

Les indices des prix des logements anciens

Cédric Cailly
Insee



Mesurer pour comprendre



Objectif

- › Objectif : Mesurer l'évolution des prix à qualité constante pour avoir une connaissance du marché immobilier
- › Pas de prévision, pas de volonté de détections de bulles immobilières
- › Cet indice n'est pas intégré à l'Indice des prix à la consommation car l'achat d'un logement est un investissement (et pas un bien de consommation).

Le partenariat Notaires-INSEE

Le système institutionnel

- Les notaires détenteurs de tous les actes de ventes
- Une nécessité d'indicateurs robustes pour l'Insee

Une collaboration efficace

- Les services statistiques des notaires (IdF et Province) collectent les actes et calculent les indices
 - L'Insee reçoit les données déjà compilées
- L'Insee apporte son expertise
- Un conseil scientifique responsable de la méthodologie

Des échanges réguliers

- Réunions trimestrielles
- Revue complète de la méthode tous les 5 à 7 ans

Les bases de données notariales

Normalisation et codification des actes de ventes avec un délai de 2 à 3 mois.

Deux bases de données : BIEN (IdF – CINP) et Perval (Province - CSN)

Taux de couverture de 70% en IdF et 50% en province

Sur environ :

150000 transactions annuelles en Idf

600000 transactions annuelles en Province

Pas de maîtrise de l'échantillon (toutefois conséquent)

Nouvelle loi en vigueur depuis janvier 2017

Bases de données anonymes (adresse incluse mais non publiée)



Les bases de données notariales

Prix des logements : net vendeur

Les caractéristiques principales des logements

Type du logement (AP / MA)

- Ancien : filtre selon la TVA
- Appartement ou maison : les autres biens sont exclus du champ
- A usage d'habitation

Localisation

Date de la vente

Surface (m²)

Époque de construction

Nombre de pièces

Nombre de salles de bains

Garages, parkings

AP : étage et présence d'un ascenseur

MA : nombre de niveaux et surface du jardin

État du logement

Diagnostic de performance énergétique

...



La méthode de calcul

Différentes méthodes pour mesurer des évolutions de prix :

- **Comparaison des prix moyens, prix médians**

Ne nécessite pas de traitement statistique, mais les prix sont liés à la structure des logements vendus

- **Ventes répétées**

Inadapté au marché du logement (ventes rares, bien analogues pas exactement comparables)

- **Régressions hédoniques**

Permettent de séparer les évolutions de prix « réelles » des différences de qualité des biens échangés

La méthode de calcul

La méthode hédonique :

- › Permet de séparer les évolutions de prix et les évolutions de qualité, sous les hypothèses suivantes :
 - › Chaque bien est défini par un nombre fini de caractéristiques (propres au bien)
 - › La valorisation des caractéristiques des logements est fixe pour une zone donnée et durant une période donnée (= coefficients fixes sur un modèle)
- › Doit prendre en compte l'importance de la localisation sur le marché immobilier
- › Donne lieu à une chaîne de production trimestrielle

La méthode de calcul

- › La méthode hédonique
 - › A qualité fixée
 - › Un parc de logements comprenant 2 ans de transactions = parc de référence, mis à jour tous les deux ans, supposé robuste sur 5 ans
 - › Près de 300 strates géographiques
 - › Un modèle par strate (sélection de variables), actualisé à chaque changement de parc
 - › Des indices de prix au m2 pour les appartements, de prix total pour les maisons
 - › Un indice trimestriel publié par l'Insee
 - › Révision de la méthode tous les 5-7 ans

La méthode de calcul

4 étapes majeures (et beaucoup de choix « pratiques ») :

- 1) Définir les zones ayant une évolution de prix homogène
- 2) Définir un modèle de prix hédonique (isoler les effets qualité) et estimer les coefficients sur le parc de référence pour chaque zone
- 3) Calculer les indices élémentaires
- 4) Pondérer pour calculer les indices agrégés

1) Stratification

Prendre en compte la localisation :

- Capter une partie de l'hétérogénéité des prix liée à l'emplacement
- Définir des zones au sein desquelles les prix sont formés de manière équivalente

