

Chapitre 3 : Calculs mis en œuvre

L'indice élémentaire est défini comme le rapport de la valeur courante d'un parc de logements de référence à sa valeur de la période de base. La méthode décrite en détail dans ce chapitre est la même pour tous les indices, sous deux réserves :

- les variables représentatives de la qualité des biens (caractéristiques) ne sont pas exactement identiques pour les appartements et pour les maisons, et selon que l'on considère l'Île-de-France ou la province ;
- la manière de découper le champ géographique en strates varie selon que l'on considère l'Île-de-France ou la province.

A l'occasion du passage de la version 2 à la version 3, la méthode de calcul des indices a été modifiée. Afin de mieux prendre en compte les évolutions du marché, les parcs d'estimation et de référence (cf. définitions ci-dessous) seront actualisés plus fréquemment (tous les deux ans). Conséquence de ce changement, les indices sont désormais chaînés. Trois formules de chaînage sont mises en œuvre : la première concerne les indices élémentaires et les deux autres les indices agrégés.

3.1 Parcs de référence et d'estimation

Définition 1 : le parc et la période de référence

Le parc de référence constitue le « panier de biens » ou « portefeuille » dont on va mesurer la variation de prix. On s'assure ainsi que l'indice retrace bien l'évolution du prix des mêmes logements et qu'il n'est pas sensible aux variations de la structure du marché.

Il s'agit donc de valoriser les logements du parc de référence à la date courante, alors qu'ils n'ont pas été vendus, c'est-à-dire que l'on n'a pas pu observer leur prix. Pour cela, on utilise un modèle économétrique qui relie le prix d'un logement¹⁹ à ses caractéristiques physiques et à sa localisation.

Le *parc de référence* est constitué de toutes les transactions d'une période - dite *période de référence* -, entrant dans le champ de l'indice, sauf celles situées aux extrêmes de la distribution des valeurs²⁰.

Dans la version 3 des indices, la période de référence est constituée des années $n-3$ et $n-2$ pour les indices des années n et $n+1$ (avec n pair), puisque le parc est mis à jour tous les deux ans à partir de 2008 (voir plus loin 3.5.1).

Définition 2 : le parc et la période d'estimation

Le parc d'estimation regroupe les transactions qui serviront à l'estimation des modèles. Il est constitué de toutes les transactions de la période d'estimation entrant dans le champ de l'indice, sauf celles pour lesquelles la valeur estimée par le modèle s'écarte de la valeur réelle de plus de deux écarts-types. Les parcs de référence et d'estimation sont découpés en zones élémentaires de calcul ou strates, au nombre de 22 en Île-de-France et 271 en province (tableau 3.2)²¹.

En province, certaines strates sont composées d'une seule ville, lorsque celle-ci possède un marché immobilier de taille suffisante (en pratique, plus de 110 transactions par trimestre). Dans la plupart des strates, les modèles peuvent faire apparaître, via des indicatrices, des zones plus détaillées de prix homogènes. Les autres strates de province sont définies par l'ensemble des villes appartenant au croisement des critères : région, département, unité urbaine ou d'autres distinctions géographiques (espace rural, communes balnéaires ou massifs alpins), ou encore selon le revenu moyen par habitant (année 2006) pour d'autres communes.

Les strates d'Île-de-France sont définies sur des critères géographiques et de prix²².

¹⁹ Plus exactement le logarithme du prix.

²⁰ Il s'agit du prix au mètre carré dans le cas des appartements et du prix total en ce qui concerne les maisons. On rabote $1/20^{\text{ème}}$ à chaque extrémité de la distribution.

²¹ Pour la constitution des strates, voir annexe 2.

²² Pour le détail des zones de calcul, voir l'annexe 2, page 47.

3.2 Le modèle servant à l'estimation

Un modèle économétrique est estimé pour chaque zone (strate), donc sur un marché supposé homogène, à partir des transactions du parc d'estimation.

Le modèle est de la forme²³ :

$$\log p_i = \log p_0 + \sum_{a=1}^2 \mu_a Y_{a,i} + \sum_{m=1}^{12} \theta_m M_{m,i} + \sum_{k=1}^K \beta_k X_{k,i} + \varepsilon_i \quad (3.1)$$

avec les notations suivantes :

p_i : prix au m^2 (pour les appartements) ou prix total (pour les maisons) du bien i

p_0 : prix du logement de référence du parc d'estimation

$Y_{a,i}$: indicatrice de l'année de mutation du bien i

$M_{m,i}$: indicatrice du mois de mutation du bien i

$X_{k,i}$: caractéristiques physiques du bien i , indexées par $k=1$ à K

Les $X_{k,i}$ sont des variables calculées à partir des modalités des caractéristiques initiales disponibles, Z_i , soit l'époque de construction, le nombre de pièces, le nombre de salles de bains, le nombre de garages ou parkings, l'état du bien ; et de plus, pour les appartements seulement, l'étage, la présence d'un balcon, d'une cave, d'un ascenseur, la surface moyenne par pièce ; et pour les maisons seulement, la présence d'un sous-sol, le nombre de niveaux, la surface du terrain et la surface habitable...²⁴

Le cas échéant, la localisation du logement à l'intérieur de la strate (quartier), est considérée comme une de ses caractéristiques, au même titre que le nombre de pièces ou sa surface.

