

5 octobre 2022

RAPPORTS

Julien **LICHERON**

avec la collaboration de

Patrick **BOUSCH**

Antoine **PACCOUD**

Konrad **SKOCZYLAS**

Karolina **ZIEBA-KULAWIK**

PROJET DE RÉFORME DE L'IMPÔT FONCIER

Évaluation réalisée par le LISER

Projet de Réforme de l'Impôt Foncier

Évaluation réalisée par le LISER¹

¹ Les auteurs du rapport remercient le Professeur Alain Trannoy (EHES & AMSE Université d'Aix Marseille) pour ses conseils et ses remarques constructives au cours de la réalisation de ces analyses. Ils présentent également leurs remerciements aux représentants des différents Ministères et Administrations qui ont fourni des données utiles à la réalisation de ces travaux.

SOMMAIRE

Synthèse des résultats	5
1. Introduction	6
2. La formule de la valeur unitaire selon un modèle spatial.....	7
3. Les données utilisées pour évaluer la formule	10
4. Les critères utilisés pour évaluer l'adéquation de la formule aux données	13
4.1. L'indicateur de référence	13
4.2. Des indicateurs complémentaires	13
5. L'évaluation de la formule de calcul	14
5.1. Enseignement n°1 – Le pouvoir prédictif de la formule	14
5.2. Enseignement n°2 – Le rôle de l'indice de centralité	15
5.3. Enseignement n°3 – L'accessibilité au(x) pôle(s) d'emploi	15
6. La calibration et l'estimation des paramètres	17
7. Tests de robustesse autour de la formule de calcul	19
7.1. Régressions sur sous-échantillons géographiques	19
7.2. Régressions par quantiles	19
7.3. Tests de sensibilité à la définition de la distance.....	20
8. Quelle stratégie pour ajuster les paramètres de la formule de calcul ?.....	20
9. Conclusions et perspectives.....	22
Bibliographie	23

Synthèse des résultats

Le présent rapport visait à évaluer la qualité de la formule de la valeur unitaire foncière, initialement proposée par le ministère de l'Intérieur, en la comparant à des données réelles de transactions immobilières portant sur des parcelles (ou ensemble de parcelles) vendues sur la période 2010-2021. Les données des transactions immobilières, initialement fournies par l'Administration de l'Enregistrement et des Domaines, sont géolocalisées et analysées dans le cadre de l'Observatoire de l'Habitat du ministère du Logement.

Il s'avère que la formule développée par le ministère de l'Intérieur est plutôt performante pour expliquer les différences de prix du foncier entre les localités : elle explique ainsi près de **81% de la variabilité entre localités** sur la période récente, 2018-2021. Parmi les variables retenues dans la formule de calcul de la valeur unitaire, le temps d'accès à Luxembourg-Ville joue un rôle déterminant dans l'adéquation de la formule de calcul des prix des terrains à bâtir entre localités.

Une **estimation des paramètres** (sur base des données réelles de transaction) a pu être réalisée pour l'ensemble des paramètres importants de la formule. Elle a été complétée par des simulations/calibrations pour les paramètres additionnels.

La formule s'est avérée **très robuste** sur données réelles :

- Elle reste pertinente à un niveau géographique relativement fin : les différences des rapports de prix entre localités sont très correctement modélisées dans la grande majorité des cas. Certaines différences locales restent toutefois difficiles à modéliser à Luxembourg-Ville et Esch-sur-Alzette, entre quartiers. Mais le faible nombre de transactions à ce niveau des quartiers rend l'exercice d'évaluation particulièrement difficile ;
- La formule de calcul est très pertinente pour retracer les prix de vente des parcelles du milieu de la distribution, un peu moins efficace pour les transactions aux prix les plus faibles *et/ou* les plus élevés.

Pour conserver une bonne pertinence de la formule de calcul, une **mise à jour de l'ensemble des paramètres tous les 3 ans**, couplée à une réévaluation annuelle de la valeur unitaire en utilisant la variation annuelle moyenne des prix au niveau national, se serait avérée une stratégie pertinente sur la période 2010-2021 si la formule avait été utilisée pour une approximation des différences de prix dans le foncier luxembourgeois.

1. Introduction

Dans le cadre de la réforme de l'impôt foncier et l'introduction d'un impôt à la mobilisation du terrain inerte, le ministère de l'Intérieur a confié au *Luxembourg Institute of Socio-Economic Research* (LISER) la réalisation d'une évaluation de la qualité de la formule de la valeur unitaire initialement proposée par un groupe interministériel sous la direction du ministère de l'Intérieur. Le présent rapport vise à retranscrire les principaux résultats de cette évaluation, qui a également été alimentée par les travaux du groupe de travail scientifique mis en place par le ministère de l'Intérieur pour associer différents acteurs publics dans la collecte et l'analyse des données pertinentes pour la réalisation d'une telle étude.

L'objectif des travaux du LISER était d'évaluer la qualité de la formule de calcul de la valeur unitaire foncière, initialement développée par le ministère de l'Intérieur, en utilisant des données réelles de transactions immobilières portant sur des parcelles (ou ensemble de parcelles) vendues sur la période 2010-2021. Les données des transactions immobilières, initialement fournies par l'Administration de l'Enregistrement et des Domaines, sont géolocalisées et analysées dans le cadre de l'Observatoire de l'Habitat du ministère du Logement.

