

CONSTRUCTION DU MODELE NON LINEAIRE « LOGARITHMIQUE-LINEAIRE » POUR L'EVALUATION IMMOBILIERE A MADAGASCAR

Henipanala Mampionona
Docteur de l'Université d'Antananarivo – Ecole Supérieure Polytechnique d'Antananarivo
Mention Urbanisme Architecture et Génie Civil
Lot: J 141 Ter Mahitsy Alasora 102 Madagascar - hmampi@yahoo.fr

Rambintsoa Tahina
Professeur de l'Université d'Antananarivo – Ecole Supérieure Polytechnique d'Antananarivo –
rambintsoatahina@gmail.com

Résumé

Lorsque le modèle linéaire ne suffit pas pour modéliser l'immobilier, il faut construire le modèle non linéaire. Il s'agit des modèles dont la spécification n'est pas linéaire, mais ses paramètres sont linéaires. Pour cette recherche, nous présentons les procédures de transformation pour les rendre linéaires. La fonction ainsi transformée n'est pas linéaire malgré le passage au logarithme, ainsi l'on ne peut pas recourir aux MCO pour estimer ses paramètres ; dans ce cas, l'on devra utiliser une méthode d'estimation des modèles non linéaires.

Un modèle semi-log est caractérisé par la présence de logarithme dans l'un de ses membres. Plus la stratification est détaillée, plus l'indice neutralise les changements intervenus dans les caractéristiques des biens couverts par l'indice immobilière.

Mots-clés : Modèle non-linéaire, Aménagement, immobilier, logarithme-linéaire.

Abstract

When the linear model is not sufficient to model real estate, it is necessary to build a non-linear model. These are models whose specification is not linear, but whose parameters are linear. For this research, we present the transformation procedures to make them linear. The transformed function is not linear despite the logarithmic transformation, so OLS cannot be used to estimate its parameters; in this case, a method for estimating non-linear models must be used.

A semi-log model is characterised by the presence of a logarithm in one of its members. The more detailed the stratification, the more the index neutralises changes in the characteristics of the properties covered by the property index.

Keywords: Non-linear model, Development, real estate, log-linear.

1- INTRODUCTION

Les hypothèses de « La Théorie de la Consommation et modèle d'équilibre » et de « Sources de Données » sont valables dans cette publication. Lorsque le modèle linéaire ne peut pas simuler le cas étudié, il faut utiliser les modèles non linéaires. En utilisant le modèle Logarithme, il n'est plus possible d'utiliser les Matrices car nous ne pouvons construire des matrices qu'avec des relations linéaires.

Toutefois, accroître le nombre de strates réduit le nombre moyen de prix observés par strate et, en fait, peut rapidement conduire à des cellules vides. Les strates ou cellules vides créent à leur tour un manque de correspondance lorsque les moyennes de prix et de quantité au sein de chaque strate sont comparées d'une période à l'autre. Une stratification très détaillée peut également faire augmenter l'écart-type de l'indice global. En outre, il peut être difficile d'identifier les principales caractéristiques déterminant le prix aussi bien que peut le faire une méthode fondée sur la régression hédonique.

2- METHODES

21- La stratification des données

La stratification (une pondération) d'un échantillon est une technique générale pour obtenir des résultats plus stables ou atténuer tout biais dû à la sélection de l'échantillon, y compris aux non-réponses. Dans le contexte de l'indice des prix de l'immobilier, l'échantillon de propriétés vendues est subdivisé en un nombre de strates (ou de cellules) relativement homogènes selon un nombre (limité) de caractéristiques qui déterminent le prix. Le prix moyen (valeur

unitaire) ou la médiane des prix peut ensuite être utilisé pour calculer un indice de prix pour chaque strate. Dans une deuxième étape, ces indices de strate sont agrégés en utilisant les pondérations provenant des ventes ou des stocks. Cette méthode, dite post-stratification, a souvent été utilisée pour ajuster les changements de composition de l'échantillon, ou les changements de composition de la qualité des maisons vendues.

La stratification peut aussi être utilisée avec d'autres méthodes d'ajustement des changements de qualité, comme les régressions hédoniques.

Les principaux avantages de la méthode de la stratification ou post-stratification sont les suivants :

- En fonction du choix des variables de stratification, la méthode permet de tenir compte des changements intervenus dans la composition de l'échantillon de logements.
- La méthode est reproductible, pour autant que l'on dispose d'une liste convenue de variables de stratification.
- Elle ne risque pas de faire l'objet de révisions.
- On peut élaborer des indices de prix pour différents types et emplacements de logements.
- La méthode est relativement facile à appliquer et à expliquer aux utilisateurs.

