

UNIVERSITÉ DE LIÈGE  
Faculté des Sciences  
Sciences Géographiques

# Étude comparative entre les préférences révélées et déclarées

Les maisons d'habitation en région liégeoise

*Mémoire présenté par*  
**Perrine DETHIER**  
*Pour l'obtention du titre de*  
**Master en Sciences Géographiques**  
*Orientation générale*

*Année académique*  
**2011 - 2012**

Je tiens à remercier mes promoteurs, Monsieur Halleux et Monsieur Cornet. Merci à Monsieur Halleux pour la proposition de ce sujet, son encadrement et sa disponibilité tout au long de cette année. Merci à Monsieur Cornet pour son suivi et son expérience.

Je voudrais remercier également les lecteurs de ce mémoire, Monsieur Billen et Monsieur Schmitz.

Mes remerciements vont aussi à Yannick Grignet pour son travail effectué l'année dernière sans quoi ce sujet n'aurait pas été réalisable. Merci à Julie Strée et Christelle Viaud-Mouclier pour leurs conseils pour la réalisation des enquêtes.

Je voudrais remercier les communes qui ont répondu à ma lettre et plus particulièrement les communes de Liège, Juprelle, Remicourt et Fexhe-le-Clocher qui m'ont fourni les données nécessaires à mon travail.

Je n'oublierais pas les habitants de Liège, Juprelle et Remicourt qui ont accepté de répondre à mon enquête. Merci pour leur temps et leur accueil.

Mes remerciements vont également à mes parents qui m'ont encouragé tout au long de ce mémoire et ont relu mon travail, à mes sœurs qui m'ont apporté leurs conseils et à mes amis.

## Table des matières

<b>INTRODUCTION GÉNÉRALE .....</b>	<b>7</b>
1. ETAT DE L'ART .....	7
1.1. <i>Les préférences révélées – la méthode hédonique .....</i>	<i>7</i>
1.2. <i>Les préférences déclarées - l'approche par contingence .....</i>	<i>8</i>
1.3. <i>Inefficacité du marché immobilier.....</i>	<i>9</i>
1.4. <i>Question de recherche.....</i>	<i>9</i>
2. HYPOTHÈSE DE TRAVAIL .....	9
<b>PARTIE 1: LES PRÉFÉRENCES RÉVÉLÉES PAR LA MÉTHODE HÉDONIQUE.....</b>	<b>10</b>
INTRODUCTION.....	11
1. L'APPROCHE HÉDONIQUE : MÉTHODE.....	12
1.1. <i>Formulation de la méthode hédonique .....</i>	<i>12</i>
1.2. <i>La méthodologie.....</i>	<i>13</i>
2. MÉMOIRE DE GRIGNET (2011) : DÉMARCHE ET RÉSULTATS .....	14
2.1. <i>Données de Grignet (2011) : segmentation et représentation .....</i>	<i>14</i>
2.2. <i>Le modèle de Grignet (2011).....</i>	<i>16</i>
2.3. <i>L'analyse des résidus .....</i>	<i>17</i>
3. CRÉATION DES VARIABLES EXPLICATIVES .....	20
3.1. <i>Données intrinsèques .....</i>	<i>20</i>
3.2. <i>Données extrinsèques .....</i>	<i>21</i>
3.3. <i>Résumé des variables explicatives construites .....</i>	<i>30</i>
4. ANALYSES PRÉLIMINAIRES.....	34
4.1. <i>Les logiciels Statistica et Gretl .....</i>	<i>34</i>
4.2. <i>Statistiques descriptives des variables .....</i>	<i>34</i>
4.3. <i>Analyse de corrélation.....</i>	<i>39</i>
5. DÉVELOPPEMENT DU MODÈLE .....	41
5.1. <i>Premier modèle .....</i>	<i>41</i>
5.2. <i>Deuxième modèle : choix parmi nos variables .....</i>	<i>43</i>
5.3. <i>Modèle 3 : correction de l'hétéroscédasticité.....</i>	<i>49</i>
5.4. <i>Modèle 4 : suppression des observations marginales .....</i>	<i>50</i>
5.5. <i>Modèle final .....</i>	<i>54</i>
6. ANALYSE DES RÉSULTATS .....	56
7. COMPARAISON AVEC LE MODÈLE DE GRIGNET (2011) .....	62
8. CONCLUSION ET PERSPECTIVES.....	64
<b>PARTIE 2: LES PRÉFÉRENCES DÉCLARÉES PAR L'APPROCHE PAR CONTINGENCE.....</b>	<b>65</b>
INTRODUCTION.....	66
1. PRÉSENTATION DE LA MÉTHODOLOGIE .....	67
1.1. <i>Echantillonnage.....</i>	<i>67</i>
1.2. <i>Le type d'enquête.....</i>	<i>68</i>

1.3.	<i>La méthode d'évaluation du CAP</i> .....	69
1.4.	<i>Le questionnaire</i> .....	70
2.	BIAIS DE LA MÉTHODE.....	75
2.1.	<i>Biais liés à la méthode</i> .....	75
2.2.	<i>Biais liés aux scenarii</i> .....	76
2.3.	<i>Biais liés à l'interviewé</i> .....	76
2.4.	<i>Biais liés à l'enquêteur</i> .....	77
3.	LE QUESTIONNAIRE .....	77
4.	DESCRIPTION DES TROIS COMMUNES.....	80
5.	ANALYSE DES ENQUÊTES.....	82
5.1.	<i>Population interrogée</i> .....	83
5.2.	<i>Les consentements à payer</i> .....	88
6.	DISCUSSION GÉNÉRALE DES RÉSULTATS .....	100
<b>PARTIE 3: COMPARAISON DES DEUX MÉTHODES .....</b>		<b>103</b>
INTRODUCTION.....		104
1.	DES ÉTUDES COMPARATIVES .....	104
2.	LES COMPARAISONS .....	105
2.1.	<i>La méthodologie</i> .....	105
2.2.	<i>Les résultats</i> .....	106
3.	DISCUSSIONS DES COMPARAISONS.....	109
3.1.	<i>Pour les variables de contrôle</i> .....	109
3.2.	<i>Pour les variables exploratoires</i> .....	109
4.	CONCLUSIONS .....	110
<b>CONCLUSIONS GÉNÉRALES .....</b>		<b>112</b>
<b>BIBLIOGRAPHIE.....</b>		<b>114</b>
<b>ANNEXES .....</b>		<b>118</b>
ANNEXE 1 : CODES DES PROGRAMMES PYTHON .....		118
ANNEXE 2 : MATRICES DE CORRÉLATION.....		121
ANNEXE 3 : LES MODÈLES HÉDONIQUES.....		127
ANNEXE 4 : COLINÉARITÉ .....		130
ANNEXE 5 : TEST DE BREUSCH-PAGAN .....		131
ANNEXE 6 : MODÈLES HÉDONIQUES POUR DIFFÉRENTES CATÉGORIES DE PRIX .....		132
ANNEXE 7 : ANALYSE DE L'UTILITÉ MARGINALE DÉCROISSANTE .....		135
ANNEXE 8 : LETTRE ÉCRITE AUX COMMUNES .....		137
ANNEXE 9 : LETTRE DISTRIBUÉE AUX HABITANTS DE NOUVELLES MAISONS .....		138
ANNEXE 10 : L'ENQUÊTE .....		139

## Table des illustrations

Carte 1 : Région urbaine de Liège, son complexe résidentiel urbain.....	15
Carte 2 : Localisation des observations.....	16
Carte 3 : Accessibilité à l'emploi.....	38
Carte 4 : Typologie des communes de l'enquête.....	80
Figure 1 : Evolution de la variable dépendante Y en fonction de la variable explicative X.....	45
Figure 2 : Variation des résidus en fonction de la variable explicative X.....	45
Figure 3 : Variation de la variable dépendante Y selon la variable explicative X pour plusieurs populations.....	46
Figure 4 : Echelle de valeurs croissantes.....	70
Figure 5 : Planche illustrative du questionnaire.....	71
Graphique 1 : Distribution des prix.....	35
Graphique 2 : Distribution des superficies de terrain.....	36
Graphique 3 : Distribution des logarithmes des superficies de terrains.....	36
Graphique 4 : Distribution des superficies habitables.....	37
Graphique 5 : Distribution des âges des biens immobiliers.....	37
Graphique 6 : Distribution de la variables sur l'accessibilité à l'emploi.....	39
Graphique 7 : Résidus du modèle 2 en fonction de la superficie terrain.....	48
Graphique 8 : Résidus du modèle 2 en fonction de la superficie habitable.....	48
Graphique 9 : Résidus du modèle 4 en fonction de la superficie habitable.....	50
Graphique 10 : Résidus du modèle 4 en fonction du logarithme de la superficie du terrain.....	51
Graphique 11 : Distribution des résidus du modèle 4.....	52
Graphique 12 : Diagramme quantile.....	52
Graphique 13 : Distribution des résidus par rapport aux prix.....	53
Graphique 14 : Variation du prix du bien en fonction de la superficie du terrain.....	57
Graphique 15 : Variation du prix du bien en fonction de l'âge du bien.....	59
Graphique 16 : Répartition des échantillons par tranche d'âge.....	84
Graphique 17 : Répartition des échantillons par tranche de revenus.....	85
Graphique 18 : Répartition des échantillons par diplôme.....	85
Graphique 19 : Répartition des échantillons selon la taille du ménage.....	86
Graphique 20 : Répartition des échantillons selon le nombre de voitures.....	87
Graphique 21 : CAP pour la situation 2.....	90
Graphique 22 : CAP pour la situation 3.....	91
Graphique 23 : CAP pour la situation 4.....	93
Graphique 24 : CAP pour la situation 5.....	94
Graphique 25 : CAP pour la situation 6.....	95
Graphique 26 : CAP pour la situation 7.....	96
Graphique 27 : CAP pour la situation 8.....	98
Graphique 28 : CAP pour la situation 9.....	99
Graphique 29 : Solde migratoire (%) par tranche d'âges pour les régions urbaines belges (1988-1996).....	102

Tableau 1 : Indicateurs globaux de la qualité du modèle de Grignet (2011).....	17
Tableau 2 : Points assignés aux différentes affectations.....	29
Tableau 3 : Statistiques descriptives de variables de l'échantillon étudié.....	34
Tableau 4 : Statistiques descriptives de l'accessibilité à l'emploi.....	39
Tableau 5 : Indicateurs globaux de la qualité d'ajustement du modèle 1 .....	43
Tableau 6 : Indicateurs globaux de la qualité d'ajustement du modèle 2 .....	43
Tableau 7 : Indicateurs globaux de la qualité d'ajustement du modèle 3 .....	49
Tableau 8 : Indicateurs globaux de la qualité d'ajustement du modèle 4 .....	50
Tableau 9 : Coefficients du modèle 4 .....	55
Tableau 10 : Statistiques basées sur les données initiales du modèle 4.....	54
Tableau 11 : Populations des communes.....	81
Tableau 12 : Revenus totaux des communes.....	81
Tableau 13 : Evolution des terrains résidentiels par commune entre 2001 et 2006 .....	81
Tableau 14 : Sexe du répondant .....	83
Tableau 15 : Statistiques descriptives de l'âge du répondant (en année).....	83
Tableau 16 : Catégorie de revenus .....	84
Tableau 17 : Statistiques descriptives de la surface habitable (m <sup>2</sup> ).....	87
Tableau 18 : Statistiques descriptives de la surface du terrain (m <sup>2</sup> ) .....	88
Tableau 19 : Statistiques descriptives sur le nombre de façades.....	88
Tableau 20 : Classification des variables selon le test statistique effectué.....	88
Tableau 21 : Statistiques descriptives sur le CAP pour la situation 2 .....	89
Tableau 22 : Statistiques descriptives sur le CAP pour la situation 3 .....	91
Tableau 23 : Statistiques descriptives sur le CAP pour la situation 4 .....	92
Tableau 24 : Statistiques descriptives sur le CAP pour la situation 5 .....	94
Tableau 25 : Statistiques descriptives sur le CAP pour la situation 6 .....	95
Tableau 26 : Statistiques descriptives sur le CAP pour la situation 7 .....	96
Tableau 27 : Statistiques descriptives sur le CAP pour la situation 8 .....	97
Tableau 28 : Statistiques descriptives sur le CAP pour la situation 9 .....	99
Tableau 29 : Etudes comparatives entre méthode hédonique et approche par contingence reprises dans Carson et al. (1996) .....	104
Tableau 30 : Résultats finaux des deux modèles pour la situation 2 .....	107
Tableau 31 : Résultats finaux des deux modèles pour la situation 3 .....	107
Tableau 32 : Résultats finaux des deux modèles pour la situation 4 .....	108
Tableau 33 : Résultats finaux des deux modèles pour la situation 5 .....	108
Tableau 34 : Résultats finaux des deux modèles pour la situation 6 .....	108
Tableau 35 : Coefficient pour les catégories de superficie habitable.....	135
Tableau 36 : Coefficient pour les catégories de superficie de terrain .....	136
Tableau 37 : Coefficient pour les catégories de distance-temps à Liège.....	136

# Introduction générale

De nombreuses études actuelles monétarisent les aménagements publics tels que de nouveaux espaces verts (Choumert & Travers, 2010; del Saz Salazar & Garcia Menendez, 2007), de nouvelles lignes de métros (Gibbons & Machin, 2005) ou encore l'instauration du quartier en zone 30 (Bureau & Glachant, 2010).

Pour ce faire, deux grandes méthodes sont utilisées. Dans un premier temps, nous réalisons un état de l'art de ces méthodes pour introduire notre question de recherche.

## 1. ETAT DE L'ART

Dans la littérature relative à la monétarisation de l'environnement (Hardelin et al. 2010), deux grands types de méthodes sont utilisées (Allin & Henneberry 2010) : la première se base sur les préférences révélées (Revealed preferences), la seconde sur les préférences déclarées (Stated preferences). Les méthodes se servant des préférences révélées utilisent des marchés existant pour déterminer la valeur de l'environnement. On retrouve entre autres comme méthode le coût de déplacement ou l'approche hédonique. Les méthodes faisant référence aux préférences déclarées estiment le prix en interrogeant les ménages sur des scénarii hypothétiques. Une des méthodes est alors l'approche par contingence.

La méthode hédonique et l'approche par contingence sont les deux grandes méthodes mises en application dans la littérature. Nous allons vous présenter plus en détail ces deux grandes méthodes. De plus ces méthodes portant sur le marché immobilier, nous allons présenter un principe important qui est l'inefficacité du marché immobilier.

### 1.1. Les préférences révélées – la méthode hédonique

Le principe fondamental de la méthode hédonique considère que « *les consommateurs tirent leur utilité non pas des biens eux-mêmes mais des caractéristiques de ces biens* » (Halleux 2010). Ce principe repose sur le travail de Lancaster (1966). Un bien, immobilier ou autre, n'est qu'un vecteur de caractéristiques et il est acheté pour ces caractéristiques. Ce principe implique que le prix d'échange d'un bien est la rémunération pour l'ensemble des caractéristiques qui le composent.

Rosen (1974) va introduire la théorie des marchés implicites. Un marché implicite est un marché où s'échangent des biens dont on peut observer le prix des caractéristiques. C'est le cas du marché des caractéristiques de l'immobilier. Seul le prix du bien immobilier est connu.

Le prix de chaque caractéristique n'est pas connu mais on sait qu'il est compris dans le prix du bien.

La méthode hédonique est largement utilisée. On retrouve de nombreuses études dans la littérature portant sur des sujets divers, réalisées dans de nombreux pays : France (Choumert & Travers, 2010), Angleterre (Gibbons & Machin, 2005), Pays-Bas (Theebe, 2004), Espagne (Morancho 2003), Etats-Unis (Zabel & Guignet, 2012), Japon (Gao & Asami, 2007), Chine (Kong, Yin, & Nakagoshi, 2007), Indonésie (Yusuf & Resosudarmo, 2009)...

Par contre, en Belgique, très peu d'études ont été développées. Cependant, à l'Université de Liège, un mémoire appliquant cette méthode a été réalisé en 2011 : Grignet (2011). Ce mémoire sera un support important pour la réalisation de ce travail car nous réutilisons certaines de ses données. Nous en discuterons d'avantage dans la Partie 1: Les préférences révélées par la méthode hédonique.

## **1.2. Les préférences déclarées - l'approche par contingence**

Les fondements théoriques de la méthode ont été conçus par Ciriacy-Wantrup (1947) mais la première application a été faite par Davis (1963). Depuis, elle est largement utilisée dans les analyses de coûts-bénéfices et ce dans de nombreux pays.