Les $X_{k,i}$ peuvent inclure des effets croisés. Elles peuvent être des variables indicatrices, ou des variables continues dans le cas des surfaces. Chacune des caractéristiques a un prix, évalué à la période de base, par rapport à celles d'un bien de référence.

Définition 3 : le logement de référence du parc d'estimation

Le *logement de référence* est celui qui possède les caractéristiques prises comme références dans les régressions (tableau 3.1).

Les caractéristiques non observables ou inobservées des logements sont supposées non corrélées avec ces caractéristiques de qualité observées. Le marché des maisons et celui des appartements sont supposés distincts : par exemple, on suppose qu'une pièce supplémentaire peut ne pas valoir le même prix en appartement et en maison. Mais elle vaut le même prix dans un appartement, quel que soit l'étage par exemple. Dans les strates, les prix relatifs des caractéristiques sont fixes. En d'autres termes, si un appartement au troisième étage vaut x% de plus qu'au rez-de-chaussée, ce rapport est le même pour tous les appartements de la strate de calcul. En revanche, il est susceptible d'être différent d'une strate à l'autre.

²³ Pour simplifier les notations, on omet l'indice s de la strate.

²⁴ Toutes ces variables ne sont pas présentes dans chacune des deux bases ni toujours exploitées dans les modèles. Pour le détail des modèles mis en œuvre, voir les chapitres 3 et 4.

Tableau 3.1 - Caractéristiques du logement de référence (parc d'estimation)²⁵

Indice	Logement de référence en Île-de-France	Logement de référence en province
Appartements	<ul style="list-style-type: none"> - 3 pièces, - rez-de-chaussée, - surface moyenne par pièce pour les studios comprise entre 20 et 30 m², pour les deux pièces entre 17 et 24 m², pour les trois pièces entre 18 et 22 m² et pour les quatre pièces et plus entre 17 et 21 m², - sans garage, - sans cave, - 1 salle de bains, - construit entre 1948 et 1969, - vendu le 12^e mois de la deuxième année de la période de référence, 	
		<ul style="list-style-type: none"> - sans terrasse ou balcon**, - en bon état**.
Maisons	<ul style="list-style-type: none"> - 4 pièces, - 2 niveaux, - époque de construction inconnue, - 1 garage, - 1 salle de bains, - 100 m² de surface habitable et 610 m² de surface de terrain, - vendue le 12^e mois de la deuxième année de la période de référence, 	
	<ul style="list-style-type: none"> - 1 bâtiment*, - pas de cave*. 	<ul style="list-style-type: none"> - pas de sous-sol**, - état du bien inconnu**,

* Variable non présente dans la base en province

** Variable non présente dans la base en Île-de-France

Le plus souvent, le modèle fait intervenir une variable indicatrice de quartier à l'intérieur de la strate. Dans ce cas, le logement de référence appartient par construction au quartier omis dans la régression. Le quartier est introduit lorsque la strate concerne une agglomération ou une ville, et que les volumes de transactions sont suffisants (au moins 110 transactions par an). Les quartiers ne correspondent pas nécessairement à des découpages administratifs.

Le tableau 3.2 présente le nombre de strates de la version 3 : 22 en Île-de-France contre 62 en version 2, 271 en province contre 234 en version 2. Le nombre de quartiers pris en compte a été réduit en Île-de-France (de 230 à 102) et augmenté en province (de 1125 à 1509).

Tableau 3.2 - Nombre de strates élémentaires de calcul et nombre de quartiers

Indice	Nombre de strates	Nombre de quartiers
<i>Île-de-France</i>	22	102
Appartements	15	66
Maisons	7	36
<i>Province</i>	271	1 509
Appartements	97	516
Maisons	174	993
Total	293	1 611

3.3 Prix courant du logement de référence

Le même type de modèle peut être utilisé en période courante. Le bien de référence possède les caractéristiques décrites au paragraphe précédent, mais correspond à la période de mutation courante t . Son prix est noté $p_{0,t}$.

L'évolution du prix du bien de référence est le fondement de la construction de l'indice. Pour cette raison, elle doit incorporer les effets saisonniers, les tendances et les cycles. Ceci explique pourquoi les coefficients saisonniers et annuels ne figurent plus dans le modèle courant.

²⁵ Les variables sont décrites au chapitre 4, paragraphe 4.4.2.

