NEW OR OLD, WHY WOULD HOUSING PRICE INDICES DIFFER?

Thomas Balcone[1] (*), Anne Laferrère[23] (**)

(*) Insee, Methodology, Statistical Coordination and International Relations Directorate
(**) Insee, Household Resources, and Living Conditions Department - CREST

Version 1 June 30, 2015

Résumé

Un indice de prix des logements est censé suivre leur valeur « à qualité constante ». En pratique, il suit la valeur d'un parc de logements ou d'un logement fictif représentatif d'un parc, estimée à partir des prix de transactions sur un territoire donné. On calcule ainsi au moyen de modèles dit hédoniques des indices de prix de logements anciens, et, plus récemment en France, des indices pour les logements neufs. Par ailleurs une bonne partie de la qualité d'un logement est sa localisation. On peut donc aussi vouloir comparer le prix d'un même logement dans des territoires différents. Face à ces différents indices nous posons trois questions principales. Pourquoi les évolutions des indices temporels de prix des logements neufs et anciens différeraient-elles ? Quelle est la bonne échelle géographique pour calculer un indice de prix des logements ? Qu'est-ce que la qualité d'un logement ? Ceci nous amène à réfléchir à la décomposition du prix d'un logement entre construction et terrain. Utilisant des sources de données très différentes nous proposons une première exploration empirique de ces questions, en particulier des effets des différences de méthode de calcul des indices. Nous proposons aussi quelques pistes de réflexions théoriques.

Abstract

An index of housing prices is supposed to follow the value of a "constant quality" dwelling. In practice, it follows the value of the housing stock or a representative fictional dwelling, estimated from transaction prices in a given territory. This is done using hedonic models. In France, it has been done for second hand dwellings since 1996, and more recently for new dwellings. A large part of the quality of a property is its location. One may want to compare the price of a dwelling in different locations. Faced with these various indices we ask three main questions. Why would price indexes for new and existing dwellings differ? What is the right geographical scale to calculate a price index for housing? What makes the quality of housing? This leads us to reflect on the decomposition of the price of a dwelling between construction and land. Using very different data sources we propose a first empirical exploration of these issues, particularly the effects of differences in methodology for computing indices. We also propose new theoretical reflections.

Mots-clés
Indices, économétrie, modèles hédoniques, économétrie spatiale

[1] thomas.balcone@insee.fr
[2] anne.laferrere@insee.fr
[3] We are grateful to Ronan le Saout and Benjamin Vignolles for helpful discussions. The latter also provided his computation of the distance variables in EPTP. We are also grateful to Cédric Cailly for his help with the EPTB data and to INSEE and SOES (Service Statistique du ministère de l'Ecologie, du Développement Durable et de l'Energie) for providing access to restricted data (convention n°2014065NF). We alone are responsible for errors.

1. Introduction

There is a wide consensus that differences in wages and amenities lead to differences in land prices across cities (Roback, 1982; Glaeser et al., 2000). Hence, what makes the value of a dwelling is for a large part its location. A "similar" apartment is more expensive in Paris than in Lyon or Munich; in London than in Paris.[4] Within a city, the price gradient is predicted to be negative as one lives further from the city center (Fujita, 1980). This is verified on French data (see Gofette-Nagot, 2009; Combes et al. 2011). Static *spatial* comparisons between cities are made from the computation of geographic or spatial housing price or rent "indices". They are usually computed once a year and make sure best sellers for magazines because they bring buyers and sellers information on local markets, even if often in the form of crude means or medians. On the other hand, various statistical institutes compute housing price indices to measure how the value of a typical "constant quality" home evolves over *time.* Many such indices are for second-hand dwellings, and this for obvious reasons. Their price is more often and easily observed than for new dwellings. For second-hand dwellings, either the property is put to sale by real estate agents who advertise the price, or when the transaction actually occurs, a deed of property is registered, or a transaction tax is paid and the price is known and registered.[5] For new dwellings the observation of prices is made more complicated by the process of construction. Part of new constructions are individually built on a piece of land that can be purchased or inherited, sometimes long before the house is built.[6] Recording the final price of the house is tricky. Even if the house is part of a development and sold by the developer, the final price is not always recorded.[7]Finally, the sheer number of observations differs: in France, for one new dwelling being built and sold, around six second-hand homes are sold. Hence, even in the best of worlds, six times less observations are available for new than for old dwellings.[8]

Nevertheless, price indices for new dwellings are beginning to be regularly computed, as part of new discussions on what should or should not be included in the consumer price index or for the needs of National accountants[9]. Such statistical development leads to new questions, or to the reformulation of old questions. What makes the quality of a dwelling, and why would the price evolution of new

[4] Laferrère (2012) computed from a hedonic model on notaries data that, compared to Creuse, the cheapest French department, the same house was 3 times as expensive in Rhône, 4.9 times in Alpes-Maritimes, and 7.3 times in Hauts de Seine.
[5] Another reason is that when the repeat sales method is used, it gives by definition a price index of second-hand dwellings.
[6] Among those who purchased a new home between 1988 and 2002, 8.9% had built on inherited land (Laferrère, 2007).
[7] Sales of new dwellings may be recorded for tax reason. For instance, in France new homes are submitted to VAT and not to stamp duty; the deed of property is registered by the tax authorities but not the price and the characteristics.
[8] Over the period 2000 to 2010, 800 000 homes were sold each year (Friggit, 2014). During the same period, an average of 368 000 homes per year were built (http://developpement-durable.bsocom.fr/Statistiques/TableViewer/tableView.aspx). Among them, 44.4% are pure individual houses where no price is usually registered. Among the rest, some 60 000 to 80 000 are social public rental dwellings that are not sold. Hence, the real ratio is likely be one observed new home price for 6 observed second-hand home price. 210 000 homes were given or transmitted by bequest from a parent to a child every year (Friggit, 2014). In case of bequest or *inter-vivos* transfers, a value is registered but not centrally available for statistics.
[9] For Europe this is done through the O*wner-Occupied Housing project*.

and old dwellings differ? Can spatial and time indices be reconciled? What is the proper geographical scale for computing such indices? Let us tackle briefly each question in turn.

What is the proper geographical scale for computing such indices? In practice, in France, the price index for new homes is computed only at the national level when the price indices of second-hand dwellings are more disaggregated.[10] This is linked to the smaller number of observations available for new homes. In principle, the more disaggregated the better, in order to have homogenous markets, if supply and demand differ by location. Also only on fairly local market can a consumer be assumed to make comparisons and substitutions between dwellings. Then studies of consumer choices need local indices. National accountants rely on aggregated indices, hence the issue of how to aggregate the sub-indices.[11]

Can spatial and time indices be reconciled? The problem is linked to aggregation and transitivity question as has been explained by Hill (2004). We leave aside this question here as *spatial* price comparisons are not our main purpose, even if related to our last question below about what makes the quality of a dwelling. We suggest below that the term “index” without further qualification for spatial price comparisons may be a misnomer.

Why would the price evolution of new and old dwellings differ? The literature on this question is not abundant.[12] Bostic et al. (2007) relate price appreciation to land leverage, i.e. the share of land in the home value. They test whether house price appreciation and house price volatility are related directly to land leverage. Then new houses would rise at a *lesser* rate than second-hand houses if the share of land under them is smaller. The prediction is similar in Davis, Heathcote and Palumbo or Davis and Palumbo (2008).

The price evolution of old dwellings is sometimes akin to that of a price index of a stock or basket of existing dwellings whose locations are fixed, and whose current prices are estimated by hedonic methods, hence a Laspeyres formula. Sometimes the “basket” is made of some of the current transactions and so changes regularly (e.g. when the hedonic model estimated on two adjacent periods includes a time dummy).[13] The price evolution of new dwellings is by definition the price of a flow of ever changing new dwellings.

What makes the quality of a dwelling? It is a unique combination of characteristics of land and structure. Taking care of the physical characteristics of the structure, and defining a “reference

[10] In other countries, it seems also the case that price indices for new dwellings are less disaggregated than for second hand ones.

[11] Linneman (1980) argued for the existence of a national housing market in the US because of the high mobility of suppliers and demanders.

[12] The only indirect statement we found is the following, on volatility: “Fluctuations in the price of new homes are known to be stronger than those of second-hand homes” (Richmond and Roehner, 2012). No reference was provided.

[13] We leave aside the issue of price measurement. Dipasquale and Somerville (1995) report that price indices based on transaction prices and on owner-reported values have similar time-series patterns (even if the levels are different) but can differ at market turning points. Some indices rely on assessed values, not on observed transaction prices.

dwelling” of “constant quality” for making price comparisons may seem rather straightforward.[14] But taking care of location characteristics and quality both at a given date, and over time is less so. The well-known problem that all the combinations of characteristics are not available, is exacerbated when considering location. Each location is unique. In a world where new dwellings are built at the previously non-built outer ring of cities, there is no way a new and a second-hand dwelling cab be compared, except maybe with a continuity assumption and by taking into account the effect of amenities and distance to the city center, or if, at the same location, both a new dwelling is built and an old dwelling is sold… But is there something as a “constant quality” new dwelling? The question is similar to asking what makes the quality of land. In a given city, of fixed size and population, if all dwellings were put to sale at each date (i.e. prices are known), and with no mobility costs, people choose their dwelling so that they are indifferent between the various bundles of characteristics (the substitutability we mentioned above takes place) and differences in prices reflects differences in the valuation of characteristics. Then at a given date a “constant quality” dwelling has exactly the same price everywhere within the city, because it provides the same utility. In such a city the spatial “pure price” index is flat. Which is perfectly compatible with a distance to city center gradient, and amenities gradients. In that sense spatial indexes are not “constant quality” indexes. Various hedonic models will be used depending on the type of spatial price comparison to be conducted. Cheshire and Sheppard (1995) underline that the price of residential land reflects both accessibility (the land as “pure-space-with-accessibility”) and amenities. They write that “many land prices can be built” depending on the variety of amenities that are taken into account. Depending on the purpose of the analysis a price inclusive of local amenities, or a price of pure space will be chosen.

This paper revolves mostly around question 3 (why would the price evolution of new and old dwellings differ?). It succinctly presents the current methods used in France to compute new and existing dwellings price indices (section 2). Then, we rule out two possible sources of differences between the evolutions of these indices. First the difference in method (section 3.1). Then, the difference in locations of new versus old dwellings (section 3.2). The difference in evolution between indexes is still far from fully explained.

Section 4 offers some thoughts on theoretical models linking the price of new and existing dwellings, and adds some dynamics. This suggests comparing not only the price indices of new and old dwellings, but also a price index for constructible land, and the construction costs indices for dwellings (section 5). For this, we use a survey providing both the purchase price and size of land plots, and the estimated value of the construction of individually built single family units. The data are used to build a land price index, a structure price index, and a new single-family houses price

[14] At least superficially. In old countries, where dwellings have been built over time, with very different materials and techniques, varying architectural beauty, the number of characteristics can increase, and become intractable.

index (section 5.1.). This new individually built house price index is compared to the one in section 2.2. covering new dwellings built as part of developments (section 5.2) and we aggregate the two new house price indices in section 5.3. Sections 5.4 and 5.5 compare our land index with the notaires-Insee index, and our structure index with the official construction cost index. Section 6 concludes.

2. Current methods of computation of new and existing dwellings price indices in France

2.1. Existing dwellings: the Notaires-Insee index

The existing dwellings index (or Notaires-Insee index) is currently computed by using a hedonic method based on estimation of disaggregated models in homogeneous zones, run separately for houses (181 zones) and apartments (112 zones). The data are 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 of the 293 zones. The reference basket is made of two years of transactions; the reference basket is updated every two years. In principle, all transactions are included and the price is the actual transaction price. The basic model is the following (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 m^2 for apartments) of dwelling *i*, $Y_{a,i}$ is a dummy for the year of sale of dwelling *i*, $M_{t,i}$ a dummy for the month of sale of dwelling *i*, $X_{k,i}$ are various characteristics of dwelling *i*, indexed by *k=1* à *K,* computed from the modalities of initial characteristics. They are physical characteristics (size, number of rooms, of bathrooms, of levels, age, plot size for houses, etc…) and location characteristics proxying for amenities (neighborhood dummies). Estimating the model gives an estimated price for each characteristic, referring to a « reference dwelling » (its characteristics are the omitted categories). p_0 represents the price of the reference dwelling. 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.[15] The ratio of the estimated values of the reference baskets between *t* and *t-1*, provides the index. They are then chained from period to period.

The sub-indices by zones and type of dwellings are aggregated by geometric means when the geographic level is infra-département, that is where the consumer is assumed to make residential choices, and by arithmetic means at higher geographic levels, with weights reflecting transaction

[15] For instance the model estimated over the period 2009-2010, was used to compute the indices for the period 2012-2013.

values. Note that 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.

2.2. New dwellings index

The new dwellings index is based on an adjacent periods time dummy variables model. Each quarter, the model is estimated on two successive quarters of data. The data source is ECLN[16], a survey on the commercialisation of new dwellings. This survey covers a subset of new constructions: all building permits of 5 dwellings or more, called “individuel groupé” (for single family units that are part of a development) or “collectif” (for multi-family units), built by developers, sold to private persons and for which a transaction price is recorded. “Individually built” homes, i.e. single family units built one by one, called “individuel pur” are left out. According to estimations for 2013, “individuel groupé” is about 13.3% of all new built homes[17] and “collectif” is 54.3% (including 2.1% residences).[18]

The data are not available at the dwelling level, but at the construction program level. Each quarter, the following information are provided, by type of program (houses or apartments)[19]for each category of number of rooms: the total number of dwellings, the average living surface in m² and the average price of the sold dwellings. Hence, only the mean characteristics of the dwellings of a program are available. For similar houses, it is not a big issue. For apartments of a given number of rooms, some important characteristics, such as the level in the building, are unknown.

The hedonic model is the following :

$$\ln(\overline{p}_{i,t}) = \alpha + \beta_{shab}\ln(shab\_moy_{i,t}) + \sum_{k=1}^{K}\beta_k I_{i,t,k} + \delta_{t_1+1}D_{i,t,t_1+1} + \varepsilon_{i,t} \qquad (2)$$

where $\overline{\mathrm{p}}_{\mathrm{i,t}}$ is the average price of dwellings of program[20] i in quarter t, $\mathrm{shab\_moy}_{\mathrm{i,t}}$ is the average floor space in m² of dwellings of program i in quarter t, $I_{i,t,k}$ is a vector of the characteristics of the sold dwellings, $D_{i,t,t_1+1} = \begin{cases} 1\ if\ \ t = t_1 + 1 \\ 0\ otherwise \end{cases}$.

[16] « Enquête sur la commercialisation des logements neufs » conducted by the Ministry for Ecology, Sustainable Development and Energy.

[17] Equivalent-logement.

[18] Figure 16.2 p. 87, in Rapport de la Commission des Compte du Logement, 2013. Residences can be nursing homes, students accommodations, or other sheltered housing. Social housing construction is also excluded. Excluding “individuel pur” (32.4% of new surface) is consistent with the definition of the European new dwellings index (see Owner-Occupied Housing regulation (Commission regulation (EU) No 93/2013)).

[19] And by kind of dwellings: private ordinary dwelling or residences offering specific services (see footnote 18).