L'évaluation par contingence se base sur les préférences déclarées. Elle utilise les enquêtes pour déterminer la propension d'un individu à payer pour un bien souvent non-marchand. Il est demandé aux personnes interrogées leur consentement à payer (CAP), *Willingness to Pay* (WTP), pour une situation hypothétique après leur avoir expliqué le scénario.

Contrairement à la méthode hédonique, les données utilisées peuvent ne pas être réelles, on présente aux interviewés une situation hypothétique. Cette méthode permet de mieux cerner le choix de la valorisation grâce aux questions posées pendant l'enquête. Cependant, elle présente de nombreux biais liés à l'enquête (biais conceptuel, biais d'ancrage, biais hypothétique, biais d'inclusion, biais stratégique, biais de l'enquêteur, biais de perception) (Hardelin et al., 2010). Ces points seront abordés dans la Partie 2.

Cette méthode a été utilisée dans certains mémoires de l'Université de Liège. On notera notamment celui de Loiseau (2011) et de Strée (2010).



### **1.3. Inefficacité du marché immobilier**

Les biens immobiliers sont des biens hétérogènes. De par leur localisation mais aussi leurs caractéristiques (superficie, l'âge, équipement de la cuisine, isolation...), les biens immobiliers sont tous différents. Ce caractère hétérogène des biens immobiliers ne permet pas de déterminer un prix unique admis par l'ensemble des acteurs du marché.

Cette inefficacité a été mise en avant dans l'article d'Evans (1995). Il déclare qu'on n'a pu identifier aucun marché immobilier où les prix des évaluations ne s'écartent pas en moyenne d'au moins 10 % du prix des transactions. L'expertise immobilière n'est donc pas une science exacte. Que ce soit par des experts professionnels ou des évaluations statistiques, l'évaluation ne permet pas de déterminer la valeur exacte du bien immobilier.

Les deux méthodes que nous allons mettre en place dans ce mémoire ne permettront pas d'établir la valeur authentique du bien immobilier. Nous pourrions être satisfaits si les valeurs obtenues se situent dans une fourchette de 10 % du prix de ce bien.

### **1.4. Question de recherche**

La méthode hédonique semble plus efficace que l'évaluation par contingence dont on souligne souvent les nombreux biais. Or, l'évaluation par contingence est très utilisée. Nous voulons vérifier ces résultats en nous posant cette question : **est-ce que les valeurs obtenues par l'évaluation par contingence sont vérifiées par la méthode hédonique ?**

## **2. HYPOTHÈSE DE TRAVAIL**

Dans ce travail, nous partons de l'hypothèse que les valeurs obtenues par l'évaluation par contingence sont inférieures aux valeurs obtenues par la méthode hédonique. Cette hypothèse a été construite suite à la lecture d'articles (Carson et al., 1996 et Hardelin et al., 2010)

Pour valider cette hypothèse, notre travail se déroulera en trois temps. Premièrement, nous mettrons en application la méthode hédonique sur la région liégeoise. Quand les résultats seront obtenus, nous réaliserons une évaluation par contingence auprès de Liégeois. Enfin nous comparerons les résultats des deux méthodes.

Cette démarche structure notre travail en trois parties qui constituent les trois premières parties de notre mémoire. La dernière partie de notre mémoire est consacrée à la conclusion générale de notre travail.

# **Partie 1: Les préférences révélées par la méthode hédonique**

## **INTRODUCTION**

La première partie consistera en la mise en application du modèle hédonique pour la région urbaine de Liège.

Dans un premier temps, nous exposerons l'énoncé de la méthode hédonique et la démarche pour la mise en application de cette méthode. La méthode hédonique réalisée dans ce mémoire s'inspire largement du mémoire de Grignet (2011).

Dans le chapitre 2, nous résumerons son travail afin de pouvoir, par la suite, nous en servir comme point de comparaison. Nous expliquerons plus en détail les données produites car nous utiliserons une de ces mêmes données.

Dans le chapitre 3, nous expliquerons la construction de l'ensemble des caractéristiques des biens immobiliers utilisées dans la méthode. Cette étape est une partie plus technique. Nous expliquerons les outils utilisés pour la création des variables.

Dans le chapitre 4, des analyses descriptives des caractéristiques et une analyse des corrélations seront réalisées. Elles seront utiles pour appréhender les variables et envisager des transformations mathématiques.

Dans le chapitre 5, nous réaliserons notre modèle. Plusieurs tests seront accomplis, notamment sur les résidus, pour s'assurer de la qualité du modèle. Comme certains des tests ne seront pas satisfaisants, nous devrons transformer notre premier modèle pour arriver au modèle définitif.

Le chapitre 6 se concentrera sur l'analyse des résultats de notre modèle hédonique. Nous passerons en revue les résultats obtenus et nous discuterons de leur valeur.

Dans le chapitre 7, nous comparerons les résultats obtenus avec ceux de Grignet (2011). De plus nous étudierons l'apport de notre mémoire au sien.

Pour terminer, nous résumerons les principaux résultats du modèle et nous envisagerons une poursuite de l'étude par un contrôle et un nuancement de nos résultats par l'approche par contingence.

# 1. L'APPROCHE HÉDONIQUE : MÉTHODE

## 1.1. Formulation de la méthode hédonique

Un bien immobilier,  $X$ , est donc une combinaison de  $k$  caractéristiques. Cela peut s'écrire de la manière suivante :

$$X = (x_1, x_2, x_3, \dots, x_k)$$

Ces caractéristiques sont de deux types : intrinsèques ou extrinsèques. Les caractéristiques intrinsèques, appelées aussi variables structurelles, sont relatives au bien immobilier. Elles définissent le nombre de pièces, la superficie habitable, l'état de rénovation du bien... Les caractéristiques extrinsèques, elles, sont relatives à la localisation du bien. Il en existe deux grandes catégories. La première catégorie concerne les variables représentant l'accessibilité au logement. Ces variables diffèrent selon le mode de transport utilisé ou selon la destination. La deuxième catégorie concerne l'environnement du logement, c'est-à-dire la qualité du milieu voisin du bien. Cette catégorie regroupe des variables mesurant le niveau socio-économique du quartier, les nuisances sonores ou de pollution du site, la qualité paysagère ou encore la distance à un espace vert.

De la l'explication ci-dessus, on peut formuler la fonction hédonique,  $F(x)$ , qui permet d'associer le prix d'un bien,  $P$ , à chacune des caractéristiques qui le composent :

$$P = F(x) = F(x_1, x_2, x_3, \dots, x_k)$$

On peut ainsi déterminer le prix de chaque caractéristique grâce à une régression multiple.

Cette modélisation proposée par Rosen (1974), suppose que l'échange se réalise dans un marché concurrentiel où le vendeur mais aussi l'acheteur ne peuvent pas contrôler le prix. Il faut aussi que les acteurs aient une connaissance parfaite des caractéristiques du contenu.

Une fonction hédonique n'est valable que pour un marché donné. Un modèle qui voudrait estimer des prix d'échange de même caractéristique mais sur des marchés différents sera mal ajusté car, pour une même caractéristique, les prix pratiqués sur deux marchés sont différents. Il est donc important d'étudier un seul marché, mais cette considération est difficilement applicable vu la continuité de ceux-ci.

## 1.2. La méthodologie

La méthodologie suivie dans ce mémoire suit la procédure proposée par Des Rosiers (2001). Celle-ci a été choisie car elle offre l'avantage d'être clairement décrite étape par étape. Elle ne présente pas une originalité particulière et est fort semblable à d'autres (Terra, 2005). Elle se structure selon ces points :

1. Définition des objectifs de la modélisation et approche analytique
2. Choix et description du secteur d'analyse et nature de l'échantillon
  - Le type de segmentation
  - La représentation cartographique
3. Collecte de l'information et définition des variables
  - Le traitement de la banque de données
  - La sélection d'un sous-échantillon pour valider ultérieurement le modèle
  - La définition opérationnelle des variables retenues
4. Description et analyse de la banque de données
  - L'application des statistiques descriptives
  - La transformation mathématique de variables
  - L'application des statistiques descriptives après correction et épuration
5. Analyse de corrélation
  - Le test de corrélation simple
  - Le test de fiabilité
6. Analyse de régression
  - La procédure de régression standard
  - La réduction du nombre de variables
  - La transformation mathématique des variables indépendantes
  - La transformation mathématique de la variable dépendante
  - La procédure de régression « pas à pas » (« stepwise »)
7. Analyse des résidus
  - L'identification des résidus marginaux
  - La représentation graphique des résidus
8. Mise au point du modèle final
9. Validation du modèle final à l'aide du sous-échantillon retenu à l'étape 3
10. Production d'indications de la valeur

Vu que nous reprenons les données récoltées par Grignet (2011), certaines étapes ont déjà été expliquées dans son mémoire. Nous rappellerons brièvement ces étapes dans le chapitre suivant.

## **2. MÉMOIRE DE GRIGNET (2011) : DÉMARCHE ET RÉSULTATS**

Notre mémoire fait suite au mémoire de Grignet (2011) qui avait collecté ses données sur Immoweb ([www.immoweb.be](http://www.immoweb.be)) pour ensuite ajuster un modèle hédonique. Il nous semble donc important d'en rappeler les résultats principaux car, d'une part, notre mémoire réutilise de nombreuses données produites par Grignet (2011) et, d'autre part, nous suivons la même démarche. De plus, nous souhaiterions améliorer certains aspects de son travail. Nous voudrions approfondir les résultats obtenus sur les caractéristiques d'accessibilité et aussi prendre en compte les problèmes d'hétéroscédasticité.

### **2.1. Données de Grignet (2011) : segmentation et représentation**

Les données produites par Grignet (2011) ont été collectées sur Immoweb. Nous sommes contraints par ses choix (source, segmentation sectorielle, segmentation typologique) que nous allons vous présenter.

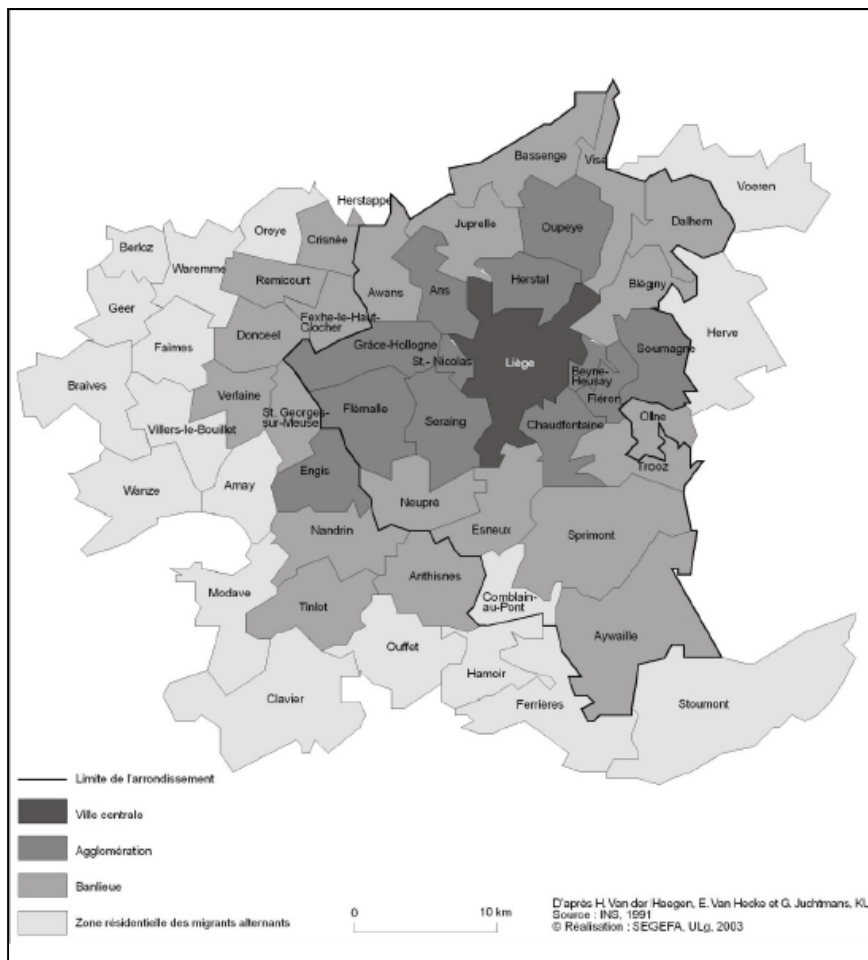
La source est Immoweb, un site internet d'annonces immobilières belge ([www.immoweb.be](http://www.immoweb.be)). Le site se proclame comme le premier en la matière. Cette source est sans conteste un leader dans les ventes immobilières en Belgique. Tant les privés que les professionnels l'utilisent. Le site est visité en moyenne par 140.000 personnes par jour et offre plus de 120.000 annonces. Le choix de cette source est dû à plusieurs raisons. Premièrement, en Belgique, il semble impossible d'obtenir les prix des transactions, que ce soit via les notaires ou le cadastre (Halleux, 2005). L'approche hédonique belge ne peut que se reporter aux annonces immobilières. Dans ce cadre, Immoweb ressort comme la source majeure. La collecte de ces données demande un temps important (plusieurs mois). Pour cette raison, nous avons préféré réutiliser des données déjà récoltées pour se concentrer sur d'autres aspects. Nous partons du même postulat que Grignet (2011) qui est que le prix donné sur Immoweb est proche du prix de transaction. D'autant plus que dans notre cas nous pouvons justifier ce postulat par les résultats de son mémoire et aussi par cette étude : « Etudes et documents logement 2, marché fonciers et immobiliers acquisitifs et locatifs : 1994-1995 » (Vandermotten C. et al. 1996, p. 39).

## Partie I : La méthode hédonique

Pour effectuer des recherches limitées à un marché unique, il faut réaliser une segmentation spatiale (critères spatiaux) et une segmentation typologique (critères sur les caractéristiques structurelles) des données disponibles.

Grignet (2011) a fait le choix de travailler avec une zone d'étude couvrant la région urbaine liégeoise (carte 1). Il a choisi de travailler sur un ensemble de communes pour augmenter l'information disponible. Le nombre d'annonces disponibles pour une commune ne permet pas de faire une approche économétrique significative.

**Carte 1 : Région urbaine de Liège, son complexe résidentiel urbain**

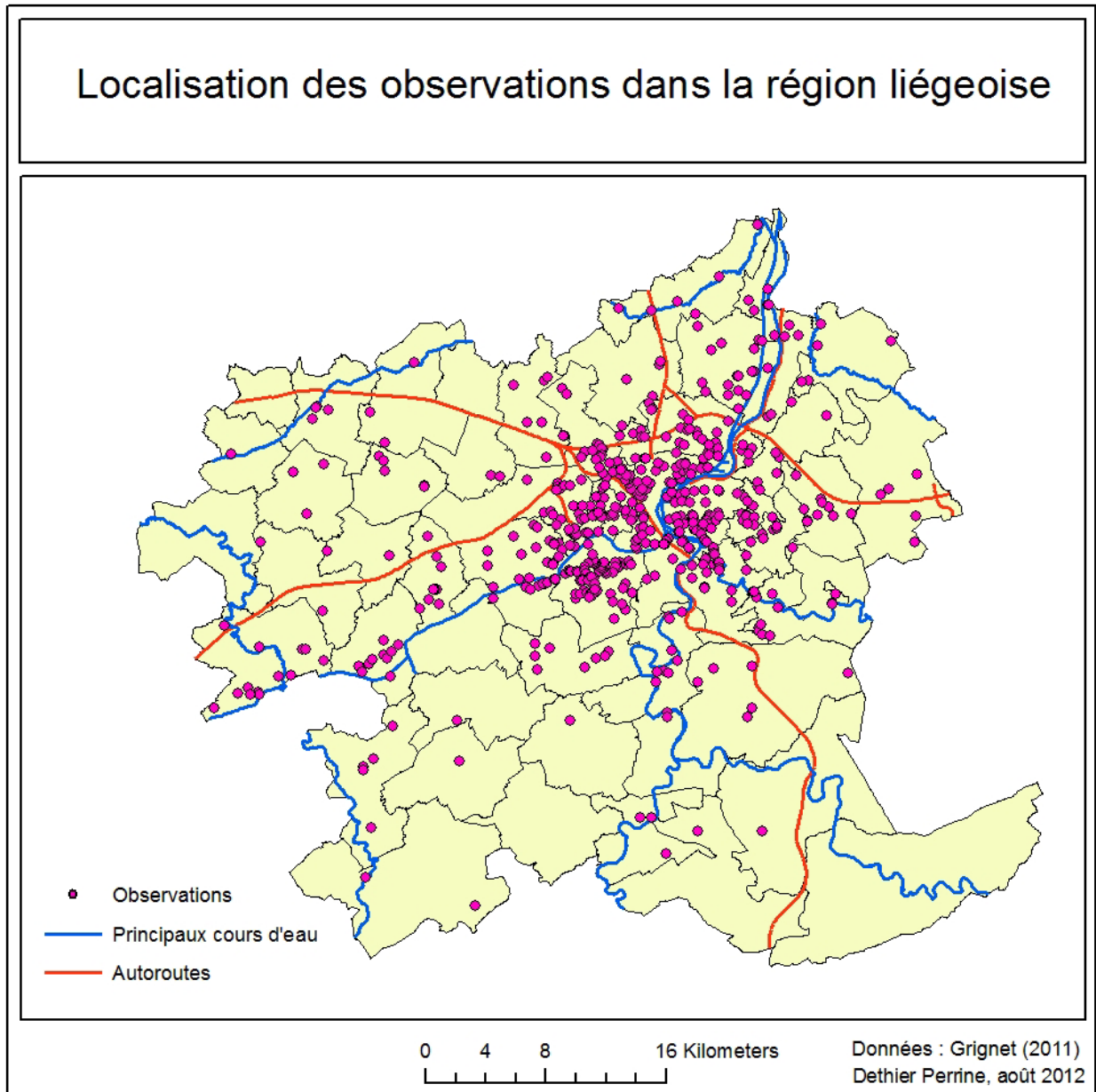


Source : Mérenne-Schoumaker, 2003, p.12.