Le prix au m² d'un bien j vendu au cours d'une période t peut s'exprimer à l'aide du modèle :

$$\log(p_{j,t}) = \log(p_{0,t}) + \sum_{k=1}^K \beta_{k,t} X_{k,j,t} + \varepsilon_{j,t}$$

On remarquera que le modèle (3.1) utilisé dans la phase d'estimation est compatible avec ce dernier modèle, utilisé pour les valeurs courantes. L'introduction des indicatrices d'années et de mois vient du fait que le parc d'estimation couvre une période de plusieurs mois. De façon équivalente, le prix de référence pour le parc d'estimation et le mois (a,m) serait :

$$\log(p_{0,a,m}) = \log p_0 + \mu_a + \theta_m$$

Il faut maintenant expliquer comment reconstituer le prix du bien de référence à partir des données de transactions courantes.

Plaçons-nous d'abord dans l'hypothèse où les coefficients $\beta_{k,t}$ sont connus et introduisons $\tilde{p}_{j,t}$ le « prix équivalent bien de référence » du bien j,t :

$$\log(\tilde{p}_{j,t}) = \log(p_{j,t}) - \sum_{k=1}^K \beta_{k,t} X_{k,j,t}$$

Le modèle peut être réécrit :

$$\log(\tilde{p}_{j,t}) = \log(p_{0,t}) + \varepsilon_{j,t}$$

Donc, si on connaît les coefficients $\beta_{k,t}$ le logarithme du prix du bien de référence $\log(p_{0,t})$ sera estimé par les moindres carrés ordinaires, c'est-à-dire par la moyenne :

$$\log(\hat{p}_{0,t}) = \frac{1}{J_t} \sum_{j=1}^{J_t} \log(\tilde{p}_{j,t})$$

où J_t est le nombre de transactions de la période t .

Si on suppose que le modèle estimé sur la période de référence est stable dans le temps (pour la période du calcul d'un indice, qui est maintenant de deux ans), on peut remplacer dans le calcul des prix équivalents bien de référence, les coefficients $\beta_{k,t}$ par les valeurs $\hat{\beta}_k$ estimées sur la période de référence :

$$\log(\tilde{p}_{j,t}) \cong \log(p_{j,t}) - \sum_{k=1}^K \hat{\beta}_k X_{k,j,t} = \log \left[\frac{p_{j,t}}{\exp \left(\sum_{k=1}^K \hat{\beta}_k X_{k,j,t} \right)} \right]$$

On estime alors le prix au m² du bien de référence à la période t , $p_{0,t}$, par une moyenne géométrique des prix équivalents bien de référence des J_t biens vendus au cours de la période t :

$$\log \hat{p}_{0,t} = \frac{1}{J_t} \sum_{j=1}^{J_t} \log \tilde{p}_{j,t} = \frac{1}{J_t} \log \left(\prod_{j=1}^{J_t} \tilde{p}_{j,t} \right)$$

soit :

$$\hat{p}_{0,t} = \left(\prod_{j=1}^{J_t} \tilde{p}_{j,t} \right)^{\frac{1}{J_t}}$$

Dans la suite nous notons $\hat{\alpha}_{0,t} = \log \hat{p}_{0,t}$.

3.4 Valeur courante du parc de référence

Connaissant une valeur approchée du bien de référence (donnée par l'estimation du prix équivalent du logement de référence à la période t), on peut alors reconstituer les valeurs approchées des biens figurant dans le parc de référence et, par agrégation, la valeur du parc lui-même. Les calculs sont menés strate par strate. Pour cette raison, nous réintroduisons l'indice s de la strate.

On estime la valeur d'un bien i du parc de référence de la strate s à la période courante t à partir de ses caractéristiques $X_{k,i,s}$, qui, rappelons-le, sont indépendantes de la date t , par définition même du parc de référence dont la composition est stable dans le temps²⁶.

Dans le cas des appartements, par exemple, la valeur approchée est :

$$\hat{p}_{i,s,t} = \exp \left(\hat{\alpha}_{0,s,t} + \sum_{k=1}^K \hat{\beta}_{k,s} X_{k,i,s} \right) A_{i,s}$$

où $A_{i,s}$ désigne la surface du bien i,s .

En sommant la valeur courante estimée des N_s biens du parc de référence de la strate s , on obtient la valeur du parc de référence à la date t :

$$\hat{W}_{s,t} = \sum_{i=1}^{N_s} \hat{p}_{i,s,t} = \sum_{i=1}^{N_s} \exp \left(\hat{\alpha}_{0,s,t} + \sum_{k=1}^K \hat{\beta}_{k,s} X_{k,i,s} \right) A_{i,s}$$

On estime de la même façon, la valeur du parc de référence dans la strate s , à la période 0 , dite *période de base* de l'indice :

$$\hat{W}_{s,0} = \sum_{i=1}^{N_s} \exp \left(\hat{\alpha}_{0,s,0} + \sum_{k=1}^K \hat{\beta}_{k,s} X_{k,i,s} \right) A_{i,s}$$

3.5 La méthode de calcul

Trois changements majeurs sont introduits par rapport à la version précédente des indices.

