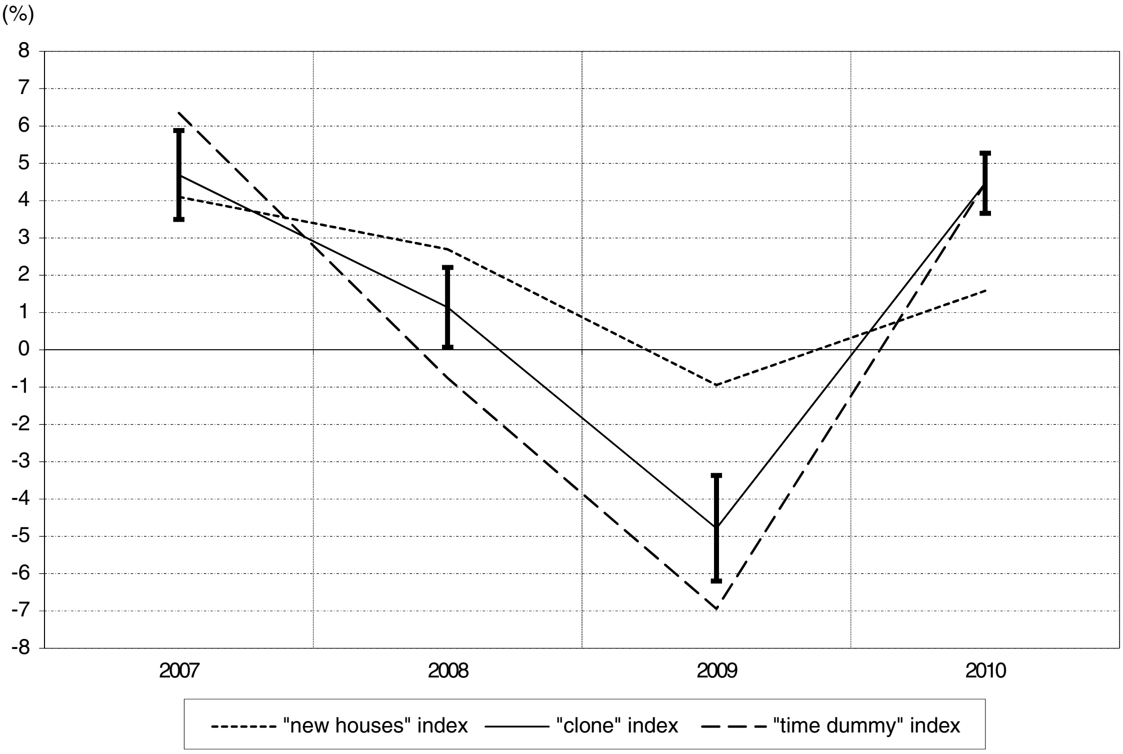
The three annual growth rates together with the two standard deviations confidence interval linked to the annual growth rate of the "clone" index for houses are plotted in Figure VIII. Over the 2006-2010 period, the difference in absolute value between the growth rate of the "geographical clone" price index and that of the new houses is almost 1.6 times lower on average than the gap between the growth rate of the second-hand "time dummy" index and that of the new houses index (2.22 versus 3.65 percentage points). The drop during the crisis period is smaller in the "clone" locations. Thus, the difference in the municipality distributions of new and second-hand houses seems to explain some of the gap between the second-hand "time dummy" index and the new houses index over the period 2006-2010. However, this must be put into perspective, because the average growth rate of the new houses price index (1.86%) is within the two standard deviations confidence interval of the "clone" price index growth rate [0.81% ; 1.95%] and that of the second-hand "time dummy" price index is very close to the lower bound (0.78%).

The results for flats show an even smaller location effect, which seems plausible as new flats are built in denser areas, hence in locations more similar to those of second-hand flats (Appendix 2).

Table 3
The methods and the samples used to compute the three indices

Index	Method	Sample
"Clone" index	Adjacent two-period time dummy hedonic model	Second-hand dwellings (from notaries databases) geographical "clones" of the new dwellings (from the *ECLN* database)
Second-hand "time dummy" index		All second-hand dwellings used to compute the *Indice Notaires-Insee*
New dwellings index (*IPLN*)		New dwellings of the *ECLN* database

Figure VIII
Annual growth rates: new houses, "clone" second-hand houses, and "time dummy" indices, 2006-2010



Reading note: For the "clone" index, the error bars are equal to two standard deviations. The closer the "clone" index to the new houses index, the more the difference in the municipality distributions of the new and second-hand houses can explain the gap between the second-hand houses index ("time dummy" index) and the new houses one.
Sources: *ECLN*, *BIEN* and *Perval* data bases (cf. Table 2); authors' computation.

We conclude that over the 2006-2010 period, the differences in municipality location have a negligible influence in 2007 and 2010 and that, even in 2008 and 2009, they do not fully explain the difference in the evolution of the two indices. It must be underlined that the correction for municipality location ignores the differences of locational distribution within a municipality. This difference is likely to be important because of the location of available land. If this could have been taken it into account, the correction might have been more effective. This leads us to the next step of our investigation.

Decomposing a house price into land price and structure price

Our second experiment draws on the *EPTB* and computes separate indices for land (building plots – *terrains à bâtir*) and structure for new "individually built" single-family houses ("*individuel pur*"). We compare all the indices to a construction cost index, in order to better understand the price dynamics of new homes.

In what was presented above (see section Separating structure from land), the evolution in value was not distinguished from the evolution in price. In other words, we abstracted from potential changes in home quality[25]. If we only consider newly built houses at each date, we can formally write the same suite of equations as above but now g_S is the evolution of the cost of the (same quality) new structure. Equation (5) can be interpreted as giving the evolution of a constant quality new house. With the sign of g_S depending on the evolution of construction cost (for a constant quality house) over time, and that of g_L depending on the evolution of the price of plots for new construction. As above, the price evolution of a new house between t and $t + 1$ is the weighted sum of the price evolution of land and that of the structure. The land share α_t can be written as follows:

$$\alpha_t = (g_v - g_S) / (g_L - g_S) \qquad (6)$$

25. We are very grateful to a referee for pointing that we had overlooked this delicate issue.

It can be computed at each date.

We have to check that $0 < \alpha_t < 1$ for the validity of the computation. It can be easily shown that this condition is met if and only if:

$$g_S < g_v < g_L \qquad g_L < g_v < g_S \qquad (7)$$

We check below (Figure IX) that conditions (7) are met for successive cross sections of new homes.

To study how a house price is decomposed into the price of land and that of the structure we use the *EPTB* survey (see above). We now compute two quality-adjusted indices: a land price index and a structure price index. The adjacent period year dummy method is similar to that of the new "individually built" single-family houses price index (*IPLN(IP)*) computed with model (3). Only location variables and building plots characteristics are used for the land price index; building plot characteristics are dropped from the hedonic model for the structure price index, while structure characteristics are included. As for the "individually built" single-family houses index above, the models are computed at the regional level. The national indices for a year *A* are the weighted average of the 21 regional indices, where the weights are respectively the shares of plot expenses and construction expenses in region *r* in year *A* – 1 (see Appendices 3 and 4). Then the indices are chained to get national annual price indices for land and structure from 2006 to 2012.