[20] More exactly is defined by a program and the number of rooms.

The characteristics of the property are its type (house or apartment), physical characteristics (floor space, number of rooms), the characteristics of the municipality (size of the urban unit, ZEAT[21], sea, ski or hiking resorts, etc.). Geographical location is also taken into account through 14 zone dummies corresponding to geographical areas that are homogeneous in terms of selling price per square meter.

The time dummy serves to adjust the average price change between the two quarters, all else being equal (i.e. at constant dwelling characteristics). The quarterly change in the index is provided by the exponential of the coefficient associated with the quarter dummy. Contrary to the existing dwellings, there is only **one model** for the whole of metropolitan France (excluding Corsica) and for all new dwellings (houses and apartments are not separated) because of the small number of quarterly observations in ECLN[22]. Balcone (2013, 2015) describes the method in details. The main differences between the new and existing dwellings price indices are summed up in table 1.

Table 1. Main differences between the new and existing dwellings indices

	New dwellings	Existing dwellings
Data	Survey on the commercialising of new dwellings	Transactions registered in the Notaries' databases
Method	Adjacent periods time dummy hedonic model	Hedonic "re-pricing" method
Geographical effect	Dummy variables in the model (only one hedonic model)	One hedonic model for each 293 area + neighbourhood dummies

By construction this new dwelling price index does not vary with location, and is the same for houses and apartments.

2.3. Comparing price indices of new and existing dwellings

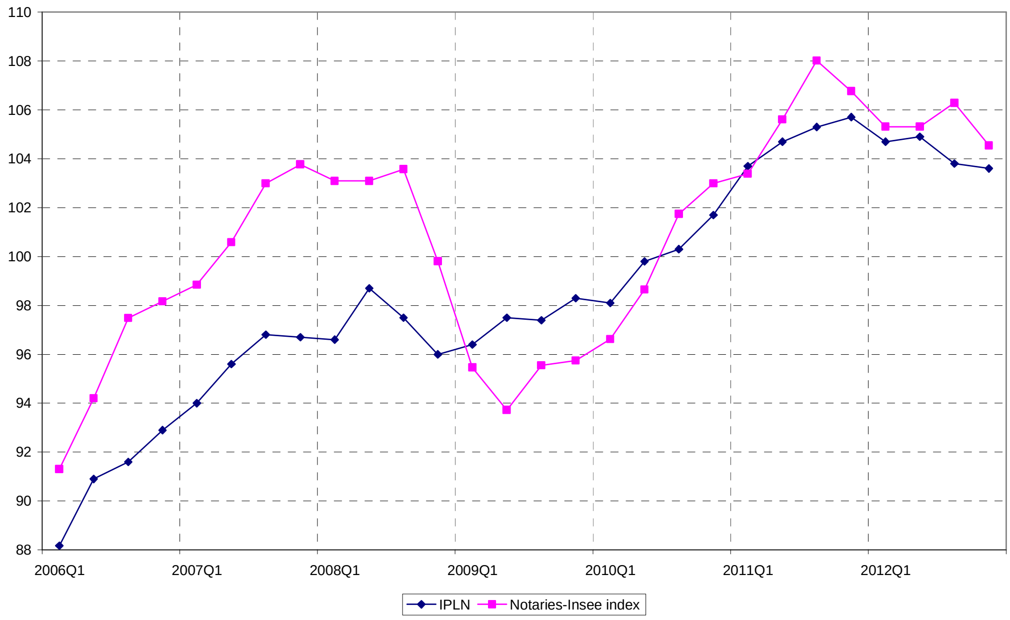
Over the 2006-2012 period, the new and existing dwellings indices evolve differently (graphs 1, 2 and 3): the average difference in absolute value between the quarterly growth rates of the two indices is not negligible (1.4 percentage points). The same is true for the annual growth rates (2.4 percentage points). Actually both indices follow the same trend except over the crisis 2008Q4-2010Q1 period (graph 3.). In 2008Q4 the Notaire-INSEE index falls by 4 points, by another 4 points in 2009Q1, and by

[21] Zone d'Études et d'Aménagement du territoire (Zone for study and development)

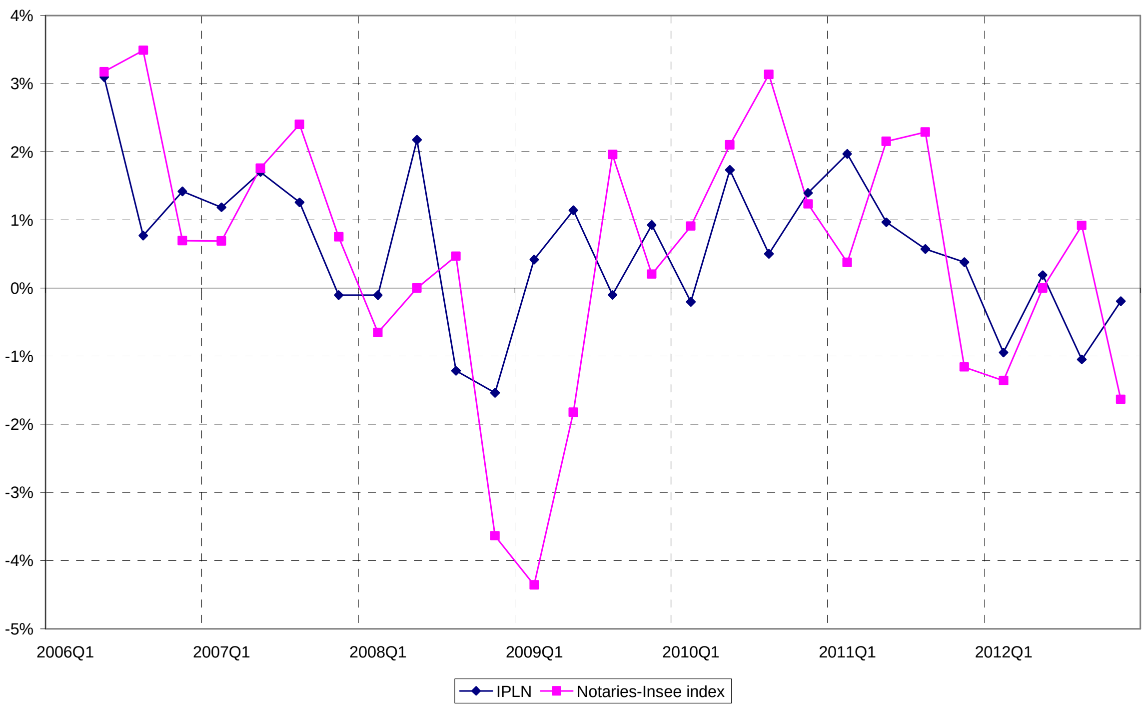
[22] Over the period 2006Q1-2012Q3, the average quarterly number of observations is 8 194 programs, corresponding to 26 105 new dwellings

2 points in 2009Q2. The new dwellings index only drops by 2 points in 2008Q4. The index for second-hand dwellings appears more volatile than the index for new dwellings (IPLN).

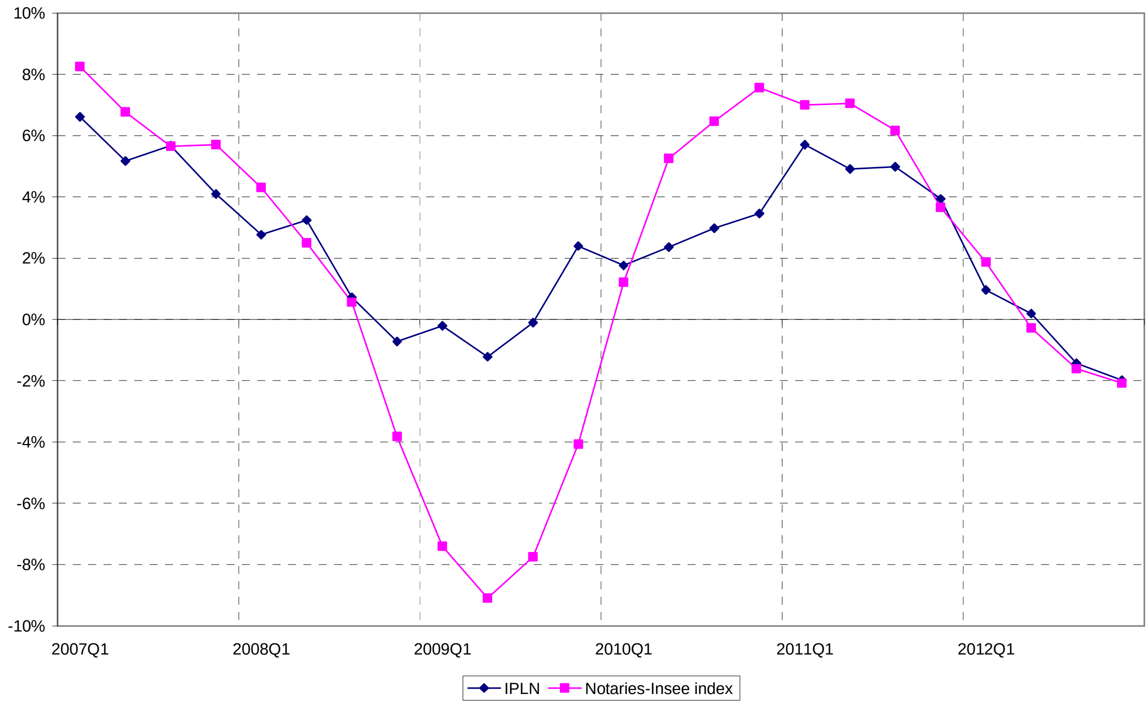
Graph 1. The new dwellings (IPLN) and the existing dwellings (Notaires-Insee) indexes (2006-2012 – 2010 = 100)



Graph 2. Quarterly growth rate of the new dwellings (IPLN) and the existing dwellings (Notaires-Insee) indexes (2006-2012)



Graph 3. Annual growth rate of the new dwellings (IPLN) and the existing dwellings (Notaires-Insee) indexes (2006-2012)



3. Ruling out two sources of differences between the evolutions of new and old housing price indices

Two main possible sources of differences between the evolutions of the two indices have been explored in details by Balcone (2013, and 2015). This section summarizes his results. First the bias due to different methods: adjacent period time dummy method on a single model (new dwellings) versus hedonic estimations of the price of a basket of dwellings (existing dwellings) at a disaggregated geographical level is explored. Then, we test if the differences in the location of new and existing dwellings are enough to explain the gap between the two indices.

3.1. First test: sensitivity of the existing dwellings price index to the choice of methods

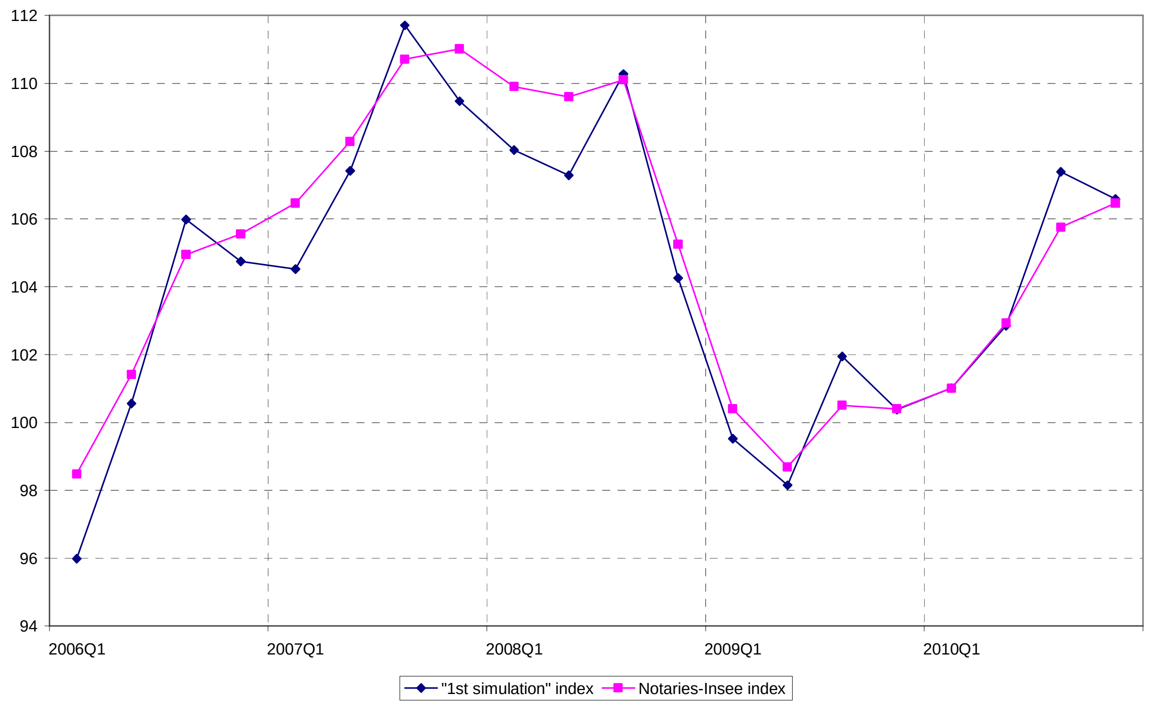
We first focus only on the computation method. Taking the samples[23] of the Notaries-Insee indices, we use the same adjacent periods time dummies hedonic model as the one used for the new dwellings index. Two indices are computed, over the period 2006-2010, one for houses, and one for apartments. To get closer to the smaller number of explanatory variables in ECLN than in the Notaries databases, the number of explanatory variables is reduced compared to model (1). The

[23] The data come from two databases. The “BIEN” database covers the Île de France whereas the “Perval” France database covers the Province.

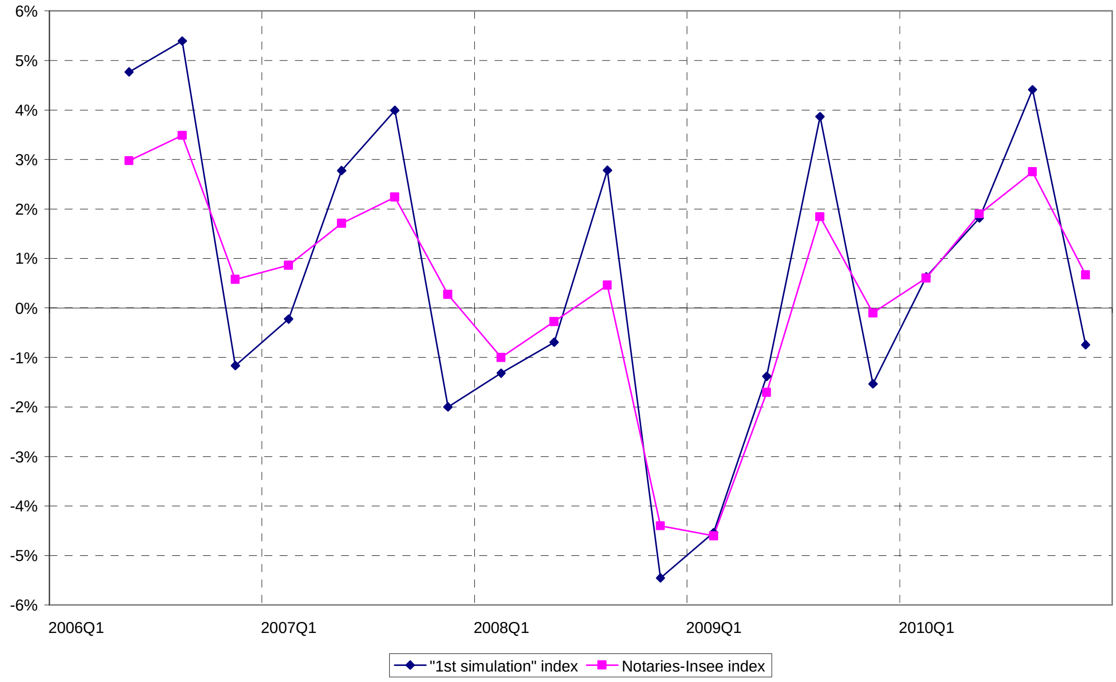
hedonic models for existing dwellings (estimated separately for houses and apartments) are similar to model (2). The geographical dummies are the same as in model (2). The physical characteristics include dummies for the number of rooms (one, or 2-4, more than five rooms) and for the presence of a balcony or terrace for apartments. Three changes are made to the current Notaires-Insee method: adjacent period versus basket of dwellings value estimation, reduction in the number of variables, and estimation of a single model for the whole country.