Pour les biens immobiliers, une segmentation typologique a été effectuée pour limiter l'hétérogénéité de ceux-ci. Grignet (2011) n'a collecté l'information que pour des maisons unifamiliales. Ces données ont été récoltées entre novembre 2010 et juin 2011. Seulement 15 % des annonces étaient exploitables car peu d'annonces possèdent une adresse et suffisamment de caractéristiques structurelles les concernant. En plus d'Immoweb, les données ont été croisées avec les données cadastrales. Ce croisement a permis une

vérification des annonces. L'ensemble des biens collectés, 548 observations, est représenté sur la carte 2 Carte 2.

Carte 2 : Localisation des observations



## 2.2. Le modèle de Grignet (2011)

Nous n'allons pas présenter l'ensemble des résultats obtenus dans le mémoire de Grignet (2011) mais il nous semble important d'en montrer une partie pour pouvoir effectuer des comparaisons avec notre modèle.



Dans son modèle final, il a pris en compte 30 variables explicatives et a utilisé 548 observations. Son modèle est caractérisé par les indicateurs de synthèse repris dans le tableau 1.

**Tableau 1 : Indicateurs globaux de la qualité du modèle de Grignet (2011)**

<b>Multiple R</b>	0,89
<b>Multiple R<sup>2</sup></b>	0,80
<b>Adjusted R<sup>2</sup></b>	0,79
<b>F(30,517)</b>	69,35
<b>P</b>	0,00
<b>Std. Err.of</b>	37479,81

Source : Grignet, 2011, p.43

La corrélation multiple dans l'échantillon R, vaut 89 %. Ce coefficient équivaut à la racine carrée du coefficient de détermination ( $R^2$ ) et illustre la force de la relation linéaire entre les variables indépendantes et la variable dépendante, le prix. Le coefficient de détermination,  $R^2$ , renseigne que 80 % de la variance du prix est expliquée par les variables indépendantes. Le  $R^2$  ajusté prend en compte les degrés de liberté, la différence entre le nombre d'observations et le nombre de variables explicatives. Cet indicateur permet de voir si l'explication marginale d'une nouvelle variable est suffisamment importante par rapport à la perte d'un degré de liberté. Dans son modèle, Grignet (2011) est arrivé à le maximiser à 79 %.

F(30,517) est la valeur statistique du F de Snedecor avec une probabilité de dépassement de 0,00 (p). On peut rejeter l'hypothèse nulle qui est  $\rho^2=0$  et conclure que les variables explicatives ont un effet significatif sur le prix.

Enfin l'erreur type est de 37.480 € ce qui équivaut à 19,8 % du prix moyen. Pour Grignet (2011) ce chiffre « *confirme par ailleurs l'inefficacité du marché immobilier, il est en effet impossible de trouver un prix qui soit admis par l'ensemble des offreurs et demandeurs* ». D'ailleurs, Evans (1995) déclare qu'« *il n'y a pas un prix du marché arrêté, seulement une étendue de valeurs possibles* ».

### **2.3. L'analyse des résidus**

L'analyse des résidus est une étape importante dans l'approche hédonique. Elle permet de vérifier le bon ajustement de la régression. Dans cette étape, il faut vérifier s'il n'y a pas de hétéroscédasticité ni d'autocorrélation spatiale des résidus.

### 2.3.1. Hétéroscédasticité

L'hétéroscédasticité se présente quand la variance des résidus n'est pas constante mais qu'elle est fonction des variables explicatives. Cette situation s'oppose à la situation d'homoscédasticité où la variance des résidus est constante.

Dans le mémoire de Grignet (2011), les résidus présentent une hétéroscédasticité en fonction du prix, la variable dépendante. Il en conclut que son modèle est mal ajusté pour les biens supérieurs à 260.000 € et que son modèle surestime la valeur de ces biens.

En ce qui concerne les variables explicatives (superficie du terrain, superficie habitable, ancienneté des habitations), il trouve une situation d'homoscédasticité.

### 2.3.2. Autocorrélation spatiale des résidus

L'analyse de l'autocorrélation spatiale des résidus n'a pas été réalisée dans le mémoire de Grignet (2011). Pourtant, il est important de vérifier que les résidus ne sont pas autocorrélés spatialement. D'une part, il faut être certain que le modèle prenne en compte les caractéristiques d'environnement les plus importantes et d'autre part, il faut vérifier qu'il n'y pas de dépendance spatiale, c'est-à-dire que les prix ne sont pas influencés par les prix des maisons voisines. Les biens immobiliers sont souvent évalués par comparaisons avec les biens voisins. Dans ce cas, nous serions en présence de sous-marchés.

Nous avons décidé de vérifier l'hypothèse d'autocorrélation spatiale pour les résidus obtenus par Grignet (2011) afin de certifier son modèle.

La table des résidus standardisés obtenus pour l'ensemble des observations ont été joints à la table des entités spatiales sur base d'une clé de liaison dans ArcGIS. Les résidus sont ainsi spatialisés. Le logiciel permet de calculer l'indice I de Moran.

L'indice I de Moran calcule la variation entre les résidus pour les voisins les plus proches, une matrice de voisinage. Il se calcule comme suivant :

$$I = \frac{n}{W} \frac{\sum_{i=1}^n \sum_{j=1}^n w_{i,j} (x_i - \bar{x})(x_j - \bar{x})}{\sum_{i=1}^n (x_i - \bar{x})^2}$$

Avec  $w_{i,j}$  le poids spatial entre l'individu i et j

$x$  la variable qui mesure le phénomène auquel on s'intéresse

$\bar{x}$  la moyenne de la variable

W la somme des poids de la matrice de voisinage :  $W = \sum_{i=1}^n \sum_{j=1}^n w_{i,j}$

n le nombre d'observations

Plus il est proche de 1, plus les résidus sont corrélés.

De plus, ArcGIS réalise un test dont l'hypothèse nulle est qu'il n'y a pas d'autocorrélation spatiale. Pour ce faire, le logiciel construit le test sur l'indice de Moran. Il compare la valeur observée de l'indice de Moran et la valeur attendue de l'indice Moran en l'absence d'autocorrélation spatiale :

$$z_I = \frac{I - E[I]}{\sigma}$$

Avec

$$E[I] = -\frac{1}{n-1}$$

$$\sigma^2 = \frac{n(n-1)(nS_1 - S_2) - 2W^2}{(n+1)(n-1)^2W^2}$$

$$S_1 = \frac{1}{2} \sum_i^n \sum_j^n (w_{i,j} + w_{j,i})^2$$

$$S_2 = \sum_i^n \left( \sum_j^n w_{i,j} + \sum_j^n w_{j,i} \right)^2$$

Si la p-value, la probabilité critique, est suffisamment petite c'est-à-dire que  $z_I$  est suffisamment élevé (par rapport au seuil de décision), on rejette l'hypothèse nulle et il y a donc corrélation spatiale. Dans ce cas, la valeur observée de l'indice de Moran est significativement différente de la valeur attendue de l'indice Moran en l'absence d'autocorrélation spatiale.

Nous avons pris comme option que la distance utilisée devrait être la distance euclidienne entre les points et que le poids d'un point par rapport à un autre,  $w_{i,j}$ , décroissait avec la distance.

Dans ce cas-ci, nous avons obtenu les résultats suivants :

```
Global Moran's I Summary
Moran's Index:    0,002356
Expected Index:  -0,001828
Variance:        0,000260
Z Score:         0,259510
p-value:         0,795242
```

Le p-value est largement supérieur au 0,05, généralement pris comme seuil de décision, et donc, nous ne pouvons pas rejeter l'hypothèse nulle. Il n'y a donc pas d'autocorrélation

spatiale. L'indice I de Moran est d'ailleurs très proche de 0. Le modèle de Grignet (2011) intègre suffisamment de variables pour rendre compte de la variation spatiale du prix.

### **3. CRÉATION DES VARIABLES EXPLICATIVES**

La collecte des données s'est faite en deux parties : nous avons repris les caractéristiques intrinsèques des biens immobiliers et avons construit ou réutilisé les caractéristiques extrinsèques. Nous allons passer en revue l'ensemble des variables qui vont être utilisées dans l'approche hédonique pour expliquer le prix d'un bien immobilier.

#### **3.1. Données intrinsèques**

L'ensemble des données intrinsèques provient du mémoire de Grignet (2011), à la fois du site Immoweb mais aussi du cadastre. Cette liste ne peut pas être améliorée car elle est extrêmement dépendante de ces deux sources. Pourtant, il nous semble que d'autres caractéristiques structurelles des biens seraient importantes dans le prix d'une maison notamment au niveau du bilan énergétique de celle-ci (isolation, châssis, panneau solaire...). Cependant, comme nous ne pouvons pas obtenir de données suffisamment pertinentes à ce sujet, il est préférable de ne pas en tenir compte. Voici donc l'ensemble des variables retenues :

- Prix de mise en vente
- Etat de rénovation du bien
- Superficie du terrain
- Superficie habitable
- Nombre de façades
- Nombre d'étages
- Nombre de salles de bain
- Nombre de salles de douche
- Nombre de chambres
- Equipement de la cuisine
- Présence d'une cave
- Présence d'un grenier
- Présence d'une terrasse
- Présence d'un jardin
- Présence d'une véranda
- Présence d'une piscine
- Présence d'une buanderie

- Nombres d'années depuis sa construction

Pour certaines caractéristiques, nous avons décidé de construire plusieurs variables binaires s'y rapportant. Ainsi pour le nombre de chambres, nous avons 4 variables (1 chambre, 2 chambres, 4 chambres et 5 chambres ou plus). Cette technique a été appliquée car la plus-value d'une chambre supplémentaire n'est pas identique si l'on passe de 1 à 2 chambres ou de 4 à 5 chambres.

## **3.2. Données extrinsèques**

Trois types de variables ont été construites : une portant sur l'accessibilité en général, une sur l'environnement socio-économique et la dernière sur l'environnement physique. Les variables socio-économiques ont été séparées du reste des autres variables environnementales car leur construction est différente et plus immédiate.

### **3.2.1. Accessibilité**

Cet aspect devait être amélioré par rapport au mémoire de Grignet (2011). Ses variables d'accessibilité portaient plus sur une nuisance liée aux caractéristiques socio-économiques des centres que sur l'accessibilité à des pôles attractifs. Nous n'avons donc pas repris les variables sur le temps de parcours pour rejoindre les autres centres.

Nous avons d'une part calculé une variable sur l'accessibilité en voiture aux centres d'emplois et à Liège et d'autre part des variables sur l'accessibilité par d'autres modes (bus, train et modes lents). Nous avons multiplié les variables sur l'accessibilité en voiture pour améliorer au mieux le modèle produit par Grignet (2011).

#### **3.2.1.1. Accessibilité en voiture aux centres d'emplois**

Nous avons décidé de travailler avec un indicateur basé sur l'accessibilité potentielle. L'accessibilité potentielle est une mesure pondérée de l'accessibilité, « *capacité d'un lieu à être atteint* » ((Rodrigue, 1998).

Elle se mesure comme ceci :

$$A(P)_i = \sum_{j=1}^n \frac{P_j}{d_{ij}}$$

Avec  $A(P)_i$  : Accessibilité potentiel d'un endroit i  
 $P_j$  : les attributs pondérant de l'endroit j  
 $d_{ij}$  : la distance entre l'endroit i et j  
 $n$  : le nombre de destinations

Dans ce cas-ci, nous avons pondéré l'endroit j par l'emploi à la destination. Nous avons situé les pôles d'emplois et donc le nombre d'emplois au centre d'emplois. L'emploi est une des raisons majeures des déplacements et la proximité des pôles d'emplois doit être attractive pour le choix d'une maison.

Nous avons donc eu besoin de données sur les distances entre les biens immobiliers (origine) et les centres d'emplois (destination) ainsi que du nombre d'emplois en ces centres. Pour les distances, nous avons décidé de travailler en distance-temps voiture. Les distances-temps sont fortement corrélées aux distances kilométriques mais rendent compte d'un élément contraignant des déplacements qui est le temps passé dans le mode. Nous pouvions utiliser une API de Google Map ou faire un programme sur Viamichelin pour connaître ces distance-temps. Mais ces outils sont limités par un nombre de requêtes par jour (2500 pour Google map), or nous devons faire plus de 1.457.680 requêtes (548 observations et 2660 centres de secteurs statistiques) ce qui aurait pris 583 jours. Pour réduire les temps de traitement, nous avons décidé de calculer les distances en mode raster dans le logiciel ArcGIS.

Dans un premier temps, nous avons séparé le shapefile des biens immobiliers en 548 nouvelles couches avec une observation unique dans chacun d'entre eux. Pour ce faire, l'extension *Split Layer By Attributes*<sup>1</sup> a été utilisée. Cet outil permet de séparer une couche sur base d'un attribut, dans ce cas-ci nous avons utilisé l'identifiant des biens immobiliers, le FID.

Nous avons ensuite développé un petit programme en python<sup>2</sup> qui permet, pour les 548 couches, de répéter la même fonction d'ArcGIS, *cost distance*. Cette fonction permet de calculer à partir d'une image de coût de déplacement de chaque pixel, le temps pour se rendre à chaque endroit du territoire depuis le point d'origine. L'image de coût de

---

<sup>1</sup> <http://arcsripts.esri.com/details.asp?dbid=14127>

<sup>2</sup> Les principaux programmes python se trouvent à l'annexe 1

déplacement est une image de friction où en chaque pixel est défini un poids fonction de la vitesse autorisée sur le réseau. Nous avons obtenus 548 images « distance-temps », représentant en chaque pixel le temps de déplacement depuis l'habitation. Cette méthode présente plusieurs désavantages par rapport à l'utilisation de Google map : la réglementation sur la route n'est pas utilisée (on peut remonter un sens interdit) et l'ensemble du territoire peut être parcouru même si les endroits non carrossables présentent un coût de parcours extrêmement élevé. Malgré ces défauts, nous obtenons une bonne approche de la réalité.

Pour les données sur l'emploi, nous avons utilisé des données du Lepur<sup>3</sup> au niveau le plus fin disponible, les secteurs statistiques. L'emploi par secteur statistique n'est pas fourni par les enquêtes socio-économiques. Le Lepur a croisé les données l'emploi de l'INS avec l'ONSS pour obtenir une estimation de l'emploi pour chaque secteur statique. Nous n'avons gardé que les secteurs où l'emploi était relativement important. Notre zone d'étude couvre 2660 secteurs statistiques, nous avons limité ce nombre à 234 secteurs. Ceux-ci sont les secteurs dans lesquels il y a plus de 800 emplois/km<sup>2</sup>. Ce seuil correspond à la moyenne de l'emplois/km<sup>2</sup> dans la région liégeoise augmentée d'un écart type. Nous avons pris aussi en compte les régions verviétoise et hutoise car elles se situent à proximité immédiate de la zone d'étude.