Le présent rapport est organisé en plusieurs étapes, correspondant au déroulement chronologique du projet. Dans un premier temps, il s'agit de revenir sur la formule de calcul de la valeur unitaire développée par le ministère de l'Intérieur, puis de présenter les données fournies par l'Observatoire de l'Habitat, qui seront utilisées pour tester la formule de calcul de la valeur unitaire « sur données réelles ».

Ensuite, nous présenterons les critères utilisés pour évaluer l'adéquation de la formule aux données, avant de présenter les principaux enseignements de l'évaluation sur base de ces critères.

Une fois la forme fonctionnelle de la formule de calcul validée, il a été possible d'estimer *et/ou* de calibrer l'ensemble des paramètres inclus dans cette formule.

Nous avons également apporté une attention particulière à la réalisation de tests de robustesse, afin de mieux comprendre les atouts et les limites de la formule de calcul de la valeur unitaire une fois confrontée aux données réelles de transaction.

Enfin, nous fournissons quelques recommandations quant au besoin d'un ajustement régulier des paramètres du modèle pour retracer aussi fidèlement que possible la valorisation des biens fonciers, avant d'esquisser les principales conclusions et perspectives de ce projet.

2. La formule de la valeur unitaire selon un modèle spatial

Le rapport du LISER s'est appuyé sur une formule qui avait été préalablement conçue par le ministère de l'Intérieur. Cette formule avait pour objectif de décrire le rapport de la valeur unitaire d'une parcelle à l'autre à l'aide d'un nombre assez restreint de variables.

Certaines de ces variables sont disponibles au niveau des localités : c'est le cas notamment du temps d'accès de la localité au principal pôle d'emploi (Luxembourg-Ville), ainsi que d'un indice de centralité de la localité retraçant la « quantité » d'aménités offertes sur le territoire de la localité comparativement aux autres localités.

Par ailleurs, la formule de calcul incluait également des variables spécifiques aux parcelles, issues des Plans d'Aménagement Général (PAG) des communes : il s'agissait du mode et du degré d'utilisation du sol (tel que retracé par le *Coefficient d'Utilisation du Sol* – CUS), du phasage du développement (avec un accent particulier mis sur les zones d'aménagement différé), et enfin de la protection du patrimoine.

Dans le détail, la formule de calcul du ministère de l'Intérieur était la suivante :

$$V_f = V_{réf} \cdot e^{-k_1 \cdot (t_{loc}^{k_2})} \cdot i_{loc} \cdot m_f \cdot d_f^{k_3} \cdot p_f \cdot c_f$$

avec :	V_f	la valeur unitaire d'une parcelle
	$e^{-k_1 \cdot (t_{loc}^{k_2})}$	l'impact du temps d'accès de la localité au principal pôle d'emploi
	i_{loc}	un indice de centralité de la localité
	$m_f \cdot d_f^{k_3}$	le mode et le degré d'utilisation du sol repris dans le PAG
	p_f	le phasage de développement, tel qu'il apparaît dans le PAG
	c_f	la protection du patrimoine, issue également du PAG

Par ailleurs, l'indicateur de centralité i_{loc} est défini comme suit :

$$i_{loc} = \begin{cases} 1 & \text{si } [Centralité]_{loc} = 0 \\ 1,1 & \text{si } [Centralité]_{loc} > 0,4 \\ [1; 1,1] & \text{entre les deux} \end{cases}$$

avec $[Centralité]_{loc} = \sum \beta_i \cdot [Domaine]_{i,loc} / \max_{loc} [Domaine]_{i,loc}$

et $[Domaine]_{i,loc} = \sum_j Services_{i,j,loc} / (Pop. loc)^\alpha$

Comme noté ci-dessus, le degré d'utilisation du sol s'appuie sur le Coefficient d'Utilisation du Sol (CUS), qui a été calculé par le ministère de l'Intérieur en utilisant les règles suivantes :

a) Pour les fonds sis en zone soumise à l'élaboration d'un plan d'aménagement particulier « nouveau quartier » établie sur base de la loi du 19 juillet 2004 concernant l'aménagement communal et le développement urbain et des dispositions et mesures d'exécution de celle-ci, telle que modifiée par la loi du 28 juillet 2011 et les lois ultérieures, la valeur du degré d'utilisation du sol correspond au coefficient d'utilisation du sol fixé par le plan d'aménagement général.

b) Pour les fonds sis dans une zone soumise à l'élaboration d'un plan d'aménagement particulier « quartier existant » en vertu des dispositions citées à l'alinéa qui précède, la valeur du degré d'utilisation du sol est déterminée comme suit :

1° Si au moins dix pour cent de la surface totale des zones d'habitation et zones mixtes de la localité sont sises en zone soumise à l'élaboration d'un plan d'aménagement particulier « nouveau quartier », elle correspond à la moyenne harmonique de l'ensemble des coefficients d'utilisation du sol fixés pour les fonds d'une localité sis en zone d'habitation ou en zone mixte et couverts d'une zone soumise à l'élaboration d'un plan d'aménagement particulier « nouveau quartier ».

2° Si moins de dix pour cent de la surface totale des zones d'habitation et zones mixtes de la localité sont sises en zone soumise à l'élaboration d'un plan d'aménagement particulier « nouveau quartier », elle correspond à la moyenne harmonique de l'ensemble des coefficients d'utilisation du sol fixés pour les fonds d'une commune sis en zone d'habitation ou en zone mixte et couverts d'une zone soumise à l'élaboration d'un plan d'aménagement particulier « nouveau quartier ».

c) Pour les fonds sis dans les zones établies sur base de la loi du 19 juillet 2004 concernant l'aménagement communal et le développement urbain et des dispositions et mesures d'exécution de celle-ci, avant sa modification par la loi du 28 juillet 2011, la valeur relative au degré d'utilisation du sol d'un fonds résulte du coefficient maximum d'utilisation du sol défini par le plan d'aménagement général, réduit de 25 pour cent.