We get experimental existing houses and apartments indexes which we call "1st simulation" indexes. Those indices are compared to the Notaires-Insee ones. Since the only source of difference between these two sets of indices is the method, the gap between the indices is used to assess the potential bias due to different computation methods**.** The absolute value of the difference between the two indices is not higher than 2.6 index points for houses (graph 4), even if the seasonal changes in the "1st simulation" index are more pronounced than in the Notaires-Insee index (graph 5). The annual profiles are similar, and the differences are usually no more than 1 percentage point (graph 6). The difference is larger (two percentage points) at the turning point during the quarter of the crisis. The difference in computation clearly cannot account for the difference between old and new houses indexes.

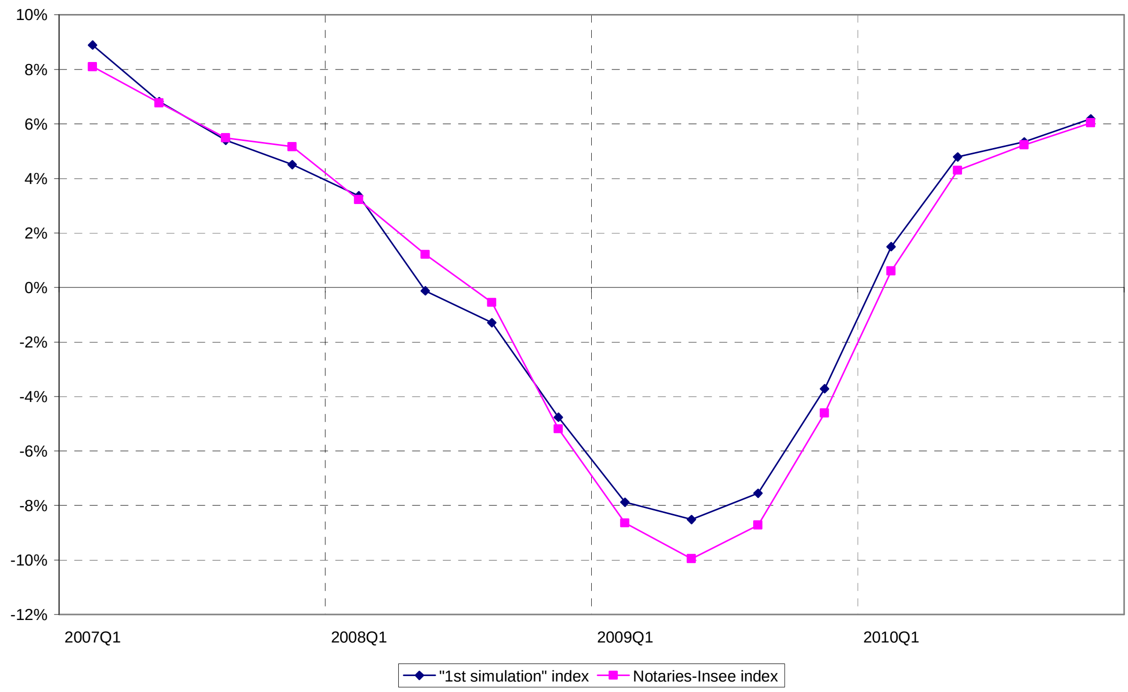
Graph 4. Existing houses: "1st simulation"and Notaires-Insee indexes (2006Q1-2010Q4 –100=2009)



Graph 5. Existing houses: "1st simulation"and Notaires-Insee indexes , quarterly growth rates (2006Q1-2010Q4)



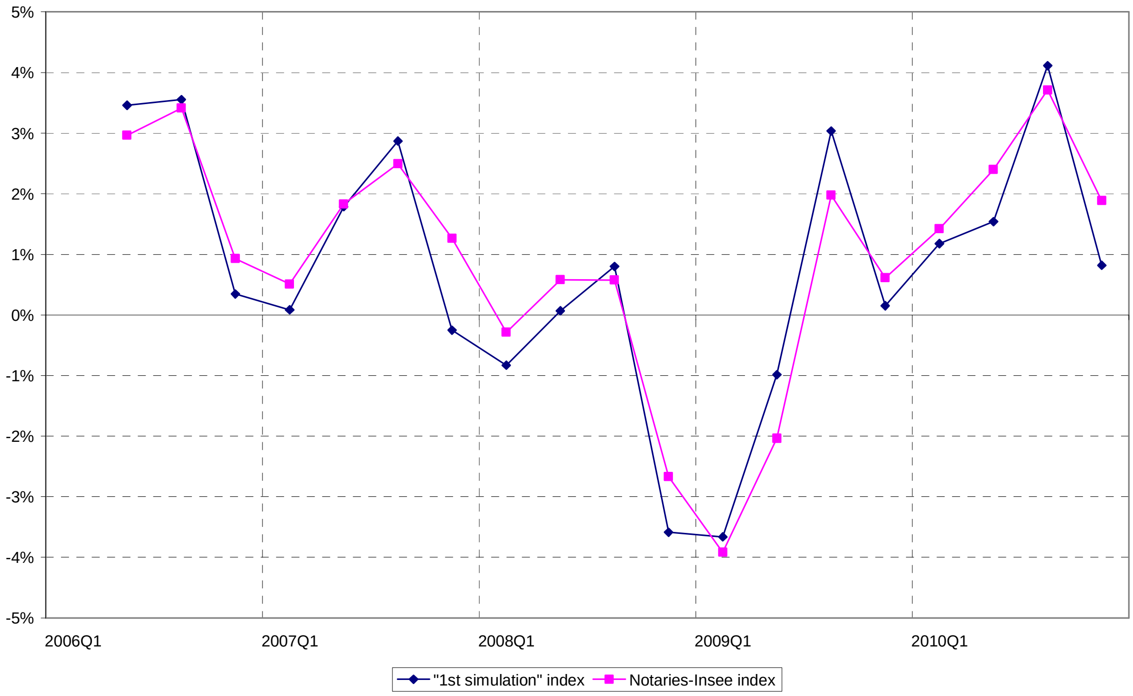
Graph 6._ Existing houses: "1st simulation"and Notaires-Insee indexes, annual growth rates (2006Q1-2010Q4)



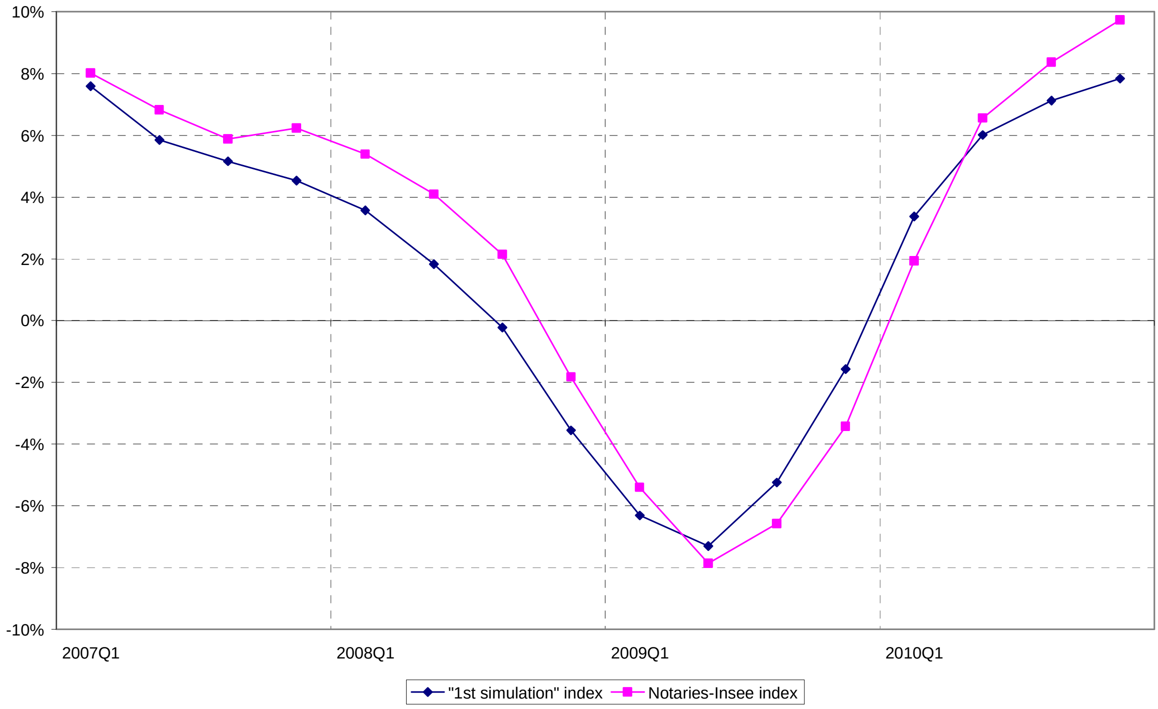
For apartments, the absolute value of the difference between the two indices is not higher than 2.6 index points (graph 7). Moreover, the evolution rates of both indices are also close (graphs 8 and 9).

Graph 7 : Existing apartments: "1st simulation"and Notaires-Insee indexes (2006Q1-2010Q4 –100=2009)

Graph 8 : Existing apartments: "1st simulation"and Notaires-Insee indexes, quarterly growth rates (2006Q1-2010Q4)



Graph 9 : Existing apartments: “1st simulation”and Notaires-Insee indexes , annual growth rates (2006Q1-2010Q4)



The aim of this first methodological work was to check if different computation methods could explain the gap between the new dwellings price index and the existing dwellings one. The results both for houses and apartments, lead to the conclusion that even if the Notaires-INSEE method is very different from this simplified adjacent period dummy method, the final difference at the national level is small. Moreover it does not explain the gap that was observed during the 2008Q4-2010Q1 crisis period.

3.2. Second test: sensitivity of the price index of existing dwellings to the location of the dwellings

Balcone (2013) then focused on the differences in location between the new and existing dwellings. New dwellings are usually not built in the same areas as the existing ones. They are often located in the suburbs of cities, where land is available for construction. If the housing markets are geographically segmented, it could lead to differences in the time evolution of prices. In order to correct as much as possible for the different locations of new versus old dwellings, the observed second-hand transactions were resampled to mimic new dwellings locations. The only difference between such “geographical clones” and the original database is then the geographical distribution of the dwellings. Thus, indices calculations carried out with the same methodology on these two databases would allow assessing if the difference in the location of the new and existing dwellings can explain the differences between their price evolutions.

The clone database is built up in the following way. The number of new dwellings transactions in ECLN and the number of existing dwellings transactions in the Notaries' databases are computed for each triplet (year[24]; type of dwelling[25]; municipality[26]) over the 2006Q1 - 2010Q4 period. Only triplets for which sales of both new and existing dwellings exist are kept. Then, for each triplet, a sample of "clones" of new dwellings among the corresponding existing dwellings was 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 was computed over the period 2006-2010 using the same adjacent periods time dummies hedonic model as that used for new dwellings (cf. 2.2).[27] As before, two indices are computed one for "clones" houses, another for "clones" apartments. The year is now the elementary time level.

To assess the impact of the difference in the locations of the new and existing dwellings, we focus, for houses and apartments on the annual growth rates of the three following indices: the "clone" existing dwellings index, the "1st simulation" existing dwellings index and the new dwellings index (recomputed separately for houses and apartments) (table 2).

Table 2. The methods and the samples used to compute the three indices

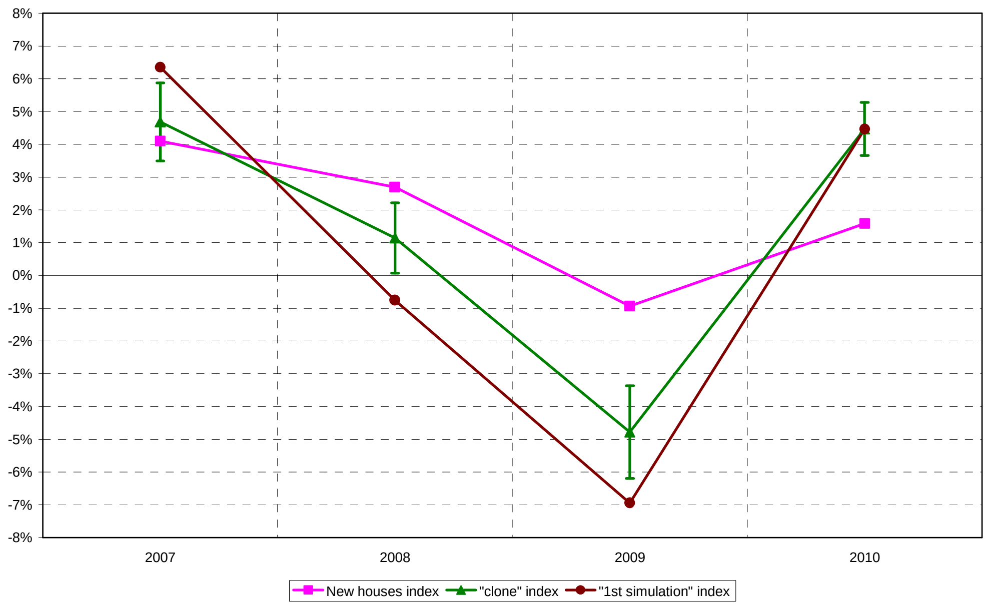
Index	Method	Sample
"Clone" index	Adjacent periods	Existing dwellings[28] "clones" of the new dwellings[29]
"1st simulation" index	Adjacent periods	Existing dwellings used to compute the Notaire-Insee indices
New dwellings index (IPLN)	Adjacent periods	New dwellings of the ECLN database

We plot 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 (graph 10). The difference in absolute value between the growth rate of the "clone" index and that of the new houses index is almost 1.6 times lower on average than the gap between the growth rate of the "1st simulation" index and that of the new houses index over the 2006-2010 period (2.22 versus 3.65 percentage

[24] We selected the year instead of quarter to have enough observations
[25] Houses (only "Individuel groupé" in ECLN) and apartments (collectif). See above for details.
[26] Arrondissement for Paris, Marseille and Lyon
[27] To get a better estimation of the annual growth rate of the "clone" index, 50 different samples of "clones" dwellings were drawn. Thus, the mean and the standard deviation of the annual growth rates of 50 "clone" indices is considered. However, to simplify the writing, we will speak below about the annual growth rate of the "clone" index.
[28] From the Notaries databases.
[29] From the ECLN database.

points). The drop during the crisis period is less important in the "clone" locations. Thus, the difference in the municipality distributions of the new and existing houses seems to explain on average part of the gap between the "1st simulation" index and the new houses index over the period. However, this must be put into perspective because the average rate of the new houses index (1,86%) is within the two standard deviations confidence interval of the rate of the "clone" index [0,81% ; 1,95%] and that of the "1st simulation" index is very close to the lower bound (0,78%).

