

## 2.2 UN INDICE DES PRIX DES APPARTEMENTS AU LUXEMBOURG BASÉ SUR LA MÉTHODE DES VENTES RÉPÉTÉES <sup>27</sup>

La progression rapide des prix immobiliers au Luxembourg suscite bien des inquiétudes (voir la section 3 du chapitre 1 de la Revue de Stabilité financière de la BCL). Dans ce contexte, il est indispensable que les décisions de politique économique se fondent sur des indicateurs fiables. Au Luxembourg, le STATEC publie chaque trimestre l'indicateur rapide C concernant les prix de vente des logements<sup>28</sup>. Cette publication présente les prix moyens par transaction, les prix moyens par mètre carré ainsi que des indices dits « hédoniques » qui cherchent à éliminer les différences dans les caractéristiques des logements. Les indicateurs de prix sont ventilés en différentes catégories, distinguant les maisons anciennes des appartements anciens et neufs, ce qui permet de comparer la dynamique des prix sur ces segments du marché. Or, chacune de ces mesures de prix a ses avantages et ses désavantages, comme expliqué plus bas.

Cette analyse présente un nouvel indice des prix en appliquant la « méthode des ventes répétées » pour comparer sa fiabilité à celle des indices déjà disponibles. L'analyse montre en quoi la mesure des prix des logements peut poser des difficultés en pratique. Si la méthode des ventes répétées produit un indice de prix qui confirme l'accélération des prix au cours des dernières années, cet indice alternatif est moins volatil que les indices hédoniques qui sont utilisés comme indices de référence au Luxembourg. En fournissant une image plus claire des évolutions à court terme, cet indicateur alternatif pourrait apporter une aide aux choix des politiques publiques.

La section suivante donne un aperçu des différentes méthodes utilisées pour mesurer l'évolution des prix des logements. Ensuite, nous présentons les données utilisées dans cette analyse avant de décrire la méthode des ventes répétées. La quatrième section compare l'évolution des prix des appartements mesurés par les différents indices et la cinquième section évoque les inconvénients potentiels de l'indice basé sur les ventes répétées.

### 2.2.1 MÉTHODES

Les indices des prix les plus simples sont les moyennes ou les médianes<sup>29</sup> des prix observés pour l'ensemble des transactions d'une période donnée, typiquement un mois ou un trimestre. Un indice peut facilement être construit en répétant ce calcul pour chaque période et en le divisant par la valeur obtenue pour la période de base. L'avantage d'un tel indice réside dans la facilité de son mode de calcul et de son interprétation. Le calcul ne nécessite que des informations sur les transactions individuelles (prix et date) et l'indice qui en résulte mesure une tendance générale des transactions à travers le temps. Or, comme plusieurs auteurs l'ont noté (voir Englund, Quigley et Redfearn, 1999 ; Jansen, 2008 ; Thion, Favarger et Hoesli, 2001), les indices des prix moyens ou des prix médians peuvent donner une impression trompeuse de l'évolution des prix sur une période donnée. En effet, les transactions sont différentes à chaque période, ce qui modifie la composition de l'échantillon utilisé pour calculer les prix moyens ou médians. Ainsi, une période durant laquelle les transactions portent surtout sur des logements haut de gamme pourrait être suivie par une période où les transactions concernent surtout des logements bas de gamme, créant une fausse impression de détérioration du prix pour l'ensemble du stock de logements (y compris ceux qui n'ont pas été vendus). Différentes méthodes existent pour pallier les effets des

<sup>27</sup> Analyse rédigée par Bob Kaempff et David Kremer, économistes au département Économie et Recherche.

<sup>28</sup> Voir <https://statistiques.public.lu/>

<sup>29</sup> La médiane est la valeur qui partage une population en deux parties égales (50<sup>ème</sup> percentile). Elle peut diverger de la moyenne si la distribution est asymétrique ou si l'échantillon inclut des observations atypiques.



changements de composition, dont notamment les indices de prix hédoniques et les indices fondés sur la méthode des ventes répétées (« *Repeated sales index* » en anglais, RSI par la suite).

Au Luxembourg, le STATEC applique la méthode hédonique qui s'appuie sur une relation estimée entre le prix d'un logement et plusieurs de ses caractéristiques, telles que sa surface, sa localisation géographique, le nombre d'emplacements de parking, etc. De cette façon, il est possible d'ajuster l'indice des prix pour tenir compte du fait que les caractéristiques des logements vendus varient d'une période à l'autre. Bien que la méthode hédonique soit souvent utilisée pour mesurer les prix immobiliers à l'étranger comme au Luxembourg (Lamboray, 2010 ; Licheron et Ferring, 2018), cette méthode nécessite des choix quant aux caractéristiques à inclure et à la forme fonctionnelle à estimer qui peuvent influencer les résultats (Wang et Zorn, 1997). Or, le choix des variables explicatives n'est pas toujours évident en pratique, soit à cause d'un manque de données, soit en raison de la spécification de la forme fonctionnelle. Pour donner un exemple, la crise du COVID-19 et le déploiement rapide du travail à distance ont occasionné des discussions concernant la valeur supplémentaire représentée par un balcon ou un jardin. Or, les informations disponibles au niveau des transactions individuelles au Luxembourg ne permettent pas de déterminer si un appartement a été vendu avec balcon, même si l'impact positif de cette caractéristique semble plausible.

Une alternative qui permet de réduire ces inconvénients est la méthode des ventes répétées, qui se base sur l'observation à différents points dans le temps du prix d'un même objet immobilier vendu à plusieurs reprises. Aux États-Unis, les indices de référence publiés par Standard&Poors ("*S&P Case-Shiller Home Price Indices*")<sup>30</sup> en sont un exemple. L'approche des ventes répétées repose sur l'hypothèse que le mouvement de prix d'un même objet s'explique par la variation du niveau général des prix pour l'ensemble des logements. En effet, il est plausible que les caractéristiques d'un objet restent les mêmes en comparant les prix de vente d'un même bien à deux périodes différentes. Cette approche évite ainsi de devoir choisir un sous-ensemble de caractéristiques communes à travers les transactions ou d'estimer une forme fonctionnelle pour relier ces caractéristiques aux prix observés lors des transactions.