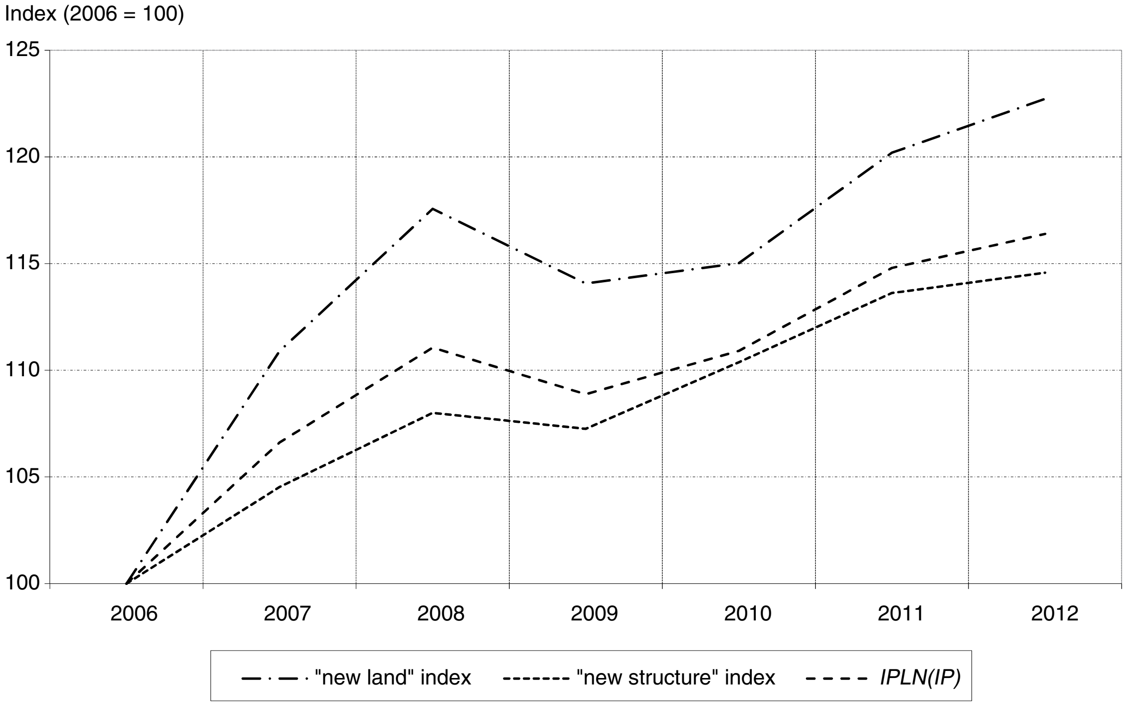
The annual evolutions of the indices for the price of land and for the price of structure are roughly similar over the period (Figures IX and X). However, the changes in land price are more pronounced at each date than in structure price, except in 2010, after the crisis. For instance in 2007, the increase is 10.9% for land, and only 4.5% for structure. In 2009, the structure prices hardly decline (-0.7%) when land prices decline by 3%. In terms of volatility, defined as above as the standard deviation of the annual rates of change in price indices, land prices are 2.5 times more volatile than structure prices (4.34% versus 1.73%). Oikarinen (2013) also found that land prices appeared to be more volatile than construction costs in the Helsinki Metropolitan Area between 1988 and 2008.

Nevertheless, the changes in structure prices appear to be closer than expected to those in land price. It may be because we have included the location characteristics of the house in the hedonic model for structure prices. The reason for this inclusion is that the cost of construction can vary with location, for instance with distance to providers of material. The location variables are proxies for such variation. As a robustness check we removed all location variables from the hedonic models of the structure price index; the decline in the structure price index in 2009 was unchanged (not shown). As Davis and Palumbo (2008) noted in their footnote 18, there is a positive covariance over time between real land prices and construction costs that also affects home prices.

Using (6) we compute the estimated land leverage α_t from the evolution of the prices of land, structure and houses (Figure XI). The land leverage increases before the crisis of 2009 to much more than 50% (+77.4% between 2007 and 2008), which we interpret as a sign of the impending price bubble. In 2007 and 2008 (and probably for years before, that unfortunately we cannot observe) the price evolution of construction is less than that of land ($g_S<g_v<g_L$). Then the crisis hits and the price of residential land declines by 3%. However, the price evolution of construction is also influenced by the crisis, and it also declines in 2009, then rebounds in 2010, with the result of increasing the estimated (i.e. ex post) land share in 2009. At the end of the period, the mean share does not rise much and stays around 31-32% while the estimated α is 36% in 2010, and 43% in 2011.

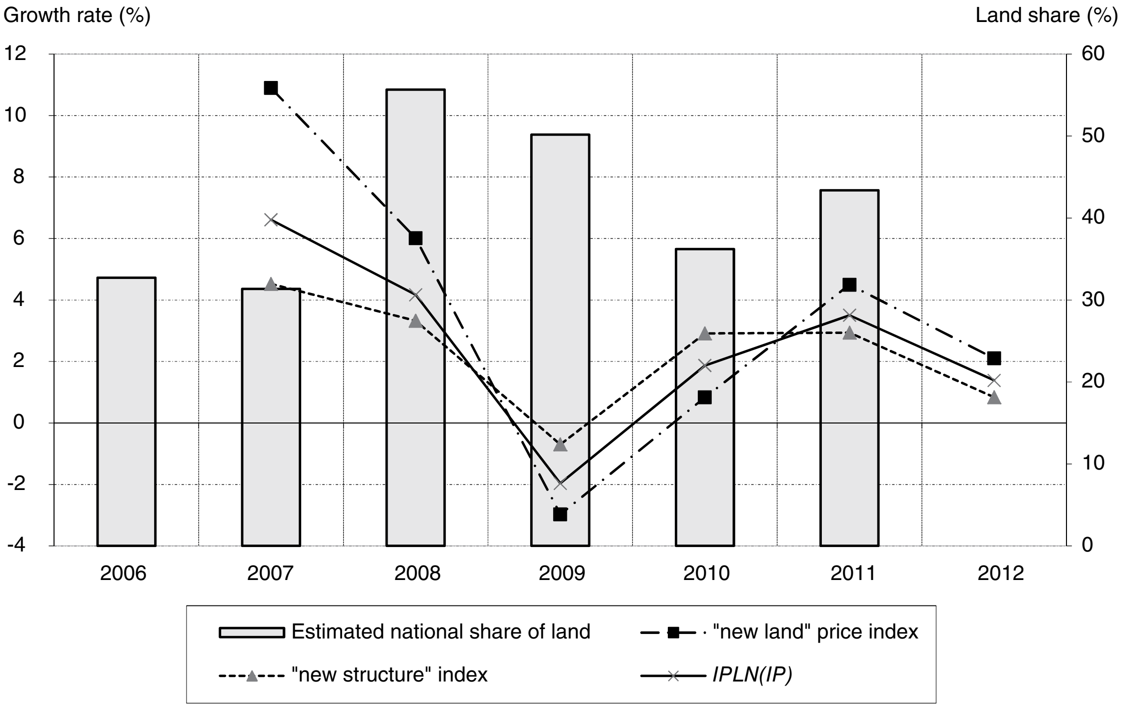
Our method also allows computing the evolution of land share α_t by region, since hedonic models were computed in each region. The peak in land share is strikingly important in Île-de-France, reaching 74% in 2008. It was 56% in Provence-Alpes-Côte d'Azur, 52% in Rhône-Alpes (Figure XI). They are the three richest French regions. The bursting of the price bubble on second-hand homes was more important in those regions: between 2008_Q4 and 2009_Q2 prices dropped by 7.8% in Ile-de-France and Rhône-Alpes, by 7.2% in Provence-Alpes-Côte d'Azur and only by 6.6% in the whole of the Province. By contrast no bubble appears in Basse-Normandie. At the end of the period, in 2011, land share is higher in Provence-Alpes-Côte d'Azur – it kept rising after 2008 – than in Rhônes-Alpes.