Pour les parcelles sises dans des zones soumises à l'élaboration d'un plan d'aménagement particulier « nouveau quartier », l'utilisation du Coefficient d'Utilisation du Sol (CUS) est parfaitement cohérente avec les analyses réalisées précédemment dans le cadre de l'Observatoire de l'Habitat : l'indice des prix des terrains à bâtir en zone à vocation résidentielle s'appuie également sur le CUS pour mesurer la densité de construction autorisée (Observatoire de l'Habitat, 2021). En l'absence d'un tel indicateur pour les zones soumises à plan d'aménagement particulier « quartier existant », la solution retenue par le ministère de l'Intérieur apparaît comme une solution pragmatique : elle repose sur la densité de construction autorisée dans les autres zones situées en plan d'aménagement particulier « nouveau quartier » de la localité *et/ou* de la commune. L'utilisation de la moyenne harmonique nous apparaît également une alternative indiquée à la moyenne arithmétique dans le cas présent.

Le présent rapport s'est efforcé d'analyser l'adéquation de cette formule de calcul aux données réelles de transaction portant sur des parcelles (ou ensemble de parcelles) entre 2010 et 2021. Une fois la forme fonctionnelle de cette formule validée, l'objectif a été de fournir des valeurs plausibles aux différents paramètres (k_1 , k_2 et k_3 , notamment) sur base d'estimations, de calibrations ou de simulations.

Il a également été possible de réaliser des simulations pour tester la qualité de l'ajustement statistique avec différentes versions des indicateurs de centralité i_{loc} (utilisant notamment des pondérations β_i identiques ou différentes des différents domaines dans l'indicateur de centralité, et différentes valeurs du paramètre α intégrant la population dans cet indicateur).

3. Les données utilisées pour évaluer la formule

Pour analyser la qualité d'ajustement de la formule, nous nous sommes appuyés sur les données des prix du foncier publiées par l'Observatoire de l'Habitat depuis 2015.

Il s'agit de statistiques basées sur un fichier reprenant l'ensemble des ventes de parcelles figurant sur la Publicité Foncière (qui est un registre des actes notariés), transmis par l'Administration de l'Enregistrement, des Domaines et de la TVA. Ces données incluent un prix de transaction qui est bien le prix réel, tel qu'enregistré sur l'acte notarié.

L'Observatoire de l'Habitat a localisé ces parcelles dans un Système d'Informations Géographiques à partir du Plan Cadastral Numérisé (PCN) fourni par l'Administration du Cadastre et de la Topographie (ACT).

Les parcelles potentiellement constructibles sont ensuite identifiées à partir des Plans d'Aménagement Général (PAG) des communes mis à jour pour 2010, 2013 et 2016 (et 2019 pour les communes ayant un nouveau PAG).

Un contrôle est réalisé pour vérifier que ces parcelles sont libres de construction au moment de la vente. Les ventes de parcelles utilisées sont donc exclusivement des ventes de terrains libres, sans aucune construction. Il ne s'agit pas de ventes de maisons incluant une parcelle, par exemple.

La méthodologie est décrite plus en détail dans trois Notes de l'Observatoire de l'Habitat (Observatoire de l'Habitat, 2015, 2019, 2021), ainsi que dans un article de recherche co-écrit par des chercheurs du LISER (Glumac, Herrera-Gomez, Licheron, 2019).

Une attention particulière est portée aux terrains situés en zones à vocation résidentielle, que sont les zones d'habitation et les zones mixtes. Un **terrain à bâtir en zone à vocation résidentielle** est ici défini comme :

- une parcelle ou un ensemble de parcelles contigües...
- ... libre de toute construction (sans aucune construction existante sur ce terrain)...
- ... dont au-moins 1 are se situe à l'intérieur du périmètre d'agglomération...
- ... et dont au-moins 90% de la surface se situe en zone d'habitat (zone d'habitation ou zone mixte).