## 2.2.2 SOURCE DES DONNÉES ET IDENTIFICATION DES VENTES RÉPÉTÉES

Comme pour l'indice hédonique du STATEC, nous construisons l'indice des ventes répétées à partir des informations sur les transactions immobilières contenues dans les extraits des actes notariés transmis à l'Administration des Domaines, de l'enregistrement et de la TVA (AED). Celles-ci sont complétées par des informations provenant de l'Administration du Cadastre concernant la surface des appartements et des annexes. Notre échantillon initial comprend toutes les ventes d'appartements<sup>31</sup> enregistrées par l'AED entre le premier trimestre 2007 et le premier trimestre 2021, soit 86 184 transactions. Un premier traitement de ces données est effectué pour, entre autres, harmoniser le traitement des prix des ventes en futur état d'achèvement et éliminer des transactions non pertinentes. Ce traitement est réalisé selon une méthodologie commune développée par un groupe de travail composé de membres du STATEC, de l'Observatoire de l'habitat, de l'AED et de la BCL<sup>32</sup>. Cet échantillon, utilisé pour la compilation de l'indice hédonique, comprend 61 850 transactions.

30 Disponible sous : <https://www.spglobal.com/spdji/en/index-family/indicators/sp-corelogic-case-shiller/sp-corelogic-case-shiller-composite/#overview>

31 Cette analyse se focalise sur les appartements puisque ces derniers sont saisis par le cadastre vertical, ce qui fournit des variables nécessaires pour le calcul des indices de prix comme la surface de l'appartement. Ces variables ne sont pas disponibles pour les maisons.

32 Une note méthodologique est publiée par le STATEC sous le lien suivant : <https://statistiques.public.lu/fr/methodologie/methodes/economie-finances/Prix/prix-logements/note-prix-de-vente.pdf>

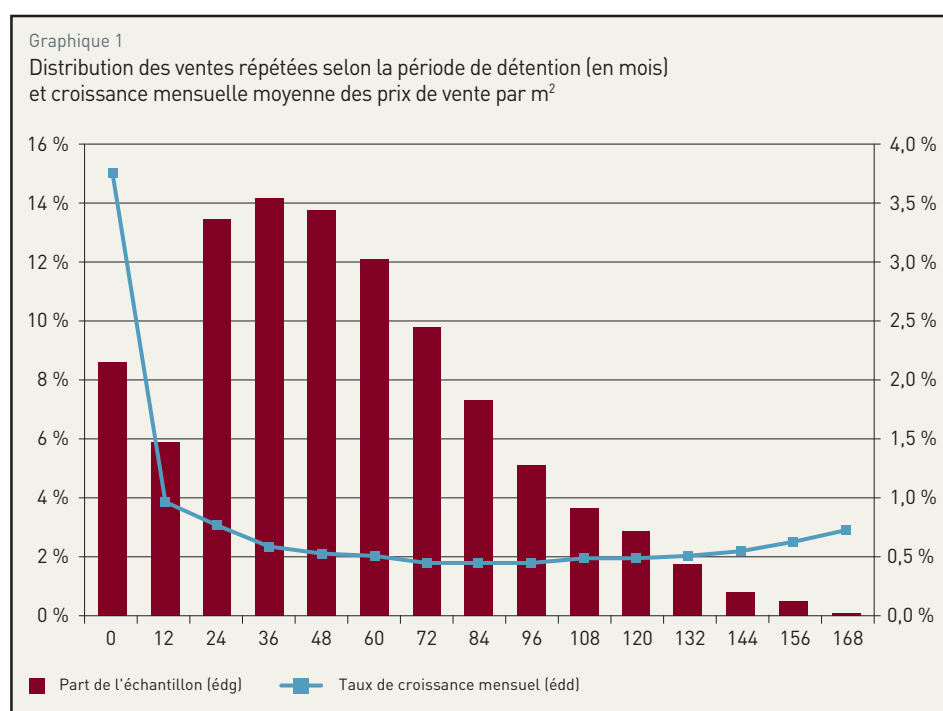
Ensuite, nous identifions les transactions portant sur le même appartement à différents points dans le temps. Un identifiant unique par appartement est construit à partir des informations suivantes :

- La section cadastrale (localité)
- La surface du terrain de la résidence
- La surface de l'appartement
- La description cadastrale de l'appartement
- La surface totale des emplacements de parking
- La surface totale des caves

Ainsi, deux transactions distinctes sont associées au même bien si toutes ces variables correspondent pour les deux transactions. De cette façon, nous identifions 23 648 transactions (38 % de l'échantillon utilisé pour l'indice hédonique) qui sont des ventes répétées, c'est-à-dire des transactions portant sur des appartements qui ont été vendus plusieurs fois entre 2007 et 2021.

Ensuite, nous procédons à trois traitements supplémentaires pour éliminer des ventes répétées qui sont susceptibles de biaiser l'indice. Premièrement, nous éliminons les appartements revendus en quelques mois, parce qu'ils réalisent généralement de très fortes augmentations en valeur, comme démontré sur le graphique 1, ce qui peut introduire un biais dans l'indice. Vraisemblablement, ces appartements sont achetés avec l'intention de les rénover et revendre rapidement (voir également section 4). Suivant l'approche proposée par Jansen (2008), nous avons exclu les transactions dont la période de détention (entre l'achat et la revente) est inférieure à 12 mois. Toutefois, nous convenons qu'il n'existe pas de séparation naturelle entre périodes de détention « courtes » et « typiques » et qu'en définitive, ce choix est arbitraire (Clapham, 2006). Ce traitement élimine 3,3 % de l'échantillon utilisé pour l'indice hédonique ou 8,6 % de l'ensemble des ventes répétées, comme l'illustre le graphique 1.

Deuxièmement, nous avons décidé d'exclure les observations comportant une augmentation de prix qui peut être qualifiée d'extrême, ce qui est susceptible de représenter des rénovations majeures qui modifient les caractéristiques du bien. Nous éliminons les transactions portant sur le même bien qui comportent une croissance annualisée entre les deux ventes dépassant un certain seuil<sup>33</sup>. Nous éliminons ainsi 1,2 % de l'échantillon utilisé pour l'indice hédonique.



Sources : Administration des Domaines, de l'enregistrement et de la TVA, calculs des auteurs

Note : La distribution (barres rouges) est calculée à partir de l'échantillon complet des ventes répétées

<sup>33</sup> Dans une première étape, nous stratifions les ventes répétées par région (canton de Luxembourg et reste du pays) et par année de la dernière vente. Ensuite, nous calculons, par strate, la croissance annuelle médiane et le premier et le troisième quartiles de la croissance annuelle. Enfin, nous éliminons toutes les transactions dont la croissance absolue dépasse la croissance médiane de 1,5 fois la différence entre le premier et le troisième quartiles de croissance de la strate. Un traitement similaire est utilisé par le STATEC pour l'indice hédonique (Lamboray, 2010).

Troisièmement, nous excluons les appartements qui ont été vendus plus de cinq fois entre 2007 et 2021. Des appartements revendus à une telle fréquence sont probablement peu représentatifs. Par exemple, ils pourraient avoir des vices cachés qui deviennent apparents seulement après une certaine période d'occupation. De plus, ces appartements représentent moins de 1 % de l'échantillon utilisé pour l'indice hédonique.