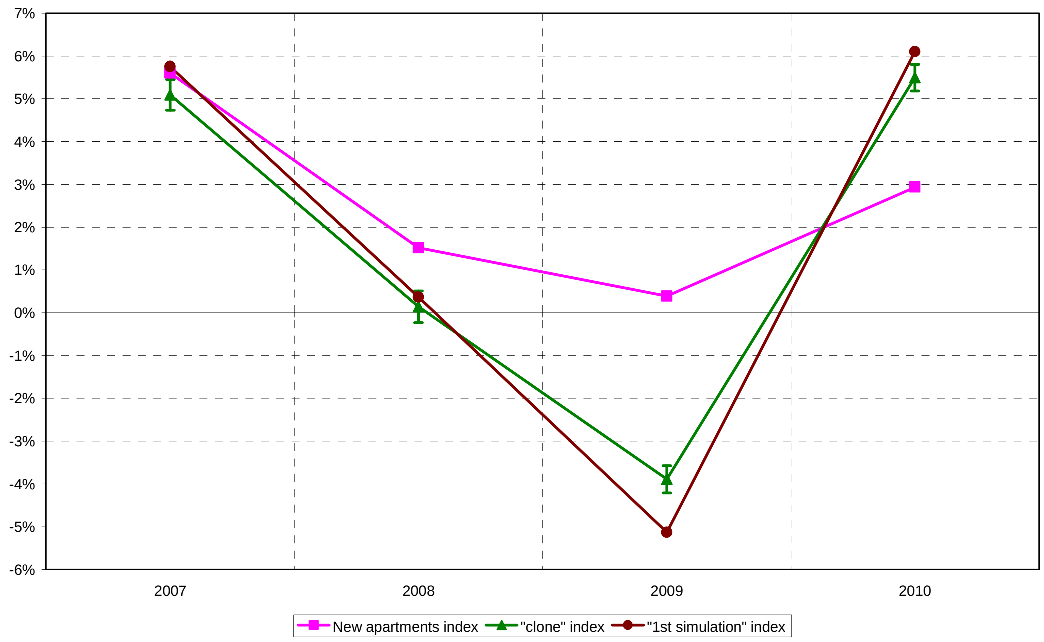
Graph 10 : Annual growth rates: New houses, "clone" existing houses, and "1st simulation" existing houses indexes (2006-2010)



N.B. : 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 existing houses can explain the gap between the existing houses index ("1st simulation" index) and the new houses index

For apartments, the "clone" index growth rate is very close to that of the "1st simulation" index. The difference in absolute value between the annual growth rates of these two indices is on average lower than 0,70 percentage points over the 2006-2010 period (0,68 percentage points, graph 11). If we put aside year 2009, the gap falls below 0.5 percentage points, even if, again, the crisis year is less marked in the "clone" locations. Thus, the difference in the municipality distributions of the new and existing apartments explains on average only a very small part of the gap between the "1st simulation" index growth rate and the one of the new apartments index over the period. Residual differences remain, suggesting differences within municipalities in the price dynamics of the two kinds of apartments.

Graph 11 : Annual growth rates : new apartments, "clone" existing apartments and "1st simulation" existing apartments indexes (2006-2010)



N.B.: For the « clone » index, the error bars are equal to two standard deviations
The closer the « clone » index to the new apartments index, the more the difference in the geographical distributions of the new and existing apartments can explain the gap between the existing apartments index ("1st simulation" index) and the new apartments index.

We can safely 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 significantly explain the difference in the two indices evolution, especially for apartments. Since we have ruled out the effect of the methodology, the fact that we do not find any significant effect is due to other unobserved differences between the markets for new and second-hand dwellings. The evolution of construction cost might influence the price of new dwellings and have less effect on that of second-hand dwellings. Also location was taken into account in a rather crude fashion. Theory predicts that, within a municipality distance and position relative to the city center and amenities play a role. If the relative prices within a city change over time this could explain the different evolutions between indices.

We pursue the reflection in two directions. One is a theoretical reflection on what is the quality of location and how it can evolve over time as a dwelling turns from new to old (section 4.1). The other direction start from the idea that the price of a new house can be easily decomposed between the price of the land on which it is built and the price of the structure (section 4.2). The evolution of the two prices might differ. This leads us to build a price index of residential land (*terrains à bâtir*) from the *Enquête sur le prix des terrains à bâtir* (EPTB) of SOeS. Then we compare this index with that of new homes, and a construction price index. That is order to better understand the new homes price dynamics (section 5).

4. Theoretical thoughts

As a new building becomes old, the quality of the structure declines while the quality of the land might improve (if, say, new public services and transportation are opened in the neighbourhood) or decline (if new sources of pollution, congestion or noise appear). Depending on the rates of depreciation and land quality evolution the value of the now “existing” dwelling will decline or increase. Supply and demand will also play a role, so will improvement and rehabilitation. We suggest reflections in two directions: the role of location characteristics in the hedonic price models (4.1) and the influence of land share and land price evolution on home price indexes (4.2).

4.1. The influence of the hedonic model specification

The estimated coefficients in the hedonic regressions are used in the price index computation but typically cannot be interpreted in terms of shadow prices of a particular characteristic because of the many omitted variables, of possible correlations between variables, of model misspecification and measurement errors. Moreover when estimating the models period after period, the evolution in such estimated coefficients is not discussed. However at some point we do have to ask what does those shadow prices, and particularly that of location mean. More precisely, as we move from one market equilibrium to another, what exactly is happening in terms of price evolution?

If the hedonic function is a joint envelope of bid and offer curves of buyers and sellers (Rosen, 1974), what happens when over time one characteristic radically changes? Could it explain part of the difference between new and second-hand indexes?

Suppose a monocentric city. At a point in time, a unit of land under an existing home is closer to the city center than a unit of land at the city limit, hence, abstracting from other amenities, is more expensive. This is the negative gradient of land price in a monocentric city model (Fujita, 1989).[30] Suppose that at date 0, prices in neighbourhood *x* (25 minutes from the city center) are 10% lower than neighbourhood *y* (10 minutes from city center). Then at date t, and a new train line is built so that *x* becomes 15 minutes from the city center. What theoretical models predict is that the land price gradient becomes more flat, as the relative value of neighbourhood x becomes higher. Gofette-

[30] Combes et al. (2011) observe such gradients on cross section data on land prices from the *Enquête sur le prix des terrains à bâtir* (EPTB) in 2008 in 100 urban units.

Nagot (2009) is interested in how the price gradient evolved over time. She has a clear explanation of the flattening of land rents, and measures it on French Housing surveys data. She finds that the price gradient was 3.6% per km at the beginning of the period (1982-1995) and 2.7% over 1995-2001.

Depending on how the hedonic models are estimated different conclusions will be reached on the "constant quality" price index evolution. If, as is the case, time-distance to city center is not included in the hedonic models, the price evolution in *x* will be overstated. We shall attribute to price increase what is a quality evolution. This is an omitted variable bias. In other words the estimated "shadow price" of location evolves over time. How this is taken into account is likely to influence the price index.

How does the evolution in the shadow price of a characteristics influence the index?

In practice we are not aware of any official land or home price index using distance to city center or detailed amenities as a hedonic characteristics. Most use zone dummies. Hill and Scholz (2013) replaced postcode dummies by geospatial splines in Sidney. They find that this leads to cumulative price change from 2001 to 2011 being higher by between 15 and 30 percent. They claim that the postcode dummies fail to fully capture changes over time in the locational quality of houses sold. They conclude that " It is important therefore that index providers start using geospatial data in their house price indexes."[31] Cheshire and Shepard (1995) introduce distance to city center, and angle of deflection from the East to take into account topography, but they are not building indices. Suppose the price model for house *i* is the following:

$$Log\ P_i = a\ A + b\ B + c\ C + \Sigma_k\ b_k\ X_{ki} + e_i$$

X_{ki} are *k* characteristic of the house structure, *A, B*, and *C* are location qualities, say 3 neighbourhoods, with *A* is the city center, *B* the 1st ring, and *C* the outer ring. The land price gradient in this monocentric city implies $a>b>c$. If we make apparent the omitted variables defining the reference house of price P_0 located in the city center:

$$Log\ P_i = Log\ P_0 + b\ B + c\ C + \Sigma_{k'}\ b_{k'}\ X_{k'i} + e_i \qquad (3)$$

Thanks to this model, estimated at each date, we can compute from the observed prices of houses sold at date *t* the estimated prices of a reference house, say a second-hand house in *A*; for a house sold in *B*:

$$Est(Log\ P_i) = Log\ P_i - \Sigma_{k'}\ b_{k'}\ X_{k'i} - b$$

$$Est(Log\ P_i) = Log\ P_{0t} + e_{it}$$

We call it deflating price P_i so it becomes the equivalent of the price of the reference house. Then one estimates price as the mean of all estimated prices, independently of the characteristics of the

[31] One could wonder whether their method could overstate the decrease in location quality if new transportation means have been built in Sidney, or transportation costs has decreased.

houses sold at date *t*. This compared to the price of the reference home at date *t-1* gives the quality adjusted price index.

Suppose new houses are built in *B* and *C*. Second-hand houses are sold in *A* and *B*. Model (3) includes A and B dummies and is used to compute the price index for 2d-hand houses. There are no transactions in C so this zone is implicitly added up to B. Model (3), including *B* and *C* dummies, is used for new houses. The models are estimated in period 1 and used in period 2.

If nothing happens in terms of relative amenities in the 3 neighbourhoods *A, B*, and *C,* a common trend in quality corrected price evolution for second hand and new houses is to be expected, while new houses price evolution will also be impacted by the evolution in construction costs.

If there is a sudden change in the quality of neighbourhood *C* (for instance because of the unexpected extension of a public bus service), shadow price c will increase. What happens on the indexes?

The price increase of new dwelling will be overestimated if we still use estimated c_{hat} price < c_{new} to deflate the price in the index computation.

For old dwelling we deflate using *b* for the houses located in *C* (that are now second hand if they are sold). So we may underestimate the price evolution. This is likely be minor as long as there are few transactions in *C*, and as the number of transaction increases c_{new} is getting closer to *b*, so the bias is reduced.

What if model (3) does not include location dummies? The problem is similar to that of an omitted or unobserved characteristic. As long as the omitted quality is the same for all houses or does not change over time, it has no adverse effect on the price index. When the shadow price of this characteristic increases (from 0 to a positive value if it is a positive amenity or quality, say air conditioning), the deflated price of the houses with this characteristic will fall. What was considered a price increase is now a quality change and treated accordingly.

Such theoretical reflections suggest future work. Data permitting, it could be interesting to build sub indexes by construction dates, to measure how the relative prices of various cohorts of homes evolve. A related question is to what extent future amenities (accessibility, public services) are capitalized at the time of construction? Price indexes methods are usually mute on this question because they lack underlying geographic models. At best they take into account amenities whose price may vary over time. For instance in the Notaires-INSEE second-hand home price index the geographical level of the computation and the neighbourhood dummies would proxy for amenity levels. Gofette-Nagot (2009) is an interesting and unique example of estimation of both geographical variation and time evolution of land rents (the effect of distance to city center in the models in section 4.2 below). She finds (1) this gradient is flatter in more populated cities, which have better and cheaper transportation system and more dispersed public services; (2) a decrease in the price gradient over time between 1982 and 2001, meaning that price increase was more rapid at the urban fringes than at the city centers, probably following the construction of new transportation system, the dispersion of activities and public services away from city centers, or demographic pressure. All this points to amelioration in the transportation system that was capitalized in land prices. Or to a relative decline in city center attraction over time, or decline in close suburbs attraction compared to more distant locations. It would be interesting to replicate her work on more recent periods. If this quality evolution of suburban neighbourhood is not taken into account, the price index of new homes may be biased.

4.2. Building on land

We start now from the idea[32] that at time t the total value of a house, V, can be separated into the value of the land, L, and the value of the building, the structure S:

$$V_t = L_t + S_t. \qquad (4)$$

Let g_L, g_S, and g_V, denote the periodic 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 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) $\alpha_t = L_t / L_t + S_t$ ($0 < \alpha_t < 1$)

$$g_v = \alpha_t g_L + (1 - \alpha_t) g_S \qquad (5)$$

The price evolution of a house between t and $t+1$ is the weighted sum of the price evolution of land and that of structure, where the weights are function of the share of land in the total value of the house in period t. If we differentiate

[32] Inspired among others by Bostric et al. (2007), Diewert (2011), Davis and Heathcote (2007), Davis and Palumbo (2008).

$$dg_v / d\alpha_t = g_L - g_S$$

The difference is positive when the price of land moves faster than the price of structure. Then the higher the land leverage, the higher the price evolution. This would predict that in general indices for second hand dwellings would move faster and be more volatile than for new dwellings. This seems to be the case at the turning point in 2008 when we compare the Notaires-Insee index and the ECLN new dwelling index. Then the shock in demand affects the land prices more than the value of constructions.

Such ideas led us to build price indexes of land (terrains à bâtir), new structures, and houses (land+ structure) from the Enquête sur le prix des terrains à bâtir (EPTB) of SOeS. Then we compare those indexes with that of new homes, and a construction cost index, in order to better understand the new homes price dynamics (see section 5).

From eq. (5) we can derive α_t:

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

If we make an assumption on the depreciation rates of structure, and if we could estimate separately the price index for land, knowing the overall home price evolution gives α. For instance if we assume no depreciation, $\alpha_t = g_v / g_L$. It is however unlikely that the price index of the underlying land is known.

Note that if we only consider new homes built at each date, we can formally write the same suite of equations but the meaning is different. For a house built in *t*:

$$V_t = L_t + S_t.$$

Let g_L and g_V, denote as before the percentage change between *t* and *t+1* in the land and overall property values, respectively. Now , g_S is the evolution in the construction cost of the (same quality) new property. The value of a new property at date *t* +1 can be expressed in two ways:

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

if we compare it to the value of a new property at date t; and

$$V_{t+1} = L_{t+1} + S_{t+1} = L_t (1 + g_L) + S_t (1+ g_S) .$$

With the sign of g_S depending on the evolution *of* construction cost over time.

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

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

$$g_v = \alpha_t g_L + (1 - \alpha_t) g_S$$

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 construction cost, where the weights are function of the share of land in the total value of the house in period *t*.

As before α_t is given by eq. (6). But now we have to check if $0 < \alpha_t < 1$ for the validity of the computation. It can be easily shown that this condition is met if and only if:

$$g_S < g_v < g_L \qquad \text{or} \qquad g_L < g_v < g_S, \qquad (7)$$

Conditions (7) are not warranted for successive cross sections of new homes.