Après la sélection des secteurs, nous avons transformé ces polygones en points pour permettre de définir les destinations. Ces points ont été pris au centre de gravité des polygones des secteurs statistiques. Pour chaque bien immobilier, nous avons associé à chaque centre d'emploi généralisé par un point la distance-temps nécessaire pour les atteindre depuis la maison. La fonction *extract value to point* a été utilisée pour associer aux destinations la valeur du pixel de l'image « distance-temps » sur lequel ils se situaient. Cette fonction a du être répétée pour l'ensemble des observations, c'est pourquoi nous avons créé un script python. Suivant le modèle, nous avons divisé l'emploi des centres par la distance-temps. Pour obtenir un indice unique pour chaque observation, nous avons additionné l'ensemble des valeurs.

### 3.2.1.2. Accessibilité en voiture au centre de Liège

Nous avons aussi construit une variable sur l'accessibilité en voiture au centre de Liège (Place Saint-Lambert). Cet indicateur parait redondant par rapport à l'accessibilité à l'emploi. Si cette variable est corrélée avec l'accessibilité à l'emploi, nous préférons continuer notre modèle avec l'accessibilité à Liège. Cette variable est plus facilement interprétable car les données sont exprimées en minutes et non en emplois par minutes comme pour

---

<sup>3</sup> Lepur : Centre de Recherche en Sciences de la Ville, du Territoire et du Milieu rural

l'accessibilité à l'emploi. De plus les données concernant l'emploi ne sont qu'une estimation du Lepur.

Pour la construction de cette variable, nous avons calculé la distance-temps par rapport au centre de Liège pour chaque point du territoire avec la fonction *cost distance*. Ensuite nous avons associé la valeur aux biens immobiliers en exécutant la fonction, *extract value to point*. Pour chaque logement nous avons la distance en minutes au centre.

### 3.2.1.3. Accessibilité en transport en commun

Premièrement nous avons repris du mémoire de Grignet (2011) les variables ACC\_BUS, ACC\_ML et ACC\_TRAIN, représentant respectivement la part modale en bus, en modes lents et en train. Ces variables proviennent de l'article de Cornet et al. (2005). Elles ont pu être extraites via des fichiers raster fournis par le Lepur. Il y avait trois fichiers raster, un pour chaque mode.

De plus, nous avons décidé de créer notre propre indice d'accessibilité aux transports en commun car ceux utilisés par Grignet (2011) sont déjà assez anciens. L'article de Cornet et al. (2005) se base entre autre sur des données de INS de 1991 soit d'il y a 20 ans. Nous avons estimé l'accessibilité aux bus et aux trains.

Pour l'accessibilité aux bus, nous avons réutilisé l'accessibilité potentielle. Nous pensons qu'à la fois la proximité d'un arrêt de bus TEC et la fréquence de passage des bus constituent des facteurs importants dans le choix d'une maison. Nous avons considéré un rayon de 500 mètres autour des biens immobiliers grâce à la fonction *Buffer*. Cette distance est considérée comme la distance maximale que les personnes sont prêtes à parcourir à pied pour prendre un bus. Nous avons calculé séparément pour chaque bien la distance euclidienne aux arrêts de bus situés dans ce cercle de 500 mètres de rayon. Nous avons donc réemployé les mêmes fonctions que précédemment : *Split Layer By Attributes* pour séparer les biens immobiliers en différentes couches, *Split Layer By Attributes* pour séparer en couches différentes les arrêts de bus se trouvant dans un rayon de 500 mètres autour d'un bien, un script python exécutant la fonction *Distance Euclidienne* pour calculer la distance depuis les différentes observations et un script python exécutant la fonction *extract value to point* pour associer aux arrêts de bus la distance aux observations. Ensuite, nous avons divisé la fréquence de passages de bus par cette distance après l'avoir additionnée à toutes celles existant pour une même observation.



Pour l'accessibilité au train, nous avons calculé la distance–temps pour rejoindre la gare des Guillemins. Ces variables ont pu être calculées avec la fonction *cost distance* et l'image des coûts de déplacement.

### **3.2.2. Environnement socio-économique**

Afin d'évaluer l'environnement socio-économique des biens immobiliers, nous avons exploité les données de l'INS au niveau des secteurs statistiques, niveau le plus fin de découpage territorial pour les statistiques. Nous avons pris comme variables la moyenne des revenus médians sur la période de 2002 à 2008 et le niveau d'instruction (2001).

La première est une variable que Grignet (2011) avait déjà construite dans son mémoire. Elle permet d'avoir un indicateur de niveau économique du voisinage du bien immobilier en gommant les variations dues à la crise de 2007. Comme elle est fortement corrélée au taux de chômage, nous n'avons pas complété le modèle par cette variable.

Pour le niveau d'instruction, nous avons pris la part de la population de plus de 18 ans ne suivant plus d'enseignement de plein exercice ayant un diplôme supérieur (type court, long et universitaire). Cet indicateur a été utilisé pour prendre en compte la valorisation sociale de certains quartiers. Certains quartiers sont peuplés par de jeunes diplômés n'ayant pas encore un revenu important.

Nous avons joint nos tableaux Excel au shapefile des secteurs statistiques à l'aide du code associé à chaque secteur. Ensuite, nous avons attribué à chaque bien immobilier la valeur du secteur dans lequel il était situé grâce à une jointure spatiale.

### **3.2.3. Environnement physique**

L'environnement physique regroupe d'autres les caractéristiques du milieu dans lequel se situe le bien immobilier. Elle regroupe à la fois les caractéristiques de proximité de commerces mais aussi sur la densité du bâti la diversité du paysage ou encore les différentes nuisances. Leur développement est expliqué pour chacune d'entre elles.

#### **3.2.3.1. Densité de population**

Il nous semblait intéressant d'inclure une variable sur la densité de population dans l'environnement immédiat des biens immobiliers. Les Belges aspirent à une maison individuelle 4 façades, ce qui augmente la consommation d'espace et diminue la densité. La Belgique est un exemple de pays qui gaspille l'espace, où la possession d'une maison individuelle est plus aisée que chez nos voisins (Halleux, 2010).

La densité a été calculée par secteur statistique sur les données de population de 2008 et a été jointe aux biens immobiliers avec la fonction *join data* sur base de la localisation du bien.

### **3.2.3.2. Environnement bâti**

Nous avons créé une variable représentant le cadre bâti dans lequel se situait le bien. Nous avons utilisé l'âge des bâtiments pour évaluer cet aspect. Nos données proviennent de l'INS et datent de 2001. Nous avons calculé un indice représentant la part de bâti construit après 1981 que nous avons ajouté à nos données comme au point 3.2.3.1.

### **3.2.3.3. Proximité des commerces**

Une variable sur la proximité des commerces alimentaires a été créée. Nous avons utilisé les données produites par le SEGEFA, Service de recherche de l'Université de Liège. Nous avons uniquement considéré sur la proximité du commerce le plus proche. Il nous semble que le plus important n'est pas la proximité de supermarchés qui sont fréquentés pour des courses hebdomadaires et pour lesquelles une voiture est souvent utilisée mais plutôt des commerces pour des courses quotidiennes qui exigent une proximité importante.

Nous avons calculé la distance euclidienne entre chaque bien et le commerce le plus proche. Nous avons utilisé la fonction *Euclidean Distance* depuis les points de commerces et puis *Extract Value To Point* pour associer la distance aux observations. Dans ce cas-ci, la variable a été séparée en différentes variables binaires car à partir d'une certaine distance il n'y plus d'impact sur la proximité du commerce. Mais comme nous ne connaissons pas à priori cette distance, nous testerons différentes distances ; ainsi nous avons choisi trois variables binaires selon que le bien se situe à moins de 100, 500 ou 700 mètres d'un commerce alimentaire.

### **3.2.3.4. Accessibilité aux nodules commerciaux**

En plus de la variable sur les commerces alimentaires, nous avons ajouté une variable sur l'accessibilité aux nodules commerciaux les plus importants de la région. Pour cela, nous avons réutilisé le concept d'accessibilité potentielle en pondérant dans ce cas avec la superficie commerciale (données du SEGEFA). D'un point de vue commercial, la superficie joue un grand rôle sur l'attractivité du client. C'est pour cette raison que nous avons décidé de l'utiliser comme facteur pondérant.

Pour estimer la valeur de cette variable, nous avons réutilisé les images distance-temps réalisées au point 3.2.1.1. Nous avons ensuite associé à chaque nodule commercial le temps d'accès depuis les différents biens immobiliers à l'aide d'un script python qui répétait

la fonction *Extract value to point*. Après cette opération, nous avons divisé la surface commerciale d'un nodule par la distance temps entre ce nodule et un bien immobilier. Comme l'opération devait être refaite pour les 548 observations, nous avons aussi utilisé un script python. Nous avons rassemblé toutes les observations dans un même fichier grâce à l'outil *Merge*. Ensuite, nous avons importé notre tableau dans Excel pour faire l'addition finale. Les résultats ont été réimportés dans ArcGIS afin de les associer aux observations.

### 3.2.3.5. Espaces verts urbains

Les espaces verts urbains sont un attrait dans une ville. Il nous semblait important de prendre cette caractéristique en compte même si tous les biens immobiliers ne sont pas situés dans un milieu urbain.

De la carte d'occupation du sol de la région wallonne (COSW) (Origine de l'information : MRW-DGATLP, septembre 2011), nous avons extrait les espaces-verts urbains et les forêts (mixtes, feuillues ou résineuses) avec la fonction *Extract By Attributes*. Ensuite, nous avons utilisé, comme pour les commerces alimentaires (3.2.3.3), les outils *Euclidean Distance* et *Extract Value To Point*. Nous avons également transformé notre variable en trois variables binaires suivant la distance entre le bien et l'espace vert. Nous avons aussi choisi 100, 500 et 700 mètres, représentant une distance piétonne.

### 3.2.3.6. Les nuisances

Les nuisances peuvent provenir de diverses sources. Dans ce travail, nous avons pris en compte les nuisances de l'industrie (COSW, MRW-DGATLP), des sites d'extraction (COSW, MRW-DGATLP), des sites d'activité économique désaffectés (DGATLP, MRW-DGATLP) et des routes nationales. Nous aurions souhaité prendre en compte les nuisances sonores, mais nous n'avons trouvé aucune donnée à ce sujet.

Pour les trois premiers points de nuisances, nous avons effectué la même démarche. Premièrement, nous les avons extraits avec *Extract By Attributes*. Ensuite nous avons calculé la distance euclidienne, *Euclidean Distance*, car ce qui compte c'est la proximité et non l'accessibilité. Enfin, nous avons associé ces distances aux observations avec *Extract Value To Point*. Comme pour les deux points précédents, nous avons créé des variables binaires mais avec des distances choisies plus petites car, dans ce cas, les nuisances sont surtout visuelles. Les trois distances sont donc 100, 200 et 500 mètres.

Concernant la variable affectée à la présence de routes nationales, nous voulions tenir compte dans notre modèle des maisons situées le long de grandes routes. Cette situation entraîne des nuisances sonores, de sécurité et aussi des vibrations. Pour ce faire, nous

avons sélectionné sur le réseau dans ArcGIS les routes où la limite de vitesse autorisée sur le réseau était supérieure à 50 km/heure. Nous avons construit, autour de ces axes, une zone tampon assez limitée (100 mètres) et pour tous les biens se situant dans cette zone nous leur avons attribué la valeur 1.

### 3.2.3.7. Diversité

Pour obtenir un indice sur la diversité de l'occupation du sol, nous avons utilisé l'indice de Shannon, déjà utilisé par Grignet (2011). Il mesure la répartition de l'occupation du sol via la formule (Choumert & Travers, 2010) :

$$H = - \sum_{i=1}^k p_i \times \ln p_i$$

Avec  $p_i$  : Proportion d'occupation du territoire d'une catégorie d'occupation du sol

$k$  : nombre de catégories d'occupation du sol

L'indice est minimal quand il n'existe sur le territoire qu'une seule occupation. Dans ce cas, il vaut 0 car  $\ln 1$  égal 0. Par contre, plus il est élevé, plus l'occupation du sol est diversifiée.

La mise en œuvre a débuté en délimitant des zones de 100 mètres autour de chaque bien avec l'outil *Buffer*. Nous avons décidé de travailler sur des territoires de cette dimension car, selon l'étude de Geoghegan, Wainger, & Bockstael (1997), c'est l'environnement le plus proche qui a le plus d'influence sur le prix. Nous avons dû individualiser chaque zone tampon car en certaines se superposaient et donc nous aurions eu des conflits. Cette étape a été réalisée à l'aide de l'extension déjà utilisée précédemment, *Split Layer By Attributes*. Pour calculer les aires des différentes occupations du sol nous avons utilisé la fonction *Tabulate area* qui comptabilise le nombre de pixel de chaque valeur. Comme nous avons 548 couches différentes auxquelles on devait faire subir la même fonction, nous avons créé un script python qui permettait de répéter la même fonction. Pour rassembler les résultats du script dans une même couche nous avons utilisé l'outil *Merge*. Ensuite, nous avons importé ce tableau dans Excel et là nous avons pu calculer  $p_i$  (aire d'une occupation / aire total de la zone), les logarithmes népériens et faire la somme. Enfin, nous avons importé nos valeurs dans ArcGIS grâce à la fonction *Join*.

Nous avons recréé un indice identique mais cette fois-ci sur une zone de 500 mètres de rayon. De plus, pour la zone de 100 mètres, nous avons utilisé les données produites pour calculer un indice sur la part du bâti dans cette zone.

### 3.2.3.8. Paysage

Le paysage est un aspect important dans le choix résidentiel. Cependant, il est difficilement mesurable, surtout quand on ne dispose d'aucune donnée sur l'orientation du bien et sur les vues principales depuis la maison. Nous avons décidé quand même de calculer une variable à ce sujet car il ne peut pas être négligé. Pour ce faire nous avons utilisé un indice inspiré de KIEMSTEDT (Schmitz, 2011) :

$$I = E + D \times C = 12,5 \log_2(Alt_{max} - Alt_{min}) + \sum A \times \sqrt{N}$$

Avec I : Valeur paysagère locale

E : Energie du relief

D : Diversité

C : Complexité

N : Nombre d'îlots d'affectation

Alt<sub>max</sub> : altitude maximale dans le voisinage considéré

Alt<sub>min</sub> : altitude minimale dans le voisinage considéré

A : Poids attribués aux différentes affectations

**Tableau 2 : Points assignés aux différentes affectations**

Affectations	Points attribués
Forêts de feuillus	3,5
Forêts de résineux	2
Forêts mixtes	3
Friches et terrains inculte	2,5
Prairies permanentes	3
Cultures saisonnières et autres	2
Espaces verts urbains	4
Habitat dense	0,5
Habitat discontinu	1,5
Habitat et services	1
Industries et services	0
Carrières, sablières et terrils	2
Terrains et aérodromes militaires	1
Eau	5
Gare	0

Cette formule prend en compte l'amplitude de l'altitude dans le périmètre étudié. Comme la formule prend le logarithme de l'amplitude du relief, les différences d'altitude pour les plus faibles énergies de relief ont une plus grande importance. Nous avons choisi les poids attribués aux différentes affectations sur base des propositions du cours de Schmitz (2011).

Pour cette méthode, nous avons utilisé la carte d'occupation du sol de la région wallonne et travaillé dans un rayon de 500 mètres autour des maisons. Nous sommes partis des résultats de la diversité du point précédent (3.2.3.7) pour connaître le nombre d'îlots d'affectation. Sous Excel, nous avons calculé N et A. Les points attribués aux différentes affectations sont repris dans le tableau 2.

Pour l'altitude, nous avons utilisé un modèle numérique de terrain d'une résolution (IGN) de 30 mètres. Pour extraire les altitudes maximales et minimales de chaque zone nous avons utilisé la fonction *Zonal Statistics as Table* d'ArcGIS. Nous avons exporté les résultats dans Excel et avons pu calculer l'indice I pour chaque bien. Ensuite, nous avons réimporté les données dans ArcGIS pour les associer aux observations.