- Les parcs de référence et d'estimation sont désormais actualisés tous les deux ans ; les vecteurs de prix relatifs des caractéristiques, et donc la valeur des parcs de référence à la période de base, sont réestimés tous les deux ans également. Le Conseil scientifique des indices Notaires-Insee²⁷ en a décidé ainsi après avoir constaté que, sur la période initialement envisagée comme nouvelle référence (2002-2006), les modèles hédoniques n'étaient pas stables²⁸. Du fait de l'actualisation régulière des coefficients et des pondérations, une méthode de chaînage est introduite. La spécification des modèles et la définition des strates devraient continuer à être revues tous les six ans environ.

- L'introduction d'indicatrices mensuelles dans l'équation (3.1) permet désormais de calculer des indices mensuels. PNS et Min.not calculent ces indices mais seul PNS les publie.

- En version 2, on utilisait des moyennes arithmétiques pour agréger les indices élémentaires. En version 3, on utilise des moyennes géométriques pour les niveaux infra-départementaux et départementaux. Ce mode de calcul est utilisé dans la construction des indices statistiques lorsque l'on pense qu'il peut exister un arbitrage entre différents biens de nature voisine. Dans notre cas, il peut y

²⁶ La qualité des logements ne varie donc pas.

²⁷ La composition du Conseil scientifique des indices Notaires-Insee, chargé de la surveillance des indices, est décrite dans les conventions passées avec le notariat. Cf. annexes 8 et 9.

²⁸ Voir l'annexe 4, Tests de stabilité et durée de la période de référence.

avoir un choix d'acquisition entre des biens appartenant à des strates géographiques voisines. On utilise des moyennes arithmétiques pour les niveaux plus agrégés que les niveaux départementaux car l'acheteur a moins de chance d'arbitrer entre des strates géographiques plus éloignées pour s'installer (région, province, France entière).

3.5.1 Renouvellement des parcs de référence et d'estimation tous les deux ans

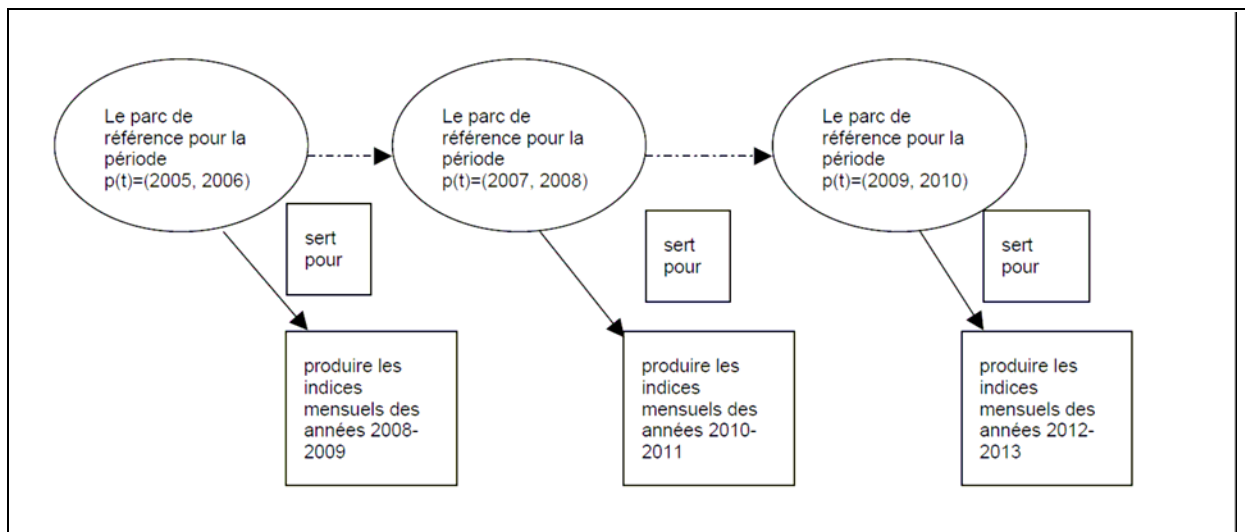
Les parcs de référence

Dans la version 1 des indices, le parc de référence incluait les transactions de trois à cinq années selon les indices (1994-96 pour les appartements de province, 1994-97 pour les maisons, et 1992-96 pour les appartements de Paris et Petite Couronne). Dans la version 2, le parc contenait les transactions des quatre années 1998-2001.

Dans la version 3, les parcs de référence sont remis à jour tous les deux ans. La période de référence $p(t)$ est également fixée à deux années de transactions, donc plus courte qu'auparavant. Elle est définie par rapport au début de la période courante de production des indices de la façon suivante : pour la production courante des indices sur le couple d'année $(n, n+1)$ avec n pair, $p(t)$ débute deux ans avant, soit $p(t) = (n-3, n-2)$ ²⁹.