Les principaux inconvénients de la méthode de la stratification ou post-stratification sont les suivants :

- Elle ne permet pas de traiter comme il se doit la dépréciation des habitations, à moins que l'ancienneté du bâtiment soit une variable de stratification. Cette situation peut entraîner des problèmes de strates ne comportant que peu d'observations de prix.
- La méthode ne permet pas de traiter comme il se doit le cas des habitations ayant fait l'objet de réparations ou de rénovations importantes (sauf si l'on dispose d'informations sur les rénovations). - Elle nécessite des informations sur les caractéristiques des habitations figurant dans les strates si l'on veut que les ventes puissent être affectées aux bonnes strates.
- Si le mécanisme de stratification est très grossier, les changements de composition vont retentir sur les indices.
- Si le mécanisme de stratification est très fin, les strates peuvent faire l'objet d'une grande variabilité d'échantillonnage du fait de la petitesse des échantillons; de la même manière, certaines strates peuvent tout simplement être vides pour certaines périodes, ce qui entraîne des difficultés pour l'élaboration de l'indice.

La stratification ou post stratification convient dès lorsque :

- On choisit pour les strates un niveau de détail adapté et applicable dans la pratique.
- La tranche d'âge du bâtiment est l'une des variables de stratification. - On ne cherche pas à décomposer l'indice en bâtiments et en terrains.

La stratification ou post-stratification est recommandée dès lors que le volume des ventes est suffisamment important et que les informations sur les caractéristiques du logement sont suffisamment détaillées pour permettre une classification détaillée des biens.

22- Relation fonctionnelle mathématique du « Modèle Log-linéaire »

Nous partons de l'hypothèse que le prix p_n^t du bien n pendant la période t est fonction d'un nombre fixe K , de caractéristiques mesurées par les «quantités» z_{nk}^t . Avec les $T+1$ périodes temporelles, comprises entre la période de référence 0 et la période T , nous avons :

$$p_n^t = f(z_{n1}^t, \dots, z_{nK}^t, \varepsilon_n^t) \quad (t=0, \dots, T)$$

Avec les techniques de régression, l'équation est représentée sous forme d'équation paramétrique avec le modèle Log-linéaire :

$$\ln p_n^t = \beta_0^t + \sum_{k=1}^K \beta_k^t z_{nk}^t + \varepsilon_n^t$$

Avec : ε_n^t : terme d'erreur de spécification

β_0^t : constante

β_k^t : paramètres estimés du modèle

La valeur d'un bien immobilier correspond généralement à la somme du prix du bâtiment et du prix du terrain, on peut inclure de façon linéaire les terrains et les bâtiments dans le modèle si l'on dispose des données nécessaires.

Toutes les sources de données ne contiennent malheureusement pas d'informations sur la taille des terrains et des bâtiments, d'où l'impossibilité de les traiter de façon linéaire. Celles concernant la taille des terrains en particulier risquent de faire défaut. Quand la taille du terrain (ou du bâtiment) n'est pas incluse comme variable explicative, le recours à des modèles log-linéaires est relativement satisfaisant.

Les paramètres des caractéristiques β_k dans l'équation (5.15) peuvent varier dans le temps. Cela cadre avec l'idée selon laquelle les conditions du marché du logement déterminent les contributions des caractéristiques : il n'y a en effet aucune raison de penser a priori que ces contributions sont constantes quand les conditions de l'offre et de la demande changent. Les conditions du marché ont cependant de très fortes chances de changer de façon progressive. On peut donc formuler l'hypothèse simplificatrice de la stabilité dans le temps des paramètres des caractéristiques. Par ailleurs, la constante β_0 peut être convertie en un indice des prix à qualité constante. Dans le cas du modèle log-linéaire, cela conduirait à l'équation :

$$\ln p'_n = \beta_0 + \sum_{k=1}^K \beta_k z'_{nk} + \varepsilon'_n$$

Dans les hypothèses d'erreur classiques, en particulier celles de moyenne nulle et de variance constante, on peut estimer les paramètres du modèle hédonique à l'aide d'une régression effectuée selon la méthode des moindres carrés ordinaires (MCO) sur les données des échantillons de chaque période, séparément. L'équation peut être estimée à partir des données regroupées des différentes périodes à condition d'inclure des variables indicatrices des périodes temporelles (moins une pour éviter une colinéarité parfaite). L'équation d'estimation pour le modèle log-linéaire, généralement appelé modèle hédonique à variables indicatrices, devient ainsi :

$$\ln p'_n = \beta_0 + \sum_{\tau=1}^T \delta^\tau D_n^\tau + \sum_{k=1}^K \beta_k z'_{nk} + \varepsilon'_n$$

Avec : D_n^τ : variable indicatrice temporelle. Elle prend la valeur 1 si l'observation correspond à la période τ et la valeur 0 dans le cas contraire.