Figure IX
"New land" and "new structure" price indices for new "individually built" single-family homes and *IPLN*(*IP*)



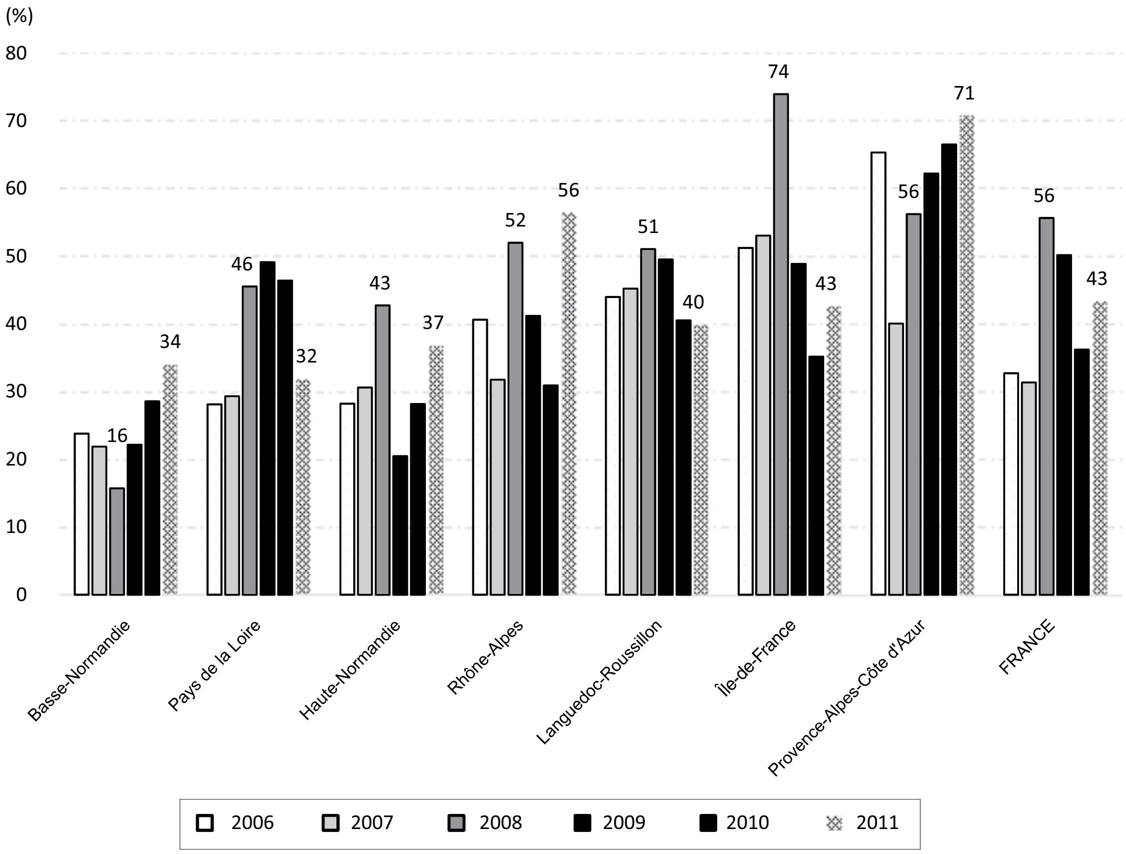
Sources: Authors computation from *EPTB*.

Figure X
Growth rates of "new land" and "new structure" price indices for new "individually built" single-family homes and of *IPLN*(*IP*), and estimated national share of land



Reading note: Land share went as high as 56% in 2008, then dropped by about 20 percentage points in 2010 (right axis).
Sources: Authors' computation from *EPTB*.

Figure XI
Estimated share of land under new houses in selected regions



Sources: Authors' computation from *EPTB*.

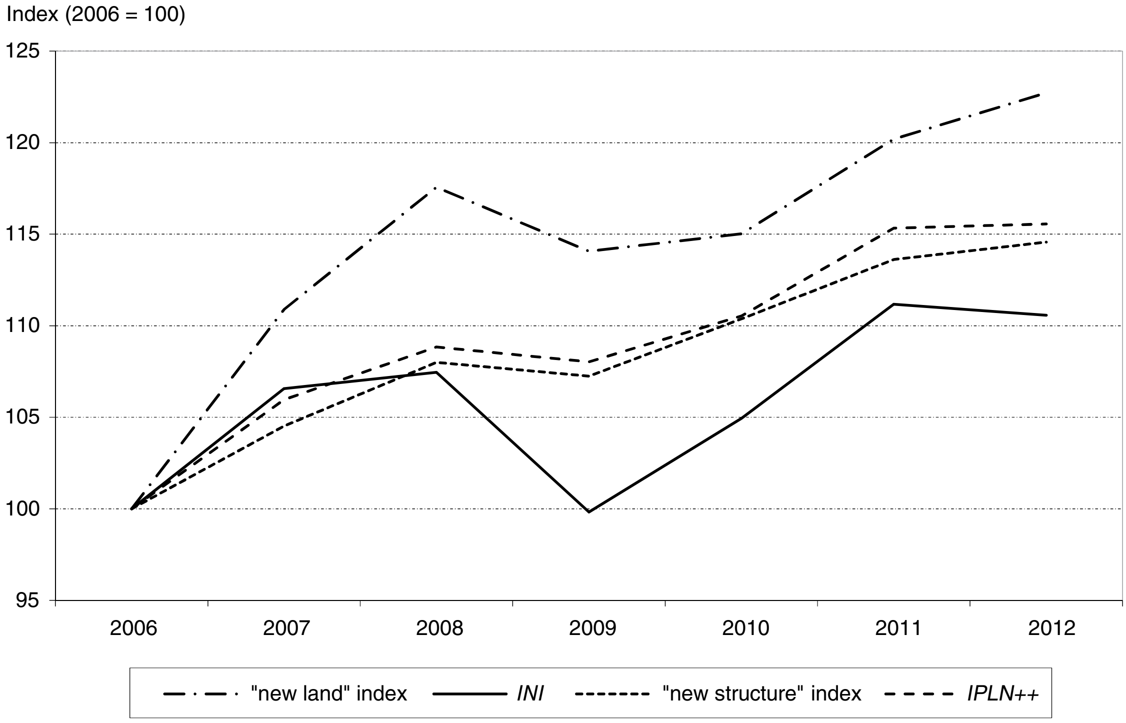
Land price index and the *Indice Notaires-Insee* for existing houses vs structure price index and the comprehensive price index

The four price indices, namely the "new land", "new structure", the *IPLN++* for new dwellings and the *INI* for existing dwellings, have roughly similar evolution over the whole period 2006-2012 (Figure XII). The "new land" price index evolution and that of the *INI* have more marked and synchronized evolutions during the bust period of 2009, respectively -3% and -6% than the structure and comprehensive new dwellings indices (only -1%) (Figure XIII).