La figure 3.1 permet de visualiser la succession des différents parcs de référence utilisés pour le calcul des indices mensuels publiés sur les années 2008-2013.

Figure 3.1 - les parcs de référence de la période 2008-2013



Période de base et base de publication

Les parcs de référence sont valorisés au dernier trimestre $q(t)$ de la période de référence, donc au quatrième trimestre de l'année paire $n-2$ pour les indices diffusés sur la période $(n, n+1)$ ³⁰. Les poids servant à l'agrégation des indices élémentaires sont calculés à partir de cette valorisation du parc.

Le trimestre de valorisation du parc de référence est appelé *période de base* des indices sur la période $(n, n+1)$. Il ne doit pas être confondu avec la *base 100* de la série publiée, qui a été fixée au premier trimestre 2010.

Notations

La formule de calcul de la valeur du parc de référence doit être généralisée pour tenir compte du renouvellement du parc de référence tous les deux ans par l'introduction d'un indicage supplémentaire, - noté $p(t)$, le parc de référence utilisé - : la valeur estimée du parc devient $\widehat{W}_{s,t}^{p(t)}$. De même, on introduit la valeur du parc de

²⁹ Par exemple, la première période courante de production des indices est le couple d'années 2008 et 2009 et $p(t) = (2005, 2006)$. Au moment de l'estimation, $n-1$ n'est pas disponible.

³⁰ Le quatrième trimestre 2006 sera donc par exemple le trimestre de valorisation du parc de référence 2005-2006.

référence à la période de base $q(t)$, servant au chaînage des indices sur la période de référence $p(t)$. Selon les étapes, elle sera notée $\hat{W}_{s,q(t),t}^{p(t)}$ ou plus simplement, si le contexte le permet : $\hat{W}_{s,t}$ ou $\hat{W}_{s,q(t),t}$ selon les cas.

En période courante, la valeur du parc de référence devient :

$$\hat{W}_{s,t} = \sum_{i=1}^{N_s} \exp\left(\hat{\alpha}_{0,s,t} + \sum_{k=1}^K \hat{\beta}_{k,s} X_{k,i,s}\right) A_{i,s}$$

avec les mêmes notations qu'au paragraphe 3.4 :

$\hat{W}_{s,t}$: la valeur courante du parc de référence dans la strate s ,

$\hat{\alpha}_{0,s,t}$: l'estimation du logarithme du prix au m² du bien de référence de la strate s à la date t ; les coefficients sont estimés sur la période $p(t)$,

N_s : le nombre de transactions de la strate s au cours de la période de référence,

$\sum_{k=1}^K \hat{\beta}_{k,s} X_{k,i,s}$: la correction à appliquer à l'estimation du logarithme du prix du bien de référence de la strate s à

la date t , pour estimer à cette même date le logarithme du prix du bien i de la strate s échangé au cours de la période de référence,

$A_{i,s}$: la surface du bien i de la strate s échangé au cours de la période de référence.

3.5.2 Calcul des indices élémentaires en période courante

Calcul des évolutions de prix

Les prix « équivalents bien de référence » sont calculés à l'aide du vecteur de prix des caractéristiques, comme indiqué au paragraphe 3.3. A l'aide de ces prix estimés, on valorise le parc de référence au trimestre t .

L'introduction d'indicatrices mensuelles dans les modèles (3.1) à la place des indicatrices trimestrielles autorise désormais la publication d'indices mensuels. Cependant on ne dispose pas de suffisamment d'observations pour calculer des indices mensuels « purs ». Les transactions mises à contribution dans ce calcul restent relatives à un trimestre. L'indice mensuel est calculé avec les données du dernier trimestre disponible incluant ce mois : il s'agit d'un indice mensuel en trimestre « glissant ».

Chaînage

L'évolution des prix entre les trimestres $t-1$ et t est obtenue en rapportant la valeur du parc de référence du trimestre t à la valeur de parc de référence du trimestre $t-1$ observé trois mois plus tôt ; t inclut les mois $m-2$, $m-1$ et m et $t-1$ inclut les mois $m-5$, $m-4$ et $m-3$. La formule de calcul de l'évolution des prix entre les trimestres $t-1$ et t est donc la suivante :

$$I_{t/0}(s) = \frac{\hat{W}_{s,t}^{p(t)}}{\hat{W}_{s,t-1}^{p(t)}} \times I_{t-1/0}(s)$$

avec t , le trimestre glissant et 0 le trimestre de référence.

L'évolution des prix entre les trimestres 0 et t est obtenue en chaînant les évolutions des valeurs du parc de référence entre 0 et t , selon la formule :

$$I_{t/0}(s) = \left[\prod_{u=1}^t \frac{\hat{W}_{s,u}^{p(u)}}{\hat{W}_{s,u-1}^{p(u)}} \right]$$

L'introduction de ce chaînage est rendue nécessaire par l'actualisation du parc de référence tous les deux ans, qui constitue une nouveauté par rapport à la précédente version des indices. En version 3, les indices antérieurs à 2008 ont été calculés en appliquant les évolutions des indices de la version 2.