Tableau 1

**Comparaison des échantillons pour l'indice hédonique et celui des ventes répétées**

	ÉCHANTILLON INDICE HÉDONIQUE	DONT		VENTES RÉPÉTÉES	ÉCHANTILLON INDICE VENTES RÉPÉTÉES (APRÈS TRAITEMENTS)
		APPARTEMENTS NEUFS	APPARTEMENTS EXISTANTS		
Nombre de transactions	61 850	23 046	38 804	23 648	20 616
Prix moyen par m <sup>2</sup>	5 101	5 709	4 740	4 781	4 752
Surface moyenne	81,5	81,9	81,4	79,8	81

Sources : Administration des Domaines, de l'enregistrement et de la TVA, calculs des auteurs

Note : Les échantillons comprennent les transactions enregistrées entre 2007T1 et 2021T1 après les traitements décrits dans la section 2.

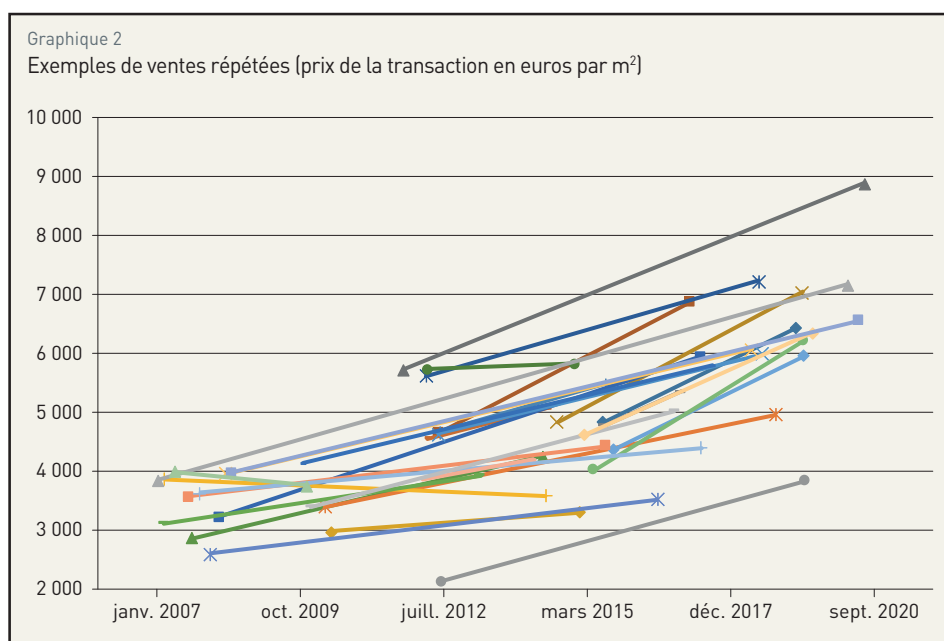
L'échantillon retenu pour le calcul de l'indice des ventes répétées comprend 20 616 transactions (33 % de l'échantillon de l'indice hédonique) ou 360 transactions en moyenne par trimestre. Le tableau 1 compare le nombre de transactions, le prix moyen par mètre carré et la surface moyenne des appartements vendus pour les différents échantillons. Bien que le nombre d'observations pour l'indice des ventes répétées (colonne de droite) soit inférieur à celui de l'indice hédonique (colonne de gauche), les prix moyens par mètre carré et les surfaces moyennes sont similaires.

Pour illustrer la structure des données, le graphique 2 montre l'évolution des prix pour un nombre limité de ventes répétées dans la commune d'Esch-sur-Alzette. Une simple comparaison visuelle suggère que les courbes ont une pente plus raide vers la fin de la période d'observation, ce qui indique une croissance plus forte des prix. Si les prix ont généralement augmenté, le graphique inclut tout

de même quelques ventes répétées pour lesquelles le prix a baissé entre les deux ventes. En effet, lors d'une vente forcée, un propriétaire en difficulté pourrait être obligé de vendre le bien en dessous de sa valeur de marché. Les sections 2 et 5 expliquent de façon plus détaillée comment nous avons cherché à minimiser le biais de sélection pour l'estimation du RSI.

### 2.2.3 ESTIMATION DE L'INDICE

L'estimation de l'indice des ventes répétées se fonde sur la méthodologie proposée par Bailey, Muth et Nourse (BMN, 1963) et reprise par le manuel Eurostat sur les indices des prix résidentiels (voir



Source : Administration des Domaines, de l'enregistrement et de la TVA

Eurostat, 2013). En pratique, le modèle utilise la méthode des moindres carrés ordinaires pour estimer une relation entre la progression du prix (logarithme du rapport des prix de la première et de la deuxième vente) et un ensemble de variables indicatrices prenant la valeur de -1 à la période de première vente ( $\tau$ ), la valeur 1 à la période de deuxième vente ( $t$ ) et 0 pour toutes les autres périodes jusqu'à la fin de l'échantillon ( $S$ ). Ainsi, l'équation de régression est égale à :

$$\ln\left(\frac{P_{it}}{P_{i\tau}}\right) = \sum_{s=0}^S \Phi_s D_{is} + \varepsilon_{it} \text{ avec } (0 < \tau < t < S)$$

$D_{is}$  correspond à la variable indicatrice à la période  $s$  pour le bien  $i$  et  $\varepsilon_{it}$  au terme d'erreur qui suit les hypothèses conventionnelles d'une distribution normale avec moyenne 0 et variance constante  $\delta^2$ . La progression de l'indice des prix entre la période de base et la période  $s$  peut être estimée à partir du coefficient  $\Phi_s$  issu de la régression.