### 3.3. Résumé des variables explicatives construites

Nom	Description	Source	Type
PRIX	Prix de mise en vente	Immoweb	Quantitatif continu
A_RENOVER	Bien à rénover	Immoweb	Qualitatif binaire
BON_OU_EXC	Bon ou excellent état	Immoweb	Qualitatif binaire
R_A_NEUF	Bien remis à neuf	Immoweb	Qualitatif binaire
SUP_TERRAI	Superficie du terrain	Immoweb	Quantitatif continu
SUP_HAB	Superficie habitable	Immoweb	Quantitatif continu
NB_FACADE	Nombre de façades	Immoweb	Quantitatif discret
NB_ETA	Nombre d'étages	Immoweb	Quantitatif discret
N3FAC	Maison à 3 façades	Immoweb	Binaire
N4FAC	Maison à 4 façades	Immoweb	Binaire
NOETA	Maison de type bungalow	Immoweb	Binaire
N2_ETA	Maison à 2 étages ou plus	Immoweb	Binaire
N1_GAR	Présence d'un garage	Immoweb	Binaire
N0SDB	Pas de salle de bain	Immoweb	Binaire
N2_SDB	Présence de 2 salles de bain ou plus	Immoweb	Binaire

*Partie I : La méthode hédonique*

N1_SDD	Présence d'une salle de douche	Immoweb	Binaire
N1CHB	Présence de 1 chambre	Immoweb	Binaire
N2CHB	Présence de 2 chambres	Immoweb	Binaire
N4_CHB	Présence de 4 chambres	Immoweb	Binaire
N5_CHB	Présence de 5 chambres ou plus	Immoweb	Binaire
S_EQ	Cuisine semi-équipée	Immoweb	Qualitatif binaire
EQ	Cuisine équipée	Immoweb	Qualitatif binaire
CAVE	Présence d'une cave	Immoweb	Binaire
GREN	Présence d'un grenier	Immoweb	Binaire
TERR	Présence d'une terrasse	Immoweb	Binaire
JARD	Présence d'un jardin	Immoweb	Binaire
VERANDA	Présence d'une véranda	Immoweb	Binaire
PISCINE	Présence d'une piscine	Immoweb	Binaire
BUANDERIE	Présence d'une buanderie	Immoweb	Binaire
AGE	Age du bien	Cadastre	Quantitatif continu
MOY_R_MED	Moyenne des revenus médians du secteur statistique entre 2002 et 2008	INS	Quantitatif continu
NIVEAU_INSTRUC	Part des diplômés du supérieur en 2001	INS	Quantitatif continu
Const_81 (SS)	Pourcentage de maisons construites depuis 1981 dans le secteur statistique	INS	Quantitatif continu
DENSITE(SS)	Densité de population dans le secteur statistique	INS	Quantitatif continu
ACC_EMPLOI	Accessibilité potentielle au secteur statistique ayant le plus d'emplois	Lepur	Quantitatif continu
ACC_EMPL2	Accessibilité potentielle au secteur statistique ayant le plus d'emplois (divisé par la distance au carré)	Lepur	Quantitatif continu
ACC_BUS	Part modale du bus	Lepur	Quantitatif continu
ACC_ML	Part modale des modes lents	Lepur	Quantitatif continu

*Partie I : La méthode hédonique*

ACC_TRAIN	Part modale du train	Lepur	Quantitatif continu
ACC_POT_BUS	Accessibilité potentielle au bus dans un rayon de 500m	Lepur et TEC	Quantitatif continu
B1-BUS	Buffer de 100 mètres autour des arrêts de bus	Lepur et TEC	Binaire
B5-BUS	Buffer de 500 mètres autour des arrêts de bus	Lepur et TEC	Binaire
B7-BUS	Buffer de 700 mètres autour des arrêts de bus	Lepur et TEC	Binaire
DIST_GUILLEMINS (min)	Distance à la gare TGV des Guillemins	MRW-DGATLP**	Quantitatif continu
DIST_COM (m)	Distance à un commerce alimentaire	SEGEFA	Quantitatif continu
B1_COM	Buffer de 100 mètres autour des commerces alimentaires	SEGEFA	Binaire
B5_COM	Buffer de 500 mètres autour des commerces alimentaires	SEGEFA	Binaire
B7_COM	Buffer de 700 mètres autour des commerces alimentaires	SEGEFA	Binaire
ACC_NODULE	Accessibilité potentielle aux nodules commerciaux	SEGEFA	Quantitatif continu
B1_EXTRACT	Buffer de 100 mètres autour des sites d'extractions	COSW*	Binaire
B2_EXTRACT	Buffer de 200 mètres autour des sites d'extractions	COSW*	Binaire
B5_EXTRACT	Buffer de 500 mètres autour des sites d'extractions	COSW*	Binaire
B1_INDU	Buffer de 100 mètres autour des industries	COSW*	Binaire
B2_INDU	Buffer de 200 mètres autour des industries	COSW*	Binaire
B3_INDU	Buffer de 500 mètres autour des industries	COSW*	Binaire
B1_SAED	Buffer de 100 mètres autour des sites d'activé économique désaffectés	MRW-DGATLP**	Binaire
B2_SAED	Buffer de 200 mètres autour des sites d'activé économique désaffecté	MRW-DGATLP **	Binaire



*Partie I : La méthode hédonique*

B5_SAED	Buffer de 500 mètres autour des sites d'activé économique désaffectés	MRW-DGATLP**	Binaire
B1_ROUTE	Buffer de 100 mètres autour des principaux axes routiers	MRW-DGATLP **	Binaire
B1_VERT	Buffer de 100 mètres autour des espaces verts	COSW*	Binaire
B5_VERT	Buffer de 500 mètres autour des espaces verts	COSW*	Binaire
B7_VERT	Buffer de 700 mètres autour des espaces verts	COSW*	Binaire
SHANNON_10	Diversité de l'occupation du sol dans un rayon de 100m	COSW*	Quantitatif continu
SHANNON_50	Diversité de l'occupation du sol dans un rayon de 500m	COSW*	Quantitatif continu
PART_BATIT_10	Part de zone bâtie dans un rayon de 100m	COSW*	Quantitatif continu
PAYSAGE	Indice paysagé de Kiemsted	COSW*	Quantitatif continu

\* MRW-DGATLP, septembre 2011

\*\* date inconnue

## 4. ANALYSES PRÉLIMINAIRES

### 4.1. Les logiciels Statistica et Gretl

Pour la réalisation des traitements statistiques nécessaires à la méthode hédonique, deux logiciels ont été utilisés : Statistica (version 10) et Gretl (version 1.9.9).

Statistica est un logiciel développé par la firme StatSoft. Il permet la gestion, l'analyse et la représentation des données. De nombreux outils statistiques sont disponibles dont une partie porte sur les modèles de régression. Ce logiciel payant est fourni aux étudiants de l'ULg.

Gretl (Gnu Regression, Econometrics and Time-series Library) est un logiciel libre développé par Allin Cottrell (<http://gretl.sourceforge.net>). Il présente de nombreuses fonctionnalités pour les régressions dont certaines (le test de Whitet ou de Breusch-Pagan) ne sont pas disponibles sous Statistica. Ce logiciel est conçu pour des analyses économétriques et est notamment utilisé dans le cours d'économétrie de 3<sup>e</sup> baccalauréat des sciences économiques de l'ULg. L'entièreté du travail aurait pu se réaliser via ce logiciel mais l'interface est moins conviviale.

### 4.2. Statistiques descriptives des variables

Cette analyse statistique descriptive permet une première approche des variables et du marché immobilier étudié. Elle pourrait apporter des suggestions de transformations de nos variables pour mieux rendre compte de leur implication dans le prix. C'est aussi à cette occasion que l'on peut repérer des observations de valeurs aberrantes dans l'échantillon qui peuvent fausser le modèle et les éliminer.

Comme nous avons repris les mêmes variables intrinsèques que Grignet (2011), nous pouvons reprendre ses analyses (tableau 3).

**Tableau 3 : Statistiques descriptives de variables de l'échantillon étudié**

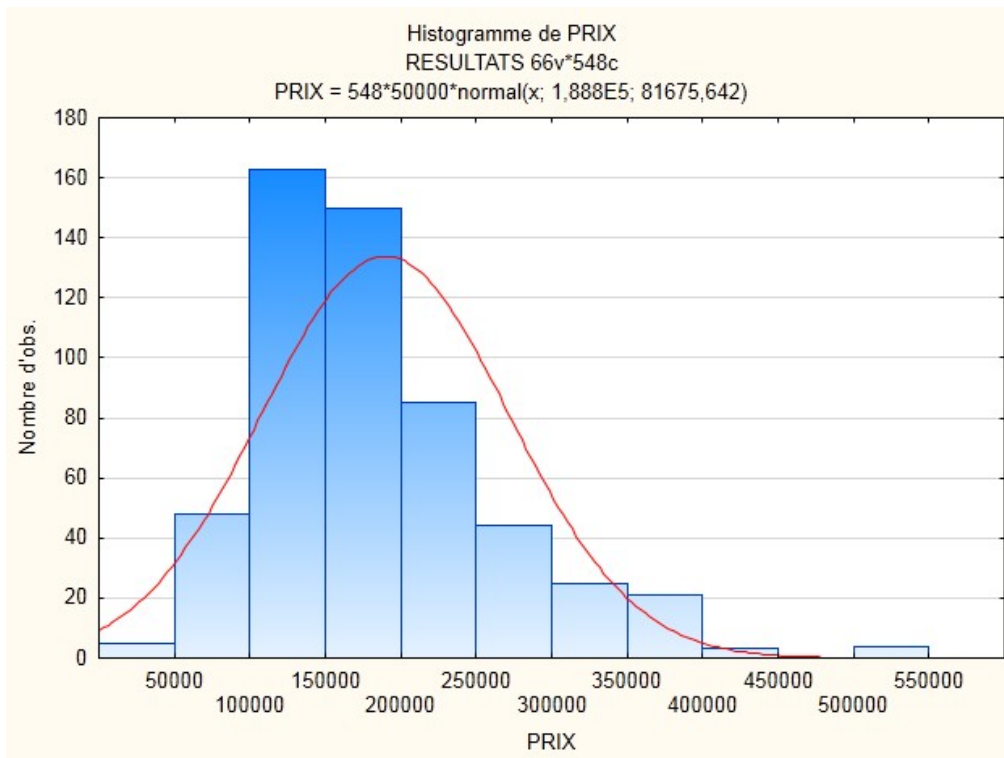
n=548	Moyenne	Minimum	Maximum	Ecart-type
Prix (€)	188802	45000	550000	81676
Superficie du terrain (m <sup>2</sup> )	627	25	5800	848
Superficie habitable (m <sup>2</sup> )	152	50	400	59
Age (année)	84	3	186	45

Source : Grignet, 2011, p. 37

## Partie I : La méthode hédonique

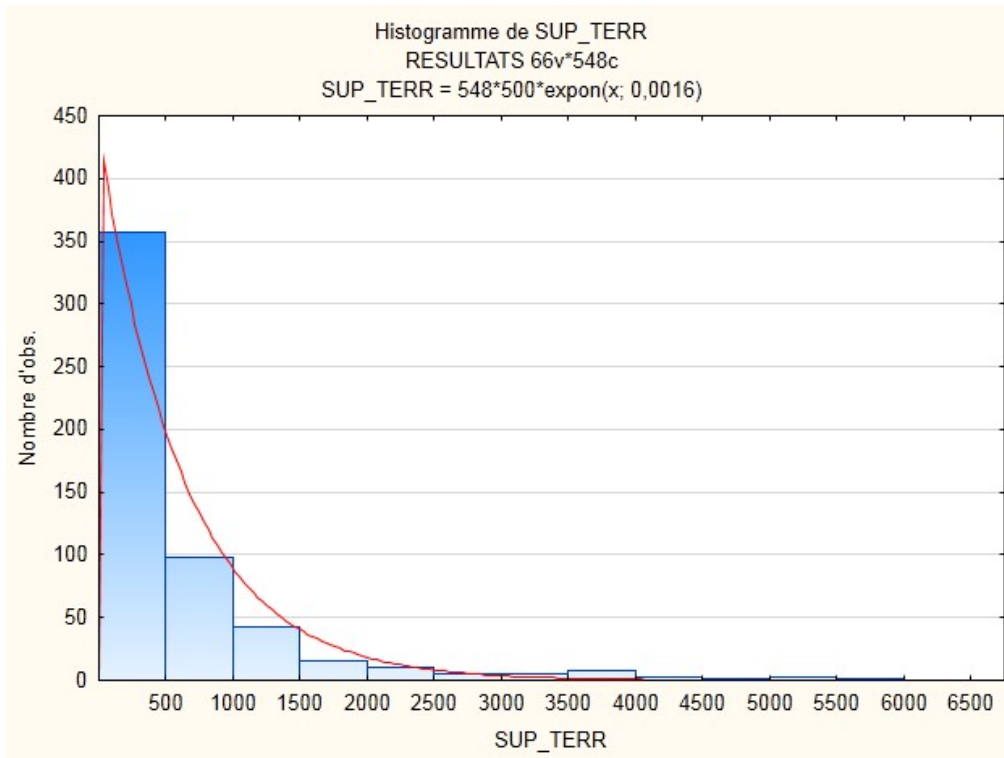
Concernant le prix, sa moyenne est de 188.800 €, avec un écart type de 81.676 €. L'échantillon est dispersé puisqu'il s'étend entre 45.000 € et 550.000 € : aucune restriction n'avait été faite sur le prix lors de la collecte des données sur Immoweb et ce pour maximiser le nombre d'observations. Nous nous rendons compte de l'étendue importante des prix et sommes conscient d'être en présence de plusieurs sous-marchés qui ne peuvent pas être distingués. Cependant nous décidons de réaliser l'étude sur l'ensemble des observations pour bénéficier d'un échantillon suffisamment important et parce qu'il est difficile de délimiter les sous-marchés. Nous envisagerons par la suite de réduire notre échantillon en supprimant les biens immobiliers avec des prix importants pour lesquels le modèle de Grignet (2011) était mal ajusté (valeurs de plus de 260.000 €). Pour ce qui est de la distribution de l'échantillonnage, on constate que le graphique 1 suit approximativement une loi normale avec quelques prix élevés marginaux.

**Graphique 1 : Distribution des prix**

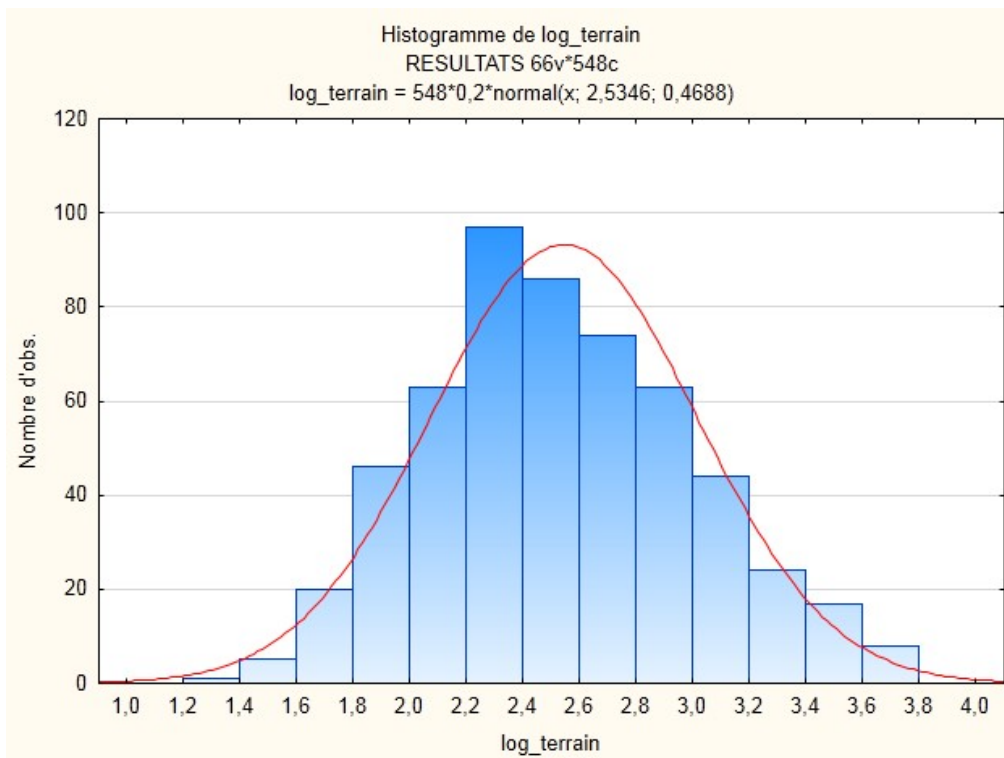


Pour les autres variables structurelles quantitatives, on peut constater grâce aux histogrammes que la distribution de la superficie de terrain suit la loi exponentielle. Nous avons donc principalement des observations avec des terrains inférieurs à 500 m<sup>2</sup> (graphique 2). Nous décidons de transformer cette variable en logarithme base 10. Après cette transformation, les observations suivent une loi normale (graphique 3). La distribution de la superficie habitable (graphique 4) suit une loi normale même si la catégorie 100-150 m<sup>2</sup> est fortement représentée et qu'elle se prolonge vers les grandes superficies.