Mise en œuvre du chaînage

Les indices mensuels étant calculés en trimestre glissant, il faut une base de début de chaînage pour les trois séries d'indices élémentaires correspondant aux premier, deuxième et troisième mois des trimestres civils. La série d'indices commence en janvier 2008. La base de début de chaînage est fixée à octobre, novembre et décembre 2007.

On calcule donc trois valeurs de départ pour permettre le calcul des trois premiers indices (ceux de janvier, février et mars 2008). On estime la valeur du parc de référence $p(t)=(2007,2008)$ au mois d'octobre 2007 (trimestre glissant composé des ventes d'août, septembre et octobre 2007), au mois de novembre et au mois de décembre 2007. L'indice d'octobre 2007 est ensuite fixé à 100. Celui de novembre est obtenu en rapportant la valeur en novembre du parc de référence $p(t)=(2007,2008)$ à sa valeur en octobre. Celui de décembre est obtenu en rapportant la valeur en décembre du parc de référence $p(t)=(2007,2008)$ à sa valeur en octobre (tableau 3.3).

A titre d'exemple, le calcul des indices de janvier 2008, décembre 2009 et janvier 2010 d'une strate s sera effectué de la façon suivante :

$$I_{janvier2008/0}(s) = \frac{\hat{W}_{s,janvier2008}^{parc2005-2006}}{\hat{W}_{s,octobre2007}^{parc2005-2006}} * I_{octobre2007/0}(s)$$

$$I_{décembre2009/0}(s) = \frac{\hat{W}_{s,décembre2009}^{parc2005-2006}}{\hat{W}_{s,septembre2009}^{parc2005-2006}} * I_{septembre2009/0}(s)$$

$$I_{janvier2010/0}(s) = \frac{\hat{W}_{s,janvier2010}^{parc2007-2008}}{\hat{W}_{s,octobre2009}^{parc2007-2008}} * I_{octobre2009/0}(s)$$

Tableau 3.3 - Calcul des indices de prix pour les trois derniers mois de l'année 2007

Mois du 4 ^{ème} trimestre 2007	Valeur estimée du parc de référence 2007 - 2008 * au mois de publication	Formule de calcul	Valeur de l'indice
1 ^{er} mois : octobre	500	100,0	100,0
2 ^{ème} mois : novembre	510	510/500 X 100	102,0
3 ^{ème} mois : décembre	513	513/500 X 100	102,6

* en millions d'euros

Ensuite, on peut commencer le chaînage des indices élémentaires pour les mois de publication du premier trimestre 2008 en utilisant la formule suivante :

$$I_{t/0}(s) = \frac{\hat{W}_{s,t}}{\hat{W}_{s,t-1}} \times I_{t-1/0}(s)$$

avec t le trimestre glissant et $p(t)=(2005,2006)$ comme défini en 3.5.1.

Les indices élémentaires ainsi obtenus peuvent être maintenant agrégés pour obtenir les indices de niveaux supérieurs, tels que département, villes, régions... Pour la publication, la base 100 des indices a été fixée au premier trimestre 2010.

3.5.3 Calcul des indices agrégés en période courante

Pour corriger la non-exhaustivité de la base des notaires, on introduit un coefficient de redressement δ . Ce coefficient est estimé pour chaque département et chaque année de mutation, à partir de données fiscales (montants des droits de mutation établis par la Direction générale des Finances publiques, DGFIP). Il est obtenu en divisant le montant des transactions dans le département d pour l'année a , estimé d'après les données fiscales, par ce même montant enregistré dans la base notariale.

δ ne distingue pas les appartements et les maisons. Il est appliqué à tous les biens du croisement (département x année de mutation)³¹. Il modifie donc le terme qui pondère les indices élémentaires des strates, c'est-à-dire leur poids en valeur dans le parc de référence.

Premier niveau d'agrégation : calcul des indices infra-départementaux et départementaux

Pour les indices infra-départementaux et départementaux, la formule utilisée est une moyenne géométrique :

$$\frac{I_{t/0}}{I_{t-1/0}} = \prod_{s=1}^n \left(\frac{I_{t/0}(s)}{I_{t-1/0}(s)} \right)^{\left(\frac{\delta_s \hat{W}_{s,q(t)}}{\sum_{u=1}^n \delta_u \hat{W}_{u,q(t)}} \right)}$$

avec :

t , le trimestre

s, u , la strate

$I_{t/0}$, l'indice de l'agrégat en t par rapport au trimestre de référence 0,

$I_{t/0}(s)$, l'indice de la strate s en t par rapport au trimestre de référence 0,

$q(t)$, le dernier trimestre de l'année paire $n-2$ (période de base),

$\hat{W}_{s,q(t)}$, la valeur des biens de la strate s échangés au cours de la période $p(t)$, estimée au trimestre $q(t)$,

δ_s , le coefficient de correction pour non exhaustivité de la base des notaires.