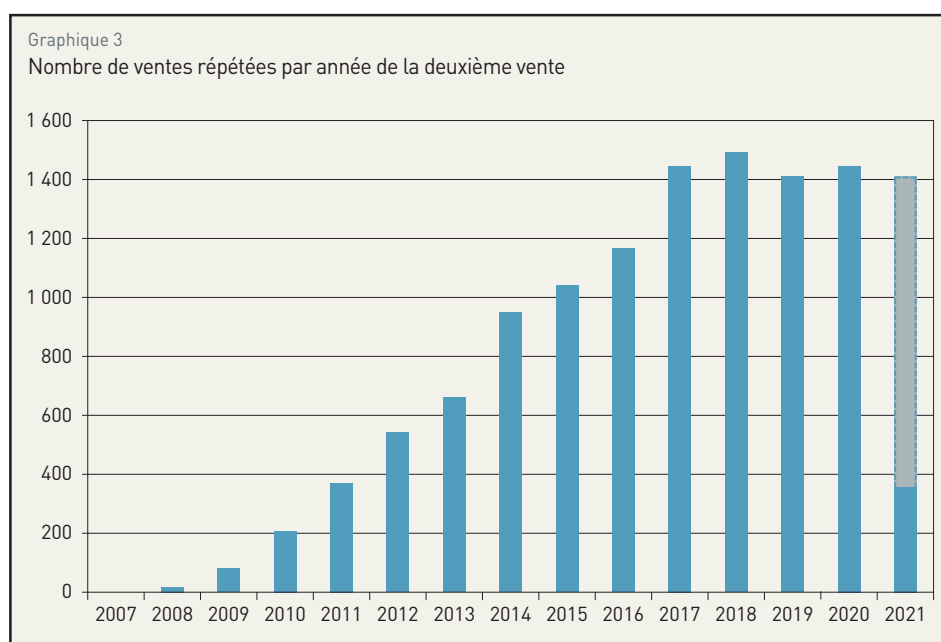
Alors que l'hypothèse d'une qualité constante est à la base de l'approche proposée par BMN, Case et Shiller (1987, 1989) affirment que cette hypothèse devient moins probable lorsque la période de détention d'un objet augmente. Afin de corriger l'indice et de prendre en compte le fait que les variations des prix des logements comprennent des composantes dont la variance augmente avec l'intervalle des ventes, ils proposent une approche à trois étapes qui sert à réduire le poids d'une vente dans l'indice lorsque l'intervalle entre deux ventes augmente. Nous appliquons également cette approche, pour laquelle une description plus détaillée est donnée en annexe.

Enfin, quand un bien est revendu plusieurs fois, nous traitons les reventes comme observations indépendantes. Selon Shiller (1991), ceci n'est pas problématique tant que les périodes de détention d'un même objet ne coïncident pas. Ainsi, un bien qui a été vendu à trois reprises pendant les périodes 1,2 et 3 figurerait dans la base de données comme deux reventes indépendantes dans les périodes {1 ;2} et {2 ;3}.

## 2.2.4 RÉSULTATS ET COMPARAISON

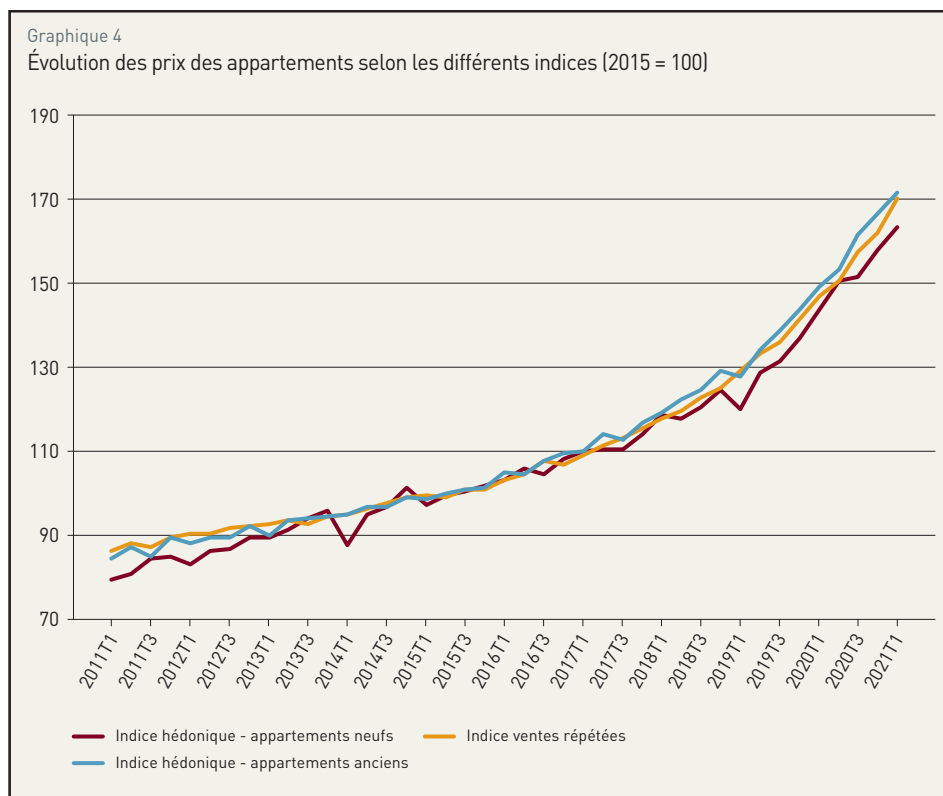
Cette section présente les résultats issus de l'estimation des indices des ventes répétées et les compare aux indices hédoniques<sup>34</sup>. Nous observons les données entre le T1 2007 et le T1 2021, mais la discussion se focalise sur la période depuis 2011 à cause de la structure temporelle de l'échantillon de ventes répétées. Comme le montre le graphique 3, du fait de sa méthode

34 Pour les calculs des différentes caractéristiques des indices hédoniques, nous avons reproduit les indices selon la méthode décrite dans Lamoray, 2010. Les indices hédoniques présentés par la suite sont donc calculés par les auteurs sur base des données fournies par l'AED.



Sources : Administration des Domaines, de l'enregistrement et de la TVA, calculs des auteurs  
Note : Pour 2021, la barre montre une simple extrapolation des observations du premier trimestre.

Graphique 4  
Évolution des prix des appartements selon les différents indices (2015 = 100)



Sources : Administration des Domaines, de l'enregistrement et de la TVA, calculs des auteurs