What partly drives the evolution of the housing market is the demand for location, i.e. the demand for land. If the land share under existing dwellings is higher than the land share under new dwellings, it could explain why land prices have grown faster than structure prices these last years. As Oikarinen (2013) wrote: "Since land prices appear to be more volatile than construction costs, it is anticipated that greater share of the land value component leads to more volatile housing prices." According to our estimation, the volatility of the price of the land under new single-family houses is more than twice that of the structure itself. This, for us, is part of the explanation of the greater volatility of the price index for second-hand dwellings *(INI)* compared to the *IPLN* (Table 4).

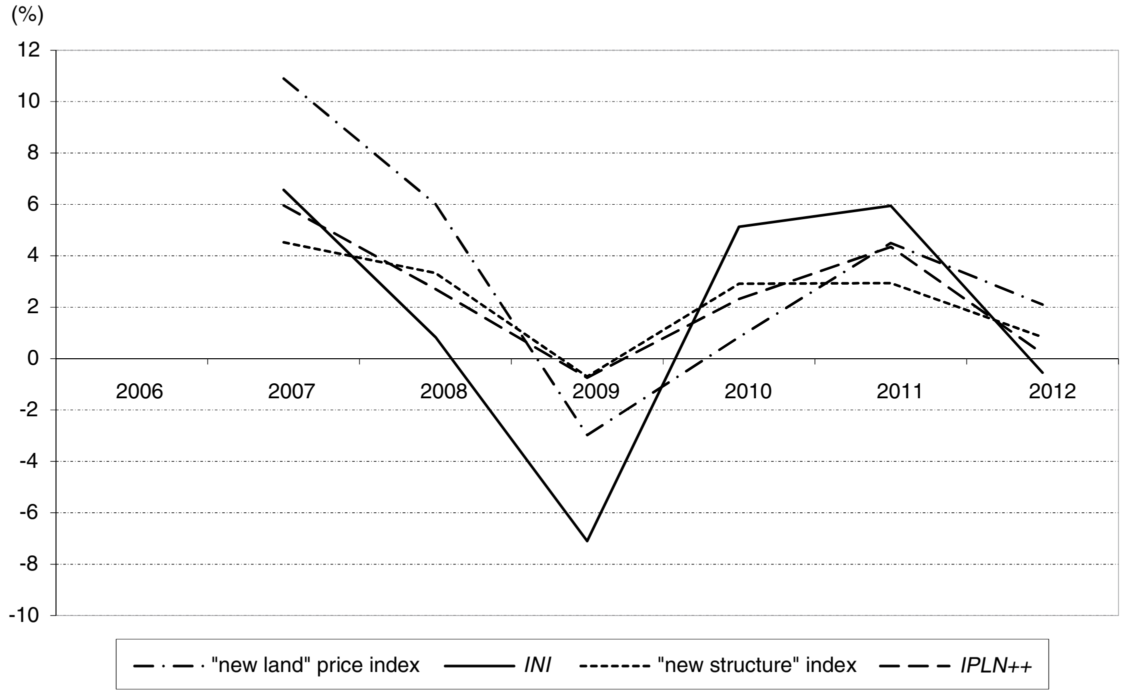
It can also be noticed that when the land price index is around 123 in 2012, the structure price index is around 115. It may seem surprising that the structure price index has increased so much over a period of six years. It could be that productivity did not improve in construction or, as some have argued, that wages, traditionally low in that sector have improved. It is also probable that the quality of homes has improved and that this is not taken into account fully in our hedonic model for lack of detailed information on house characteristics. New stringent norms of construction also play a role. Note however that our structure price index is very similar to the index of

Figure XII
"New land" index, "new structure" index, *IPLN++* and *INI*



Sources: *EPTB*, *ECLN*, *BIEN* and *Perval* data bases (cf. Table 2); authors' computation.

Figure XIII
Growth rates of the "new land" index, "new structure" index, *IPLN++* and *INI*



Sources: *EPTB*, *ECLN*, *BIEN* and *Perval* datsa bases (cf. Table 2); authors' computation.

Table 4
Volatility of the various housing price indices over 2007-2012

Index	Volatility 2007-2012 (in %)
INI	4.77
"New land"	4.34
IPLN(IP)	2.65
IPLN++	2.28
IPLN	2.22
"New structure"	1.73

Note: Volatility is defined as the standard deviation of the annual rates of evolution in price indices.

construction cost (*BT01*), which is supposed to be a constant quality index (see Appendix 5).

Demand and supply of new construction

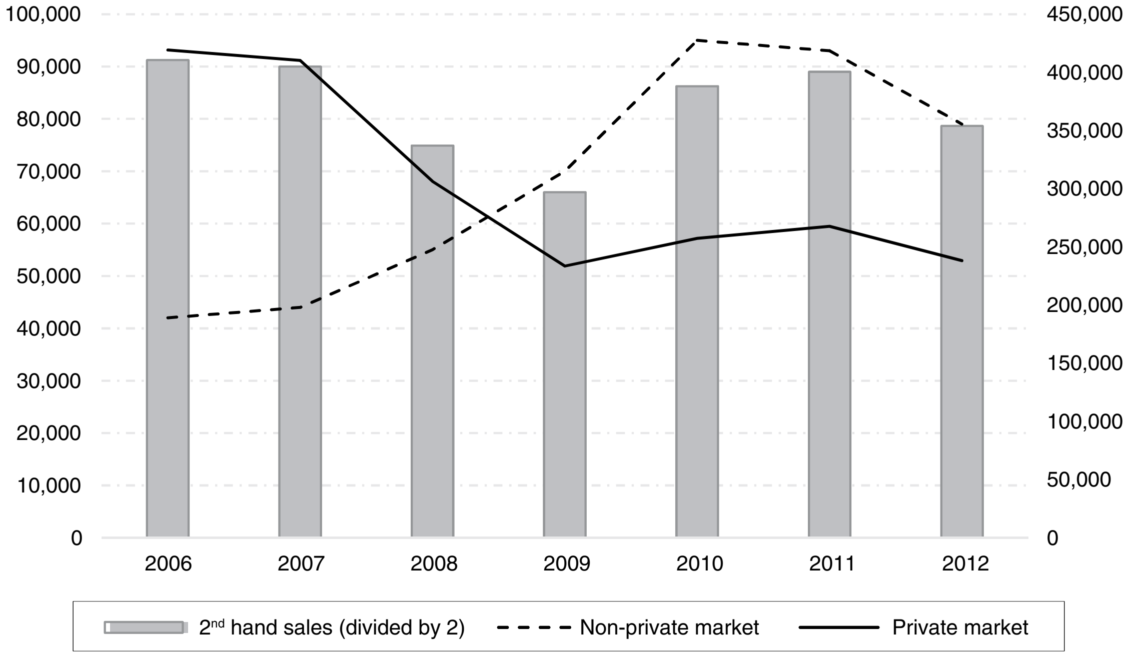
The price index for new dwellings is less volatile than the price index for second-hand dwellings. We did not find any related study on the relative volatility of new versus second-hand housing price indices. Richmond and Roener (2012) about US data wrote just the reverse of what we find: "Fluctuations in the price of new homes are known to be stronger than those of second-hand homes". They do not provide references. Volatility may be linked to the variability in new housing investment. For Topel and Rosen (1988), the short run supply price elasticity of investment in single-family homes is smaller than the long run elasticity, which is as high as 3, and the two elasticities converge quickly. The low volatility in France would mean that the construction cycle is longer and much less reactive to prices than in the US.