23- Résolution utilisant la « Méthode des variables indicatrices temporelles »

La méthode des variables indicatrices temporelles est souvent utilisée pour construire un indice hédonique des prix des logements. L'un de ses avantages est sa simplicité; l'indice des prix découle directement de l'équation estimée de la régression sur les variables indicatrices temporelles regroupées. En effectuant une régression globale sur les données regroupées des échantillons $S(0), S(1), \dots, S(T)$ concernant les périodes $t = 0, \dots, T$ (avec les superficies $N(0), N(1), \dots, N(T)$), on obtient les coefficients $\hat{\beta}^0$, $\hat{\delta}^t$ ($t=1, \dots, T$) et $\hat{\beta}_k$ ($k=1, \dots, K$). Le paramètre de la variable indicatrice temporelle fait monter ou descendre la superficie hédonique et mesure l'effet du « temps » sur le logarithme du prix. L'exponentiation des coefficients des variables indicatrices temporelles permet donc de neutraliser les effets des variations quantitatives des caractéristiques et donne une indication de l'évolution des prix des logements ajustés pour tenir compte des changements de qualité entre la période de référence 0 et chaque période de comparaison t . Autrement dit, l'indice à variables indicatrices temporelles entre la période 0 et la période t est donné par :

$$P_{TD}^{0t} = \exp(\hat{\delta}^t)$$

Les coefficients de régression $\hat{\beta}_k$ seront généralement affectés de moindres écarts-types que les coefficients $\hat{\beta}_k^t$ qui seraient obtenus en estimant le modèle séparément à l'aide des données des échantillons $S(0), S(1), \dots, S(T)$.

Quand on utilise les moindres carrés ordinaires (MCO), l'indice hédonique à variables indicatrices temporelles peut être exprimé sous la forme suivante :

$$P_{TD}^{0t} = \frac{\prod_{s=0}^{t-1} (p_s)^{1/N(s)}}{\prod_{s=0}^{t-1} (p_s)^{1/N(0)}} \exp \left[\sum_{k=1}^K \hat{\beta}_k (\bar{z}_k^0 - \bar{z}_k^t) \right]$$

où $\bar{z}_k^s = \sum_{n \in S(s)} z'_{nk} / N(s)$ est la moyenne empirique de la caractéristique k pendant la période s ($s = 0, t$).

Cette équation nous dit que l'indice à variables indicatrices temporelles est essentiellement le produit de deux facteurs. Le premier facteur est le ratio des prix exprimé sous forme de moyenne géométrique pendant les périodes t et 0.

Le deuxième facteur, $\exp \left[\sum_{k=1}^K \hat{\beta}_k (\bar{z}_k^0 - \bar{z}_k^t) \right]$, ajuste ce ratio des moyennes empiriques brutes pour tenir compte des différences entre les caractéristiques moyennes \bar{z}_k^0 et \bar{z}_k^t ; il fait fonction de facteur d'ajustement pour la qualité qui prend en compte à la fois l'évolution des composantes de la qualité et les changements observés dans la qualité des biens immobiliers (à condition que tous les attributs pertinents déterminant la qualité soient inclus dans le modèle

hédonique). Il convient de noter que l'indice des prix à variables indicatrices temporelles se réduit au ratio des prix exprimé sous forme de moyenne géométrique $\bar{z}_k^0 = \bar{z}_k^t$, c'est-à-dire s'il se trouve que les caractéristiques moyennes sont égales pendant la période t et la période 0.

Les échantillons de logements achetés ou vendus, $S(0)$ et $S(t)$, peuvent ne pas être représentatifs du stock total de logements (ou du nombre total de logements vendus). Une solution pourrait être de regrouper les échantillons de manière à les rendre représentatifs. Effectuer une régression des MCO sur l'ensemble de données regroupées (ou pondérées) équivaut à réaliser une régression selon la méthode des moindres carrés pondérés (MCP) sur la série initiale de données. Il convient de noter que l'utilisation de la méthode des variables indicatrices temporelles avec recours aux MCP permettra aussi d'obtenir un indice reposant sur la moyenne géométrique qui, dans ce cas, sera pondéré.

Un problème, ou une limite, que pose la méthode des variables indicatrices temporelles est la révision qu'elle implique. Si la série temporelle est étendue de T à T + 1 avec de nouvelles données d'observations ajoutées, les coefficients des caractéristiques s'en trouveront modifiés. Les nouveaux chiffres de l'indice des prix calculés pour les périodes $t = 1, \dots, T$ différeront en conséquence de ceux antérieurement obtenus. Lorsque de nouvelles données sont disponibles, l'efficacité liée au regroupement des données augmente et de meilleures estimations peuvent être effectuées, ce qui peut être en fait considéré comme un avantage plutôt que comme un inconvénient de la méthode.

La méthode des variables indicatrices temporelles couvrant plusieurs périodes semble donc présenter un intérêt limité pour la production d'indices des prix des logements, même si le problème des révisions peut être surmonté de plusieurs façons. L'une d'elles serait d'estimer les indices à variables indicatrices temporelles pour les périodes adjacentes t-1 et t, puis de les multiplier pour obtenir une série temporelle ne nécessitant pas de révisions. Ce chaînage des données de fréquence élevée présente en outre l'avantage d'assouplir l'hypothèse des paramètres fixes.

3- RESULTAT : APPLICATION DU MODELE NON LINEAIRE «Log-Linéaire» AVEC PLUSIEURS VARIABLES ENDOGENES ET DANS LE TEMPS

3.1 La Stratification des données et calcul de moyennes- médian- première approche de régression

Les calculs de la moyenne et de la médiane des prix, font partie des méthodes les moins coûteuses en données. La moyenne simple et la médiane n'ont besoin que des prix de vente de l'immobilier d'un endroit donné pour calculer un indice de prix. L'information géographique sera en conséquence nécessaire.