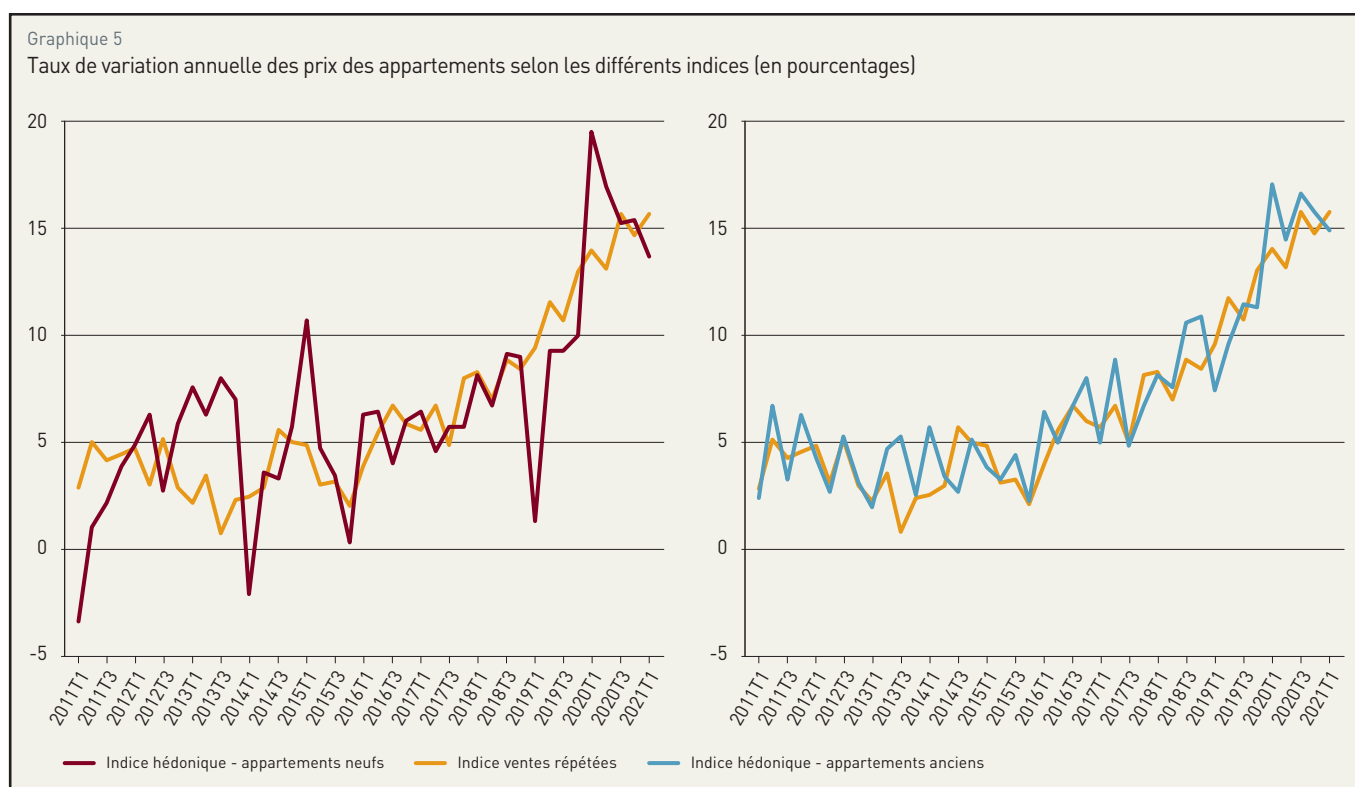
tements en état de futur achèvement, puisqu'une vente répétée est forcément la vente d'un bien existant. Pour cette raison, le RSI est plus comparable à l'indice hédonique pour les appartements anciens qu'à l'indice hédonique pour les appartements neufs. Le graphique 4 compare ces trois indices en niveau, ce qui permet de tirer les premières conclusions. Entre le T1 2011 et le T1 2021, la croissance du RSI est comparable à celle des deux indices hédoniques. Ce n'est qu'après 2015 que la croissance de l'indice hédonique des appartements neufs est légèrement plus faible, de sorte que le niveau de ce dernier est légèrement inférieur à celui du RSI à la fin de la période d'observation. Enfin, le RSI semble être moins volatil, tandis que les fluctuations à court terme semblent plus prononcées pour l'indice hédonique des appartements neufs.

Sur base du graphique 5, la variation annuelle de l'indice RSI est plus proche de celle des appartements anciens. Comme expliqué en haut, le RSI exclut les biens neufs jusqu'au moment de leur première revente. Entre le T1 2011 et le T1 2021, le taux de variation annuel moyen du RSI est de 6,5 %, très proche de celui de l'indice hédonique des appartements neufs (6,6 %) et légèrement en dessous de celui de l'indice hédonique des appartements anciens (6,9 %).

Alors que les indices hédoniques peuvent présenter de fortes fluctuations de prix à court terme, l'évolution du RSI semble être moins volatile, tout en confirmant la tendance haussière à long terme. Une volatilité élevée pourrait être le signe que l'indice est contaminé par des erreurs de mesure qui perturbent les signaux de prix. En mesurant la volatilité par l'écart-type de la croissance trimestrielle, le RSI est environ 2,2 fois moins volatil que l'indice hédonique pour les appartements neufs et 1,5 fois moins volatil que l'indice hédonique pour les appartements anciens (tableau 2).

de construction, notre base de données est caractérisée par une augmentation graduelle du nombre d'observations pour les périodes plus récentes. Ainsi, en fixant une période de détention minimale de 12 mois, nous commençons à retenir les premières ventes répétées seulement à partir de l'année 2008. Or, après le nettoyage des données précédemment décrit, nous n'observons qu'une vingtaine de paires de ventes où la deuxième a eu lieu en 2008. Ainsi, nous avons décidé de mettre l'accent sur l'analyse des résultats à partir de 2011, en supposant qu'environ 370 observations par année (92 par trimestre) sont suffisantes pour être informatives, sachant que ce chiffre continue à augmenter graduellement pour se stabiliser à partir de 2017.

Par définition, un RSI ne permet pas d'analyser les prix des appartements



Sources : Administration des Domaines, de l'enregistrement et de la TVA, calculs des auteurs

Tableau 2

Comparaison de la croissance annuelle moyenne et de la volatilité pour les différents indices

(en %)	INDICE VENTES RÉPÉTÉES		INDICE HÉDONIQUE - APPARTEMENTS NEUFS		INDICE HÉDONIQUE - APPARTEMENTS ANCIENS	
	CROISSANCE ANNUELLE	VOLATILITÉ	CROISSANCE ANNUELLE	VOLATILITÉ	CROISSANCE ANNUELLE	VOLATILITÉ
2011-2017	4,2	1,01	4,6	3,03	4,6	1,89
2018-2021T1	11,6	1,05	11,1	2,65	11,9	1,61
<b>2011-2021T1</b>	<b>6,5</b>	<b>1,36</b>	<b>6,6</b>	<b>3,01</b>	<b>6,9</b>	<b>2,00</b>

Source : Administration des Domaines, de l'enregistrement et de la TVA

Note : La volatilité est mesurée comme l'écart-type des taux de croissance trimestriels.

Le tableau 2 divise l'échantillon en deux, une période de croissance modérée entre 2011 et 2017 et une période de croissance plus prononcée depuis 2018. Pendant la première période, la croissance des prix des appartements mesurée par le RSI était légèrement plus faible que celle mesurée par les deux indices hédoniques. Pendant la période de croissance plus prononcée, le RSI se situe en moyenne légèrement au-dessus de l'indice des appartements neufs et en dessous de l'indice des appartements anciens. Sur les deux sous-périodes, le RSI est moins volatil que les deux indices hédoniques.