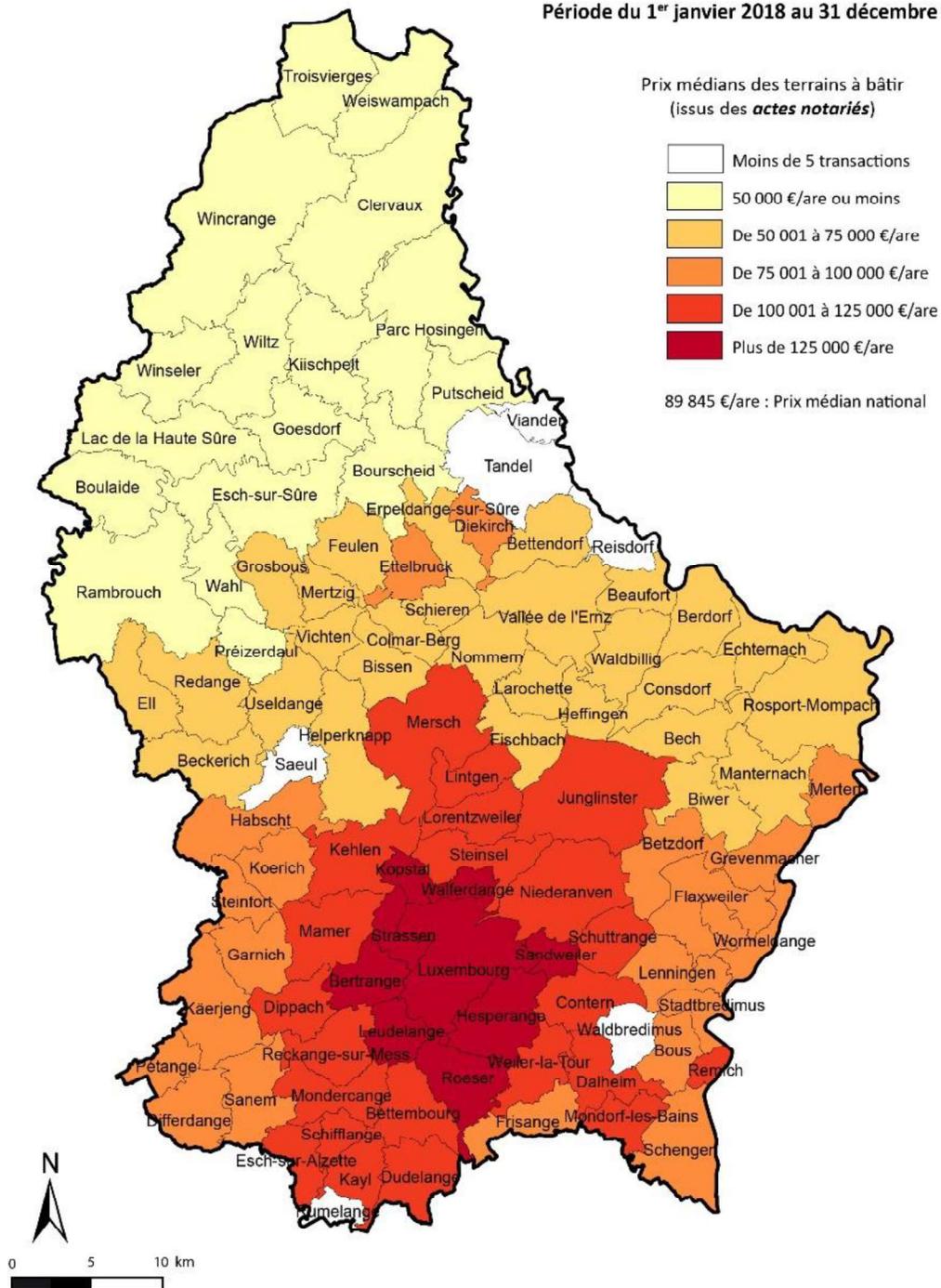
Les statistiques fournies par l'Observatoire de l'Habitat montrent une structuration très monocentrique du marché du foncier au Grand-Duché de Luxembourg, avec des prix médians par commune qui sont fortement influencés par le temps d'accès vers Luxembourg-Ville, principal pôle d'activités, d'emplois et de services, comme l'illustre la Carte 1. Sur cette carte, certaines communes ne sont pas associées à des catégories de prix compte-tenu du faible nombre d'observations sur la période 2018-2020 : c'est le cas des six communes de Reisdorf, Rumelange, Saeul, Tandel, Vianden et Waldbredimus. Cela traduit la faible activité sur le marché foncier de ces six communes sur la période récente. En revanche, les quelques

transactions associées à ces différentes communes ont bien entendu été utilisées pour la calibration de notre modèle dans la suite de ce rapport, et des simulations suggèrent que la formule de calcul est cohérente également pour retracer les prix observés sur ces quelques transactions dans ces 6 communes : la localisation de ces 6 communes dans un milieu rural où la formule semble particulièrement pertinente a même pour conséquence que l'erreur moyenne de prédiction sur ces six communes est inférieure à celle obtenue en moyenne sur les 96 autres communes.

En définitive, les données retenues portent sur un total de 204 localités qui intègrent au-moins une transaction portant sur un terrain à bâtir en zone à vocation résidentielle entre 2018 et 2020. Ces 204 localités représentent environ 96% de la population du Grand-duché de la Luxembourg en 2020, ce qui suggère que notre échantillon de parcelles ayant fait l'objet de transactions préserve une bonne représentativité spatiale. Les localités non représentées sont, là encore, exclusivement situées en zone plutôt rurale, où la formule semble particulièrement pertinente.

Carte 1 : Prix médians du foncier à vocation résidentielle par commune, de 2018 à 2020

Période du 1^{er} janvier 2018 au 31 décembre 2020



Auteur : J. Licheron, LISER, Octobre 2021
 Fonds de carte : ACT, LISER, 2018
 Données statistiques : Observatoire de l'Habitat, en collaboration avec l'Administration de l'Enregistrement, des Domaines et de la TVA
 Note : Un minimum de 5 ventes de terrains après sélection, rejet des extrêmes et traitements est requis pour afficher le prix médian communal.

Cf. Observatoire de l'Habitat (2021).

4. Les critères utilisés pour évaluer l'adéquation de la formule aux données

4.1. L'indicateur de référence

Pour évaluer la qualité de l'adéquation de la formule de calcul de la valeur unitaire développée par le ministère de l'Intérieur sur base de données réelles de transaction, il est indispensable de choisir au préalable un ou plusieurs critères d'évaluation de la qualité.

Dans le cadre de la présente étude, l'indicateur de référence utilisé est assez simple et standard en économétrie appliquée. Il s'agit du **coefficient de détermination (R^2) modifié** de la régression linéarisée de la formule de calcul sur les données entrantes de prix des terrains.