Dans un premier temps, nous construisons un indice en utilisant la moyenne des prix. La moyenne simple des prix relevés d'un échantillon de maisons est calculée pour une période et un lieu (ou endroit) données. Cet indicateur, qui peut être exprimé en valeur monétaire ou par un indice, mesure en pourcentage, la variation du prix moyen de l'échantillon entre deux périodes. Les observations aberrantes doivent aussi être identifiées et retirées de l'échantillon si l'on pense que cela pourrait fausser ou biaiser le résultat.

Pour simplifier la compréhension, prenons encore le cas d'Ivato, utilisant 5 observations de prix en 2018 et 7 observations en 2019, et mesurons l'évolution de la moyenne simple des prix d'un appartement de 80m² de Surface à cet endroit :

- Pour l'année 2018, le prix du logement et la moyenne est : $(350.000 + 352.000 + 378.000 + 366.000 + 402.000)/5 = 370.000$ Millions d'Ariary.
- Pour l'année 2019, le prix du logement et la moyenne est : $(360.000 + 350.000 + 382.000 + 395.000 + 380.000 + 400.000 + 450.000)/7 = 388.000$ Millions d'Ariary.

Une fois que les prix moyens pour chaque année sont obtenus, nous pouvons calculer l'évolution d'une période à l'autre, exprimée en pourcentage. Pour le cas d'Ivato, entre 370.000 Millions d'Ariary et 388.000 Millions d'Ariary, le prix moyen des maisons a augmenté de 5 % entre les deux périodes 2018 et 2019.

L'influence des observations aberrantes est réduite lorsque la médiane des prix de l'échantillon est utilisée à la place de la moyenne. Par exemple, si plusieurs maisons ou appartements très chères sont vendues pendant une période donnée, la moyenne des prix ne sera pas représentative des prix du marché à cette période. La médiane est un indicateur très utilisé des prix de l'immobilier, principalement en raison de sa simplicité, son économie en données et donc la rapidité de sa disponibilité.

En prenant les mêmes données du calcul de la moyenne des années 2018 et 2019, la médiane des prix de ces périodes est de 366.000 Millions d'Ariary pour l'année 2018 et de 382.000 Millions d'Ariary pour l'année 2019. Le prix médian a donc augmenté de 4,4 % entre les deux périodes.

32- Cas réel sur Ivato

Prenons le cas réel étudié pour Ivato. Nous prenons un échantillon de 587 prix de ventes effectives de maisons individuelles pendant plusieurs années. Dans les calculs, on constate que la moyenne des prix du logement est supérieure à la médiane, selon le tableau 1 suivant :

Tableau 1 - Tableau des Ecart-type ; Moyenne et Médiane d'échantillons de vente à Ivato

	2013	2014	2015	2016	2017	2018	2019
Observations	77	84	89	82	84	92	79
Écart-type	64.130	62.042	73.405	76.432	84.587	96.559	96.152
Moyenne (en Millions d'Ariary)	249.702	270.174	290.686	299.087	315.099	347.009	365.190
Variation en %		8,2%	7,6%	2,9%	5,4%	10,1%	5,2%
Médiane (en Millions d'Ariary)	236.000	255.000	273.000	280.000	292.000	319.000	340.600
Variation en %		8,1%	7,1%	2,6%	4,3%	9,2%	6,8%

Nous savons que l'emplacement est un facteur déterminant du prix d'une maison et de son évolution dans le temps. Ainsi, pour améliorer la fiabilité d'un indicateur, il faut stratifier ou post stratifier l'estimateur ou l'information pour segmenter le marché (ou l'échantillon de ventes) est disponible. La stratification géographique consiste à regrouper les maisons selon un critère géographique (dans notre cas, nous prenons la zone d'Ivato), assurant ainsi une certaine homogénéité des logements de la strate. La stratification permet aussi de donner aux utilisateurs des informations sur les prix de l'immobilier selon différents segments du marché, comme le type de maison ou d'appartement ou le voisinage.

- a) Résultats de la régression MCO de la modélisation « Log-Linéaire » pour l'évaluation immobilière à Ivato, sans tenir compte du temps

Après calcul, nous obtenons les résultats suivants pour l'année 2019 (Tableau 2):

Tableau 2- Tableau de résultats de la modélisation Log-linéaire pour Ivato sans tenir compte du temps