## 2.2.5 INCONVÉNIENTS ÉVENTUELS DE LA MÉTHODE DES VENTES RÉPÉTÉES

### Perte de données

Une limitation de la méthode des ventes répétées est la perte des données. Par définition, l'échantillon des ventes répétées exclut toute transaction jusqu'au moment d'une deuxième vente du même bien. Ainsi, une partie non négligeable des données disponibles n'est pas prise en compte. Comme le montre

le tableau 1, cela représente environ deux tiers des transactions utilisées pour l'indice hédonique. Dans ce contexte, il convient de rappeler que le nombre d'observations de ventes répétées augmente graduellement au début de l'échantillon (graphique 3) et que pour les périodes plus récentes, le RSI devient plus efficace en termes d'utilisation des données collectées. Un arbitrage s'impose entre le nombre d'observations qui sont exclues pour l'estimation d'un RSI et le nombre de variables additionnelles à collecter pour l'estimation d'un indice hédonique.

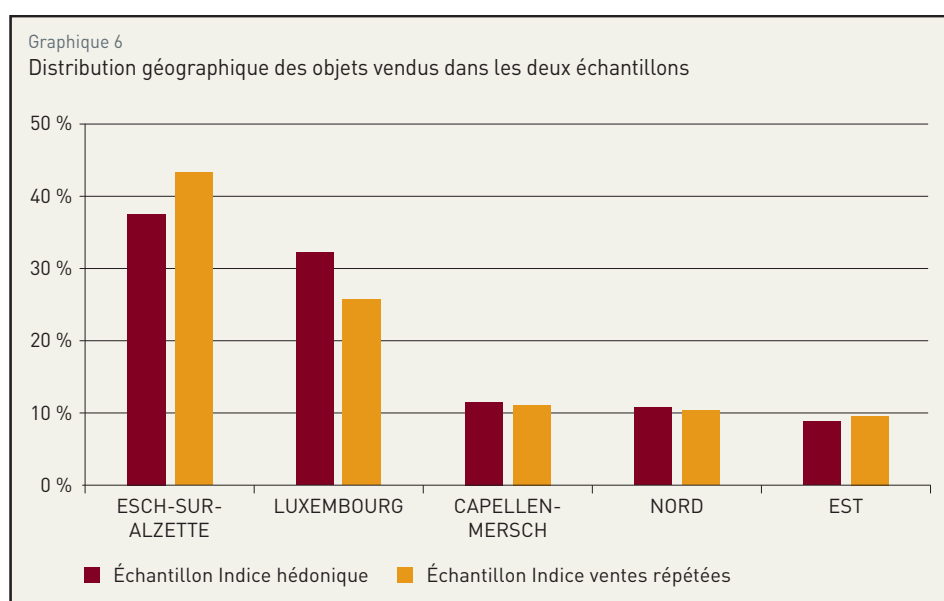
### Biais de sélection

Une autre limitation de la méthode des ventes répétées est le possible biais de sélection qui peut résulter du traitement des données. Premièrement, certaines propriétés sont achetées expressément pour être rénovées et revendues. Cette possible source de biais peut être limitée en imposant une durée minimale entre le moment des deux ventes, 12 mois dans notre cas (voir section 2). Deuxièmement, les propriétés qui ne répondent pas aux attentes initiales des acheteurs en raison de défauts cachés peuvent être revendues plus souvent. Pour cette raison, nous avons aussi éliminé les objets vendus plus de cinq fois dans notre échantillon (voir section 2). Troisièmement, les « starter homes », c'est-à-dire les logements destinés aux primo-acquéreurs, se vendent plus fréquemment car les acheteurs ont tendance à passer plus rapidement à d'autres logements plus grands ou de meilleure qualité (Jansen, 2008).

Le fait d'écartier les ventes uniques peut donc introduire un biais de sélection si un certain type de logement se revend plus fréquemment, ce qui fait que l'indice reflète moins bien l'évolution du niveau général des prix sur le marché immobilier.

Dans ce contexte, nous comparons la distribution géographique des ventes dans les deux échantillons. Comme le montre le graphique 5, l'échantillon des ventes répétées est plus concentrée sur les transactions dans la région Esch-sur-Alzette et moins concentrée sur les transactions dans la région autour de la capitale. Dans les trois autres régions du pays, la distribution reste comparable.

À part la localisation, il est difficile de comparer la distribution des caractéristiques des appartements à travers les deux échantillons. D'ailleurs, cette difficulté à observer les caractéristiques des objets vendus est précisément un des désavantages des indices hédoniques. Sur base des données disponibles, nous pouvons comparer la surface des appartements et des caves ainsi que le nombre d'emplacements de parking. Le tableau 3 montre que les surfaces des appartements et des caves ainsi que le nombre d'emplacements sont, en moyenne, très similaires entre les deux échantillons.



Sources : Administration des Domaines, de l'enregistrement et de la TVA, calculs des auteurs  
Note : La région « Nord » comprend les cantons de Clervaux, Diekirch, Redange, Vianden et Wiltz.  
La région « Est » comprend les cantons d'Echternach, Grevenmacher et Remich



Tableau 3

## Comparaison de différentes caractéristiques pour les échantillons des deux indices

	ÉCHANTILLON INDICE HÉDONIQUE	ÉCHANTILLON INDICE VENTES RÉPÉTÉES
Surface totale (m <sup>2</sup> )	81,54	81,03
Surface de la cave (m <sup>2</sup> )	7,60	6,83
Nombre emplacements parking	0,88	0,87

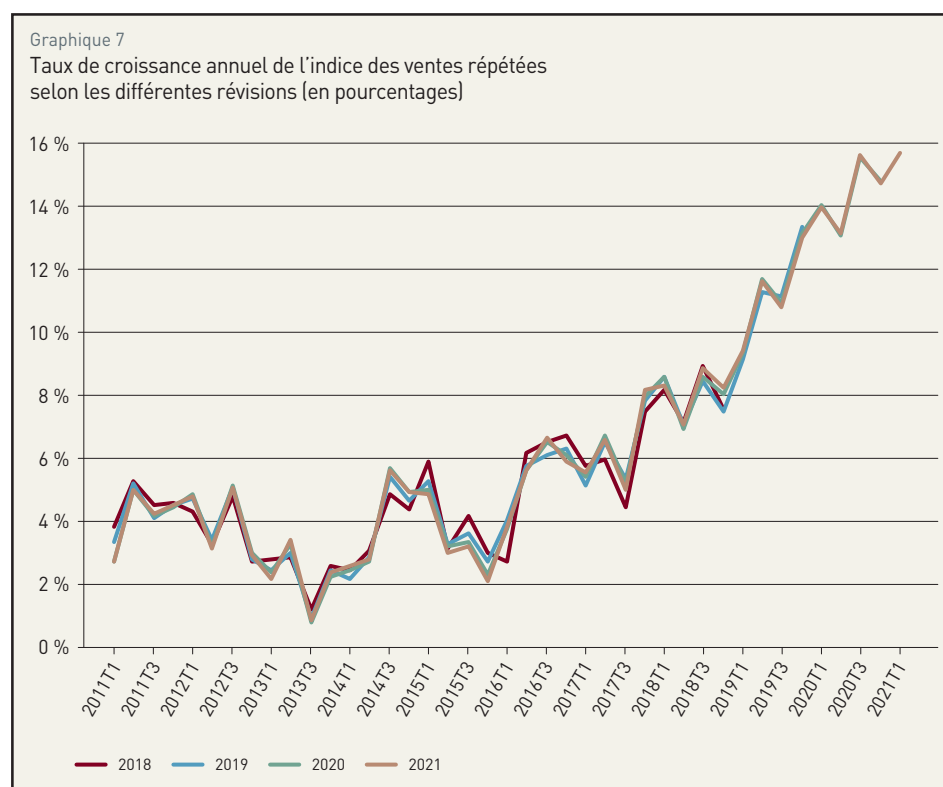
Sources : Administration des Domaines, de l'enregistrement et de la TVA, calculs des auteurs