Pour cela :

- Un gros découpage géographique :
 - 97 strates pour les appartements
 - 174 strates pour les maisons
- Un modèle par strate
- Ainsi qu'un second zonage détaillé au sein des strates (utilisé dans les modèles)

1) Stratification

- Avec des méthodes distinctes :
 - Par analyse de données en IdF (CAH)
 - D'après un lourd travail manuel en Province
 - Une autre stratification par analyse de données testée actuellement
- Des strates suffisamment grandes
 - Au moins 110 ventes par trimestre dans une strate
- Marchés spécifiques
 - Paris, villes-strates
 - Stations balnéaires
 - Stations de ski
- Contiguïté géographique au sein d'une strate ?
 - Question ouverte

2) Les modèles de prix hédoniques

Objectifs :

- Etablir une relation entre les prix des logements et leurs caractéristiques qui capte au mieux la variance des prix (sous hypothèse de stabilité)
- Pouvoir observer l'évolution des prix au cours du temps une fois ces effets qualité gommés

Les modèles :

On régresse :

- Le logarithme du prix au mètre carré pour les appartements
- Le logarithme du prix total pour les maisons

2) Les modèles de prix hédoniques

Modèles estimés sur un parc de 2 ans de transactions.

On estime un modèle par strate.

Les modèles sont de la forme :

$$\log p_i = \log p_0 + \alpha_1 Y_{1,i} + \sum_{t=1}^{11} \theta_t M_{t,i} + \sum_{k=1}^K \beta_k X_{k,i} + \varepsilon_i$$

avec les notations suivantes :

p_i prix du bien i ,

p_0 prix du logement de référence dans parc d'estimation

$Y_{1,i}$ indicatrice de l'année de mutation du bien i ,

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

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



2) Les modèles de prix hédoniques

Les modèles :

- Variables très majoritairement catégorielles
- Définition de modalités de référence
- Travail sur les données :
 - Traitement des valeurs manquantes
 - Choix des modalités et croisement de variables
 - Equilibre entre robustesse et précision
- Sélection automatique des variables (par strate)
 - Procédure forward
 - Impact quasi-nul du type de sélection et des références choisies
- Suppression des valeurs aberrantes

→ Obtention des coefficients

3) Calcul des indices élémentaires

On a calculé sur le parc de référence les coefficients de valorisation des caractéristiques du bien β_k .

N	N+1	N+2	N+3	N+4
Parc de référence			Période courante	
Coefficients fixes				

On va maintenant les appliquer sur la « période courante » afin de calculer l'évolution des prix (= évolution de la « constante »).

3) Calcul des indices élémentaires

Sur le parc de référence, on estime :

$$\log p_i = \log p_0 + \alpha_1 Y_{1,i} + \sum_{t=1}^{11} \theta_t M_{t,i} + \sum_{k=1}^K \beta_k X_{k,i} + \varepsilon_i$$

Avec p_0 le prix du logement de référence sur le parc.

Sur un trimestre t de la période courante, on aurait :

$$\log p_i = \log p_{0,t} + \sum_{k=1}^K \beta_k X_{k,i} + \varepsilon_i$$

Avec $p_{0,t}$ le prix du logement de référence en t .

Comme on a déjà estimé β_k , on recalcule :

$$\log p_{0,t} = \text{Moyenne}_t \left(\log p_i - \sum_{k=1}^K \beta_k X_{k,i} \right)$$

$$p_{0,t} = \exp \left[\text{Moyenne}_t \left(\log p_i - \sum_{k=1}^K \beta_k X_{k,i} \right) \right]$$

3) Calcul des indices élémentaires

On peut enfin calculer une évolution de prix au sein d'une strate donnée :

$$I_{t/t-1} = p_{0,t} / p_{0,t-1}$$

Les indices de prix des strates sont appelés indices élémentaires.

4) Pondération – calcul des indices agrégés

Une fois les indices élémentaires calculés, on les agrège à différents niveaux (département, région..).