	SS	Df	MS	Résultat économétrique	
Modèle	20,0634692	4	5,0158673	Nombre d'observations = 79 $F(4, 75) = 156,02$ $\text{Prob}>F = 0,0000$ $R^2 = 0,441$ $R^2 \text{ ajusté}_{\text{par construction}} = 0,4382$ $\sqrt{\text{MSE}} = 0,1793$	
Résidu	25,4293063	75	0,032148301		
Total	45,4927755	79	0,057223617		
Prix	Coef	Erreur-type	T	Intervalle de confiance 95%	
Chambres	0,1156791	0,0098159	11,78	0,0964108	0,1349473
Salle de Bains	0,0999522	0,0095996	10,41	0,0811086	0,1187958
Ancienneté	-0,002561	0,0004173	-6,14	-0,0033801	-0,001742
Terrain	$9,39 \cdot 10^{-6}$	$1,28 \cdot 10^{-6}$	7,31	$6,87e-06$	0,0000119
Construction	12,0647	0,0383342	314,72	11,98945	12,13995

➤ Lecture du résultat :

Le modèle hédonique exprime le prix d'un bien comme une fonction des caractéristiques qui le déterminent. Le prix du logement (variable endogène) étudié à Ivato est en fonction des variables exogènes suivantes :

- nombre de chambres
- nombre de salles de bains

- l'âge du logement ou son ancienneté
- la superficie du terrain
- la qualité générale de la construction (pour l'ajustement de R²)

Comme échantillon, nous avons observé 79 prix.

La valeur du test de Student qui est non nulle montre que les variables explicatives ont un effet sur le résultat attendu. Le coefficient de détermination R² vaut 0,441. Cela signifie que les variations des 5 variables exogènes ci-dessus expliquent 44% de la variance du prix des logements. Si nous ajoutons d'autres variables explicatives ou variables exogènes (comme la présence d'une cheminée et d'un garage,... et on prend la qualité de la construction pour notre cas), la valeur de R² augmente pour s'approcher de 1.

Les résultats des régressions Logarithmique-Linéaire peuvent s'interpréter comme suit :

- Chaque chambre en plus augmente le prix du logement de 11,56 %, toutes choses étant égales par ailleurs.
- Un logement avec une salle de bain supplémentaire coûte presque 9,99 % plus cher qu'une maison sans salle de bain supplémentaire, toutes choses étant égales par ailleurs.
- Une année d'ancienneté de plus fait baisser le prix d'un logement (ou le logement se déprécie) de 0,256 %, toutes choses étant égales par ailleurs.
- Un mètre carré de terrain en plus augmente le prix du logement de 0,000939 %, toutes choses étant égales par ailleurs. (nous nous considérons des appartements, mais non des villas individuelles).

« Toutes choses étant égales par ailleurs » signifient que les deux logements ne diffèrent que par la variable exogène considérée.

- b) Calcul des indices de prix hédoniques par la Méthode des variables indicatrices temporelles (tenant compte du temps)

Reprenons l'équation pour la Méthode des variables indicatrices temporelles :

$$\ln p_n^t = \beta_0 + \sum_{\tau=1}^T \delta^\tau D_n^\tau + \sum_{k=1}^K \beta_k z_{nk}' + \varepsilon_n^t$$

D_n^t représente l'indicatrice qui est égale à un si l'observation provient de la période τ (τ = 1,..., T) et sinon elle vaut zéro. La variable indicatrice temporelle pour la période de référence 0 est laissée de côté pour éviter toute ambiguïté de notation avec la constante β₀.

L'exponentielle du coefficient estimé par la régression δ^τ mesure la variation en pourcentage «à qualité constante» du prix des propriétés entre la période de référence et la période t. L'estimation de δ^t peut être interprétée comme la variation du logarithme du prix due au passage du temps, les autres variables étant constantes. L'exponentielle de δ¹ donne l'estimation de l'indice des prix de la période 1 :

$$P_{TD}^{01} = \exp(\hat{\delta}^1)$$

Le même raisonnement est mené pour toutes les autres périodes. Ainsi, l'indice des prix par la méthode des variables indicatrices temporelles, de la période de référence à la période de comparaison est :

$$P_{TD}^{0t} = \exp(\hat{\delta}^t)$$

NB : la période de référence est ainsi l'indice 1.

Nous utilisons les données de ventes pendant deux années (années 2017 et 2018) dans cette régression groupée (on utilise le terme « groupée » car la régression regroupe des années):

$$\ln p_n^t = \beta_0 + \beta_1 \text{Taille du terrain}_n + \beta_2 \text{Chambre}_n + \beta_3 \text{Bain}_n + \beta_4 \text{Ancienneté}_n + \delta^1 D_n^1 + \varepsilon_n^t$$

P_n^t : prix de la maison n pour les années comme variable endogène.

Les variables exogènes ou explicatives sont : la taille du terrain, le nombre de chambres, le nombre de salles de bain et l'ancienneté (age);

B₀ à β₄ : les paramètres estimés du modèle.

ε_n^t : terme d'erreur

Le coefficient d'intérêt est celui associé à l'année 2017 vers 2018, δ^{2018} . Sa valeur est de 0,0781548. Ce coefficient est ensuite utilisé pour estimer la variation en pourcentage des prix pour les maisons entre les années 2017 et 2018. Cette transformation consiste à prendre l'exponentielle du coefficient δ^{2018} : $P_{TD}^{2017/2018} = \text{expo}(0,0781548) = 1,08129$. Ainsi, en gardant constantes toutes les caractéristiques, la variation en pourcentage du prix des logements entre 2017 et 2018 est de 8,129 % par la méthode de la régression hédonique. Notons que la moyenne et la médiane ont conduit à des augmentations respectives de 10,1 % et de 9,2 %, pour la même période.