## Révisions

Une dernière limitation du RSI est qu'il est susceptible d'être révisé chaque fois que de nouvelles transactions sont rajoutées à l'échantillon. La stabilité est souvent négligée en tant que caractéristique désirable des indices de prix, notamment pour les indices des prix des logements, étant donné leur large utilisation (Clapham, 2006). Alors qu'on pourrait supposer que pour une période donnée l'échantillon de ventes utilisé pour les indices hédoniques est définitif à la fin de la période<sup>35</sup>, l'échantillon de ventes répétées varie à travers le temps en raison des ventes de nouveaux biens, initialement exclues de l'estimation et rajoutées graduellement quand ils sont revendus. Ces nouvelles transactions contiennent des informations non seulement pour les prix actuels mais également pour les prix passés.

Afin d'évaluer la stabilité du RSI, nous avons suivi l'approche proposée par plusieurs auteurs (Baroni, 2005 ; Jansen, 2008) en ré-estimant l'indice sur différentes sous-périodes. En comparant ces différentes estimations de l'indice, on obtient une première indication de l'étendue possible des révisions susceptibles d'affecter le RSI. Le graphique 7 montre quatre versions de l'indice obtenues en rajoutant à l'échantillon 2005-2018 les transactions devenues des ventes répétées au cours des années 2019, 2020 et 2021 respectivement. Comme le montre le tableau 4 ci-dessous, 2 543 transactions en 2019 sont des deuxièmes ventes qui représentent des observations rajoutées à l'échantillon. Si l'indice est effectivement révisé au fur et à mesure que de nouvelles ventes répétées sont rajoutées, ces révisions ont relativement peu d'impact sur les taux de variations annuels. D'ailleurs, l'amplitude des révisions au RSI peut être réduite en écartant davantage des transactions avec des périodes de détention courtes. (Clapham, 2006)

<sup>35</sup> Au Luxembourg, l'estimation de l'indice hédonique s'appuie également sur les extraits des actes notariés transmis à l'AED. Or, quelques actes étant transmis avec plusieurs mois de délai, l'indice hédonique est aussi susceptible d'être révisé lorsque de nouvelles données sont mises à disposition.



Sources : Administration des Domaines, de l'enregistrement et de la TVA, calculs des auteurs

Tableau 4

**Révisions de l'indice des ventes répétées**

DERNIÈRE ANNÉE PRISE EN COMPTE	TAUX DE CROISSANCE ANNUEL MOYEN (2011 - 2018)	NOMBRE D'OBSERVATIONS	NOUVELLES OBSERVATIONS AVEC PREMIÈRE VENTE AVANT 2019
2018	4,71 %	14 619	0
2019	4,70 %	17 162	2 543
2020	4,68 %	19 904	2 721
2021	4,68 %	20 616	692

Sources : Administration des Domaines, de l'enregistrement et de la TVA, calculs des auteurs

**2.2.6 CONCLUSION**

Alors que les indices des prix immobiliers jouent un rôle important pour les décisions de politique économique, leur estimation peut poser des difficultés pratiques. Cette analyse présente un indice des prix des appartements au Luxembourg estimé par la méthode des ventes répétées. Cette approche cherche à pallier certaines limitations d'autres indices mieux connus en se fondant uniquement sur les immeubles qui ont été vendus à plusieurs reprises. Partant du principe que la qualité de ces objets individuels est constante à travers le temps, l'indice repose sur l'hypothèse que la différence de prix d'un même bien entre sa première vente et sa revente reflète l'évolution du niveau général des prix pour l'ensemble des appartements. Depuis 2011, l'indice par ventes répétées a progressé de 6,5 % par an en moyenne, ce qui est très proche de la croissance de l'indice hédonique pour les appartements anciens (6,9 %) et de celle de l'indice hédonique pour les appartements neufs (6,6 %). De plus, l'indice par ventes répétées confirme la croissance plus rapide des prix des appartements depuis 2018 (de 4,2 % entre 2011 et 2017 à 11,6 % entre 2018 et 2021). Notre analyse a également montré que l'indice par ventes répétées est moins volatil que les deux indices hédoniques, ce qui simplifie l'interprétation des signaux de prix et donc, l'orientation des politiques économiques. En conclusion, bien que l'estimation d'un indice des prix reste incertaine et que différents indices puissent diverger d'un trimestre à l'autre, l'indice par ventes répétées confirme les tendances à long terme visibles dans les indices hédoniques, ainsi que la nette croissance moyenne des prix des appartements au Luxembourg depuis 2018.

**2.2.7 RÉFÉRENCES**

Bailey, M. J., Muth, R. F., & Nourse, H. O. (1963) : *A regression method for real estate price index construction*, Journal of the American Statistical Association, 58(304), 933-942.

Baroni, M., Barthélémy, F., & Mokrane, M. (2005) : *Real estate prices : A Paris repeat sales residential index*, Journal of Real Estate Literature, 303-322.

Case, K. E., & Shiller, R. J. (1987) : *Prices of single family homes since 1970 : New indexes for four cities*, New England Economic Review, issue Sept/Oct 1987, 45-56.

Case, K. E., & Shiller, R. J. (1989) : *The Efficiency of the Market for Single-Family Homes*, American Economic Review, vol. 79(1), 125-137.

Clapham, E., Englund, P., Quigley, J. M., & Redfearn, C. L. (2006) : *Revisiting the past and settling the score : index revision for house price derivatives*, Real Estate Economics, 275-302.

Englund, P., Quigley, J. M., & Redfearn, C. L. (1999) : *The choice of methodology for computing housing price indexes : comparisons of temporal aggregation and sample definition*, The journal of real estate finance and economics, 19(2), 91-112.