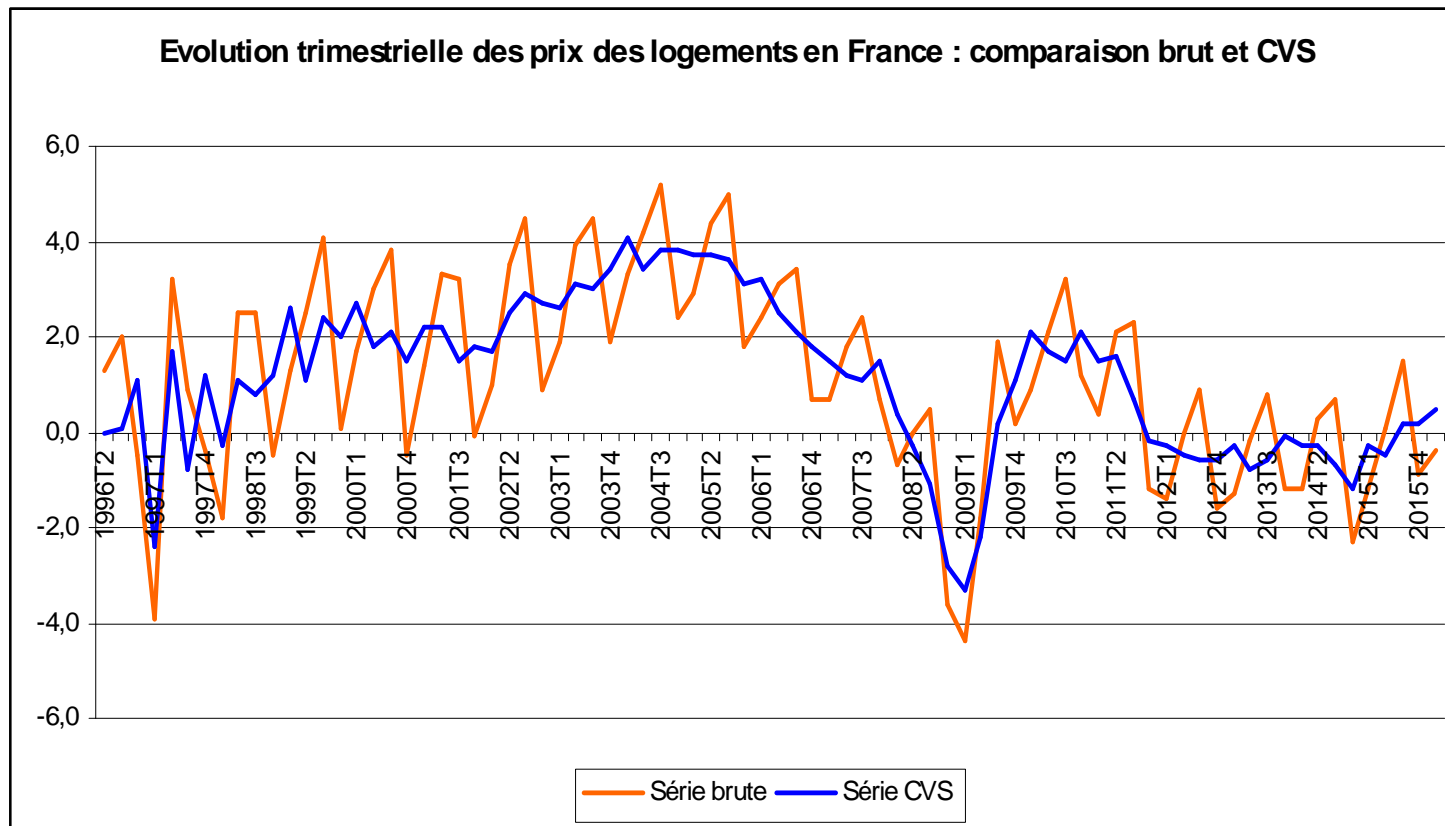
Pour cela, on pondère les strates au sein de l'agrégat selon la somme du prix des biens de chaque strate.