We relate the variation in prices to the variation in the number of sales, relying on annual numbers of sales and constructions as quarterly numbers are not available. Figure XIV presents the estimated number of sales of second-hand homes and compares it to the total number of new constructions, separately for private homes (defined as those sold at market price) and non-private homes (social housing, either not sold or sold below market price thanks to various subsidies). It is not straightforward to compute such numbers because social housing is more and more built in "VEFA" (that is, dwellings sold before they are built – *vente en l'état futur d'achèvement*) by private developers who sell them at a discount, either to social housing agencies or to individual low-income buyers. As in Table 1, we combine information from building permits (*logements commencés*) in *Sit@del2* and that of subsidised operations from CDC (2015) on the number of social dwellings produced.

More interesting here than the exact number of homes built are the diverging trends in construction of private and non-private homes. In 2008, the number of sales declined by 17% for second-hand homes when the price evolution started to slow down; the number of homes built for the private market also declined markedly (-24%) but the number of publicly financed new constructions increased by 16% (25% when the "VEFA" sales are included). The prices evolution of new homes slowed down less than second-hand homes. In 2009, the crisis year, when second-hand home prices declined by 9% in annual rhythm in quarter 2, the number of sales declined for second-hand homes (-12%), the number of homes built for the private market also declined (-24%); the number of publicly financed new construction increased significantly (+22%, +27% including "VEFA") and increased even more in 2010 (cf. Figure XIV). Such countercyclical movement in subsidized construction does not compensate the overall decline, since publicly financed housing represents only 18% of new constructions (cf. Table 1). But the rather sustained demand for new homes may have helped maintain their prices: one element of explanation could be that builders, since the SRU law of 2000, must include a certain percentage of subsidized homes in a new project, so the price they can sell a home on the subsidized sector market influences the price on the private market. This could contribute to explain the more moderate decline in new homes prices compared to second-hand homes prices during the crisis[26].

26. Developing this interesting topic is left to future research.

Figure XIV
Number of new constructions built for the private and for the non-private market, and of second-hand sales



Note: "Non-private market" includes housing owned in private-public partnerships (*logements non conventionnés appartenant à une Société d'économie mixte*).
Reading Note: While new constructions of homes for the private market (line, right vertical axis) declined in 2008 and 2009, as did the number of second-hand home sales (bars, right axis - numbers are divided by 2 in order to fit in the picture), the construction for the subsidized sector (line, left vertical axis) increased.
Sources: SOeS, Survey on social rental housing (*Enquête sur le parc locatif social*, EPLS) and Register of rental social housing (*Répertoire des logements locatifs des bailleurs sociaux*, RPLS).

Another potential factor of the evolution of new homes prices is the new fiscal incentive to invest in "buy to rent", called *dispositif Scellier* created in 2009[27]. Contrary to former public policy schemes (known as *Robien*, *Borloo* and *Périssol*, from the names of the person who promoted them) the *Scellier* was focused on new homes, with an explicit aim of promoting construction and increasing housing supply. It ended in 2012. In 2009 the *Scellier* represented two-third of new homes construction (Rapport de la Commission des Finances[28]) and 70% in 2010. However we are not aware of any study of its net effect on housing construction compared to former schemes[29]. These schemes have been shown to have price effects (Bono & Trannoy, 2012), and they may have contributed to the reduced decline in new homes prices in 2009. Concluding on causality is beyond the scope of this paper.

Many buyers of homes are also sellers of their former home. This explains why the demand for housing and prices evolve in the same direction. This suggests another non-exclusive explanation for the more modest evolution of new homes prices compared to second-hand housing prices. Homeowners who want to move have a tendency to wait in periods of price decline because of down-payment constraints or loss aversion (Stein, 1995). This partly explains the huge decline in the number of sales of second-hand dwellings in 2009, in line with the decline in prices. Home builders cannot afford to lower their prices, and the buyers of new homes are more often first-time buyers than homeowners. They are less influenced by the decline in prices of second-hand homes and may have settled on the price at the time the construction was started. Indeed, according to the *Enquête Logement* 2013 (Insee), 61% of those who recently purchased a newly built home[30] were first-time buyers (39 % already owned their home), while they were only 52% among those who bought a second-hand home (and 48 % already homeowners). First-time buyers are more likely to buy newly built homes than buyers

27. We are grateful to a referee for the suggestion.
28. http://www.assemblee-nationale.fr/13/rap-info/i3631.asp. Quoted by Levasseur (2011).
29. Grislain-Letrémy and Trevien (2016) conclude to an absence of effect of rent subsidy to tenants on rental supply.
30. Defined here as a home built less than 5 years ago.

who already owned a home (32.5% versus 29.2%) (Insee, 2017, p. 117). Hence, new home buyers may be less likely than second-hand homes buyers to wait for a period of price decline for their purchase, because they do not suffer from a decline in the price of their own home. For them, the purchase may be a more long-term operation. This may contribute to explain the more moderate reaction of new home prices compared to second-hand homes.

To summarize, during the crisis years (2008-2009) the relative decline in the number of constructions was larger than the decline in the number of sales of second-hand homes, but the price shock was lower. However, among new homes, the number built as social housing has increased, especially during the crisis years. This, together with public policy schemes supporting new home construction and the behaviour of buyers and developers, may have attenuated the demand shock, hence limited the price decline. In other words, the different characteristics of the markets for new and second-hand homes may explain the lower sensitivity of new home prices in downturns, and their lower overall volatility.

* *
*

The starting motive of this paper was to look for reasons why the price evolutions of new and second-hand dwellings would differ, and more precisely, in France, why the official *IPLN* price index for new homes and the *Indice Notaires-Insee* (*INI*) for second-hand homes differ, with the second being more volatile than the first. Two first sources of differences were explored. First, the hedonic *INI* was recomputed using the more simple *IPLN* two-period time-dummy method. The difference accounts for 1 to 2 percentage points of growth rate, less than the difference to be explained (2.4 percentage points over the 2006-2015 period), especially at the turning point of 2009. Second, we extended the scope of the *IPLN* by complementing it with an index for new individually built single-family houses ("*individuel pur*"), which represent around 39% of new constructions for the private market: using data on the construction of single-family houses built individually on plots that have been purchased separately (*EPTB*), we compute an alternative comprehensive price index *IPLN++* including such individually built houses. The new comprehensive index did not differ much from the current *IPLN*.

We then turned to other sources of differences in the markets for second-hand and new homes. New homes are mostly built in the periphery of cities, where land is available and cheaper, while old ones are closer to city centres. Decomposing the value of a house between land and structure shows that the price evolution of a house, new or second-hand, is the weighted sum of the price evolution of land and that of the built structure, where the weights are function of the share of land in the total value of the property. The share of land in a home value, the land leverage, is smaller for newly built homes than for second-hand ones. When the structure depreciates over time, this might influence price evolution. We show that the higher the land leverage, the higher the volatility of the index. This would explain what is observed on French data: the index for second-hand dwellings, with a higher land leverage, is more volatile than the new dwellings index.