Deuxième niveau d'agrégation : calcul des indices supra-départementaux

Pour les indices d'un agrégat A supra-départemental, on utilise la moyenne arithmétique :

$$\frac{I_{t/0}(A)}{I_{t_0/0}(A)} = \sum_{d \in A} \left(\frac{\delta_d \hat{W}_{d,q(t)}}{\sum_{e \in A} \delta_e \hat{W}_{e,q(t)}} \right) \left(\frac{I_{t/0}(d)}{I_{t_0/0}(d)} \right)$$

avec, outre les notations précédentes :

t_0 , le dernier trimestre de l'année $n-1$,

d, e , département ou partie de département

$I_{t/0}(A)$, l'indice de l'agrégat A au trimestre t par rapport au trimestre de référence 0,

³¹ Ce coefficient est l'inverse du taux de couverture des bases notariales (présenté au chapitre 4, paragraphe 4.3). Toutefois, lorsque le taux de couverture est inférieur à 20%, le coefficient utilisé est plafonné et n'est donc jamais supérieur à 5. De même que dans la version 2 des indices, on utilise un rapport de montants et non un rapport de volumes (nombre de transactions), comme c'était le cas dans la version 1 des indices.

$I_{t/0}(d)$, l'indice du département d au trimestre t par rapport au trimestre de référence 0,

$\hat{W}_{d,q(t)}$, la valeur des biens du département échangés au cours de la période $p(t)$ estimée au trimestre $q(t)$.

3.6 Un exemple numérique pour novembre 2012

Voici un exemple de calcul d'un indice élémentaire de prix pour novembre 2012. Cet exemple concerne les appartements, dans la strate 2 d'Île-de-France. Les coefficients utilisés sur la période 2012-2013 sont estimés sur les transactions de 2009-2010.

Étape 1 : extraction de la base

L'indice de novembre 2012 est calculé à partir de l'ensemble des mutations de septembre, octobre et novembre 2012. 222 mutations, correspondant au champ de l'indice et comportant les informations nécessaires au calcul des indices, ont été enregistrées dans la base pour la strate et le trimestre considérés. Les prix au m² s'échelonnent de 1 731 € / m² à 5 225 € / m².

Étape 2 : calcul du prix équivalent bien de référence

La première des 222 mutations porte sur un trois-pièces de 64 m², avec une salle de bain, deux garages et une cave, situé au rez-de-chaussée d'un immeuble d'époque de construction inconnu ; le logement est situé dans le quartier n°2. Le prix de la vente a été de 190 000 €, soit 2 969 € / m².

Il faut enlever les coefficients des caractéristiques spécifiques au bien pour se ramener à son prix « équivalent bien de référence »³². On a donc :

$$\begin{aligned}\text{Log}(\tilde{p}_{j,\tau}) &= \text{Log}(2\,969), \text{ soit } 7,9959 \\ &+0,0598 \text{ (quartier 2} \rightarrow \text{quartier de référence)} \\ &-0,0227 \text{ (1 cave et plus} \rightarrow \text{sans cave)} \\ &-0,0670 \text{ (2 garages et plus} \rightarrow \text{sans garage)} \\ &-0,1573 \text{ (époque de construction inconnue} \rightarrow \text{1948-1969)} \\ &= 7,8088\end{aligned}$$

Le prix « équivalent bien de référence » $\tilde{p}_{j,\tau}$ de cette première mutation s'obtient comme $e^{7,8088}$, soit 2 462 € / m². On réitère l'opération pour les 221 autres mutations.

Dans chaque quartier, les mutations dont le prix équivalent bien de référence est inférieur au 2^e centile ou supérieur au 98^e centile de la distribution sont retirées. Dans notre exemple, le quartier ne compte que 36 transactions pour le trimestre. Dans ce cas, on retire le bien le plus cher et le bien le moins cher. Sur l'ensemble des cinq quartiers de la strate 2, on supprime ainsi 14 références dont les prix équivalents au bien de référence sont extrêmes (6 dans le quartier 1 et 2 dans chacun des 4 autres).

Le prix du bien de référence, égal à l'exponentielle de la moyenne arithmétique des logarithmes des 208 prix équivalents bien de référence des mutations conservées, est $\text{Log}(7,8314)$, soit 2 518 € / m².

Étape 3 : calcul de l'indice mensuel d'une strate en novembre 2012

Pour déterminer l'indice mensuel provisoire³³ d'une strate en novembre 2012, on calcule le prix au m² définitif de la strate en août 2012 (composé des transactions de juin, juillet et août 2012), comme moyenne géométrique des prix au m² obtenus à l'étape 2 pour chaque bien de la strate, soit 2 498 € / m². Ce prix est comparé au prix provisoire trois mois plus tard (2 518 € / m²) pour déterminer l'évolution trimestrielle de la strate. C'est l'évolution de ce prix au m² par rapport au prix au m² définitif estimé dans la strate en novembre 2011 (période

³² Cf. tableau 3.1 en page 19.