5. Decomposing a house price into the price of land and the price of the structure

5.1. Using a survey (EPTB) to compute price indices for land, structure and individual houses

EPTB[33] is a survey on the prices of residential building land. It covers the building permits of "individually built" single family houses in France and provides the prices and features of land plots (surface in m², whether it was purchased or not, purchase date, presence of servicing works, etc.). It also provides the expected prices of the construction and a few of its features: floor space (SHON), nature of the main coordinator of the works, type of heating system. We keep only the 344 847 observations for which land was purchased between 2006 and 2012, was located in metropolitan France (excluding Corsica)[34], and for which the surface of the purchased land is equal to the surface registered in the building permit[35] (see Appendix 1). From these data, we compute three quality-adjusted indices using the adjacent periods method: a land price index, a structure price index, a new single houses price index.

5.1.1. A land price index

To compute a quality-adjusted land price index we take into account the geographical effect by building one hedonic model per region. In each model we add dummies for additional amenities at the municipality level: coastline, estuary, touristic « arrière-pays littoral »[36] or of ski or alpine resort[37]. We also control for the urban unit category of the municipality[38]. 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

[33] « Enquête sur le prix des terrains à Bâtir » conducted annually by the Ministry for Ecology, Sustainable Development and Energy

[34] The same geographical field as for the new dwellings index.

[35] The registered surface is the total underlying soil (the surface of the ground floor + that of gardens and outhouses). We exclude cases where the individual buys just an extension of a plot which was his already, or where he buys a large plot, then divides it and uses only a part of it for the building. Then the price per square meter may be lower because only a parcel of the purchased land may have building permission (the remaining part corresponding to farmland for instance).

[36] 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).

[37] Cf the law of 14 April 2006

[38] 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

a city center or a suburb[39]. Municipalities outside urban units are called rural. We finally add the straight-line distance in km between the municipality where the land is bought and the closest urban center.[40]

The hedonic model (8) is the following in each region r :

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

$$\ln\left(\frac{p\_L_{i,a}}{s\_L_{i,a}}\right) = \alpha_{A,r} + \beta^{A,r}_{s\_L}\ln(s\_L_{i,a}) + \beta^{A,r}_{shon}\ln(shon_{i,a}) + \beta^{A,r}_{dist}dist_{i,a} + \sum_{k=1}^{K}\beta^{A,r}_{k}I_{i,a,k} + \delta_{A,r}D_{i,a,A} + \varepsilon_{i,a} \quad (8)$$

Where, for observation of plot i in year a , $\mathrm{p\_L_{i,a}}$ is the price in euros of the plot, $\mathrm{s\_L_{i,a}}$ is its size in m², $\mathrm{shon_{i,a}}$ is the size in m² of the house[41] , $\mathrm{dist_{i,a}}$ the distance to the nearest urban center. $(\mathrm{I_{i,a,1}},...,\mathrm{I_{i,a,k}},...,\mathrm{I_{i,a,K}})$ is a vector of K dummies for the characteristics of the plot. $\mathrm{D_A}$ is year A dummy defined as follows :

$$\forall a = A-1, A\ et\ \forall i = 1,..., nb\_obs(a),\ D_{i,a,A} = \begin{cases} 1\ if\ \ a = A \\ 0\ otherwise \end{cases}$$

The K characteristics of the plot are the geographic dummies described above to which we add whether the plot was serviced (“viabilisé”), or was bought through an intermediary or not and its type. From the models, we compute, for each of the 21 regions r an **annual** constant quality price index for year A, 100 = A-1, $\mathrm{I\_L_{r,A/A-1}}$:

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

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

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

$\mathrm{I\_L_{A/A-1}} = \sum_{\mathrm{r=1}}^{21}\mathrm{w\_L_{r,A-1}} * \mathrm{I\_L_{r,A/A-1}}$, where $\mathrm{w\_L_{r,A-1}}$ is the share of plot expenses in region r, in year A-1 (see Appendix 2). Then this indice is chained to get a national annual price index for plots for year A, 100=2006 , $\forall \mathrm{A} = 2007,...,2012,$

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

[39]Definitions can be found at : http://www.insee.fr/en/methodes/default.asp?page=definitions/ville-centre-et-banlieue.htm

[40] The urban center (pôle) is an urban unit offering at least 10,000 jobs and not located in the crown of another urban center. The crown from an urban cluster covers all the municipalities in the urban area to the exclusion of its urban center.

[41] Floor area(GFA) replaced the net ground area (SHON) on 1 March 2012

5.1.2. A structure price index

From the expected house price we compute an annual price index for the structure, in exactly the same fashion. The model used is similar to model (8). The only differences are that the dependent variable is $\ln(p\_S_{i,a})$ where $p\_S_{i,a}$ is the price of the structure (excluding plot), the regressor $\ln(s\_L_{i,a})$ does not appear, the K dummies include the same geographic variables as above, but the plot characteristics are replaced by that of the structure: degree of finish (totally fitted, ready to decorate, only “clos et couvert ”, heating mode (gas, electricity, renewables, etc.), type of builder (architect, developer, artisan, self-building, other). Then, we compute a national indice of structure price for year A, 100=2006, noted $I\_S_{A/2006}$ using the same methodology[42] as for plot.

5.1.3. A new single-family houses price index

An annual price index for new single family units can be computed in the same fashion from the EPTB survey, assuming that the house price is the sum of the price of plot and price of construction, noted $p\_V_{i,a}$:

$$p\_V_{i,a} = p\_S_{i,a} + p\_L_{i,a}$$

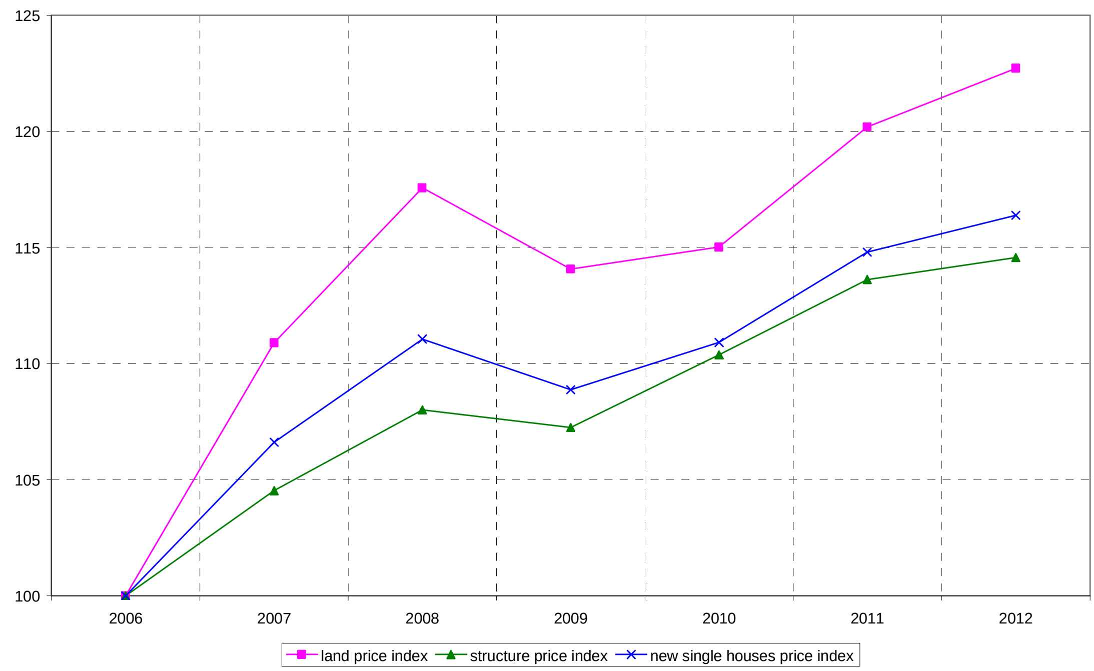
The model used is again similar to model (8). The only differences are that the dependent variable is $\ln(p\_V_{i,a})$ and that the characteristics of the structure are added in the K dummies. Then, a national index for single family homes can be computed as before for year A, 100=2006, noted $I\_new\_ip_{A/2006}$[43].

5.1.4. Comparison of the three indices derived from EPTB

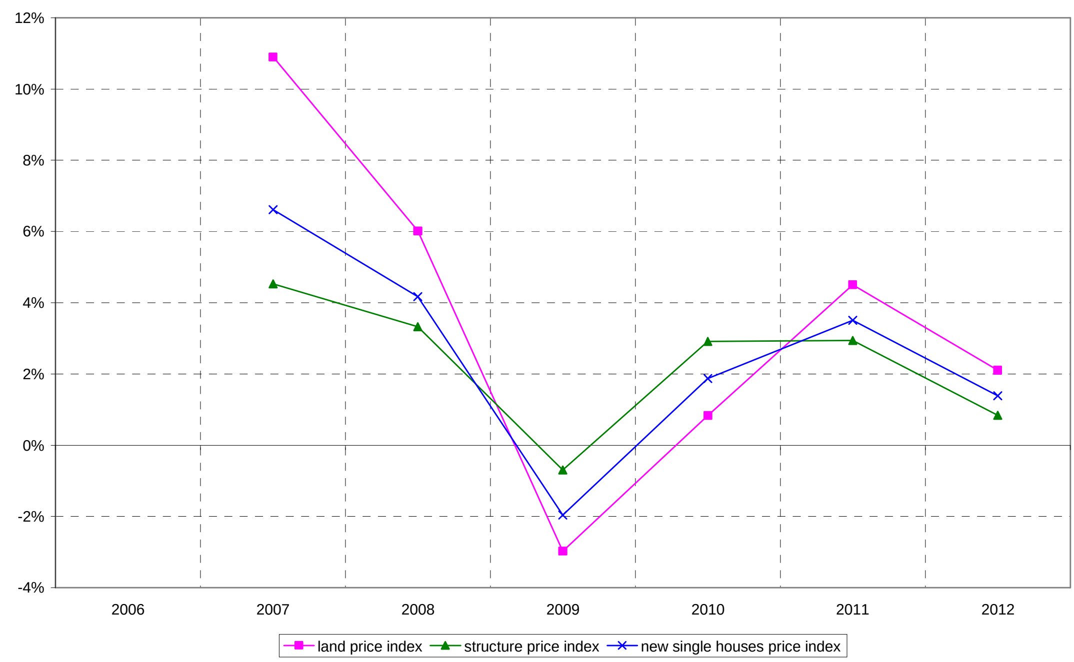
We compare the three indices derived from EPTB: that for the price of land, that for price of structure and that for the house (land + structure). Their evolutions are roughly similar over the period (Graphs 10 and 11). The land price movements are more pronounced at each date than the structure price movements. For instance in 2007 the increase for land is 10.9%, and only 4.5% for structure (it is 6.6% for the total home price). In 2009 the structure prices decline much less than the land price (-0.7% versus -3% for land).

[42] Weights are given in Appendix 3.
[43] Weights are provided in Appendix 4.

Graph 10 : The three indices derived from EPTB (100 = 2006)



Graph 11 : Growth rates of the three indices derived from EPTB



Nevertheless the structure price movements are closer than expected to the land price movements. It may be because we have included the location characteristics of the house in the hedonic model for structure. The rational for the inclusion is that the cost of construction varies with location, and could vary with distance to providers of material. Our location variables are proxies for such variation. Such inclusion may not be fully justified.

5.1.5. Land and construction shares in the price of a house

Instead of computing directly a price index for single-family units we now compute it as a weighted average of the price index of land and the price index of construction that we computed in 5.1.2 and 5.1.3 (graph 12). More precisely, the computation of the aggregated index for a year A (100 = 2006) is done in three steps :

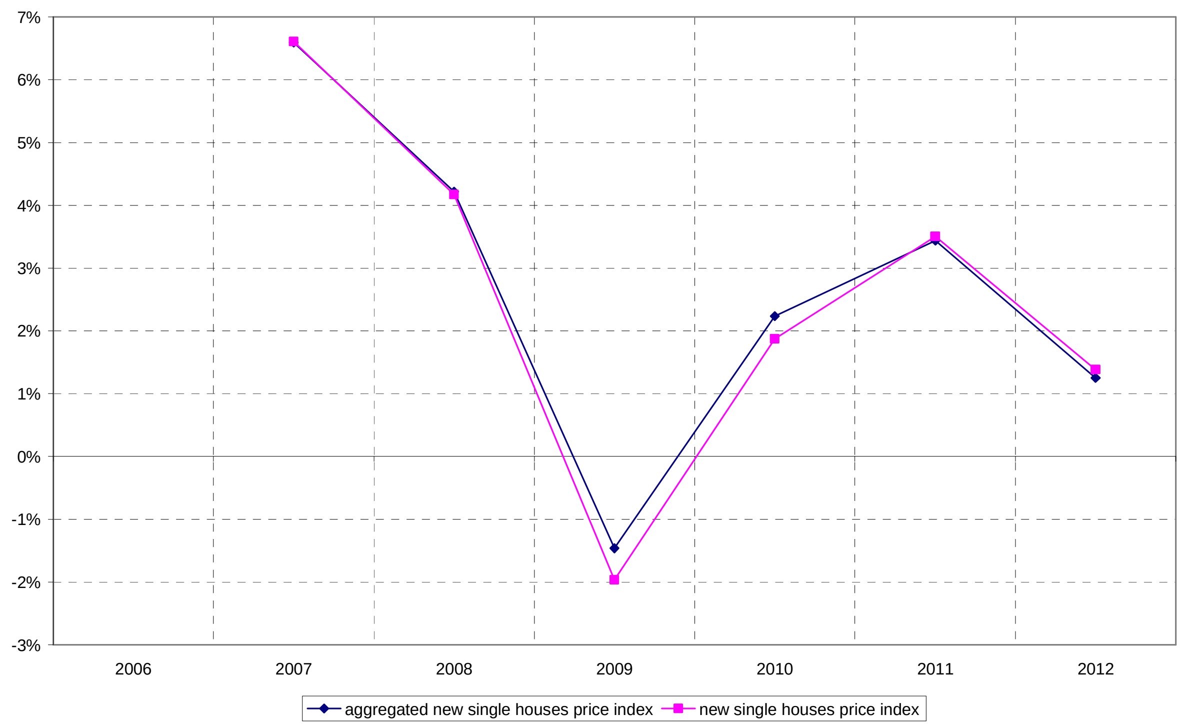
- an aggregated unchained index (i.e. 100 = A-1) is computed for each region as the weighted mean of the corresponding unchained price indices of land and of construction. The weights are given by the share of land α_r and construction $(1-\alpha_r)$ in each region in year A-1 (Appendix 5 and 6). On average, considering metropolitan France (excluding Corsica), the structure share is 69,1%, while the land share is 30,9%. There are large differences between regions. The land share is as high as 48.5% in Ile-de-France. It is slightly declining over the period. The share is 19% in Limousin, increasing from 17% in 2006 to 21% in 2012.
- Then, an aggregated unchained index for France is computed as the weighted mean of the regional aggregated unchained indices. The weights are given by the share of the expenses of each region in all the expenses for single family units built in year A-1 (cf. Appendix 4).
- Finally, we get the aggregated new single houses price index (100 = 2006) for France by chaining the unchained index.