En ajoutant l'année 2019 comme troisième période, l'équation de la régression devient :

$$\ln p'_n = \beta_0 + \beta_1 \text{Taille du terrain}_n + \beta_2 \text{Chambre}_n + \beta_3 \text{Bain}_n + \beta_4 \text{Ancienneté}_n + \delta^1 D_n^1 + \delta^2 D_n^2 + \varepsilon'_n$$

La valeur du coefficient de l'indicatrice temporelle pour l'année 2018 est de 0,1332734. Son exponentielle $\text{expo}(0,1332734) = 1,14$. Cela signifie que l'augmentation à qualité constante de l'indice du prix des maisons est de 14 % entre 2017 et l'année 2019. En revanche, les augmentations des prix données par la moyenne et la médiane pour la même période sont respectivement de 16 % et de 17 %.

Cette technique peut être généralisée à plus que trois périodes. Cela consiste à combiner les données de plus de périodes et à ajouter plus de variables indicatrices temporelles. Cependant, des régressions groupées avec plusieurs périodes ne sont pas nécessairement idéales pour construire des séries chronologiques, car l'ajout d'une nouvelle période modifiera légèrement les résultats obtenus pour les périodes antérieures. La stabilité des coefficients d'une régression groupée peut devenir problématique quand le nombre de périodes augmente.

Pour illustrer ce qui est dit ci-dessus, à partir des résultats obtenus, nous pouvons construire une série chronologique par chaînage des indices de deux périodes consécutives :

$$P_{TD, \text{chaînage}}^{2019/2017} = P_{TD}^{2019/2018} \times P_{TD}^{2018/2017} = 1,057 \times 1,081 = 1,143$$

Ce résultat est seulement légèrement différent de celui obtenu avec la régression groupée (voir le résultat dans le tableau 3) où la variation des prix entre les années 2017 et 2019 est estimée à 14,0 %. Maintenant, en chaînant les indices des variables indicatrices temporelles sur les deux périodes temporelles consécutives, la variation estimée des prix est de 14,3 %. Cette différence, assez légère, est expliquée par la linéarisation.

Tableau 3 - Résultats de la régression groupée pour les années 2017 et 2018

	SS	df	MS	Résultat économétrique	
Modèle	48,4501865	4	9,6900373000,	Nombre d'observations = 171 $F(4, 167) = 286,64$ $\text{Prob}>F = 0,0000$ $R^2 = 0,4571$ $R^2 \text{ ajusté}_{\text{par construction}} = 0,4555$ $\sqrt{\text{MSE}} = 0,18386$	
Résidu	57,5372376	167	033805663		
Total	105,987424	171	0,057223617		
Prix	Coeff	Erreur-type	t	Intervalle de confiance 95%	
Chambres	0,0840483	0,0069071	12,17	0,0705009	0,0975957
Salle de Bains	0,121815	0,0071529	17,03	0,1077855	0,1358444
Ancienneté	-0,0029137	0,0003183	-9,15	-0,0035381	-0,0022894
Terrain	0,0000137	9,24e-07	14,78	0,0000119	0,0000155
δ^{2018}	0,0781548	0,0089128	8,77	0,0606736	0,095636
Construction	11,96531	0,0273032	438,24	11,91176	12,01886

Tableau 3- Résultats de la régression groupée pour 2017 et 2018

➤ Lecture du résultat :

Le modèle exprime le prix d'un bien comme une fonction des caractéristiques qui le déterminent. Le prix du logement étudié est en fonction des variables exogènes suivantes :

- nombre de chambres
- nombre de salles de bains
- l'âge du logement ou son ancienneté

- la superficie du terrain
- le coefficient de l'indicatrice temporelle pour l'année 2018
- la qualité générale de la construction (pour l'ajustement de R²)

La valeur du test de Student qui est non nulle montre que les variables explicatives ont un effet sur le résultat attendu. Le coefficient de détermination R² vaut 0,4571. Cela signifie que les variations des 5 variables exogènes ci-dessus expliquent 45,71% de la variance du prix des logements. Si nous ajoutons d'autres variables explicatives ou variables exogènes (comme la présence de cheminée et d'un garage), la valeur de R² augmente et s'approche de 1. Les résultats des régressions Logarithmique-Linéaire peuvent s'interpréter comme suit pour 2017 et 2018:

- Chaque chambre en plus augmente le prix du logement de 12,17%, toutes choses étant égales par ailleurs.
- Un logement avec une salle de bain supplémentaire coûte presque 17,03% plus cher qu'une maison sans salle de bain supplémentaire, toutes choses étant égales par ailleurs.
- Une année d'ancienneté de plus fait baisser le prix d'un logement (ou le logement se déprécie) de 9,15%, toutes choses étant égales par ailleurs.
- Un mètre carré de terrain en plus augmente le prix du logement de 14,78 %, toutes choses étant égales par ailleurs. (nous nous considérons des appartements, mais non des villas individuelles).
- le coefficient de l'indicatrice temporelle pour l'année 2018 augmente le prix du logement de 8,77%, toutes choses étant égales par ailleurs.