Le R^2 mesure la **part de la variabilité** dans les données qui est **expliquée** par la formule de calcul. Par exemple un R^2 égal à 0,80 suggérerait que 80% de la variabilité observée dans les données pourrait être expliquée directement par le modèle utilisé.

Le coefficient de détermination dépend du nombre d'observations utilisés pour estimer le modèle, et dépend également du nombre de variables explicatives utilisées. Il est souvent admis qu'un **coefficient de détermination (R^2) élevé** (par exemple proche de 80% sur des données temporelles ou agrégées, ou supérieur à 50% sur des micro-données), serait signe d'une bonne qualité d'ajustement du modèle sous-jacent.

4.2. Des indicateurs complémentaires

En complément du coefficient de détermination, nous avons retenu deux autres types d'indicateurs permettant de comparer différentes spécifications de la formule de calcul :

1. D'une part, l'erreur moyenne *et/ou* médiane de prédiction du prix de vente observé d'un terrain à partir de la formule de calcul théorique ;
2. D'autre part, une série d'indicateurs permettant de juger de la qualité de différents modèles avec un critère de parcimonie : le Critère d'information d'Akaike (*AIC*) et le Critère d'Information bayésien (*BIC*)

Ces indicateurs complémentaires sont difficiles à interpréter par eux-mêmes. Ils sont surtout utiles pour comparer différentes spécifications d'un même modèle (par exemple des spécifications utilisant des définitions alternatives du temps d'accès vers Luxembourg-Ville dans une formule de calcul strictement identique par ailleurs).

Ces indicateurs complémentaires seront particulièrement utiles dans le cadre des tests de robustesse réalisés dans la Section 7 du présent rapport.

5. L'évaluation de la formule de calcul

5.1. Enseignement n°1 – Le pouvoir prédictif de la formule

Si l'on utilise des données récentes de ventes de terrains à bâtir en zones à vocation résidentielle (2018-2021) pour tester la formule de calcul développée par le ministère de l'Intérieur, il s'avère que cette formule s'avère :

- très performante pour expliquer la variabilité des prix entre localités...
- ... mais toutefois moins performante pour expliquer la variabilité à l'intérieur des localités

Le résultat principal obtenu montre ainsi que nous obtenons un R^2 proche 0,81 pour la formule de calcul appliquée aux données agrégées de prix de vente médians par localité (avec un nombre d'observations $N=284$). La formule de calcul du ministère de l'Intérieur permet ainsi de retracer environ 81% de la variabilité des prix des terrains à bâtir entre localités observée entre 2018 et 2021.

Dans ce modèle, c'est clairement la variable d'accessibilité à Luxembourg-Ville qui joue un rôle majeur pour expliquer la variabilité des prix entre localités. Ceci est parfaitement cohérent avec les résultats mis en avant par l'Observatoire de l'Habitat au niveau communal, comme l'illustre par exemple la Carte 1 ci-dessus.

5.2. Enseignement n°2 – Le rôle de l'indice de centralité

L'indicateur de centralité proposé dans la formule de calcul du ministère de l'Intérieur joue également un rôle significatif dans l'explication de la variabilité inter-localités des prix : les localités les mieux dotées en aménités (centres commerciaux, réseaux de transport, soins de santé, infrastructures d'éducation...) présenteraient des prix de vente des parcelles plus élevés, toutes choses égales par ailleurs.

Nous avons également développé un modèle alternatif, en décomposant cet indicateur de centralité en quatre classes, afin de mieux exploiter la richesse de cette information :

- La centralité « commerciale » ;
- La centralité en matière de transports ;
- La centralité liée aux soins de santé ;
- La centralité liée aux infrastructures d'éducation.

Il s'avère qu'il est possible d'obtenir un R^2 ajusté plus élevé, proche de 0,84 en utilisant une telle décomposition de l'indicateur de centralité, en comparaison d'un indice unique de centralité. Les indicateurs complémentaires décrits à la section précédente (erreur moyenne *et/ou* médiane de prédiction, critères AIC et BIC) suggèrent également que le modèle enrichi des quatre classes d'aménités est plus performant que le modèle initial en termes d'adéquation aux données.

La formule de calcul pourrait donc être modifiée en incorporant une telle décomposition en quatre sous-indices. Néanmoins, nous n'avons pas été en mesure pour l'instant de réaliser des estimations rétrospectives en suivant ces indicateurs dans le temps. Dès lors, il est difficile de recommander à ce stade une telle décomposition sans pouvoir s'assurer de sa stabilité dans le temps. Cela pourrait faire l'objet d'une vérification régulière et ainsi d'un ajustement ultérieur de la formule de calcul, lorsque la collecte de telles informations sur les aménités aura été systématisée.

5.3. Enseignement n°3 – L'accessibilité au(x) pôle(s) d'emploi

L'accessibilité à Luxembourg-Ville reste de loin le principal déterminant dans la variabilité intercommunale des prix des terrains à bâtir : le R^2 serait proche de 0,73 en prenant seulement l'accessibilité à Luxembourg-Ville comme variable explicative des prix médians par localité.