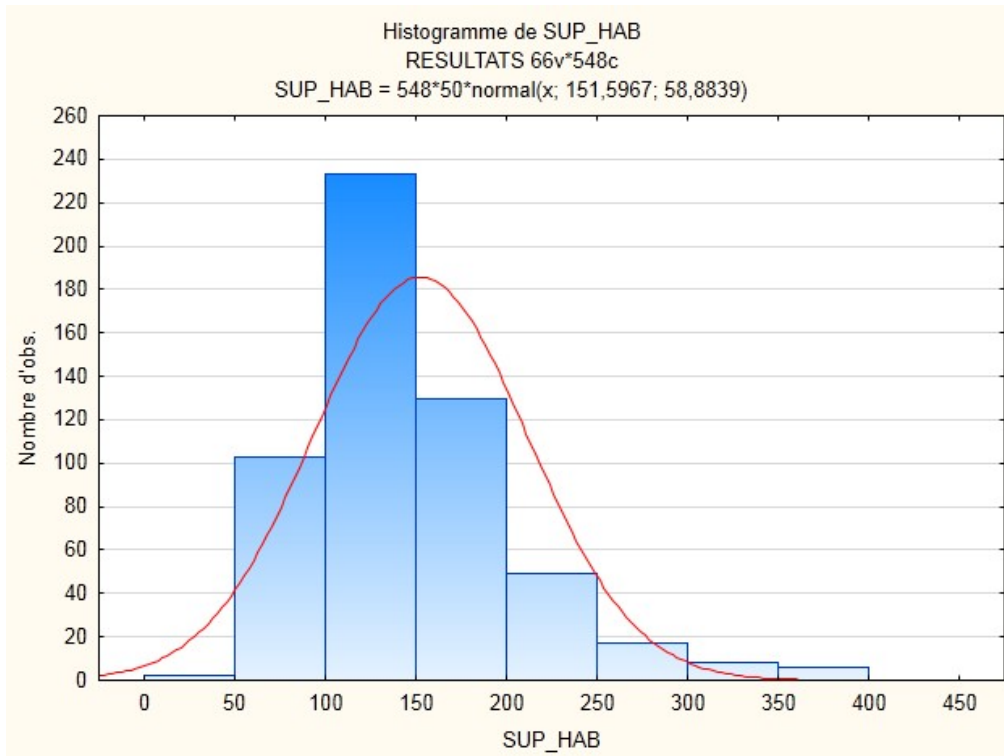
Graphique 2 : Distribution des superficies de terrain



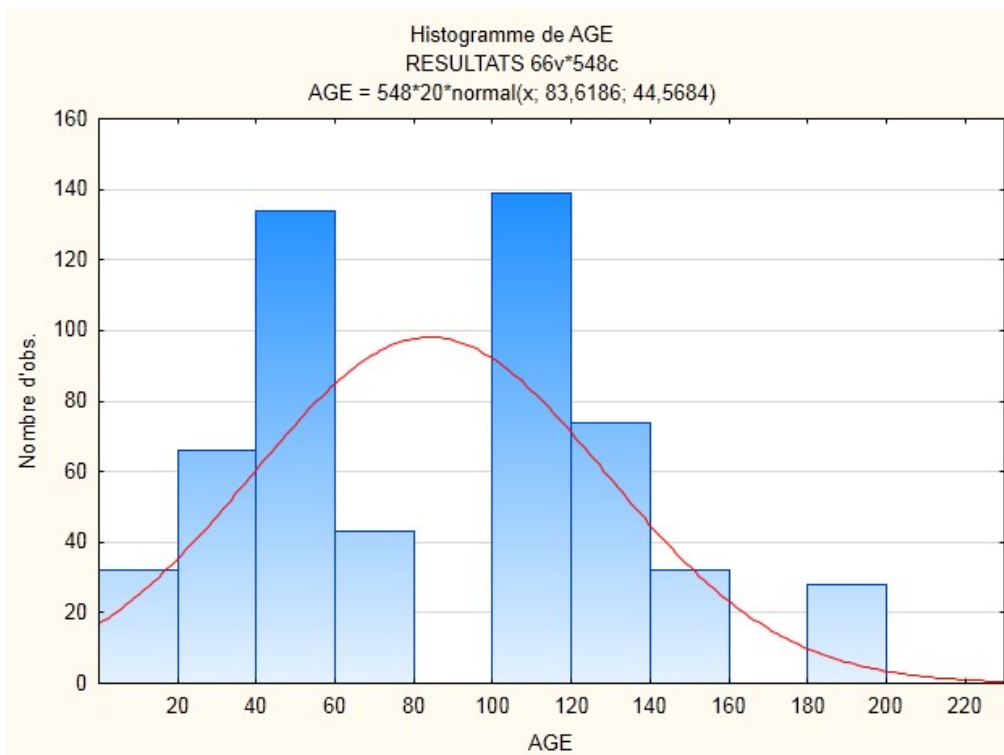
Graphique 3 : Distribution des logarithmes des superficies de terrains



Graphique 4 : Distribution des superficies habitables



Graphique 5 : Distribution des âges des biens immobiliers

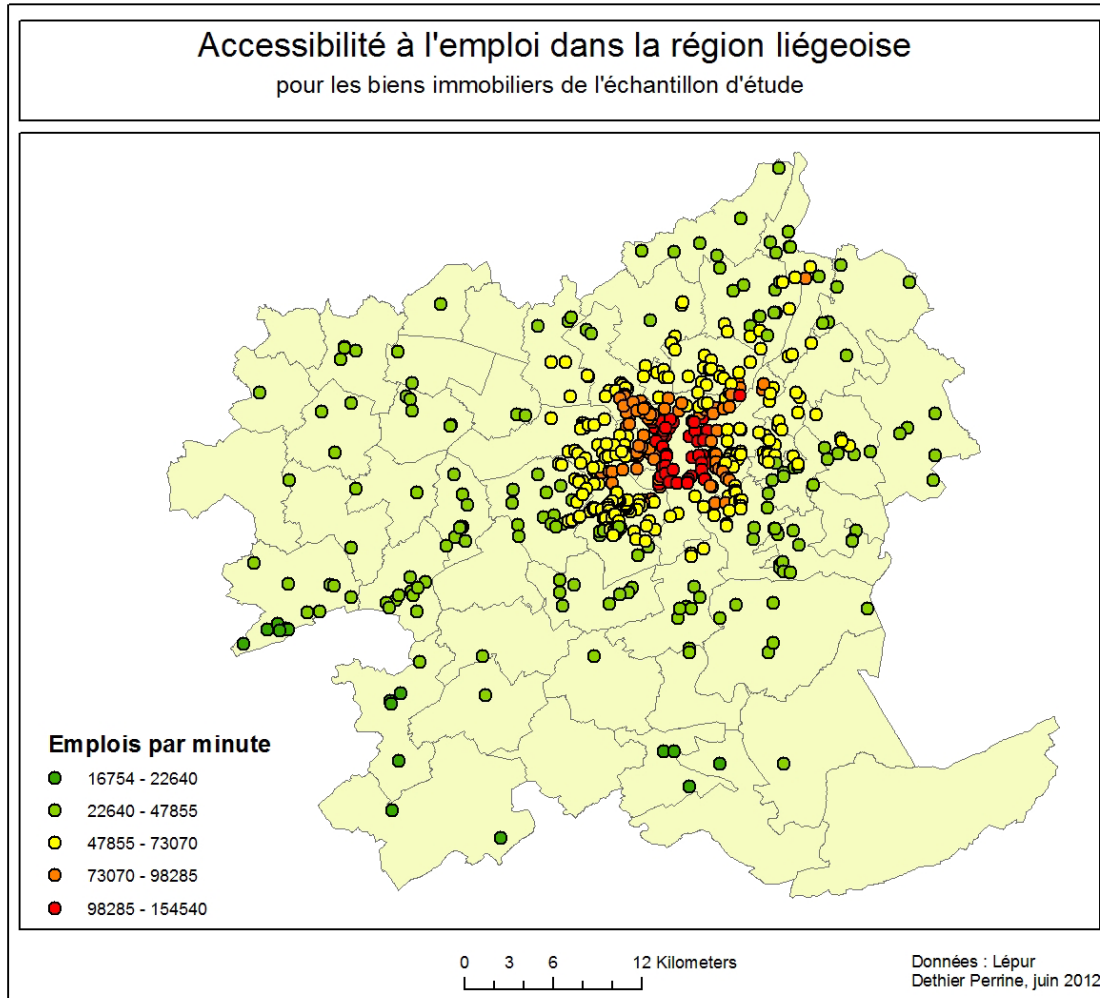


Le graphique 5 concernant les âges des habitations comporte de fortes irrégularités. Cette perturbation est due aux guerres qui stoppent la construction. La décennie des années 1910 (bien entre 100 et 90 ans) est absente et les biens entre 60 et 80 ans sont peu représentés.

D'après le mémoire de Grignet (2011), l'augmentation de l'âge n'entraîne pas une diminution linéaire du prix mais c'est lors des premières années que la diminution est la plus importante. Nous transformons aussi cette variable en calculant son logarithme en base 10.

Comme ces observations ne sont pas indispensables à la réalisation d'un modèle hédonique nous nous contenterons de présenter uniquement les résultats pour l'accessibilité à l'emploi.

**Carte 3 : Accessibilité à l'emploi**



La moyenne est de 60.462,16 emplois par minute depuis le domicile. Ce chiffre est important pour plusieurs raisons. Premièrement, cet indice ne considère pas l'espace comme un espace à deux dimensions mais à une seule dimension. Ainsi, si il y a un emploi à une minutes à l'ouest et un emploi à une minute à l'est, soit l'opposé, l'indicateur ne sera pas un emploi par minute mais deux emplois par minute vu que l'on fait une somme. Ensuite, si la distance est inférieure à une minute, cela entraîne une augmentation d'emploi. Ce cas est rare mais existe. Enfin, une part importante des biens est située dans la commune de Liège, le pôle principal d'emplois. On peut confirmer cette déclaration grâce à la carte 3. Plus le bien se situe à proximité du centre de Liège, plus l'indicateur est important. On a donc de

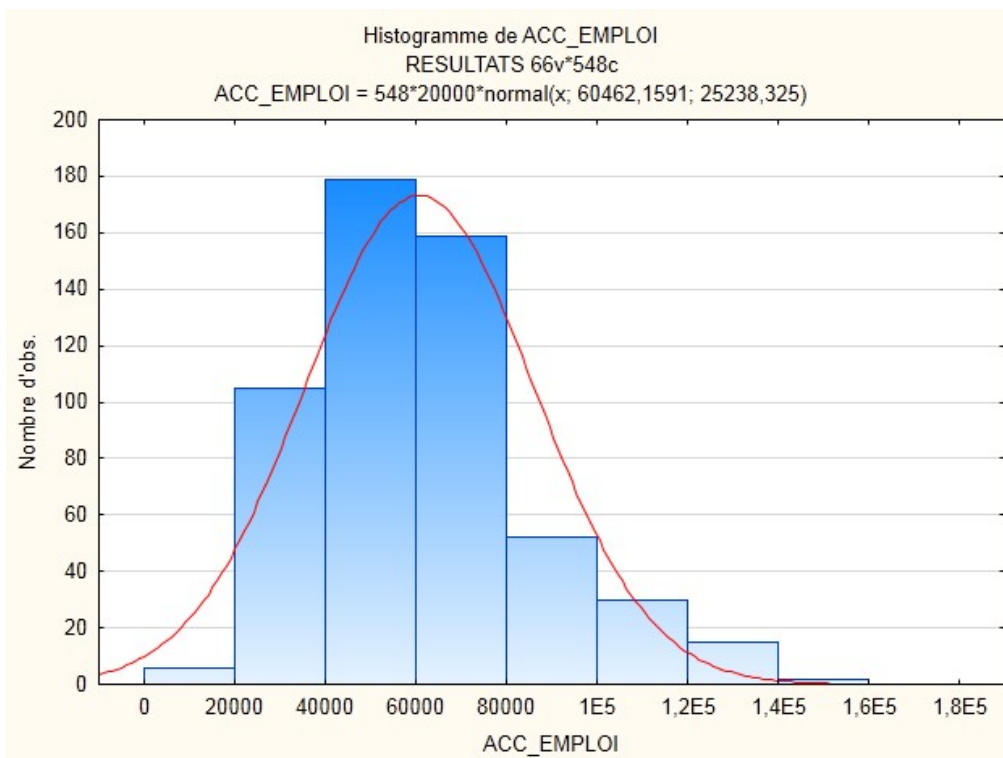
grandes chances que cet indicateur soit fortement corrélé avec la distance au centre de Liège.

**Tableau 4 : Statistiques descriptives de l'accessibilité à l'emploi**

Moyenne	Médiane	Minimum	Maximum	Ecart-type
60462,16	57497,87	16753,80	154540,3	25238,32

La distribution de cette variable suit approximativement une loi normale. Il y a peu de valeurs faibles car dans la région liégeoise, tous les endroits sont facilement accessibles d'autant plus que nous avons peu d'observations dans des communes fort éloignées.

**Graphique 6 : Distribution de la variables sur l'accessibilité à l'emploi**



### 4.3. Analyse de corrélation

La seconde étape, avant la réalisation de la régression multiple, est l'étude de la corrélation entre les variables.

La première analyse portera sur la corrélation entre les variables indépendantes et le prix pour vérifier s'il existe un lien. Ensuite, il est important d'étudier la corrélation entre les variables pour limiter la colinéarité. « On dit qu'il y a colinéarité parfaite entre les régresseurs d'une régression multiple lorsque une (ou plusieurs) des variables explicatives de la régression est une combinaison linéaire exacte des (ou d'un sous-ensemble des) autres

*variables explicatives du modèle*) » (Lejeune B., 2011). Si deux variables sont fortement corrélées (au delà de 0,8), nous déciderons de supprimer l'une d'entre elles pour éviter ce problème de colinéarité. Quand des variables sont colinéaires, les coefficients estimés par la régression sont individuellement peu ou pas significatifs et donc l'estimation du modèle est faussée.

#### **4.3.1. Corrélation avec le prix**

Nous devons calculer des corrélations entre la variable prix et les variables explicatives. La matrice de corrélations se trouve à l'annexe 2. Comme Grignet (2011° Nous décidons de garder toutes les variables et nous déciderons de les supprimer uniquement si lors de notre première régression, il ressort effectivement que cette variable n'a pas d'influence.

#### **4.3.2. Corrélation entre les variables**

Entre les variables quantitatives, une matrice de corrélation est calculée (annexe 2). Nous observons que dans la matrice plusieurs couples sont fortement corrélés.

D'une part, la corrélation entre l'accessibilité à l'emploi et l'accessibilité aux nodules commerciaux est de 0,85. Nous décidons de supprimer la variable accessibilité aux nodules commerciaux car elle présente des faibles corrélations avec d'autres variables. Cette corrélation nous paraît tout à fait logique car les pôles d'emplois sont situés aux mêmes endroits que les principaux nodules commerciaux. La variable accessibilité à l'emploi est corrélée à -0,82 à la distance aux Guillemins. Nous supprimons la variable distance aux Guillemins car nous avons déjà une variable sur l'accès au train et que la distance aux Guillemins est fortement corrélée à la proximité à Liège. Ces deux variables (accessibilité à l'emploi et proximité des Guillemins) sont corrélées car elles sont relatives à la proximité de Liège, centre d'emploi important. Nous avons aussi une corrélation de -0,83 entre l'accès à l'emploi et la distance à Liège. Comme nous le supposions, l'accès à l'emploi est lié à la proximité de Liège. La corrélation est négative car plus on est proche de Liège plus le temps pour rejoindre le centre est court mais plus il y a de l'emploi à proximité. Nous décidons quand même de garder ces 2 variables et de choisir lors de nos premières régressions celle qui produit un meilleur modèle. Nous avons quand même une préférence pour la variable sur la distance à Liège car elle est plus facile à interpréter pour l'évaluation par contingence.

Nous constatons aussi une forte corrélation (0,84) entre la part modale en bus et l'accessibilité au bus. Nous décidons de garder la première variable car elle prend en compte des indicateurs socio-économiques.



Enfin il y a une corrélation de 0,78 entre la moyenne des revenus médians et le niveau d'instruction. Nous choisissons comme seuil de rejet 0,8 mais, après avoir réalisé plusieurs tests, nous nous sommes rendu compte que ces deux variables étaient trop corrélées et que nous avions des problèmes de colinéarité. Donc nous avons supprimé la variable niveau d'instruction.

Pour les variables binaires, nous décidons de les garder toutes car, après analyse des coefficients de corrélation de points, aucun couple ne ressort corrélé.