« Toutes choses étant égales par ailleurs » signifient: logements ne diffèrent que par la variable exogène considérée. Le Tableau 4 montre les résultats de la régression groupée pour les années 2017 à 2019.

Tableau 4 - Résultats de la régression groupée pour les années 2017 à 2019

	SS	Df	MS	Résultat économétrique	
Modèle	73,4886776	4	12,2481129	Nombre d'observations = 171 F(4, 167) = 366,64 Prob>F = 0,0000 R ² = 0,4684 R ² ajusté par construction = 0,4671 $\sqrt{MSE} = 0,18277$	
Résidu	83,4154327	167	0,033406261		
Total	156,90411	171	0,0626864		
Prix	Coef	Erreur-type	t	Intervalle de confiance 95%	
Chambres	0,0942001	0,0056566	16,65	0,083108	0,1052923
Salle de Bains	0,1139931	0,0057443	19,84	0,102729	0,1252572
Ancienneté	-0,0028112	0,0002538	-11,08	-0,0033089	-0,0023135
Terrain	0,0000122	7,51e-07	16,28	0,0000108	0,0000137
δ^{2018}	0,0781257	0,008856	8,82	0,0607598	0,0954916
δ^{2019}	0,1332734	0,0090681	14,70	0,1154916	0,1510552
Construction	11,95724	0,0225891	529,34	11,91295	12,00154

➤ Lecture du résultat :

Le modèle exprime le prix d'un bien comme une fonction des caractéristiques qui le déterminent. Le prix du logement étudié est en fonction des variables exogènes suivantes :

- nombre de chambres
- nombre de salles de bains
- l'âge du logement ou son ancienneté
- la superficie du terrain
- le coefficient de l'indicatrice temporelle pour l'année 2018
- le coefficient de l'indicatrice temporelle pour l'année 2019
- la qualité générale de la construction (pour l'ajustement de R²)

La valeur du test de Student qui est non nulle montre que les variables explicatives ont un effet sur le résultat attendu. Le coefficient de détermination R² vaut 0,4684. Cela signifie que les variations des 5 variables exogènes ci-dessus expliquent 46,84% de la variance du prix des logements. Si nous ajoutons d'autres variables explicatives ou variables exogènes (comme la présence d'une cheminée et d'un garage), la valeur de R² augmente pour s'approcher de 1.

- Chaque chambre en plus augmente le prix du logement de 16,65%, toutes choses étant égales par ailleurs.

- Un logement avec une salle de bain supplémentaire coûte presque 19,84% plus cher qu'une maison sans salle de bain supplémentaire, toutes choses étant égales par ailleurs.
- Une année d'ancienneté de plus fait baisser le prix d'un logement (ou le logement se déprécie) de 11,08%, toutes choses étant égales par ailleurs.
- Un mètre carré de terrain en plus augmente le prix du logement de 16,28 %, toutes choses étant égales par ailleurs. (nous nous considérons des appartements, mais non des villas individuelles).
- le coefficient de l'indicatrice temporelle pour l'année 2018 augmente le prix du logement de 8,82%, toutes choses étant égales par ailleurs.
- le coefficient de l'indicatrice temporelle pour l'année 2019 augmente le prix du logement de 14,70%, toutes choses étant égales par ailleurs.

Tableau 5 montre les résultats de la régression groupée pour les années 2018 à 2019.

Tableau 5 - Résultats de la régression groupée pour les années 2018 et 2019 :

	SS	Df	MS	Résultat économétrique	
Modèle	45,441478	5	9,08829560,	Nombre d'observations = 1.670 F(4, 791) = 271,91 Prob>F = 0,0000 R ² = 0,4497 R ² ajusté par construction = 0,4480 √MSE = 0,18282	
Résidu	55,6172267	1.664	0,033423814		
Total	101,058705	1.669	0,060550452		
Prix	Coeff	Erreur-type	t	Intervalle de confiance 95%	
Chambres	0,1041401	0,0068861	15,12	0,0906337	0,1176465
Salle de Bains	0,1070142	0,0068881	15,54	0,093504	0,1205244
Ancienneté	-0,0026926	0,0003045	-8,84	-0,0032899	-0,0020953
Terrain	0,0000117	9,42e-07	12,42	9,85e-06	0,0000135
δ ²⁰¹⁹	0,0555370	0,0089625	6,20	0,073116	0,037958
Construction	12,07482	0,026871	449,36	12,02212	12,12753

➤ Lecture du résultat :

Le modèle exprime le prix d'un bien comme une fonction des caractéristiques qui le déterminent. Le prix du logement étudié est en fonction des variables exogènes suivantes :