L'accessibilité aux pôles d'emploi secondaires n'est jamais statistiquement significative (au seuil de 5%²) dans nos 2 modèles, qu'il s'agisse de

- L'accessibilité à Esch-sur-Alzette
- L'accessibilité à la *Nordstad* (Ettelbrück / Diekirch)

Il faut noter également que nous avons analysé l'impact d'un changement dans le mode de calcul de l'accessibilité à ces pôles d'emploi : en prenant le temps de parcours en minutes, ou la distance parcourue en kilomètres. Nous avons également comparé la qualité de l'ajustement statistique en prenant les temps d'accès en voiture ou en transports en commun (cf. infra, Section 7.3). Il s'avère que la qualité de l'ajustement statistique est nettement supérieure avec le temps d'accès en minutes, comparativement à la distance du trajet en kilomètres : le R^2 atteint 0,81 en utilisant le temps d'accès en minutes, contre $R^2 = 0,75$ avec la distance du trajet en kilomètres. De même, l'ensemble de nos indicateurs complémentaires (erreur moyenne *et/ou* médiane de prédiction, critères AIC et BIC) tranchent en faveur de l'utilisation du temps d'accès plutôt de la distance au(x) pôle(s) d'emploi. La « déformation » des distances par les infrastructures routières joue donc un rôle prépondérant dans la structuration spatiale des prix du foncier au Grand-duché de Luxembourg.

Toutefois, après mise en place de la formule de calcul de l'impôt foncier, il serait judicieux de procéder à une vérification régulière de sa pertinence (et l'ajuster si nécessaire). Ici, le rôle joué par l'accessibilité aux pôles secondaires d'emploi pourrait prendre de l'importance économique réelle dans les prochaines années, si les objectifs de l'aménagement du territoire sont réalisés. Nous reviendrons en détail sur la nécessité d'une évaluation régulière de la formule de calcul dans la Section 8 du présent rapport.

² Ce seuil est celui que nous avons retenu dans nos analyses : il correspond au risque de se tromper en rejetant l'hypothèse de nullité du coefficient associé à l'accessibilité à l'un des pôles secondaires. Ici, nous avons donc moins de 5% de « chances » de nous tromper en rejetant cette hypothèse d'absence d'effet statistique lié au temps d'accès en voiture vers Esch-sur-Alzette ou vers la Nordstad.

6. La calibration et l'estimation des paramètres

La Section précédente montrait que la formule de calcul initialement développée par le ministère de l'Intérieur permettait de retracer assez correctement la valorisation foncière telle que mesurée par les prix de transaction des parcelles sur la période 2018-2021.

Une fois la formule validée, il a été possible d'estimer une grande partie des paramètres de cette formule en utilisant les données réelles de transaction. Les résultats sont présentés dans le Tableau 1 ci-dessous.

Il s'est avéré que les estimations des paramètres liés à la variable d'accessibilité (paramètres k_1 et k_2) étaient possibles, avec des intervalles de confiance relativement étroits. En revanche, les estimations des paramètres m_f et k_3 (relatifs au mode et au degré d'utilisation du sol dans les PAG) sont particulièrement complexes, mais essentielles au bon fonctionnement de la formule. L'estimation et la calibration ont été possibles dans les zones résidentielles (*zones d'habitation / zones mixte*) et commerciales (*activités économiques*), mais plus délicates pour les zones publiques.

Par ailleurs, nos simulations ont montré que le jeu d'indicateurs de centralité le plus pertinent était obtenu en utilisant des pondérations β_i identiques des différents domaines dans l'indicateur de centralité. Le paramètre α retraçant l'importance de la population d'une localité dans la définition de sa centralité a été fixé à 0 après les simulations : des valeurs différentes n'apportaient aucune plus-value statistique pour ce qui concerne la qualité d'ajustement de notre modèle.

En définitive, les intervalles de confiance nous apparaissent réalistes, crédibles et utiles (parce que suffisamment étroits pour fournir un degré de précision correct) dans les cas suivants :

- les paramètres k_1 et k_2 liés à la variable d'accessibilité ;
- les paramètres m_{HAB-1} et m_{MIX-U} et associés à la localisation de la parcelle en zones d'habitation HAB-1 et en zone mixte MIX-U, comparativement à une localisation en zone d'habitation HAB-2 (qui est la catégorie utilisée en référence)

En revanche, les intervalles de confiance associés aux estimations se sont avérés plus larges et donc moins précis pour les variables suivantes :

- le paramètre k_3 associé au degré d'utilisation du sol dans le PAG ;
- les paramètres m_{ECO-C1} , m_{ECO-C2} , m_{HECO-N} , m_{ECO-R1} , ainsi que m_{MIX-C} , m_{MIX-R} et m_{MIX-V} , associés respectivement à la localisation de la parcelle en zones d'activités économiques ECO-C1, ECO-C2, ECO-N et ECO-R1, ou en zone mixte MIX-C, MIX-R et MIX-V, comparativement à une localisation en zone d'habitation HAB-2.

Compte-tenu de ces intervalles assez larges pour ces variables, il sera nécessaire de choisir des valeurs plausibles des paramètres sur base des intervalles de confiance produits mais aussi de la connaissance du fonctionnement du marché foncier.