Au final, nous avons 61 variables. C'est évidemment trop pour une régression multiple, nous en éliminerons durant la réalisation des régressions. A noter que de nombreuses caractéristiques sont représentées par plusieurs variables binaires.

## 5. DÉVELOPPEMENT DU MODÈLE

Après cette analyse statistique descriptive de nos variables, la régression multiple peut être réalisée. L'ajustement du modèle se fera avec le logiciel Gretl et ce en n'utilisant pas la méthode de sélection des variables « pas à pas » (« *stepwise* »). Selon Des Rosiers (2001), il est déconseillé de l'utiliser car elle ne permet pas à l'analyste de choisir des variables dans le modèle. La méthode « *stepwise* » ajoute dans le modèle seulement les variables qui sont significatives et retire celles qui ne le sont plus. De plus, elle évite des problèmes de colinéarité.

Dans un premier temps, nous allons réaliser l'ajustement par la méthode des moindres carrés ordinaires (MCO) avec les variables retenues et en sélectionnant celles qui ont une p-value inférieure à 0,10. Ensuite nous essayerons d'améliorer la part de la variance expliquée ( $R^2$ ) et les différents critères expliqués ci-après.

Quand un modèle nous conviendra, nous vérifierons s'il n'y a pas de colinéarité et analyserons les résidus pour vérifier l'hétéroscédasticité. Si c'est le cas, nous ajusterons le modèle après avoir transformé certaines variables ou en utilisant une autre méthode que la régression des moindres carrés ordinaires. Enfin, nous pourrions vérifier qu'il y a absence d'autocorrélation spatiale des résidus.

L'ensemble des modèles réalisés sont présent à l'annexe 3.

### 5.1. Premier modèle

Dans ce premier modèle, nous choisirons parmi ces variables :

- Accès à l'emploi ou la distance à Liège

## Partie I : La méthode hédonique

- 100, 500 ou 700 mètres autour des arrêts de bus
- 100, 500 ou 700 mètres autour des commerces alimentaires
- 100, 500 ou 700 mètres autour des espaces verts
- 100, 200 ou 500 mètres autour des sites d'extraction
- 100, 200 ou 500 mètres autour des industries
- 100, 200 ou 500 mètres autour des SAED

Lors de la construction de nos variables, nous n'avons pas pu choisir une des solutions et avons préféré faire ce choix lors de la première régression en prenant celles qui permettent le meilleur modèle.

La qualité de notre premier modèle (MCO, 548 observations, variable dépendante : le prix) est décrit par les indicateurs globaux du tableau 5. Il explique 80 % du prix et ce avec une p-value extrêmement petite. D'autres paramètres statistiques sont fournis par le logiciel Gretl. Le Critère d'Akaike, souvent abrégé AIC, permet de faire une sélection entre plusieurs modèles en choisissant la plus faible valeur de cet AIC. Il équivaut à (Lancelot & Lesnoff, 2005) :

$$AIC = -2\ln(L) + 2k$$

Avec  $k$  : le nombre de paramètres dans le modèle

$L$  : vraisemblance du modèle

Le critère de Schwarz ou BIC est fort similaires à l'AIC. Il prend plus en compte le nombre de variables explicatives choisies dans le modèle. Il faut choisir le modèle possédant le plus petit BIC. Il égale (Lancelot & Lesnoff, 2005) :

$$BIC = -2\ln(L) + k\ln(n)$$

Avec  $n$  : le nombre d'observations

$k$  : le nombre de paramètres dans le modèle

$L$  : vraisemblance du modèle

Enfin, l'indice Hannan-Quinn est un critère de sélection qui doit aussi être minimisé. Il est un intermédiaire entre l'AIC et le BIC. Il prend en compte le nombre d'observations mais de manière moins importante que le BIC.

**Tableau 5 : Indicateurs globaux de la qualité d'ajustement du modèle 1**

Moy. var. dép.	188802,5	Éc. type var. dép.	81675,64
Somme carrés résidus	7,46e+11	Éc. type de régression	37685,87
R <sup>2</sup>	0,795664	R <sup>2</sup> ajusté	0,787102
F(22, 525)	92,92290	p. critique (F)	5,0e-165
Log de vraisemblance	-6540,128	Critère d'Akaike	13126,26
Critère de Schwarz	13225,30	Hannan-Quinn	13164,97

## 5.2. Deuxième modèle : choix parmi nos variables

Après ce premier essai, nous essayons d'améliorer notre modèle en remplaçant l'accessibilité par la distance à Liège. Notre modèle 2, tableau 6, présente une significativité fortement similaire à notre modèle 1 et les critères (d'Akaike, de Schwarz, Hannan-Quinn) sont légèrement supérieurs. Nous décidons de prendre en compte la distance à Liège et non l'accès à l'emploi car elle ne donne pas un modèle réellement différent mais, par contre, elle est plus explicite. Pour nos variables environnementales, deux variables sont significatives à un seuil de p-value sous 0,10. Ce sont les biens situés à moins de 700 mètres d'un arrêt de bus et 200 mètres des industries. Cependant la variable de proximité d'un arrêt de bus donne un coefficient de régression contre-intuitif puisque celui-ci est négatif. La proximité d'un arrêt de bus est une moins-value.

**Tableau 6 : Indicateurs globaux de la qualité d'ajustement du modèle 2**

Moy. var. dép.	188802,5	Éc. type var. dép.	81675,64
Somme carrés résidus	7,44e+11	Éc. type de régression	37692,88
R <sup>2</sup>	0,795978	R <sup>2</sup> ajusté	0,787023
F(23, 524)	88,88465	p. critique (F)	3,3e-164
Log de vraisemblance	-6539,708	Critère d'Akaike	13127,42
Critère de Schwarz	13230,77	Hannan-Quinn	13167,81

### 5.2.1. Test de colinéarité

Nous allons maintenant vérifier la colinéarité et analysant pour chaque variable la valeur de son facteur d'inflation de variance, VIF. Cet indicateur doit être relativement faible. Souvent, on utilise le seuil de 7,5.

$$VIF_j = \frac{1}{1 - R_j^2}$$

Avec  $R_j$  : coefficient de corrélation multiple entre la variable  $j$  et les autres variables indépendantes.

Dans notre cas, aucune variable n'a un VIF supérieur à 2,6 donc nous n'avons pas de problème de colinéarité (annexe 4). Quand nous prenions en compte le niveau d'instruction, le VIF du revenu médian était proche de 7 et son coefficient de régression n'était plus significatif.

### **5.2.2. Analyse des résidus**

Le modèle fournit pour chaque bien une valeur estimée du prix de vente. Celle-ci diffère du prix observé. La différence entre le prix observé et le prix estimé donne le résidu.

Il est important d'analyser les résidus car cette analyse permet de vérifier si l'ajustement est bien ajusté. Si dans le modèle, les variables influençant le prix ne sont pas pris en compte ou pas de manière assez importante, on a un risque pour certains biens d'avoir des résidus plus importants que la normale, d'avoir de l'hétéroscédasticité ou de l'autocorrélation spatiale. Cette étape d'analyse des résidus permet de juger la qualité du modèle.

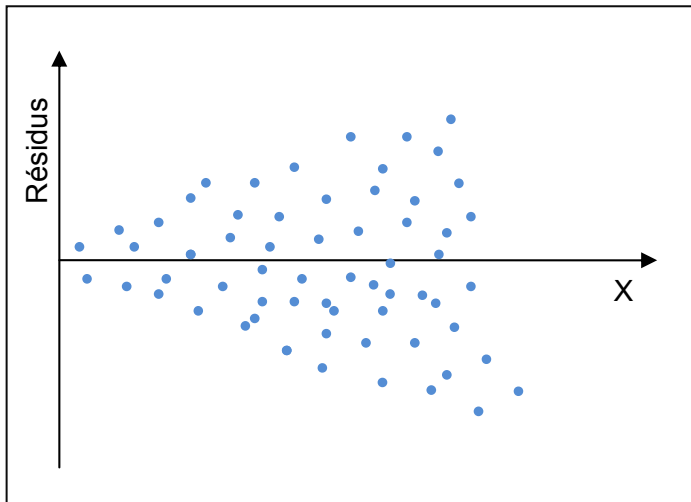
Nous allons donc faire trois analyses : un test d'hétéroscédasticité, la suppression des observations présentant un résidu excédentaire et une analyse de l'autocorrélation spatiale des résidus.

#### **5.2.2.1. L'hétéroscédasticité**

L'hétéroscédasticité se présente quand la variance des résidus n'est pas constante. Les causes sont diverses (Williams, 2011) :

- Les résidus peuvent croître en même temps qu'une variable explicative croît. Dans ce cas, la variable explicative pour des valeurs élevées n'est pas suffisante pour expliquer l'évolution de la variable dépendante.

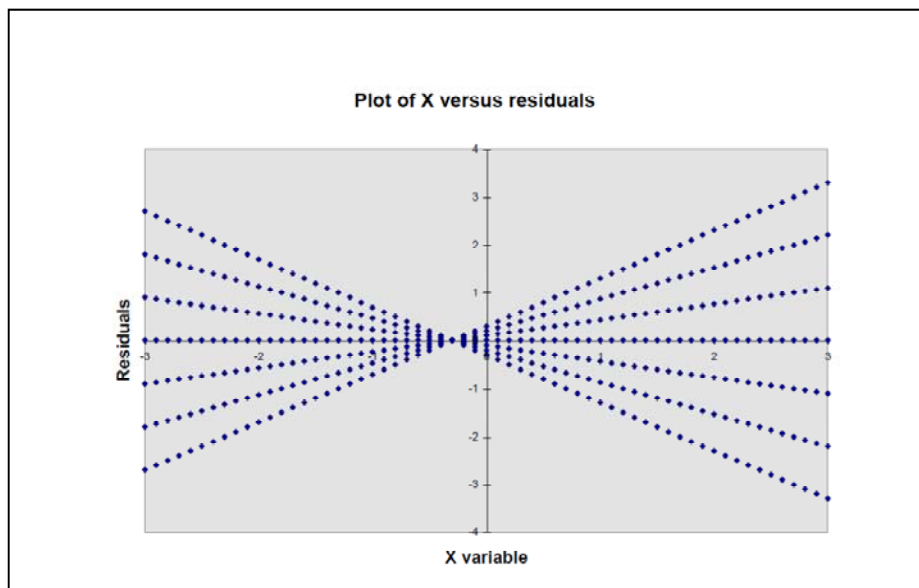
Figure 1 : Evolution de la variable dépendante Y en fonction de la variable explicative X



Source : Williams, 2011, p. 1

- Les résidus peuvent croître vers les deux extrêmes d'une variable explicative : vers des valeurs faibles et élevées.

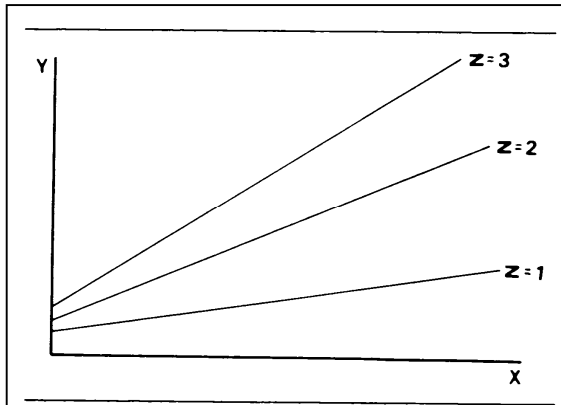
Figure 2 : Variation des résidus en fonction de la variable explicative X



Source : Williams, 2011, p.1

- Il y a de l'hétéroscédasticité quand il y a des sous-populations. Ces sous-populations ont une évolution différente de la variable dépendante en fonction d'une variable explicative.

Figure 3 : Variation de la variable dépendante Y selon la variable explicative X pour plusieurs populations



Source : Williams, 2011, p.2

- Certaines variables explicatives peuvent entraîner des erreurs dans le modèle car la relation les liant au prix n'est pas linéaire. A la place d'utiliser la variable X, il faut la transformer en  $X^2$ ,  $\log X$ ...pour mieux l'ajuster.
- Enfin, il peut y avoir des erreurs de mesure des données.

Le principe d'une régression MCO est de minimiser la somme des carrés des résidus ( $\sum e_i^2$ ). Cette démarche entraîne que les observations pour lesquelles il y a potentiellement une plus grande erreur vont obtenir un plus grande influence dans la régression. Un résidu peut être d'autant plus minimisé qu'il est initialement grand. Cette correction permettra une plus grande amélioration du modèle puisque  $\sum e_i^2$  sera nettement diminuée. Donc, dans une régression par MCO, ce sont les valeurs extrêmes qui influencent le plus le modèle. Or si ces valeurs ne sont pas bien ajustées au modèle, c'est qu'elles donnent moins d'information sur les relations étudiées. Comme dit Allison (1998) : "*The reason OLS is not optimal when heteroskedasticity is present is that it gives equal weight to all observations when, in fact, observations with larger disturbance variance contain less information than observations with smaller disturbance variance*".

Il existe plusieurs moyens pour détecter l'hétéroscédasticité : un graphique, un test de Breusch-Pagan ou un test de White.

- Le graphique représente les résidus en fonction d'une variable. S'il y a de l'hétéroscédasticité, les résidus ne seront pas similaires pour toutes les valeurs de la variable explicative. Ils dessineront un cône ou un sablier renversé. Cette technique est simple mais il est préférable d'utiliser les deux autres tests pour certifier la présence d'hétéroscédasticité.

## Partie I : La méthode hédonique

- Le test de Breusch-Pagan. Ce test effectue une régression linéaire avec pour variables dépendantes les résidus et comme variable explicative, chacune des variables du modèle. L'hypothèse nulle de ce test est l'homoscédasticité.
- Le test de White. Ce test, en plus d'effectuer la régression linéaire expliquée ci-dessus, utilise les carrés et les produits croisés des variables. Dans ce cas des relations non linéaires entre les résidus et les variables explicatives peuvent être trouvées.

Dans le cas de l'hétéroscédasticité, plusieurs solutions sont possibles :

1. La modification du modèle. Dans un premier temps, il faut envisager de transformer certaines variables qui ne sont pas en relation linéaire avec la variable à expliquer et aussi inclure de nouvelles variables explicatives importantes qui ont été oubliées
2. L'estimateur robuste à l'hétéroscédasticité. Dans une régression par MCO, les écart-types des coefficients sont biaisés. Cette méthode permet de calculer des écart-types en tenant compte de la variation des résidus. Par cette méthode, les coefficients de la régression estimés ne changent pas mais les tests statistiques (écart-type et le test de conformité) sont modifiés.
3. La méthode des Moindres Carrés Généralisés (MCG). Cette solution est plus puissante que la précédente mais plus difficile à mettre en œuvre car il faut déterminer une variance pour chaque résidu. Le problème de la méthode des MCO est de donner un poids trop important aux observations avec une grande variance. Le modèle doit donc être pondérer les observations de manière à donner une plus grande importance aux observations avec une variance faible qui apportent plus d'information.