The decomposition led to conduct two experiments. We computed a second-hand price index for dwellings situated in the same municipalities as the new ones. The difference between the two indices is somewhat reduced but not eliminated. Then separated price indices for land and structure were computed. Land price and second-hand housing price evolutions appear very similar, and their volatilities are close. Structure prices have more influence on new homes price indices. However construction prices also reacted to the demand shocks during the crisis. An output of the computation provides land shares and their evolution over time. They increased before the crisis, especially in the richest regions where demand was high.

Coming back to the differences in volatility of the two indices, particularly large during the 2008-2009 crisis, time series of sales and constructions suggest that the countercyclical building of social housing might have contributed to support new home prices and explains their lower volatility. More work is clearly to be done on the differences in the markets for new homes and for second-hand homes, to explore why the former might be less reactive than the latter in terms of prices. ☐

BIBLIOGRAPHY

Balcone, T. (2013). Mémoire de master de statistique publique. Miméo, ENSAI.

Balcone, T. (2018). Élaboration d'un indice de prix des logements neufs dans le cadre de la nouvelle réglementation Européenne 93/2013. *Insee – Document de travail* DSDS (à paraître).

Bono, P.-H. & Trannoy, A. (2012). Évaluation de l'impact du dispositif Scellier sur les prix fonciers. *Document de travail*.
https://halshs.archives-ouvertes.fr/halshs-00796188/document

Bostic, R.C., Longhofer, S.D. & Redfearn, C.L. (2007). Land leverage: decomposing home price dynamics. *Real Estate Economics,* 35(2), 183–208.
http://dx.doi.org/10.1111/j.1540-6229.2007.00187.x

CDC (Caisse des Dépôts), 2015. Logements sociaux produits en VEFA : retour d'expérience. *Éclairages, Étude* N° 7.
https://www.prets.caissedesdepots.fr/IMG/pdf/eclairages_numero_7.pdf

Clarenc, P., Côte, J.-F., David, A., Friggit, J., Gallot, P., Gregoir, S., Laferrère, A., Nobre, A., Rougerie, C. & Tauzin, N. (2014). Les indices Notaires-Insee de prix des logements anciens. *Insee Méthode*, 128.
https://www.insee.fr/fr/information/2569926

Davis, M.A. & Heathcote, J. (2007). The Price and Quantity of Residential Land in the United States. *Journal of Monetary Economics*, 54(8), 2595–2620.
https://doi.org/10.1016/j.jmoneco.2007.06.023

Davis, M.A. & Palumbo, M.G. (2008). The Price of Residential Land in Large U.S. Cities. *Journal of Urban Economics*, 63(1), 352–384.
https://doi.org/10.1016/j.jue.2007.02.003

Diewert, W. E. (2011). Alternative Approaches to Measuring House Price Inflation. WP, University of British Columbia.

Diewert, W. E., Haan, J. de & Hendriks, R. (2015). Hedonic Regressions and the Decomposition of a House Price Index into Land and Structure Components. *Econometric Reviews*, 34(1-2), 106–126.
https://doi.org/10.1080/07474938.2014.944791

Friggit, J. (2014). Évaluation du nombre de mutations de logements à titre onéreux et gratuit par statut d'occupation. Rapport du Conseil Général de l'environnement et du Développement durable N° 008187-02.
www.ladocumentationfrancaise.fr/var/storage/rapports-publics/144000669.pdf

Gouriéroux, C. & Laferrère, A. (2009). Managing Hedonic Housing Price Indices: the French Experience. *Journal of Housing Economics*, 18(3), 206–213.

Grislain-Letrémy, C. & Trevien, C. (2016). The Impact of Housing Subsidies on the Rental Sector the French example, Working paper.

Insee (2017). Propriétaires occupants, acquisitions et accession. *Insee Référence – Les conditions de logement en France*, pp. 116-117.
https://www.insee.fr/fr/statistiques/2586377

Levasseur, S. (2011). Dispositif Scellier : un bilan contrasté pour un coût élevé. Ofce, *Le Blog*.
http://www.ofce.sciences-po.fr/blog/dispositif-scellier-un-bilan-contraste-pour-un-cout-eleve/

Oikarinen, E. (2010). An Econometric Examination on the Share of Land Value of Single-family Housing Prices in Helsinki. *Research on Finnish Society,* 3, 7–18.

Richmond, P. & Roehner, B. (2012). The predictable outcome of speculative house price peaks. *Evolutionary and Institutional Economics Review,* 9(1), 125-139.
https://doi.org/10.14441/eier.A2012005

Stein, J. C. (1995). Prices and Trading Volume in the Housing Market: A Model with Down-Payment Effects. *The Quarterly Journal of Economics* 110, 379–406.
https://www.jstor.org/stable/2118444?seq=1#page_scan_tab_contents

Topel, R. & Rosen, S. (1988). Housing Investment in the US. *Journal of Political Economy,* 96(4), 718–740.

APPENDIX 1

WEIGHTS FOR THE NEW "INDIVIDUALLY BUILT" SINGLE-FAMILY HOUSE PRICE INDEX *IPLN* (*IP*)

$w\_V_{A-1}^{r}$ is the share (in %) of the expenses for single-family units in region r in the total expenses of year A-1 :

$$w\_V_{A-1}^{r} = \frac{\sum_{i=1}^{nb\_obs(A-1,r)} p\_V_{i,A-1}}{\sum_{r=1}^{21} \sum_{i=1}^{nb\_obs(A-1,r)} p\_V_{i,A-1}}$$

(In %)

Region	2006-2007	2008	2009	2010	2011	2012	Mean
ÎLE-DE-FRANCE	8.95	7.97	8.12	6.23	3.54	4.43	6.88
CHAMPAGNE-ARDENNE	2.55	2.53	2.74	2.51	1.81	1.67	2.34
PICARDIE	2.52	2.34	2.13	2.90	2.81	3.16	2.63
HAUTE-NORMANDIE	2.84	3.13	2.73	3.32	3.67	3.70	3.17
CENTRE	4.38	4.24	3.96	4.81	5.39	4.56	4.53
BASSE-NORMANDIE	3.10	3.37	3.22	3.22	3.42	3.65	3.30
BOURGOGNE	2.24	2.44	2.01	2.38	2.22	2.45	2.28
NORD-PAS-DE-CALAIS	3.91	4.01	3.46	3.55	3.72	3.92	3.78
LORRAINE	3.04	2.74	2.45	2.66	2.76	2.87	2.79
ALSACE	2.63	2.75	3.10	2.44	1.68	2.00	2.46
FRANCHE-COMTÉ	2.17	2.11	1.99	2.06	2.67	2.34	2.22
PAYS DE LA LOIRE	13.24	12.45	13.96	13.12	12.48	11.26	12.82
BRETAGNE	8.87	8.12	8.19	8.50	9.77	9.91	8.89
POITOU-CHARENTES	4.07	6.76	8.03	6.17	5.05	4.62	5.54
AQUITAINE	6.12	6.15	6.49	6.85	7.86	7.70	6.75
MIDI-PYRÉNÉES	5.86	5.45	5.25	5.57	7.02	6.53	5.93
LIMOUSIN	1.22	1.33	1.19	1.26	1.36	1.37	1.28
RHÔNE-ALPES	9.37	9.94	10.26	9.89	11.09	11.49	10.20
AUVERGNE	2.19	2.21	2.40	2.41	2.44	2.47	2.33
LANGUEDOC-ROUSSILLON	5.22	4.91	4.58	5.59	5.44	5.25	5.17
PROVENCE-ALPES-CÔTE D'AZUR	5.53	5.03	3.75	4.57	3.80	4.67	4.70
All	100.00	100.00	100.00	100.00	100.00	100.00	100.00

Note: 2006 and 2007 have the same weight because 2006 is the first year for which data are available.
Sources: *EPTB*; authors' computation.