Tableau 2 : Synthèse des paramètres estimés et calibrés

PARAMÈTRE	DESCRIPTION	ESTIMATION / CALIBRATION ?	VALEUR CENTRALE	INTERVALLE DE CONFIANCE A 95%	
				BORNE INFÉRIEURE	BORNE SUPÉRIEURE
$V_{réf}$	Numéraire	<i>Calibration</i>	250.000	-	-
k_1	Accessibilité - Paramètre principal	<i>Estimation</i>	0,0081	0,0076	0,0086
k_2	Accessibilité - Paramètre de lissage	<i>Estimation</i>	1,1260	1,0310	1,2210
α	Centralité - Effet de la Population	Calibration après simulations ³	0,0000	-	-
β_i	Centralité - Pondération des Domaines	Calibration après simulations	1,0000	-	-
$l_{loc\ min}$	Centralité - Valeur minimale de l'Indice (1 initialement)	Calibration	1,0000	-	-
$l_{loc\ max}$	Centralité - Valeur maximale de l'Indice (1,1 initialement)	<i>Simulations</i>	1,1712	1,1481	1,1943
k_3	Densité - Degré	<i>Estimation</i>	0,5813	0,3915	0,7710
m_{HAB-1}	Zoning - HAB-1	<i>Estimation</i>	0,7421	0,7168	0,7674
m_{HAB-2}	Zoning - HAB-2	Calibration	1,0000	-	-
m_{ECO-c1}	Zoning - ECO-C1	<i>Estimation</i>	0,4121	0,3618	0,4624
m_{ECO-C2}	Zoning - ECO-C2	<i>Estimation</i>	0,4102	0,3389	0,4815
m_{HECO-N}	Zoning - ECO-N	<i>Estimation</i>	0,3815	0,3219	0,4411
m_{ECO-R1}	Zoning - ECO-R1	<i>Estimation</i>	0,3928	0,3429	0,4427
m_{MIX-C}	Zoning - MIX-C	<i>Estimation</i>	0,8861	0,8351	0,9371
m_{MIX-R}	Zoning - MIX-R	<i>Estimation</i>	0,8129	0,7791	0,8467
m_{MIX-U}	Zoning - MIX-U	<i>Estimation</i>	0,8629	0,8312	0,8946
m_{MIX-V}	Zoning - MIX-V	<i>Estimation</i>	0,8017	0,7691	0,8343

³ Dans le cas des paramètres α et β_i , la calibration se base en réalité sur des simulations préalables réalisées sur la formule de calcul. Il s'agit donc pas en réalité d'une « calibration » sur base de considérations externes à notre analyse, mais plutôt d'une fixation des paramètres aux valeurs suggérées par les résultats de nos simulations.

7. Tests de robustesse autour de la formule de calcul

7.1. Régressions sur sous-échantillons géographiques

Un certain nombre de tests de sensibilité ont été réalisés pour évaluer la robustesse de la formule de calcul finalement validée, avec l'ensemble des paramètres estimés ou calibrés.

Dans un premier temps, il s'agissait de tester la robustesse spatiale de cette formule de calcul : est-elle surtout pertinente sur une partie du territoire, dans les zones urbaines ou rurales ?

Nous avons ici procédé à des régressions sur sous-échantillons géographiques, au niveau des 12 cantons, pour tester la sensibilité de la formule de calcul à des changements géographiques. Il est possible de conclure à une bonne robustesse de la formule de calcul au niveau géographique : les R^2 ajustés (pour tenir compte des différences dans les nombres d'observations par canton) générés par les modèles estimés au niveau cantonal se sont avérés très similaires.

La formule de calcul développée par le ministère de l'Intérieur, ici calibrée et estimée, est donc relativement robuste au niveau spatial, elle s'applique assez fidèlement à l'ensemble des cantons.

7.2. Régressions par quantiles

Par la suite, nous avons procédé à des régressions par quantiles, après ordonnancement des prix par are des parcelles. Il s'agissait avant tout d'analyser la capacité de la formule de calcul à retracer les prix des parcelles les plus élevés, les plus faibles et ceux du milieu de la distribution.

La formule de calcul plus efficace dans le milieu de la distribution, plus aléatoire sur les queues de distribution, comme c'est souvent le cas sur des données liées aux prix des logements et/ou du foncier : le modèle de valorisation fonctionne mieux sur les transactions « standard », moins efficacement sur les transactions à des prix assez faibles ou au contraire très élevés. Cela provient sans doute du bruit statistique plus important sur les queues de distribution, avec des prix par are qui peuvent être davantage influencés par la surface (très élevée ou au contraire très réduite) de la parcelle vendue, par les pouvoirs

de négociation de l'acheteur *et/ou* du vendeur, ou par des caractéristiques non- ou mal-observées de la transaction (avec des variables qui ne figurent pas dans la formule de calcul).

7.3. Tests de sensibilité à la définition de la distance

Enfin, nous avons analysé la sensibilité de la formule de calcul à des modifications dans la définition de la distance utilisée pour la variable d'accessibilité au principal pôle d'emplois : distance euclidienne plutôt que temps d'accès vers Luxembourg-Ville, temps de trajet en transports en commun ou en mode doux dans Luxembourg-Ville, etc.

Il s'avère que la définition utilisée dans le modèle initial, reposant sur un temps d'accès en voiture, est la plus adéquate : elle permet d'optimiser le R^2 , et les indicateurs complémentaires (erreur moyenne *et/ou* médiane de prédiction, critères AIC et BIC) suggèrent également que cette spécification est la plus performante.

8. Quelle stratégie pour ajuster les paramètres de la formule de calcul ?

Pour analyser la pertinence de la formule de calcul avec différentes stratégies d'actualisation de la formule de calcul, nous avons utilisé des données temporelles plus longues, portant sur la période 2010-2021.