On corrige également la non-exhaustivité des bases en pondérant chaque bien par un coefficient calculé au niveau du département

- Confrontation des ventes des bases notariales avec les données fiscales

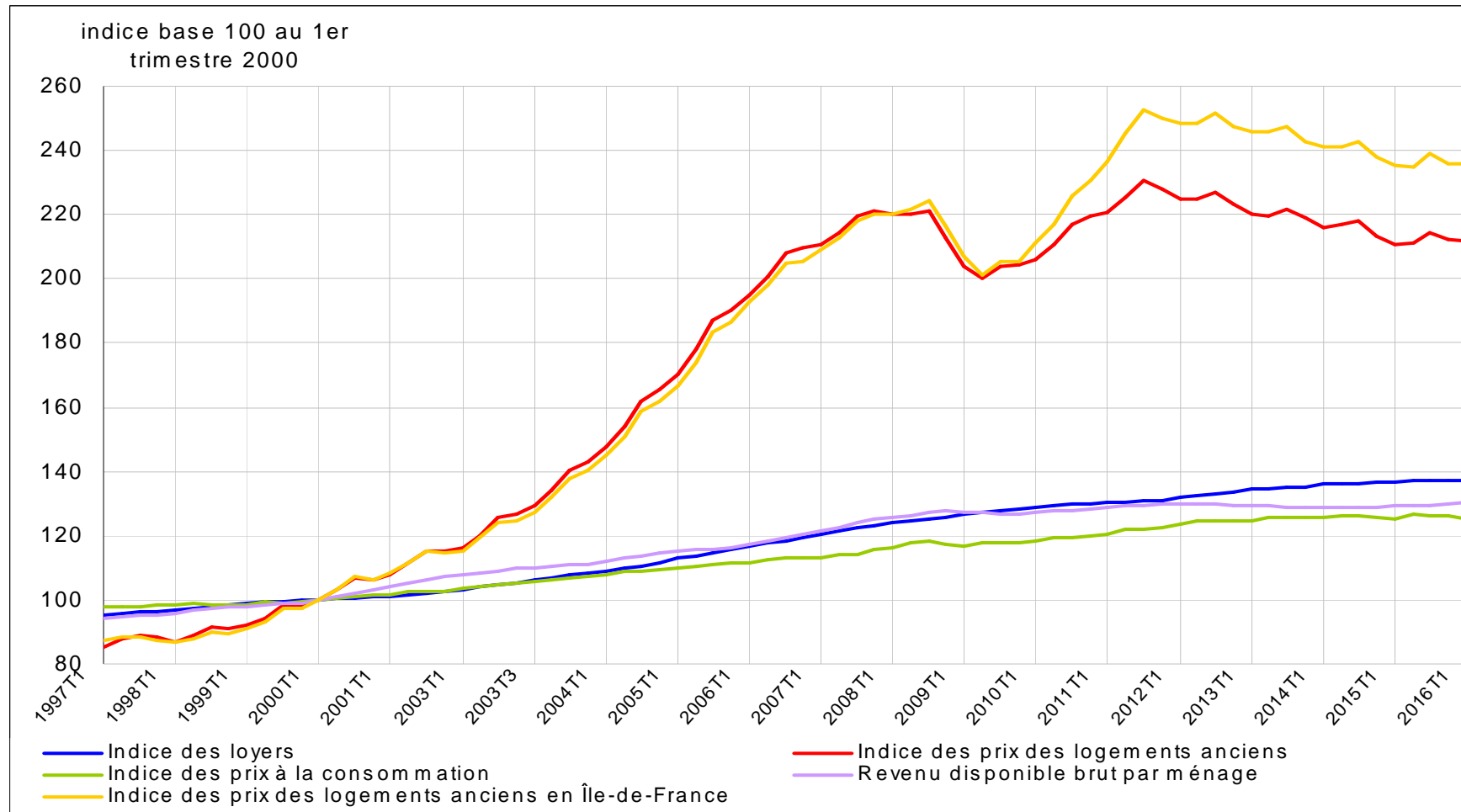
Correction des variations saisonnières

- › Profil saisonnier des prix de l'immobilier : hausse des prix au troisième trimestre (rentrée scolaire)