Eurostat (2013) : *Handbook on Residential Property Prices Indices (RPPIs)*, Methodologies and Working papers, disponible sur <https://ec.europa.eu/eurostat/>

Jansen, S. J. T., de Vries, P. A. U. L., Coolen, H. C. C. H., Lamain, C. J. M., & Boelhouwer, P. J. (2008) : *Developing a house price index for the Netherlands : A practical application of weighted repeat sales*, The Journal of Real Estate Finance and Economics, 37(2), 163-186.

Lamboray, C. (2010) : *Un indice des prix hédonique des appartements*, Collection Economie et Statistiques, 44.

Licheron, J., & Ferring, M. (2018) : *Le logement en chiffres Numéro 7*, disponible sur [http://observatoire.liser.lu/pdfs/Logement\\_chiffres\\_2017T2.pdf](http://observatoire.liser.lu/pdfs/Logement_chiffres_2017T2.pdf)

Shiller, R. J. (1991) : *Arithmetic repeat sales price estimators*. Journal of Housing Economics, 1(1), 110-126.

Steele, M., & Goy, R. (1997) : *Short holds, the distributions of first and second sales, and bias in the repeat-sales price index*. The Journal of Real Estate Finance and Economics, 14(1), 133-154.

Thion, B., Favarger, P., & Hoesli, M. (2001) : *Indices des ventes répétées et modification de l'environnement immobilier*, Revue d'Economie Regionale Urbaine, (5), 809-830.

Wang, F. T., & Zorn, P. M. (1997) : *Estimating house price growth with repeat sales data : What's the aim of the game?*, Journal of Housing Economics, 6(2), 93-118.

## 2.2.8 ANNEXE MÉTHODOLOGIQUE

Pour estimer l'indice des prix, cette analyse utilise une modification de la méthodologie proposée par Bailey, Muth et Nourse (BMN). Plus précisément, elle s'appuie sur la méthode introduite par Case et Shiller (1987). Cette méthode, dite « *Weighted repeat sales* » (WRS) est souvent reprise en pratique et sert notamment de base pour estimer l'indice des prix du marché immobilier aux États-Unis (*Standard & Poor's CoreLogic Case-Shiller U.S. National Home Price Index*).

Case et Shiller (1987, 1989) ont fait valoir que les variations des prix des logements comprennent des composantes dont la variance augmente avec l'intervalle de temps entre les deux ventes, de sorte que l'hypothèse d'une variance constante des termes d'erreurs telle que proposée par BMN est contredite. Ils ont proposé une approche à trois étapes de type « *Weighted Least Squares* » (WLS) pour corriger ce type d'hétéroscédasticité. Derrière cette critique se trouve l'idée que l'hypothèse de qualité constante d'un bien immobilier devient moins probable lorsque l'intervalle de temps entre les deux ventes augmente. Ainsi, une école qui est construite à proximité d'un appartement est un exemple de changement aléatoire qui peut influencer le prix de vente entre les deux ventes. Pour cette raison, Case et Shiller proposent d'accorder moins de poids aux transactions lorsque la période de détention du bien augmente.

Pour construire l'indice WRS, nous partons de l'hypothèse que le prix d'un logement  $i$  à un moment  $t$  ( $P_{it}$ ) est donné par :

$$P_{it} = C_t + H_{it} + N_{it} \quad (2)$$

L'équation (2) indique que le prix d'un logement  $i$  s'explique par un niveau de prix de marché ( $C_t$ ), une marche aléatoire  $H_{it}$  qui représente une dérive du prix d'un logement individuel à travers le temps et un terme d'erreur  $N_{it}$ . Comme précédemment décrit,  $H_{it}$  cherche à capturer les variations de prix aléatoires d'un bien  $i$  à travers le temps. Pour estimer l'indice, nous allons chercher à identifier le niveau de prix de marché  $C_t$ .

Par construction, nous avons :

$$E(H_{it} - H_{i\tau}) = 0$$

$$E(H_{it} - H_{i\tau})^2 = (t - \tau)\sigma_H^2$$

et  $N_{it}$  est un bruit blanc tel que :

$$E(N_{it}) = 0$$

$$E(N_{it})^2 = \sigma_N^2$$

Supposons ensuite que la valeur d'un bien est donnée par  $V_{it} = P_{it} + Q_{it}$ , où représente un indicateur de qualité. Suivant l'hypothèse de qualité constante, nous avons  $Q_{it} = Q_{i\tau}$  et :

$$V_{it} - V_{i\tau} = P_{it} - P_{i\tau} = (+1) * C_t + (-1) * C_\tau + H_{it} - H_{i\tau} + N_{it} - N_{i\tau} \quad (3)$$

L'équation (3) peut s'approximer par :

$$V_{it} - V_{i\tau} = \sum_{s=0}^S \phi_s D_{is} + H_{it} - H_{i\tau} + N_{it} - N_{i\tau} \quad (4)$$

où  $D_{is}$  est égal à -1 pour période de la vente  $\tau$ , 1 pour période de la revente  $t$  et 0 sinon.

En se fondant sur (4), la première étape consiste à estimer l'équation (1) par la méthode des moindres carrés, c'est-à-dire :

$$\ln\left(\frac{P_{it}}{P_{i\tau}}\right) = \sum_{s=0}^S \phi_s D_{is} + \varepsilon_{it}$$

Dans un deuxième temps, le carré des résidus  $\varepsilon_{it}^2$  est régressé sur un terme constant et l'intervalle de temps  $(t - \tau)$  entre les deux ventes du bien  $i$ . En conséquence directe des hypothèses sur  $H_{it}$  et  $N_{it}$ , le terme constant est une approximation de  $2\sigma_N^2$  et le coefficient de pente une approximation de  $\sigma_H^2$ .

Troisièmement, l'équation (1) est réestimée au moyen d'une régression des moindres carrés pondérée en utilisant comme poids pour chaque observation l'inverse de la racine carrée des valeurs estimées à la deuxième étape.

Tableau A.1

**Résultats issus des trois étapes de l'estimation par la méthode WLS**

MODÈLE	R2	PARAMÈTRES	VALEUR-P
Étape 1: Régression par MCO - Sans constante	0,8182		
Étape 2: Régression par MCO	0,0534		
Constante		-0,00001	0,991
Période entre les deux ventes		0,0008653	<0,01
Étape 3: Régression par WLS - Sans Constante	0,8089		

Sources : Administration des Domaines, de l'enregistrement et de la TVA, calculs des auteurs

Le tableau A.1 montre les résultats pour les différentes régressions de chaque étape. Les valeurs sont proches de celles présentées dans le papier de Case-Shiller (1987) et Jansen (2008) et présentent les signes qui suivent l'intuition. Ainsi, le coefficient de pente issu de la deuxième étape est significatif à 1 % et positif, ce qui implique que les observations dont l'intervalle de temps est plus long vont avoir moins de poids à la troisième étape.