- nombre de chambres
- nombre de salles de bains
- l'âge du logement ou son ancienneté
- la superficie du terrain
- le coefficient de l'indicatrice temporelle pour l'année 2019
- la qualité générale de la construction (pour l'ajustement de R²)

La valeur du test de Student qui est non nulle montre que les variables explicatives ont un effet sur le résultat attendu. Le coefficient de détermination R² vaut 0,4497. Cela signifie que les variations des 5 variables exogènes ci-dessus expliquent 44,97% de la variance du prix des logements. Si nous ajoutons d'autres variables explicatives ou variables exogènes (comme la présence de cheminée et de garage.), la valeur de R² augmente pour s'approcher de 1. Les résultats des régressions Logarithmique-Linéaire peuvent s'interpréter comme suit :

- Chaque chambre en plus augmente le prix du logement de 15,12%, toutes choses étant égales par ailleurs.
- Un logement avec une salle de bain supplémentaire coûte presque 15,54% plus cher qu'une maison sans salle de bain supplémentaire, toutes choses étant égales par ailleurs.
- Une année d'ancienneté de plus fait baisser le prix d'un logement (ou le logement se déprécie) de 8,84%, toutes choses étant égales par ailleurs.
- Un mètre carré de terrain en plus augmente le prix du logement de 12,42 %, toutes choses étant égales par ailleurs. (nous nous considérons des appartements, mais non des villas individuelles).
- le coefficient de l'indicatrice temporelle pour l'année 2018 augmente le prix du logement de 6,20%, toutes choses étant égales par ailleurs.

« Toutes choses étant égales par ailleurs » signifient que les deux logements ne diffèrent que par la varia

4-DISCUSSION

La pondération est l'ensemble de nombres dont la somme est égale à l'unité, utilisés pour calculer des moyennes. Dans le contexte des paramètres du prix de l'immobilier, les pondérations sont généralement des parts de dépenses (ventes) ou des parts de valeur de stock dont la somme est, par définition, égale à l'unité. Elles sont utilisées pour calculer la moyenne des rapports de prix de chaque propriété.

La méthode, qui apporte une plus de précision, consiste à observer l'évolution des prix d'une maison donnée au cours d'une certaine période, sur la base du prix de vente constaté chaque fois que la maison est vendue. L'évolution du prix d'un échantillon de biens immobiliers au cours de périodes se chevauchant peut être observée pour estimer, à l'aide d'un modèle de régression à variable indicatrice, la tendance générale de l'évolution du prix des biens résidentiels. Le fait de mesurer l'évolution des prix moyens à l'occasion de ventes répétées des mêmes biens permet une comparaison «à l'identique» (sans tenir compte du fait que la dépréciation et les rénovations apportées aux bâtiments entre les périodes de vente peuvent modifier le bien).

5- CONCLUSION

Les contributions de la recherche les plus pertinentes du point de vue de notre problématique sont les outils d'ajustements des changements de qualités. Les indices des prix fondés sur des régressions hédoniques afin de prendre en compte les variations dans la composition des échantillons au sein des strates permettent d'utiliser des valeurs relatives du stock de logements pour pondérer les indices corrigées des variations de composition. Cette spécificité de notre approche s'est donnée pour objectif de corriger les principales faiblesses des autres méthodes d'évaluation et permet de faire une utilisation optimale des données disponibles.

Avec le développement des techniques de localisation par GPS et avec l'utilisation des adresses cadastrales grâce à leur numérisation progressive et le développement des bases de données immobilières, notamment notariales, dans lesquelles les transactions sont indexées par leur adresse, il est devenu possible de comparer avec précision dans un secteur donné, les prix de plusieurs biens dont on a identifié les caractéristiques principales. La généralisation de cette nouvelle technologie qui permet de déterminer le prix d'un bien à partir de toutes les dernières transactions connues, qu'il s'agisse de contrats de vente ou de compromis, puis de le positionner sur une carte, devrait progressivement remplacer toutes les méthodes d'évaluation précédemment décrites. Les seuls freins à sa diffusion demeurent, encore pour le moment, les restrictions imposées pour la diffusion du prix des transactions.

REFERENCES BIBLIOGRAPHIQUES

- Rossini, P., and P. Kershaw (2006), "Developing a Weekly Residential Price Index Using the Sales Price Appraisal Ratio, paper presented at the Twelfth Annual Pacific Rim Real Estate Society Conference, Auckland, 22–25 January.
- Shiller, R.J. (1991), "Arithmetic Repeat Sales Price Estimators," *Journal of Housing Economics* 1(1), 110–26.
- Thibodeau, T.G. (1992), "Residential Real Estate Prices: 1974–1983," The Blackstone Company, Mount Pleasant (MI).
- Quigley, J.M. (1995), "A Simple Hybrid Model for Estimating Real Estate Price Indexes,"
- Ranaivoarimanana Njaka. Urbanisme de coalition : articulation entre infrastructure routière et plus-value foncière dans la fabrique urbaine - Le cas de la ville de Tananarive, Thèse de doctorat Paris Est, 2017.