Nous avons analysé ici trois options :

1. une stratégie de réévaluation de l'ensemble des paramètres de la formule **chaque année**
2. une stratégie de réévaluation de l'ensemble des paramètres de la formule **tous les 3 ans** + une réévaluation annuelle de la valeur unitaire en utilisant la variation moyenne des prix du foncier **au niveau national** entre chaque réévaluation ;
3. Une stratégie de réévaluation de l'ensemble des paramètres de la formule **tous les 5 ans** + une réévaluation annuelle de la valeur unitaire en utilisant variation moyenne des prix du foncier **au niveau régional** (c'est-à-dire en découpant le territoire en cinq grandes régions constituant des regroupements de cantons, telle que l'utilisent par exemple le STATEC et l'Observatoire de l'Habitat

dans leurs publications conjointes, cf. STATEC & Observatoire de l'Habitat, 2022)⁴ entre chaque réévaluation ;

Le R² ajusté moyen sur la période 2010-2021 s'élève à :

1. **R² ≈ 0,83** dans la stratégie 1
2. **R² ≈ 0,77** dans la stratégie 2
3. **R² ≈ 0,74** dans la stratégie 3

Pour conserver une bonne pertinence de la formule de calcul, une mise à jour de l'ensemble des paramètres tous les 3 ans, couplée à une réévaluation annuelle de la valeur unitaire sur base d'un indice des prix (national), se serait donc avérée une stratégie pertinente sur la période 2010-2021. Si elle reste moins qualitative qu'une réévaluation complète de l'ensemble des paramètres chaque année, elle s'avère plus efficace que des réévaluations plus espacées, même avec une prise en compte plus fine des différences régionales dans l'intervalle.

Il s'avère d'ailleurs que l'introduction d'indices régionaux pour ajuster les valeurs unitaires (en l'absence d'une réévaluation systématique des paramètres) n'est que marginalement plus efficace que l'utilisation d'un indice national, compte-tenu de la grande hétérogénéité des *évolutions* au sein des régions (même si les *niveaux* de prix au sein d'une région sont assez homogènes).

⁴ Les cantons d'Esch-sur-Alzette et de Luxembourg constituent ainsi deux régions par eux-mêmes. La zone Capellen-Mersch regroupent les cantons de Capellen et de Mersch. La zone Nord est constituée des cantons de Clervaux, Diekirch, Redange, Vianden et Wiltz. La zone Est regroupe, quant à elle, les cantons d'Echternach, Grevenmacher et Remich.

9. Conclusions et perspectives

La formule développée par le ministère de l'Intérieur est plutôt performante pour expliquer les différences de prix du foncier entre les localités (elle explique environ **81% de la variabilité entre localités** sur la période 2018-2021). Les données de base utilisées pour analyser cette performance reposent sur des prix de transactions repris dans les actes notariés, avec une connaissance très fine des caractéristiques des parcelles vendues. Il ne s'agit pas d'un échantillon des ventes, mais de l'ensemble des ventes réalisées sur la période 2018-2021 qui a été mobilisée pour tester la performance de la formule de calcul développée par le ministère de l'Intérieur.

L'accessibilité à Luxembourg-Ville joue un rôle déterminant dans l'adéquation de la formule de calcul des prix des terrains à bâtir entre localités.

Une **estimation des paramètres** (sur base de données de transaction réelles) a pu être réalisée pour l'ensemble des paramètres importants de la formule. Elle a été complétée par des simulations/calibrations pour les paramètres additionnels.

La formule s'est avérée très robuste sur données réelles :

- Elle reste pertinente à un niveau géographique relativement fin : les différences de prix entre localités sont très correctement modélisées dans la grande majorité des cas. Certaines différences locales restent toutefois difficiles à modéliser à Luxembourg-Ville et Esch-sur-Alzette, entre quartiers. Mais le faible nombre de transactions à ce niveau des quartiers rend l'exercice d'évaluation particulièrement difficile.
- La formule de calcul est très pertinente pour retracer les prix de vente des parcelles du milieu de la distribution, un peu moins efficace pour les transactions aux prix les plus faibles et/ou les plus élevés.

Pour conserver une bonne pertinence de la formule de calcul, une mise à jour de l'ensemble des paramètres tous les 3 ans, couplée à une réévaluation annuelle de la valeur unitaire sur base de l'évaluation du prix moyen au niveau national, se serait avérée une stratégie pertinente sur la période 2010-2021. Cette réévaluation périodique des paramètres de la formule de calcul devrait également inclure une investigation plus fine des effets des aménités dans l'indicateur de centralité. Une décomposition en quatre classes d'aménité pourrait apporter une plus-value au modèle, mais il est encore indispensable de tester la stabilité d'une telle décomposition au cours du temps.

Bibliographie

Glumac Brano, Herrera-Gomez Marcos & Licheron Julien (2019): « A hedonic urban land price index », *Land Use Policy*, vol. 81, pp. 802-812.

Observatoire de l'Habitat (2015) : « Les prix du foncier à bâtir en zone à vocation résidentielle entre 2010 et 2014 », *Note de l'Observatoire de l'Habitat n°21* (décembre 2015), ministère du Logement, 12 pages.

Observatoire de l'Habitat (2019) : « Les prix de vente des terrains à bâtir en zone à vocation résidentielle entre 2010 et 2017 », *Note de l'Observatoire de l'Habitat n°24* (février 2019), ministère du Logement, 10 pages.

Observatoire de l'Habitat (2021) : « Les prix de vente des terrains à bâtir en zone à vocation résidentielle entre 2018 et 2020 », *Note de l'Observatoire de l'Habitat n°28* (novembre 2021), ministère du Logement, 12 pages.

STATEC & Observatoire de l'Habitat (2022) : *Le Logement en Chiffres n°11*, mars 2022, ministère du Logement et STATEC, 22 pages.