CONSTRUCTION OF THE GEOGRAPHICAL "CLONE" INDICES FOR THE SECOND-HAND DWELLINGS

The clone database is built up in the following way. The number of new dwellings transactions in *ECLN* and the number of second-hand dwellings transactions in the Notaries' databases are computed for each triplet 'year (instead of quarter to have enough observations) – type of dwelling (houses, only "*Individuel groupé*" in ECLN, or flats) – municipality (arrondissement for Paris, Marseille and Lyon)' over the 2006_Q1-2010_Q4 period.

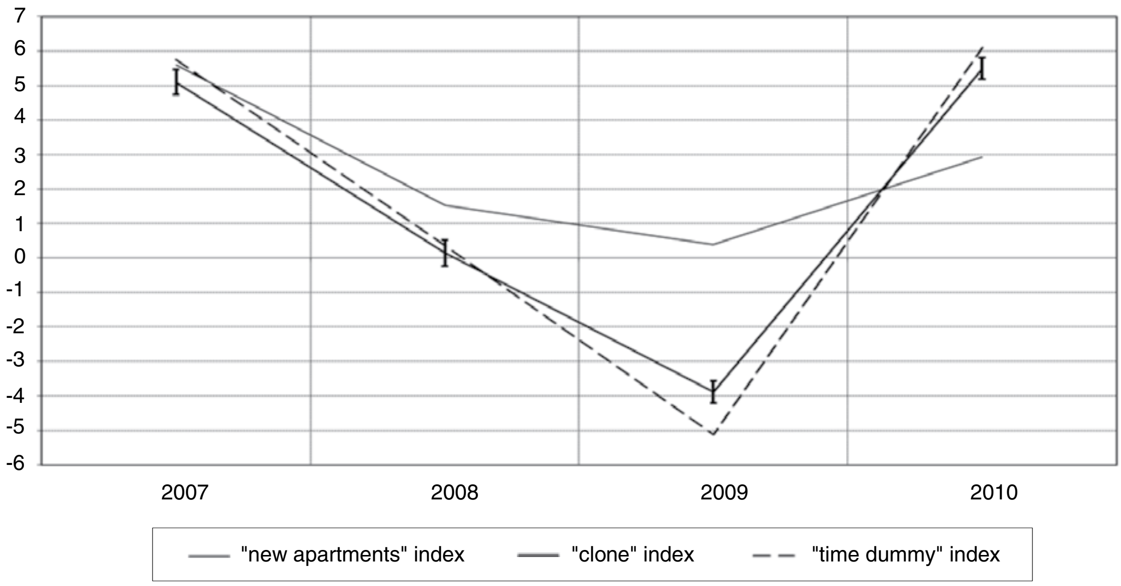
Only triplets for which sales of both new and second-hand dwellings exist are kept. We thus miss 6% of the new dwellings sold during the period, i.e. for them no "clone" second-hand dwelling is sold. To check the effect of the selection of the non-missing triplets, we re-compute the index leaving out those 6%. The annual index does not differ by more than 0.6 index point over the period. Then, for each triplet, a sample of "clones" of new dwellings among the corresponding second-hand dwellings is randomly drawn (with replacement). This "clone" population has the same municipality distribution as the new dwellings. For each triplet (year; type of dwelling; municipality), the number of "clones" is equal to the number of new dwellings. An index for the "clones" population is computed over the period 2006-2010 using the same adjacent two-period time dummies hedonic model as that used for new dwellings. To get a better estimation of the annual growth rate of the "clone" index, 50 different samples of "clones" dwellings are drawn. Thus, we focus on the mean and the standard deviation of the annual growth rates of 50 "clone" indices. However, to simplify, we speak of the "clone" index. Two indices are computed: one for "clones" houses, another for "clones" flats. The year is now the elementary time level.

We concentrate, for houses and flats, on the annual growth rates of the three following indices: the "clone", the "time dummy" and the new dwellings indices (recomputed separately for houses and flats) (see table 3).

For flats, the "clone" index growth rate is very close to that of the second-hand "time dummy" index. The difference in absolute value between the annual growth rates of these two indices is on average less than 0.70 percentage points over the 2006-2010 period (0.68 percentage points, Figure A2.I). If we put aside year 2009, the gap falls below 0.50 percentage points, even if, again, the crisis year is less marked in the "clones" locations.

Figure A2.I
Annual growth rates: new flats, "clone" second-hand flats and "time dummy" indices, 2006-2010

(In %)



Note: For the "clone" index, the error bars are equal to two standard deviations.
Reading note: The closer the "clone" index to the new flats index, the more the difference in the municipality distributions of the new and second-hand flats can explain the gap between the second-hand flats index ("time dummy" index) and the new flats index.
Sources: *ECLN*, *BIEN* and *Perval* data bases (see Table 2); authors' computation.

APPENDIX 3

WEIGHTS FOR THE RESIDENTIAL LAND PRICE INDEX

The weights $w\_L^r_{A-1}$ are the share (in %) of the expenses for plots in region r in the total expenses of year $A - 1$:

$$w\_L^r_{A-1} = \frac{\sum_{i=1}^{nb\_obs(A-1,r)} p\_L_{i,A-1}}{\sum_{r=1}^{21} \sum_{i=1}^{nb\_obs(A-1,r)} p\_L_{i,A-1}}$$

(In %)

Region	2006-2007	2008	2009	2010	2011	2012	Mean
ÎLE-DE-FRANCE	13.64	11.94	11.74	9.36	5.09	6.38	10.26
CHAMPAGNE-ARDENNE	1.93	2.02	2.20	2.11	1.53	1.46	1.88
PICARDIE	2.23	2.16	1.93	2.69	2.71	3.08	2.43
HAUTE-NORMANDIE	2.42	2.95	2.69	3.22	3.70	3.56	2.99
CENTRE	3.82	3.79	3.52	4.38	5.09	4.18	4.09
BASSE-NORMANDIE	2.17	2.43	2.34	2.47	2.74	3.02	2.48
BOURGOGNE	1.67	1.84	1.60	1.88	1.88	2.07	1.80
NORD-PAS-DE-CALAIS	3.67	3.90	3.32	3.50	3.62	3.82	3.64
LORRAINE	2.33	2.00	1.88	2.12	2.27	2.39	2.19
ALSACE	2.47	2.60	3.04	2.33	1.64	1.95	2.36
FRANCHE-COMTÉ	1.44	1.41	1.41	1.52	2.00	1.77	1.57
PAYS DE LA LOIRE	11.71	11.13	12.87	11.80	11.24	10.08	11.51
BRETAGNE	7.06	6.64	6.88	7.14	8.14	8.21	7.30
POITOU-CHARENTES	3.08	5.22	6.20	4.70	4.03	3.76	4.30
AQUITAINE	6.01	6.14	6.51	7.04	8.39	8.11	6.89
MIDI-PYRÉNÉES	5.81	5.37	5.03	5.41	7.15	6.40	5.85
LIMOUSIN	0.64	0.73	0.68	0.76	0.90	0.87	0.75
RHÔNE-ALPES	11.19	12.28	12.60	11.93	13.17	13.51	12.27
AUVERGNE	1.47	1.54	1.70	1.72	1.91	1.85	1.66
LANGUEDOC-ROUSSILLON	6.93	6.47	6.23	7.32	7.13	6.61	6.80
PROVENCE-ALPES-CÔTE D'AZUR	8.31	7.42	5.64	6.62	5.67	6.90	6.98
All	100.00	100.00	100.00	100.00	100.00	100.00	100.00