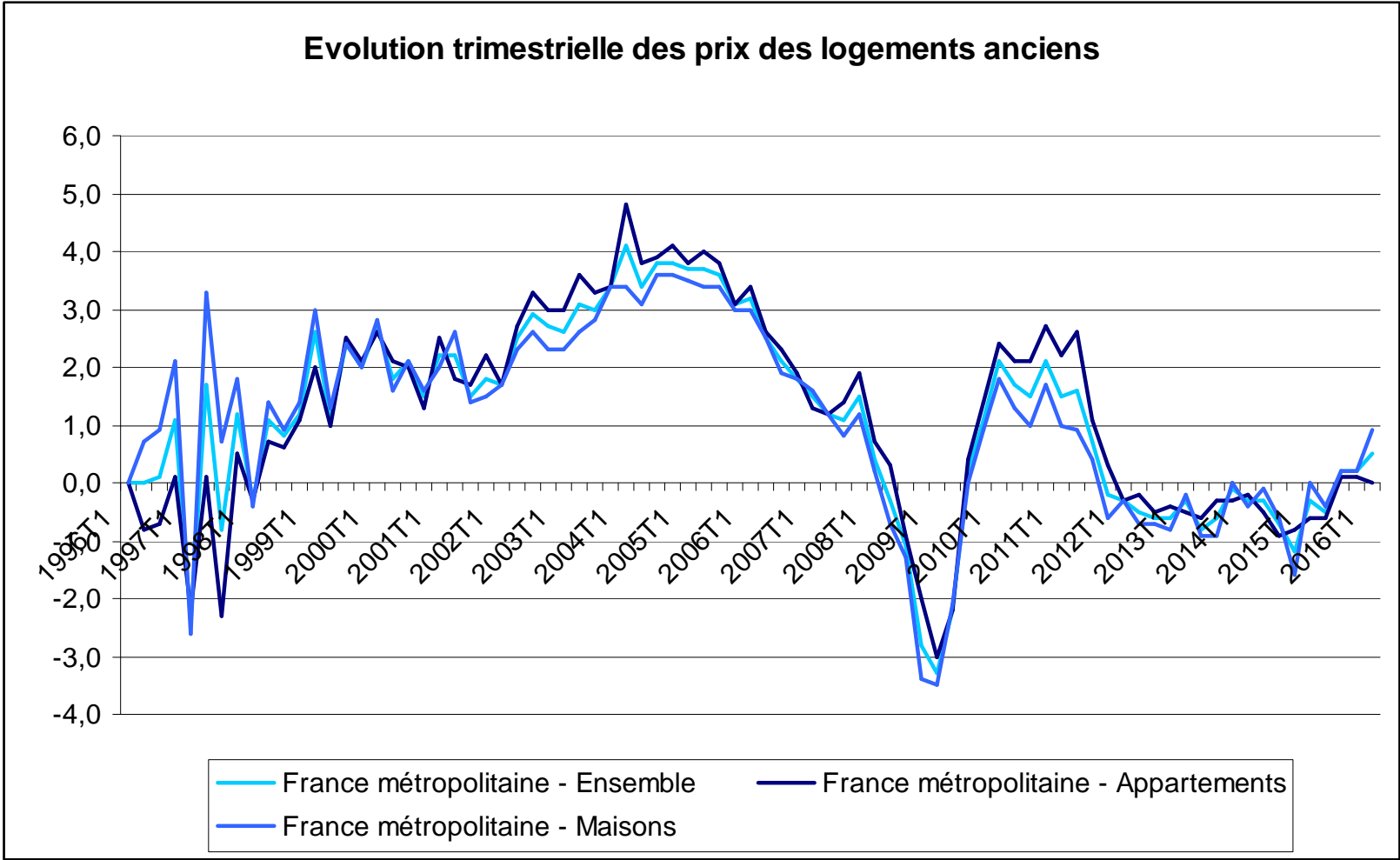


Résultats et suivi

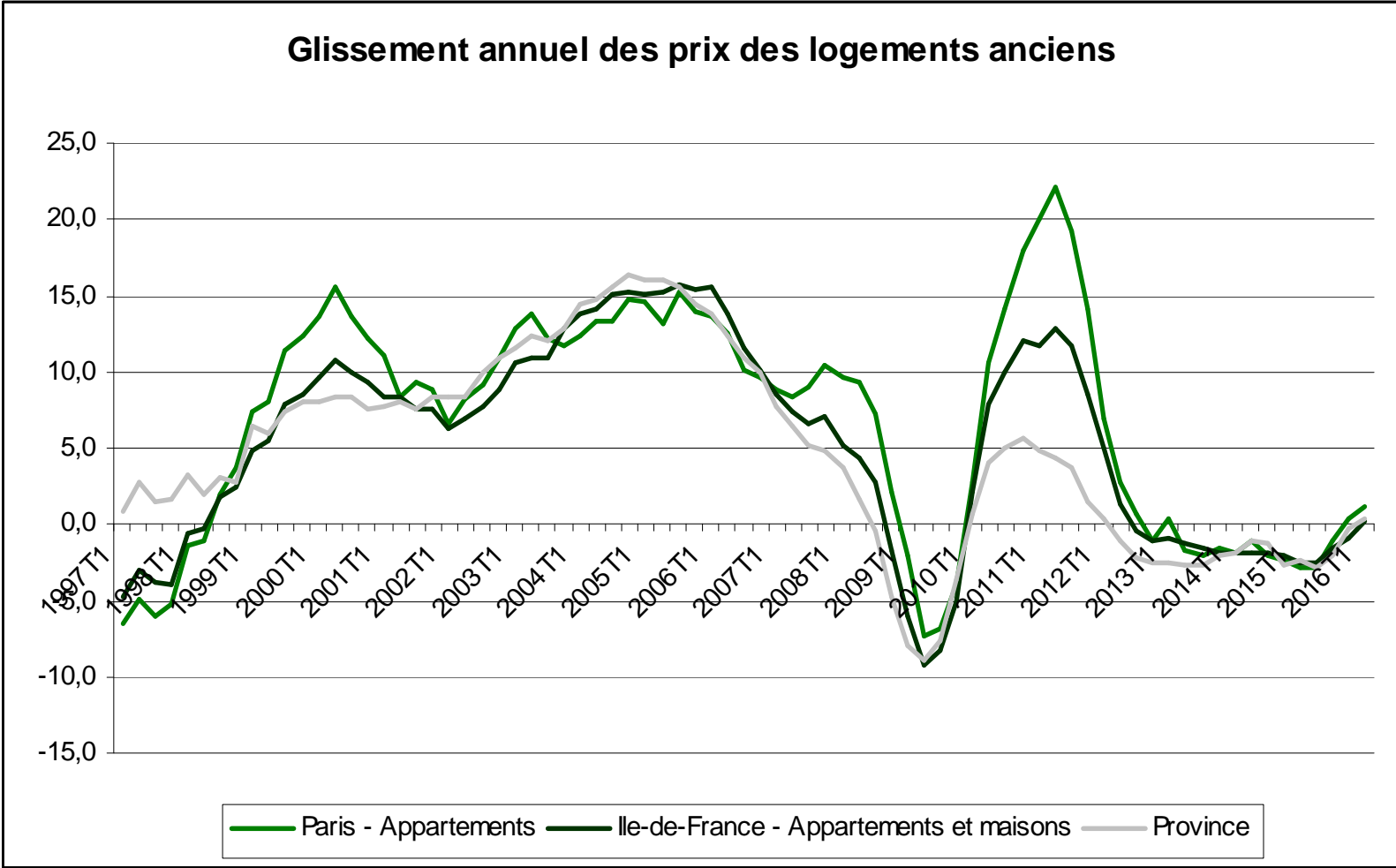
Indice non déflationné, mais mis en relief du reste régulièrement



Résultats et suivi



Résultats et suivi



Résultats et suivi

Un système de contrôle à plusieurs niveaux

- Validation trimestrielle de la qualité des données
- Révision tous les 5 ans du zonage, tous les 2 ans des modèles, des parcs d'estimation et de référence.

Veille méthodologique trimestrielle assurée par le Conseil scientifique

Une conférence de presse en Île-de-France (pas uniquement centrée sur les indices) et une diffusion commune à la même date

Perspectives

Travaux sur les promesses de vente

Indices par nombre de pièces

Indices plus localisés



Merci de votre attention !