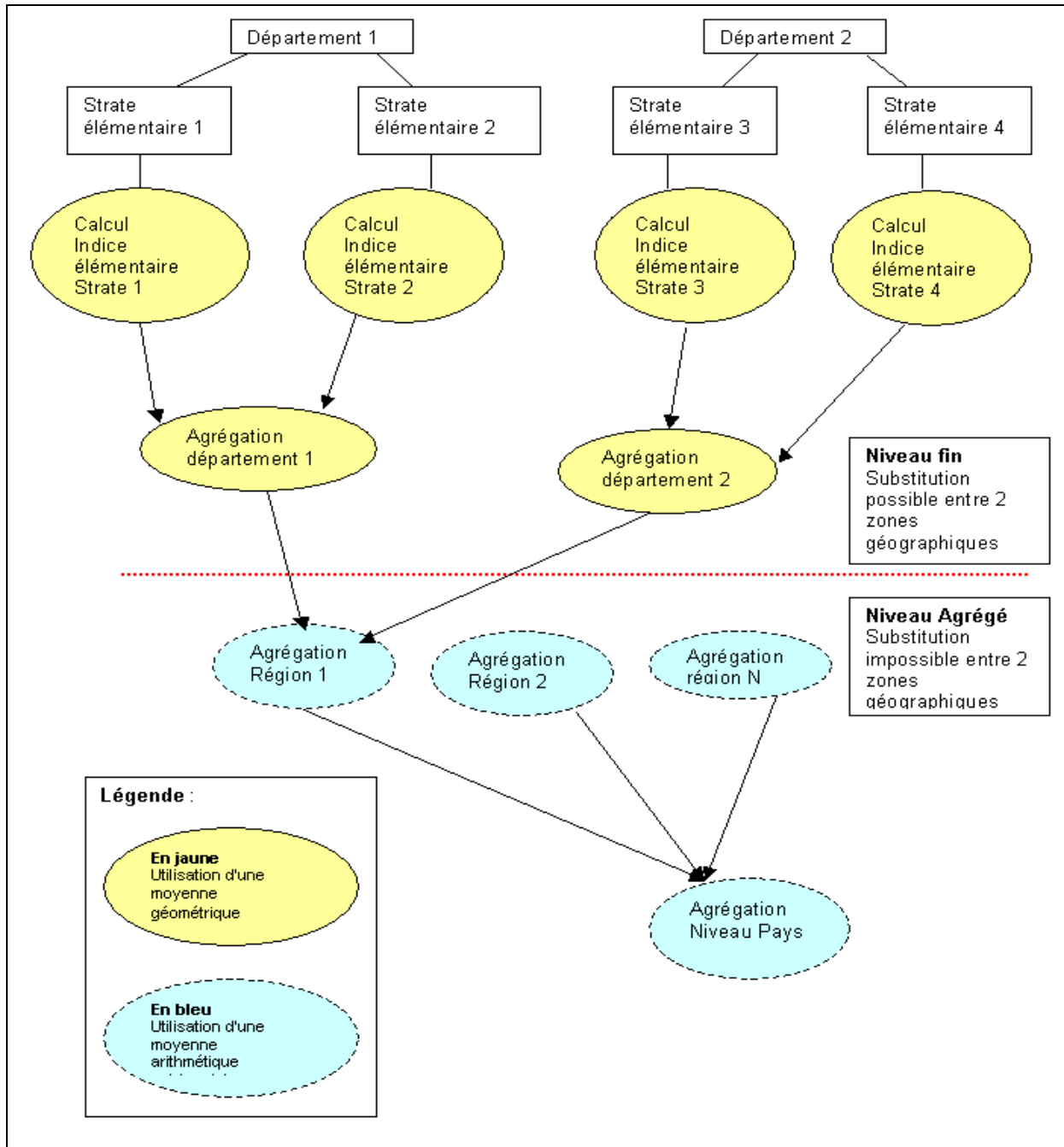
³³ L'indice définitif est publié trois mois plus tard.

de référence des évolutions en 2012 et 2013) qui sera agrégée au niveau départemental et régional pour déterminer les évolutions des prix par département et pour la région.

Des indices élémentaires à l'indice national : agrégation des indices élémentaires

La figure 3.2 ci-dessous montre l'enchaînement des étapes pour passer de l'indice élémentaire à l'indice national. Les départements 1 et 2 appartiennent à la région 1. La moyenne géométrique est appliquée pour les niveaux infra-départementaux et départementaux. On utilise la moyenne arithmétique pour les niveaux plus agrégés³⁴.

Figure 3.2 - Agrégation des indices élémentaires



³⁴ Cf. en page 21.