Note: 2006 and 2007 have the same weight because 2006 is the first year for which data are available.
Sources: *EPTB*; authors' computation.

APPENDIX 4

WEIGHTS FOR THE STRUCTURE PRICE INDEX

The weights $w\_S_{A-1}^{r}$ are the share (in %) of the construction expenses in region r in the total expenses of year A – 1 :

$$w\_S_{A-1}^{r} = \frac{\sum_{i=1}^{nb\_obs(A-1,r)} p\_S_{i,A-1}}{\sum_{r=1}^{21} \sum_{i=1}^{nb\_obs(A-1,r)} p\_S_{i,A-1}}$$

(In %)

Region	2006-2007	2008	2009	2010	2011	2012	Mean
ÎLE-DE-FRANCE	6.70	6.00	6.29	4.70	2.81	3.48	5.24
CHAMPAGNE-ARDENNE	2.85	2.78	3.01	2.70	1.94	1.77	2.56
PICARDIE	2.66	2.43	2.23	3.00	2.86	3.19	2.72
HAUTE-NORMANDIE	3.04	3.22	2.74	3.37	3.66	3.76	3.26
CENTRE	4.64	4.46	4.17	5.02	5.53	4.74	4.74
BASSE-NORMANDIE	3.55	3.84	3.66	3.59	3.74	3.95	3.70
BOURGOGNE	2.51	2.74	2.22	2.63	2.38	2.63	2.52
NORD-PAS-DE-CALAIS	4.02	4.06	3.53	3.58	3.76	3.97	3.85
LORRAINE	3.38	3.11	2.74	2.92	2.98	3.11	3.09
ALSACE	2.70	2.82	3.13	2.50	1.70	2.03	2.51
FRANCHE-COMTÉ	2.52	2.46	2.29	2.32	2.98	2.62	2.53
PAYS DE LA LOIRE	13.98	13.10	14.51	13.76	13.07	11.83	13.46
BRETAGNE	9.73	8.85	8.85	9.17	10.54	10.74	9.66
POITOU-CHARENTES	4.55	7.52	8.95	6.89	5.53	5.03	6.15
AQUITAINE	6.17	6.15	6.47	6.75	7.62	7.49	6.69
MIDI-PYRÉNÉES	5.88	5.49	5.36	5.65	6.96	6.59	5.97
LIMOUSIN	1.50	1.63	1.45	1.51	1.58	1.61	1.54
RHÔNE-ALPES	8.50	8.79	9.09	8.90	10.11	10.50	9.20
AUVERGNE	2.54	2.55	2.76	2.74	2.69	2.78	2.65
LANGUEDOC-ROUSSILLON	4.39	4.14	3.75	4.74	4.64	4.58	4.38
PROVENCE-ALPES-CÔTE D'AZUR	4.19	3.85	2.79	3.57	2.92	3.58	3.58
All	100.00	100.00	100.00	100.00	100.00	100.00	100.00

Note: 2006 and 2007 have the same weight because 2006 is the first year for which data are available.
Sources: *EPTB*; authors' computation.

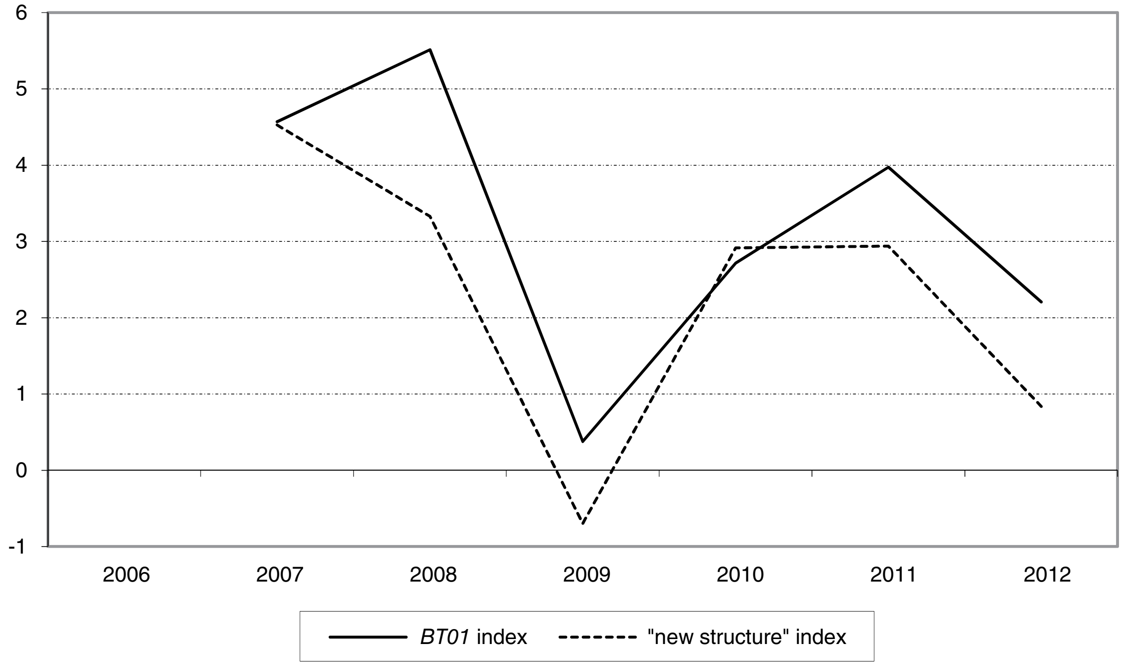
APPENDIX 5

COMPARISON OF THE "NEW STRUCTURE" PRICE INDEX AND THE CONSTRUCTION COST INDEX

To check the validity of our "new structure" price index we compare it to the *BT01* index of construction cost (*Indice national du bâtiment, tous corps d'état*). The two profiles are strikingly similar (Figure A5). The rates of evolution differ by as much as 2 percentage points in 2008, but less than 1 pct point in the other years. Even with a far from prefect hedonic model, we seem to recover a plausible structure price index from the *EPTB* survey covering only single-family homes. What drives the construction costs evolution is left for future research.

Figure A5
Growth rates of the "new structure" price index and the *BT01* index

(In %)



Sources: Insee; authors' computation from *EPTB*.