The two indices, direct or aggregated, are very close, even if the difference reaches nearly one index point value in 2012 (Graph. 12). The difference comes from 2009, the crisis year in France : the difference in the growth rates of the two indices is maximal for this year (0.5% - cf. graph 13).
With the notations of section 4.2., the growth rate of the aggregated index for each region r is :

$$g_{v,r} = \alpha_r \, g_{L,r} + (1 - \alpha_r) \, g_{S,r} \qquad (9)$$

Graph 12 : Aggregated vs "direct" new single houses price index(100 = 2006)

Graph 13 : The growth rates of the aggregated and the "direct" new single houses price indices



5.1.6. A direct computation of α

In section 5.1.5 we used average land shares computed at the regional level to aggregate the price indexes of land and construction (eq. 9). We can indirectly check the effect of such averaging for each region r by computing α_r from $\tilde{g}_{V,r}$ (the growth rate of the “direct” new single houses price index), $g_{S,r}$ and $g_{L,r}$:

$$\tilde{\alpha}_r = \frac{\tilde{g}_{V,r} - g_{S,r}}{g_{L,r} - g_{S,r}} \qquad (10)$$

The computation appears not to make much sense at the regional level. At the national level, the order of magnitude of α is the same in 2006 and 2007, around 30% for the average share and 31-33% for the estimated α. The land leverage increases during the crisis years 2008 and 2009 to more than 50% which might seem counter-intuitive. In 2009 and 2010 the price evolution of construction is larger than that of land (which decreases sharply in 2009 - cf graph 11). However the price evolution of construction is also influenced by the crisis, and also declines in 2009, then rebounds in 2010, with the result of increasing the ex post land share in 2008. At the end of the period, the mean share does not raise much and stay around 31-32% while the indirect computation of α gives 36% in 2010, and 43% in 2011.

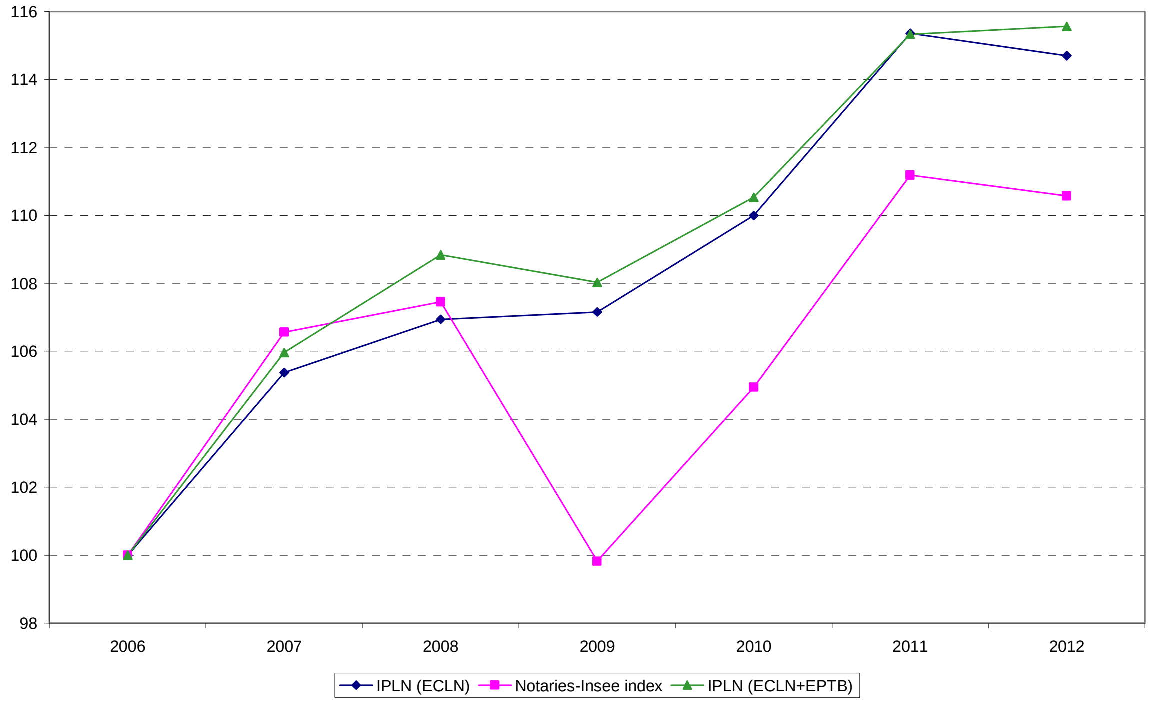
Table 3. Land leverage for France

Year *t*	2006	2007	2008	2009	2010	2011
Mean of regional land shares	29,8%	30,9%	31,7%	31,2%	31,1%	31,6%
estimated α_t	32,7%	31,4%	55,7%	50,2%	36,2%	43,4%

5.2. Towards a comprehensive new dwellings price index?

Aggregating the “individuel pur” new house price (IP from EPTB) index with IPLN (from ECLN), we get a comprehensive price index for new dwellings. This new index, Indice de Prix des Logements Neufs, noted IPLN (ECLN+EPTB) covers the three types of dwelling: “individuel pur”, “individuel groupé”, for single family units and “collectif” for apartments. This comprehensive Indice des Prix des Logements Neufs is very close to the previous one: the average absolute difference between the two indices does not exceed 0.7 index value (graph 14). As a consequence, the comprehensive index does not get closer to the Notaires-Insee index (the average absolute difference in the growth rates of these two indices is 2.3% whereas it is 2.1% if we consider the original Indice des Prix des logements Neufs).

Graph 14 : The Notaires-Insee existing dwelling, the IPLN (ECLN) and the IPLN (ECLN+EPTB) price indices

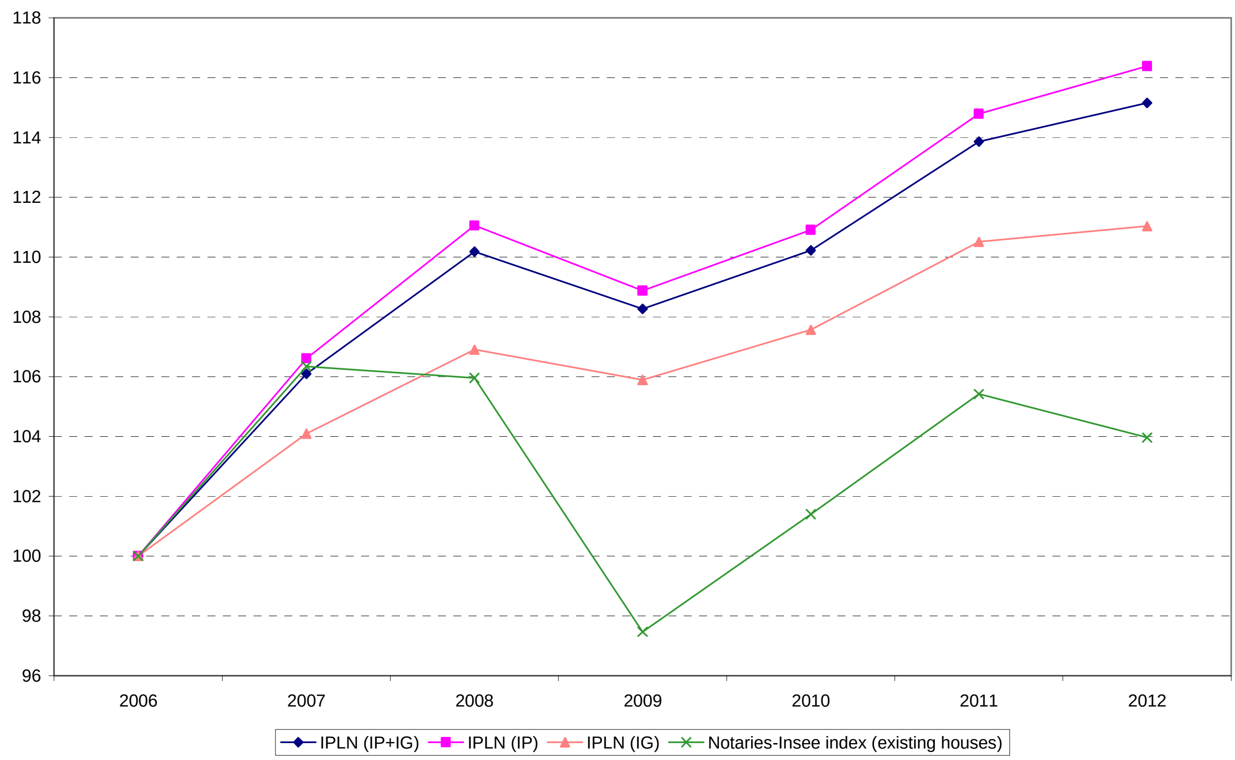


Graph 15 : Growth rates of the Notaires-Insee existing dwelling, the IPLN (ECLN) and the IPLN (ECLN+EPTB) price indices

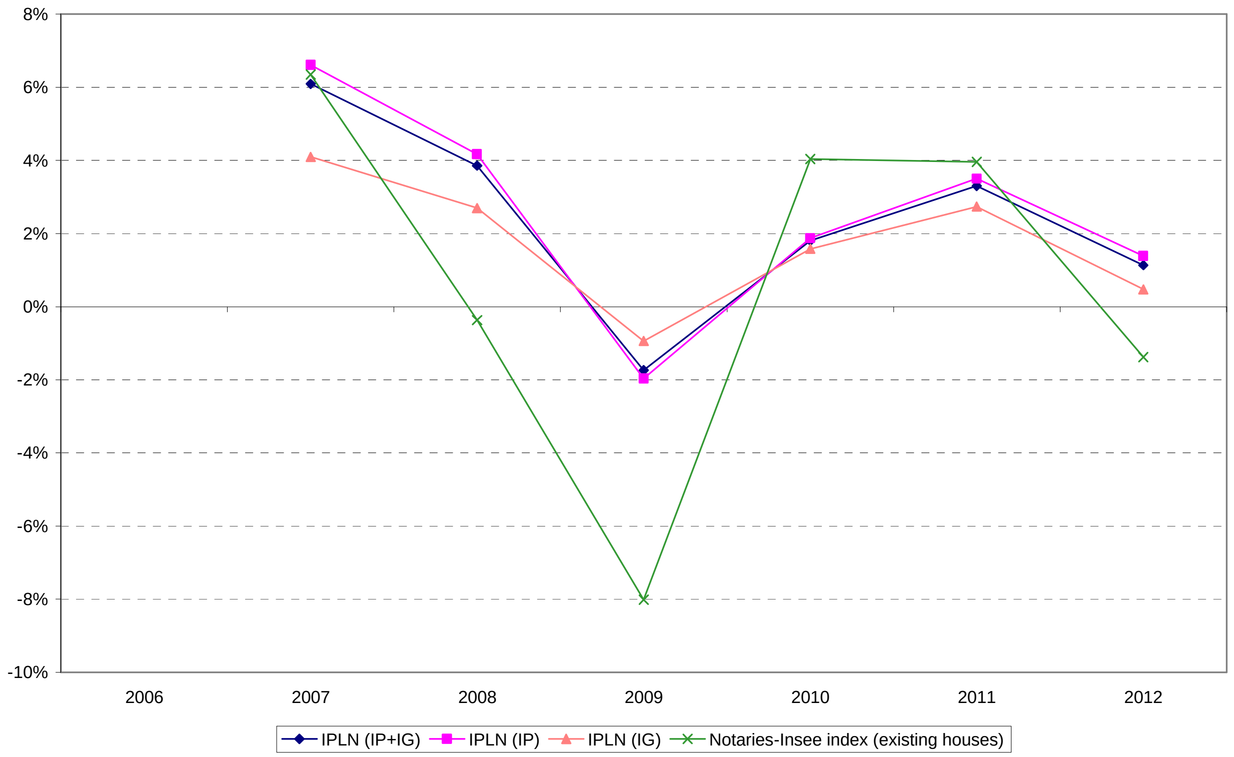
5.3. A new price index for houses

In the same fashion, since, the EPTB survey provided an index for « individuel pur ». A new indice covering all single family house constructions can be computed (individuel pur + l'individuel groupé). It is clear on Graphs 16 and 17 that the new house index differs more from the current IPLN that was seen on the index aggregating houses and apartments. This is because "individuel pur" (IP) makes up a large part of new houses constructions. It would be interesting to compare the land share between IG and IP. This new index can in turn be compared to the Indices Notaires-INSEE for second-hand houses, as was done in section 3. The differences are still there and hardly reduced by the inclusion of "individuel pur" in the index : the new index IPLN(IP+IG) is further from the Notaires-Insee index for existing houses compared to the original one (IPLN(IG)) (graph 16). However, the absolute difference in the growth rates is smaller for the new index IPLN(IP+IG) (2.7% for IPLN(IP+IG) vs 3.0% for IPLN(IG) - graph 17).

Graph 16: New house price indices (IPLN (IP+IG), IPLN(IP), IPLN(IG)) and the Notaires-Insee index for existing houses



Graph 17: The growth rates of the new house price indices (IPLN (IP+IG), IPLN(IP), IPLN(IG)) and the Notaires-Insee index for existing houses

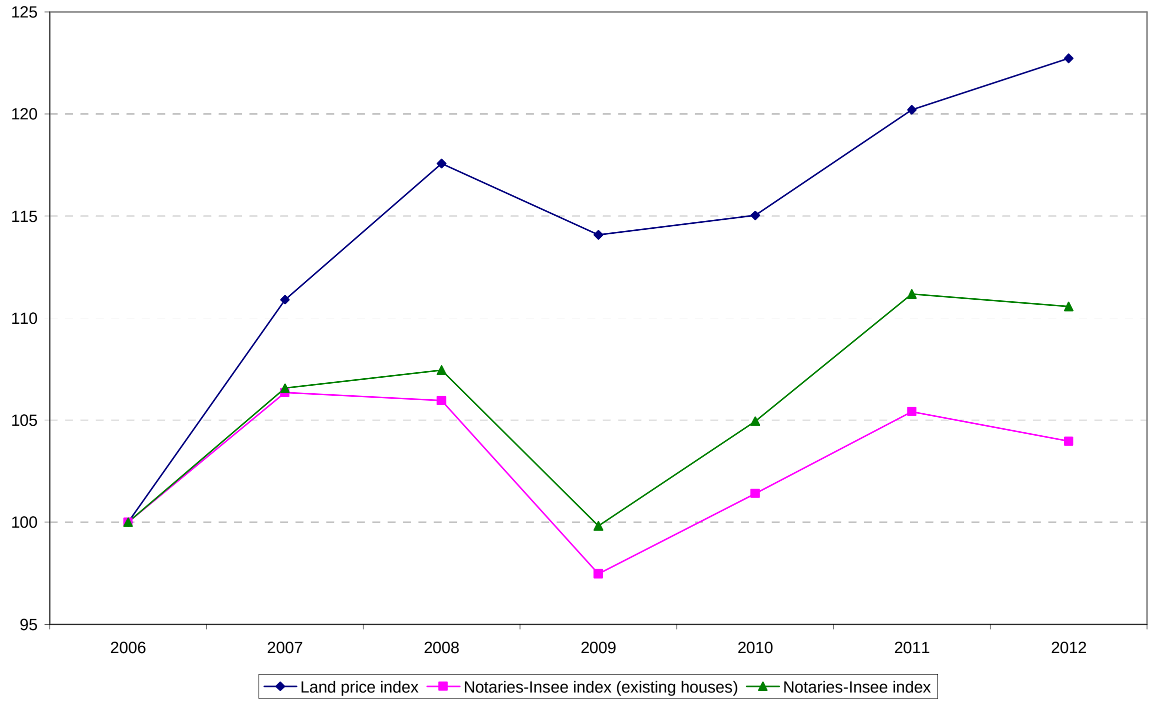


5.4. Comparison of the land price index with the Notaires-Insee second-hand houses index

The land price index is close to the Notaires_Insee index. It is slightly closer to the global Notaires-Insee index than to the Notaires-Insee index for existing houses (graph 18). It is also closer to those indices than to the indices of new dwellings (see the graphs in former subsections). The land price index follows the same trend than the Notaires-Insee indices till 2009 (graph 19). But, after the crisis, it is not longer exactly the case, even if the land price index and the Notaires-Insee index for existing houses have similar growth rate in 2011 (4.5% for the first one and 4.0% for the second one). However, the land price index is much closer to the price evolution of second-hand dwellings (and houses) than the IPLN index (Graphs 18, 19 and 15).

What drives the evolution of the housing market is the demand for space and location, that is for land. If we follow our theoretical model in section 4.2, we can safely conclude that the land share under existing dwellings is higher than the land share under new dwellings. In other words the land share in the stock is higher than the land share in the flow of new homes and land prices have grown faster than structure prices. This we check in the next section.

Graph 18: Land price index, Notaires-Insee index and Notaires-Insee index for existing houses



Graph 19: The growth rates of the land price index, the Notaires-Insee index and the Notaires-Insee index for existing houses

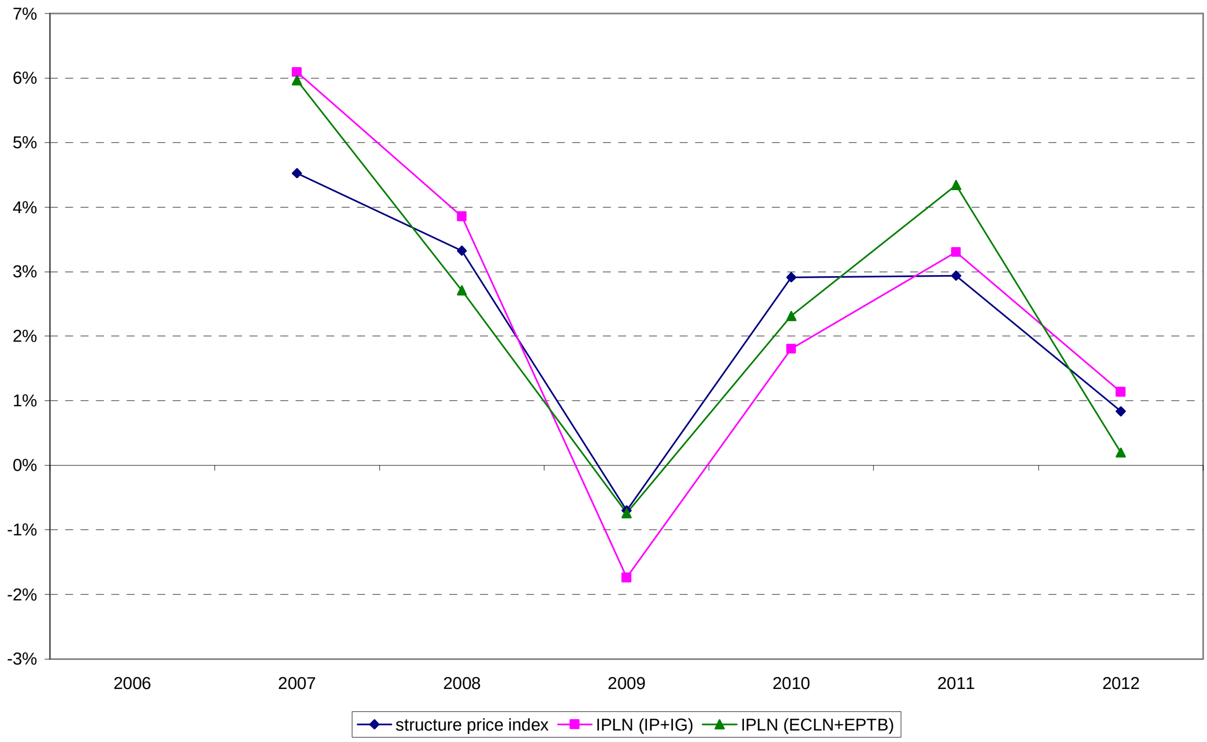
5.5. Comparison of the structure price index with the new existing dwellings index

The structure price index is below the comprehensive new dwellings index (IPLN (ECLN+EPTB)) and the comprehensive new houses index (graph 20). However, the three indices reach the same level in 2010 (110.3 for the structure price index, 110.2 for IPLN (IP+IG) and 110.5 for IPLN (ECLN+EPTB)). Moreover, the three indices follow the same trend over the whole period (graph 21) even if the decrease in 2009 is stronger for the new houses price index (IPLN (IP+IG)) (-1.7% vs -0.7% for the two other indices).

It can also be noticed on graphs 20 and 21 that the structure price index is around 115 in 2012, when the land price index is around 123 (graph 18). It may seem surprising that the structure price index has increased so much over the period. It could be that there is no productivity improvement in construction, or that wages in that sector, traditionally low have improved. It is also probable that the quality of homes has improved and this is not taken into account in our hedonic model for lack of information on house characteristics. New stringent norms of construction also play a role. Above, we also mentioned our doubt about the inclusion of geographic variables in the hedonic models for structure. They might capture the demand for land rather than construction costs.

Graph 20: Structure price index, IPLN (ECLN+EPTB) and IPLN (IP+IG)

Graph 21: The growth rates of the structure price index, IPLN (ECLN+EPTB) and IPLN (IP+IG)



To check the validity of our structure price index we compare it to the BT01 indice of construction cost (graph 22). The two profiles are strikingly similar. The rates of evolution differ by as much as 2 percentage points in 2008, but less than 1 in the other years. Even with a far from prefect hedonic model, we seem to recover a plausible structure price index from EPTB. What drives the construction costs evolution is left for future research.

Graph 22: The growth rates of the structure price index and BT01

6. Conclusion

The starting question 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 ECLN index and the Notaires-INSEE index differ. Such question provided the opportunity to ask other related questions on what makes the quality of a dwelling, as a construction, and as a location providing access to jobs and services and amenities.

First we looked for possible methodology bias. The Notaires-INSEE index was recomputed to mimic the more simple ECLN method (a single equation, less variables). The difference accounts for 1 to 2 percentage points in growth rate, less than the difference to be explained, especially at turning point of 2009.

Second, because new and old dwellings are not situated in the same areas, we tried another recomputation of the Notaires-INSEE index, keeping only the second-hand dwellings located in the same municipalities as the new dwellings at each date. This clone index is closer to the ECLN index, especially for apartments, but yet it does not fully explains the differences in evolution.

It led us to new theoretical reflections on house prices. The first questions the specification of the hedonic models when quality, and especially location quality, evolves over time. In a monocentric city model, the land gradient is negative as one goes further from the city center, but importantly for a builder of price index the gradient might get flatter over time if new amenities reach the periphery. Depending on the specification of the models this may be taken into account or not, when computing price indexes. The second strand of reflections goes back to how a new house is built when land and structure are paid for separately. We show that the price index of a house is the weighted sum of the price index of land and that of structure, where the weights are function of the share of land in the total value of the house. Then the higher the land leverage is, the higher the price index is. This would explain what we observe on French data, that the index for second hand dwellings is more volatile than the new dwellings index.

Such theoretical decomposition led us to use a new set of data on the construction of single family homes built individually on plots that have been purchased separately. Such rich data (the EPTB survey) allows building a price index for land, another for structure, one for the whole land + construction. They also provide land share and their evolution over time. We compared them to the construction costs index and all our former indexes. We even build an alternative comprehensive price index for new dwellings including the individually built houses, which represent a large share of new house constructions. The conclusion is that the price of land drives the price of dwellings more than the price of the structures. It is especially true for second-hand dwellings. However the evolution of the price index of structure is also striking, especially in its reaction to demand shocks during the crisis. More work is clearly to be done to explain this evolution.

References

Balcone, T.. 2013, Mémoire de master de statistique publique ENSAI.

Balcone, T., 2015, Élaboration d'un indice de prix des logement neuf dans le cadre de la nouvelle réglementation Européenne 93/2013, Document de travail de la DSDS (à paraître).

Bostic, R.C., Longhofer, S.D., Redfearn, C.L.2007, Land leverage: decomposing home price dynamics, *Real Estate Economics*, 35, 183-208.

Cheshire, P. and S. Sheppard, 1995, On the Price of Land and the Value of Amenities, *Economica*, New Series, 62, 246, 247-267

Clarenc, P., Côte, J-F., David, A., Friggit, J., Gallot, P., Gregoir, S., Laferrère, A., Nobre, A., Rougerie C. and N. Tauzin, 2014. *Les indices Notaires-Insee de prix des logements anciens*, INSEE Méthode, 128, 153 p..

http://www.insee.fr/fr/publications-et-services/sommaire.asp?reg_id=0&ref_id=IMET128

Combes, P.-P., G. Duranton and L. Gobillon, 2011, Le prix des terrains en France : la localisation, encore la localisation, toujours la localisation, *Le point sur*, 74, Observation et Statistiques, CGDD.

Davis M.A. and J. Heathcote, 2007. The Price and Quantity of Residential Land in the United States, *Journal of Monetary Economics*, 54, 2595-2620.

Davis M.A. et M.G. Palumbo, 2008. The Price of Residential Land in Large U.S. Cities *Journal of Urban Economics*, 63, 352-384.

Diewert, W. E., 2010, Alternative Approaches to Measuring House Price Inflation, WP, University of British Columbia

Diewert, W. E., Jan de Haan and Rens Hendriks, 2011, Hedonic Regressions and the Decomposition of a House Price index into Land and Structure Components, Discussion Paper 11-01, Department of Economics, The University of British Columbia.

Dipasquale, D. and C. Tsuriel Somerville, 1995. Do House Price Indices Based on Transacting Units Represent the Entire Stock ? Evidence from the American Housing Survey, *Journal of Housing Economics*, 4, 195-229.

Friggit, J., 2014, Évaluation du nombre de mutations de logements à titre onéreux et gratuit par statut d'occupation, Rapport n° 008187-02 du Conseil Général de l'Environnement et du Développement Durable.

Glaeser, E. M., Kolko, J. and A. Saiz, 2000, Consumer cities, NBER WP7790.

Goffette-Nagot, F., 2009. Prix fonciers et demande de sol à usage résidentiel en France (1975-2000), *Revue Économique*, 60, 853-862.

Gouriéroux, C. and A. Laferrère, 2009. Managing Hedonic Housing Price Indexes: the French Experience, *Journal of Housing Economics*, 18, 206-213.

Hill, R., 2004. Constructing Price Indexes across Space and Time: The Case of the European Union, *The American Economic Review*, 94, 5, 1379-1410.

Hill, R. J. and M. Scholz, 2013, Incorporating Geospatial Data into House Price Indexes: A Hedonic Imputation Approach with Splines, mimeo.

Laferrère, A., 2007. L'héritage de logement en France. Faits et théories, in *Vieillissement de la population et logement, Les stratégies résidentielles et patrimoniales*, La Documentation Française, 403-415.

Laferrère, A., 2012. La construction d'un indice spatial, Actes du colloque "Les marchés du logement, un nouveau cycle", 7 décembre 2011, Conseil supérieur du notariat.

Linneman, P., 1980. Some Empirical Results on the Nature of the Hedonic Price Function for the Urban Housing Market, JUE, 8, 47-68.

Oikarinen, E;, 2009, Dynamic linkages between housing and lot prices: Empirical evidence for Helsinki. Disc. paper, 53. Published in *Urban Studies*, 2013.

Rapport de la Commission des Comptes du Logement, Compte du Logement 2013, premiers résultats 2014, 2013. SoES, ministère de l'Écologie, du Développement durable et de l'Énergie.

Richmond, P. and B. Roehner, 2012. The predictable outcome of speculative house price peaks, mimeo.

Roback, J., 1982. Wages, Rents and the Quality of Life, *Journal of Political Economy*, 90, 61, 1257-1278.

Rosenthal, S . S. 1999, Residential Buildings and the cost of construction: new evidence on the efficiency of the housing market, *Review of Economics and Statistics*, 81, 288-302.

Appendix 1. Distribution of the observations in EPTB according to the fourth condition

Excluding observations for which the purchased plot has not the same size as that registered in the building permit leads to keeping 90,2% of the observations of metropolitan France (excluding Corsica) between 2006 and 2012.

Building permit surface	2006	2007	2008	2009	2010	2011	2012	Mean
= purchased land surface	87,50%	87,96%	89,22%	91,00%	91,94%	91,00%	92,94%	90,22%
> purchased land surface	6,81%	6,55%	5,44%	4,87%	4,14%	4,65%	3,64%	5,16%
< purchased land surface	5,69%	5,49%	5,34%	4,13%	3,92%	4,35%	3,43%	4,62%
All	100,00%	100,00%	100,00%	100,00%	100,00%	100,00%	100,00%	100,00%

The annual distribution of the EPTB sample is the following

Distribution of the observations in EPTB used

Year	Number of observations	Percentage
2006	61 612	17,87%
2007	57 021	16,54%
2008	42 808	12,41%
2009	27 814	8,07%
2010	34 514	10,01%
2011	52 224	15,14%
2012	68 854	19,97%
All	344 847	100,00%

Appendix 2. Weights for the land price index

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

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

Region	2006[44]-2007	2008	2009	2010	2011	2012	Mean
ILE-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-COMTE	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-PYRENEES	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%
RHONE-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-COTE 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%

[44] It is the same weights as for 2007 because 2006 is the first year for which data are available

Appendix 3. Weights for the structure price index

The weights $w\_S^{r}_{A-1}$ are the share of the construction expenses in region r in the total expenses of year A-1 :

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

Region	2006[45]-2007	2008	2009	2010	2011	2012	Mean
ILE-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-COMTE	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-PYRENEES	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%
RHONE-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-COTE 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%

[45] It is the same weights as for 2007 because 2006 is the first year for which data are available

Appendix 4. Weights for the new single houses price index

$w\_V^{r}_{A-1}$ are the share of the expenses for single family units in region r in the total expenses of year A-1 :

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

Region	2006[46]-2007	2008	2009	2010	2011	2012	Mean
ILE-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-COMTE	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-PYRENEES	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%
RHONE-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-COTE 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%

[46] It is the same weights as for 2007 because 2006 is the first year for which data are available

Appendix 5. Land share α in the total price of a house by region

The land share is defined as the ratio of total land price to total price of "land plus structure" in each region.

Region	2006[47]	2007	2008	2009	2010	2011	Mean
ILE-DE-FRANCE	49,46%	49,61%	48,55%	49,25%	46,07%	47,26%	48,37%
CHAMPAGNE-ARDENNE	24,54%	26,46%	27,02%	27,54%	27,07%	28,65%	26,88%
PICARDIE	28,78%	30,51%	30,45%	30,41%	30,86%	32,05%	30,51%
HAUTE-NORMANDIE	27,68%	31,22%	33,10%	31,81%	32,23%	31,58%	31,27%
CENTRE	28,35%	29,61%	29,91%	29,83%	30,26%	30,07%	29,67%
BASSE-NORMANDIE	22,68%	23,80%	24,40%	25,11%	25,63%	27,17%	24,80%
BOURGOGNE	24,16%	24,94%	26,69%	25,86%	27,11%	27,68%	26,07%
NORD-PAS-DE-CALAIS	30,46%	32,21%	32,21%	32,28%	31,16%	31,99%	31,72%
LORRAINE	24,91%	24,15%	25,73%	26,19%	26,36%	27,25%	25,77%
ALSACE	30,58%	31,36%	32,95%	31,27%	31,27%	32,01%	31,57%
FRANCHE-COMTE	21,58%	22,11%	23,70%	24,27%	23,96%	24,83%	23,41%
PAYS DE LA LOIRE	28,70%	29,59%	30,96%	29,51%	28,82%	29,38%	29,49%
BRETAGNE	25,85%	27,07%	28,22%	27,52%	26,66%	27,18%	27,08%
POITOU-CHARENTES	24,54%	25,56%	25,93%	24,97%	25,55%	26,70%	25,54%
AQUITAINE	31,89%	33,06%	33,72%	33,70%	34,16%	34,58%	33,52%
MIDI-PYRENEES	32,17%	32,60%	32,15%	31,87%	32,59%	32,14%	32,25%
LIMOUSIN	16,93%	18,15%	19,27%	19,78%	21,18%	20,98%	19,38%
RHONE-ALPES	38,74%	40,88%	41,20%	39,55%	38,02%	38,58%	39,50%
AUVERGNE	21,77%	23,04%	23,74%	23,40%	25,06%	24,59%	23,60%
LANGUEDOC-ROUSSILLON	43,12%	43,61%	45,64%	42,96%	41,98%	41,32%	43,11%
PROVENCE-ALPES-COTE D'AZUR	48,82%	48,79%	50,50%	47,52%	47,76%	48,44%	48,64%
Mean	29,80%	30,87%	31,72%	31,17%	31,13%	31,64%	31,05%

Appendix 6 . Structure share *(1- α)* in the total price of a house by region

Region	2006	2007	2008	2009	2010	2011	Mean
ILE-DE-FRANCE	50,54%	50,39%	51,45%	50,75%	53,93%	52,74%	51,63%
CHAMPAGNE-ARDENNE	75,46%	73,54%	72,98%	72,46%	72,93%	71,35%	73,12%
PICARDIE	71,22%	69,49%	69,55%	69,59%	69,14%	67,95%	69,49%
HAUTE-NORMANDIE	72,32%	68,78%	66,90%	68,19%	67,77%	68,42%	68,73%
CENTRE	71,65%	70,39%	70,09%	70,17%	69,74%	69,93%	70,33%
BASSE-NORMANDIE	77,32%	76,20%	75,60%	74,89%	74,37%	72,83%	75,20%
BOURGOGNE	75,84%	75,06%	73,31%	74,14%	72,89%	72,32%	73,93%
NORD-PAS-DE-CALAIS	69,54%	67,79%	67,79%	67,72%	68,84%	68,01%	68,28%
LORRAINE	75,09%	75,85%	74,27%	73,81%	73,64%	72,75%	74,24%
ALSACE	69,42%	68,64%	67,05%	68,73%	68,73%	67,99%	68,43%
FRANCHE-COMTE	78,42%	77,89%	76,30%	75,73%	76,04%	75,17%	76,59%
PAYS DE LA LOIRE	71,30%	70,41%	69,04%	70,49%	71,18%	70,62%	70,51%
BRETAGNE	74,15%	72,93%	71,78%	72,48%	73,34%	72,82%	72,92%
POITOU-CHARENTES	75,46%	74,44%	74,07%	75,03%	74,45%	73,30%	74,46%
AQUITAINE	68,11%	66,94%	66,28%	66,30%	65,84%	65,42%	66,48%
MIDI-PYRENEES	67,83%	67,40%	67,85%	68,13%	67,41%	67,86%	67,75%
LIMOUSIN	83,07%	81,85%	80,73%	80,22%	78,82%	79,02%	80,62%
RHONE-ALPES	61,26%	59,12%	58,80%	60,45%	61,98%	61,42%	60,51%
AUVERGNE	78,23%	76,96%	76,26%	76,60%	74,94%	75,41%	76,40%
LANGUEDOC-ROUSSILLON	56,88%	56,39%	54,36%	57,04%	58,02%	58,68%	56,90%
PROVENCE-ALPES-COTE D'AZUR	51,18%	51,21%	49,50%	52,48%	52,24%	51,56%	51,36%
Mean	70,20%	69,13%	68,28%	68,83%	68,87%	68,36%	68,95%

[47] These are the weights used to compute the aggregated new single houses price index of 2006 and 2007

A**ppendix 7. Indexes from EPTB computed with the time dummy method in a single national model**

We also tested the robustness of the results to the index method by using a pure time dummy method. We simplify the hedonic method and estimate only three models. All include year dummies, region dummies interacted when necessary with the 3 location qualities (seaside, estuary, arrière-pays, mountain), dummies for city size (center, suburbs, isolated town), and distance to nearest urban center (pôle) (0 if the location is within the urban center). Hence the variables are the same as those used in the adjacent period time dummy method used in section 5. We use the calibrated sample weights.

The price of land per m² of plot size is also regressed on logarithm of plot size, and characteristics of the plot (whether if was bought though an agency, the builder, or other means; its servicing), to which we add the SHON of the house (the rational is that the house size is linked to the quality of the land).

The price of the house construction is also regressed on logarithm of SHON, and construction characteristics (degree of finish - ready to inhabit or not-, heating - electricity, wood, renewable, other - use of an architect, self-built, type of builder…).

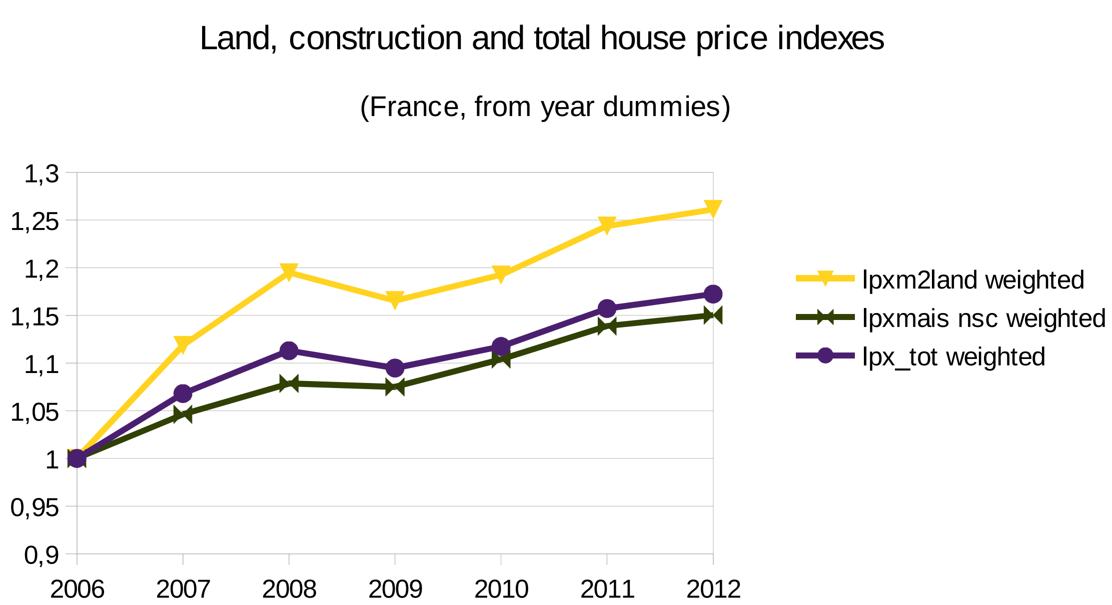
The total price of house and land is finally regressed on the logarithm of SHON, logarithm of plot size and construction and plot characteristics.

Table A71. Models of land, construction and total house price (France)

	(1)	(2)	(3)	(4)
VARIABLES	lpxm2land weighted	lpxmais weighted	lpxmais Nsc weighted	lpx_tot weighted
_Ian_2007	0.112***	0.0473***	0.0452***	0.0658***
	(0.00283)	(0.00148)	(0.00148)	(0.00138)
_Ian_2008	0.178***	0.0790***	0.0757***	0.107***
	(0.00298)	(0.00157)	(0.00157)	(0.00146)
_Ian_2009	0.153***	0.0772***	0.0723***	0.0905***
	(0.00315)	(0.00166)	(0.00166)	(0.00155)
_Ian_2010	0.176***	0.104***	0.0990***	0.111***
	(0.00299)	(0.00162)	(0.00162)	(0.00151)
_Ian_2011	0.218***	0.136***	0.130***	0.146***
	(0.00304)	(0.00165)	(0.00164)	(0.00153)
_Ian_2012	0.232***	0.148***	0.140***	0.159***
	(0.00298)	(0.00161)	(0.00161)	(0.00150)
dist	-0.0192***	-0.000535***	-0.000213***	-0.00537***
	(9.77e-05)	(5.12e-05)	(5.07e-05)	(4.77e-05)
lsurf	-0.920***	0.0380***		0.0607***
	(0.00180)	(0.000910)		(0.000889)
lshon	0.543***	0.863***	0.888***	0.751***
	(0.00353)	(0.00193)	(0.00185)	(0.00181)
Constant	7.393***	7.060***	7.212***	7.830***
	(0.0186)	(0.0107)	(0.0101)	(0.0103)
Nb obs	371,367	371,367	371,367	371,367
R^2	0.724	0.506	0.504	0.580

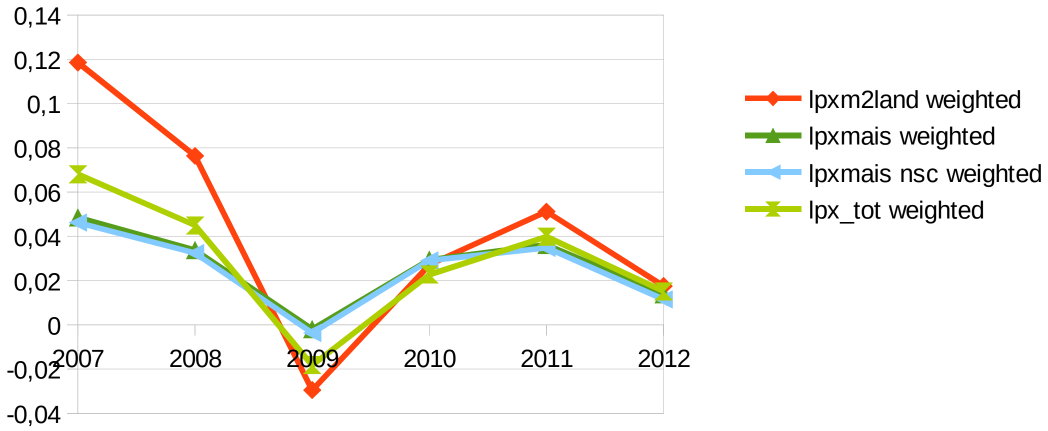
Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. All models include the variables discussed above. Model (3) does not include the land plot size. Data source : EPTB.

The price index is computed from the year dummies estimated coefficients, taking the exponential. For an increase of one km in *dist* the price per m² of the plot decreases by 1.9%. This compares with 2.7% in Gofette-Nagot (2009).

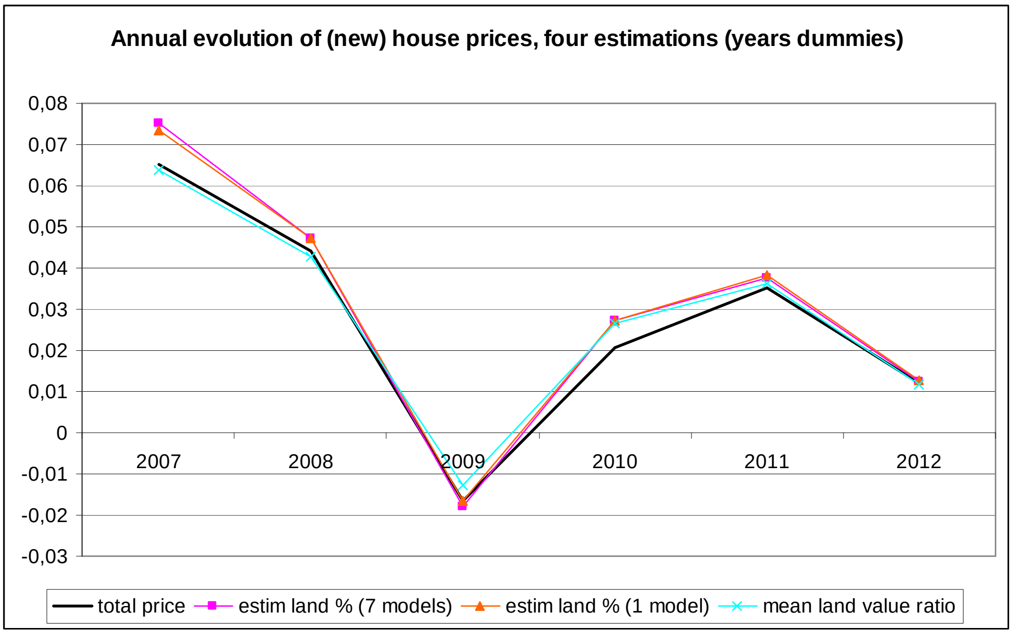


Yearly evolution of price index of land, construction and total house

(France, from year dummies models)



We also compared the total price evolution obtained with the year dummies method described above, with what is obtained by weighting the two sub-indices of land and construction price to build a house price index, using the ratio of plot value to total house price as weight. This ratio is itself obtained with three method: using the simple mean ratio computed each year; using the ratio from a hedonic regression of total price on plot price, either ran each year, or from a single model with year dummies interactions.



The indices evolution differ by a maximum of 1 percentage point. Weights used in aggregation clearly influence the indice.

Table A72. Model of the land share in the totel price of houses (Metropolitan France)

VARIABLES	(1) px_land
2006b.an#c.px_tot	0.456***
	(0.000773)
2007.an#c.px_tot	0.456***
	(0.000743)
2008.an#c.px_tot	0.464***
	(0.000767)
2009.an#c.px_tot	0.447***
	(0.000903)
2010.an#c.px_tot	0.437***
	(0.000862)
2011.an#c.px_tot	0.440***
	(0.000743)
2012.an#c.px_tot	0.439***
	(0.000678)
lshon	-36,097***
	(196.6)
lsurf	4,060***
	(89.00)
distc	-255.5***
	(1.972)
Constant	104,885***
	(1,075)
Observations	371,367
R-squared	0.790

Standard errors in parentheses, *** p<0.01, ** p<0.05, * p<0.1