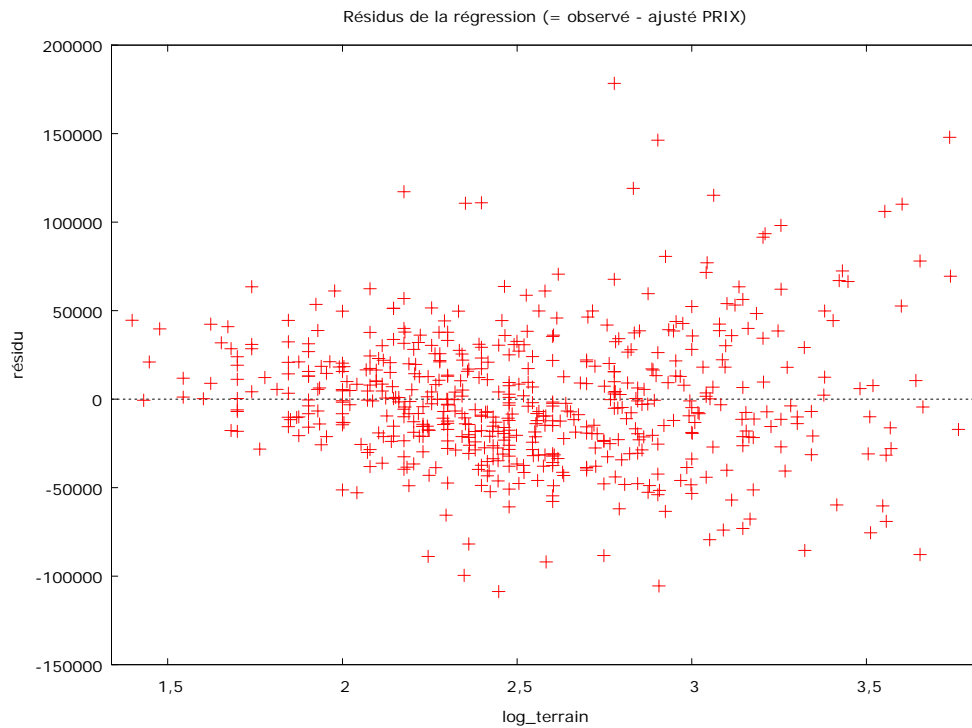
MCO.....MCG

$$\sum(Y_j - \hat{Y}_j)^2 \dots \sum \frac{(Y_j - \hat{Y}_j)^2}{VAR(\epsilon_j)}$$

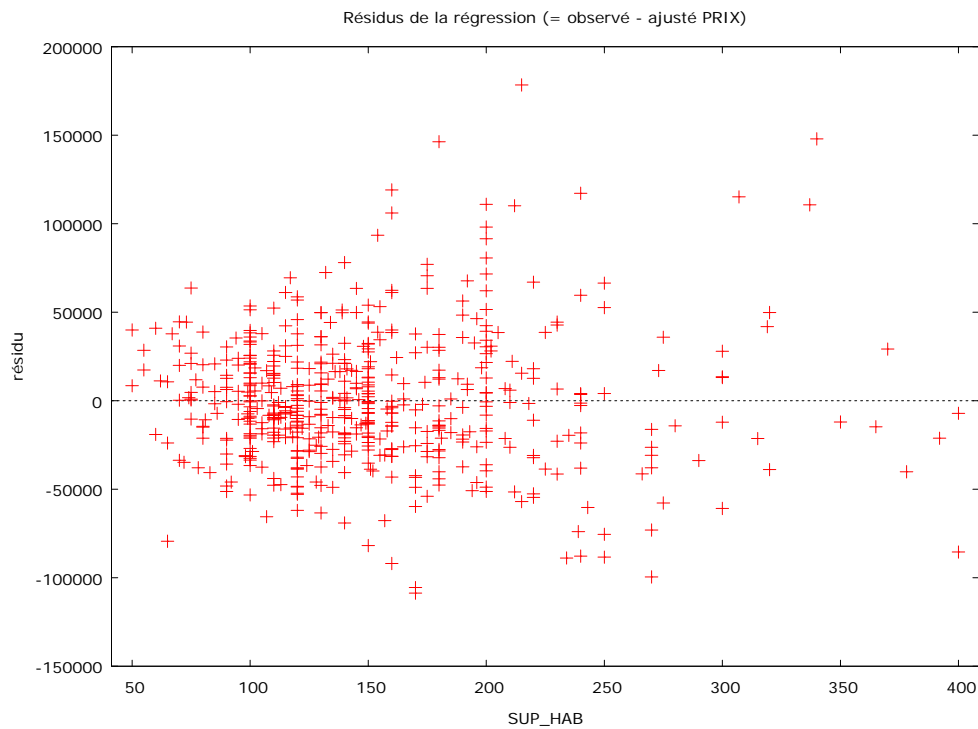
Dans notre cas, le test de Breusch-Pagan (annexe 5) montre que notre modèle présente de manière significative de l'hétéroscédasticité et cela sur trois variables : la superficie du terrain, la superficie habitable et la présence d'une piscine. Nous pouvons le confirmer en regardant les graphiques 7 et 8. Sur ces graphiques, on voit que plus la superficie augmente moins les résidus sont proches de 0. Les résidus ne sont pas stables. En conséquence, nous devons réaliser un nouveau modèle (modèle 3) en solutionnant ce problème.

## Partie I : La méthode hédonique

**Graphique 7 : Résidus du modèle 2 en fonction de la superficie terrain**



**Graphique 8 : Résidus du modèle 2 en fonction de la superficie habitable**





### 5.3. Modèle 3 : correction de l'hétéroscédasticité

Dans un premier temps, il faut effectuer des transformations sur ces variables. La superficie du terrain est déjà utilisée en logarithme. Pour la présence d'une piscine, variable binaire, aucune transformation ne permettra de résoudre le problème car elle restera binaire. Avec la superficie habitable, une transformation logarithmique (la théorie dit que plus la surface grandit, moins la valeur ajoutée d'un mètre carré supplémentaire sera importante) a été testée et elle ne permet pas une amélioration de la distribution des résidus.

Il faut donc passer à l'étape suivante qui est soit l'utilisation des erreurs standards robustes soit l'ajustement par MCG. Le logiciel utilisé, Gretl, permet de réaliser cette régression en trois étapes :

1. Une estimation du modèle par MCO
2. Une régression pour estimer la variance des erreurs. Le logiciel régresse le logarithme (pour s'assurer que la variance estimée n'est pas négative) des carrés des résidus en fonction des variables explicatives originelles et de leur carré. Les coefficients de cette régression sont nommés  $u^*$
3. La méthode des moindres carrés pondérés utilisant l'inverse la variance estimée ( $1/\exp(u^*)$ )

Nous réalisons donc un modèle d'hétéroscédasticité corrigé (548 observations). Le modèle 3 est caractérisé par les paramètres statistiques du tableau 7.

**Tableau 7 : Indicateurs globaux de la qualité d'ajustement du modèle 3**

Somme carrés résidus	2083,433	Éc. type de régression	1,986428
R2	0,812130	R2 ajusté	0,805370
F(19, 528)	120,1293	p. critique (F)	1,3e-177
Log de vraisemblance	-1143,504	Critère d'Akaike	2327,009
Critère de Schwarz	2413,134	Hannan-Quinn	2360,671

#### 5.3.1. Analyse des résidus excédentaires

Si des observations ont des résidus trop importants (2,5 fois l'écart-type des résidus), nous décidons de supprimer lors de l'ajustement du modèle car elles biaisent notre régression.

Nous supprimons 18 observations. La grande partie de ces observations correspondre à des biens qui avaient un prix supérieur à 350.000 €. Ces observations sont marginales par

rapport à notre échantillon et le mauvais ajustement du modèle pour celles-ci est dû au fait que ces biens sont très différents par rapport au reste de l'échantillon et qu'ils se situent dans un sous-marché distinct.

## 5.4. Modèle 4 : suppression des observations marginales

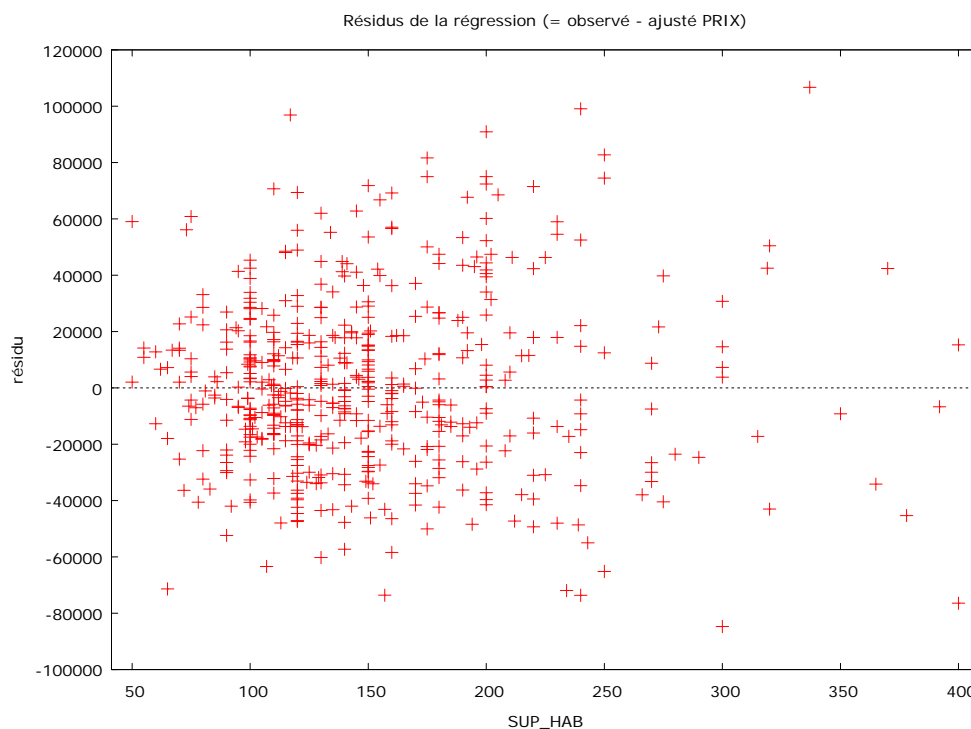
Après cette suppression, nous ajustons notre modèle 4, similaire au modèle précédent. Le dernier modèle est mieux ajusté que le précédent. Les critères d'Akaike, de Schwarz ainsi que l'indice Hannan-Quinn de sont moins importants et le modèle explique 88 % de la variance du prix avec une p-value très petite.

**Tableau 8 : Indicateurs globaux de la qualité d'ajustement du modèle 4**

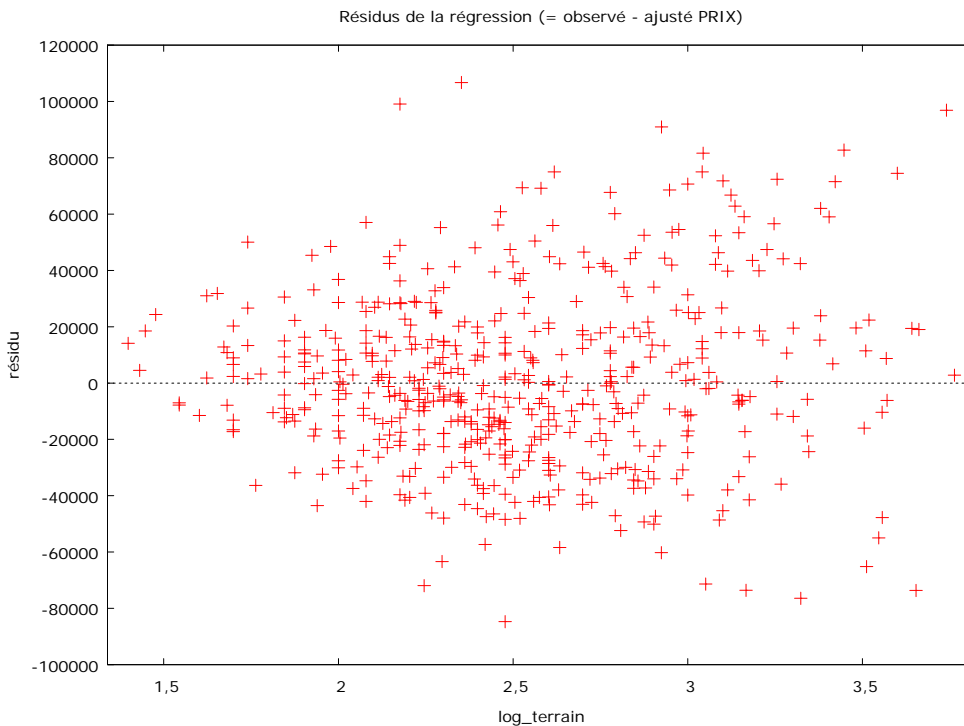
Somme carrés résidus	1816,612	Éc. type de régression	1,892898
R2	0,879950	R2 ajusté	0,874740
F(22, 507)	168,9195	p. critique (F)	4,3e-217
Log de vraisemblance	-1078,478	Critère d'Akaike	2202,956
Critère de Schwarz	2301,232	Hannan-Quinn	2241,423

Il est important de signaler que nous ne sommes toujours pas en présence de colinéarité puisque nous utilisons les mêmes variables explicatives.

**Graphique 9 : Résidus du modèle 4 en fonction de la superficie habitable**



Graphique 10 : Résidus du modèle 4 en fonction du logarithme de la superficie du terrain

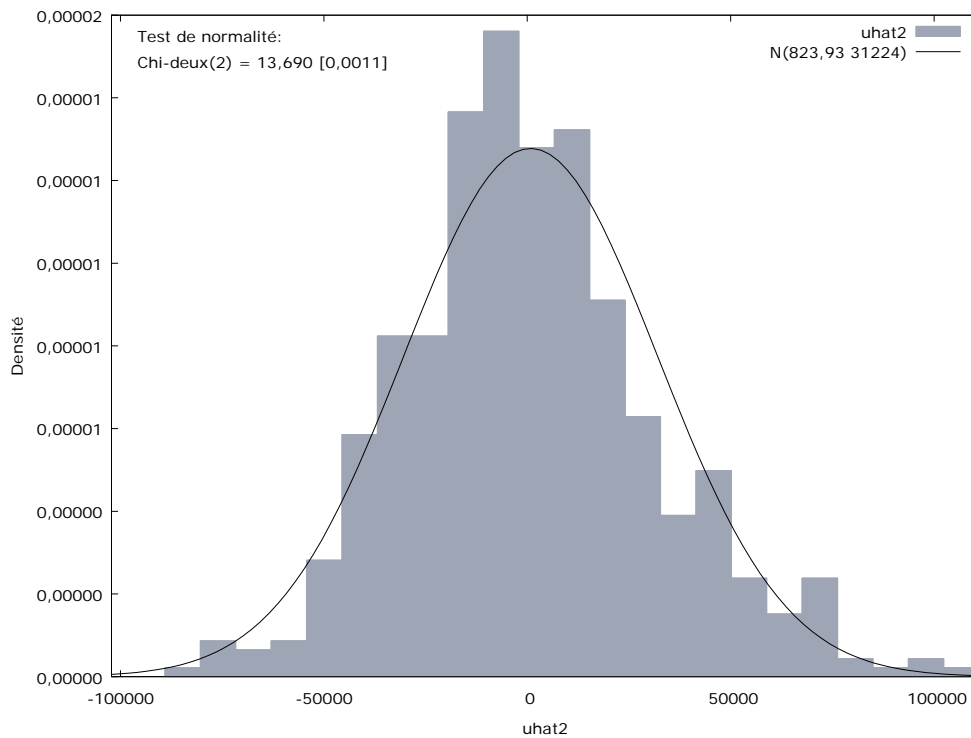


Nous constatons sur les graphiques 9 et 10 une plus grande variabilité des résidus pour les caractéristiques se situant dans la moyenne. Sur ces graphiques, la distribution des observations est moins conique que sur les graphiques 7 et 8.

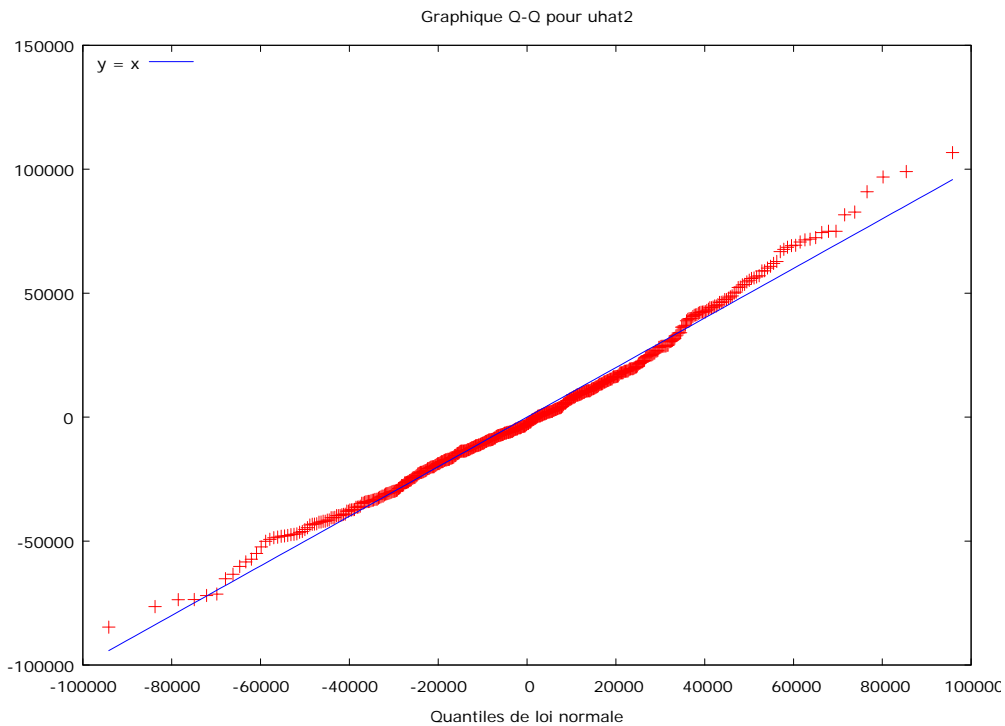
Si nous analysons les résidus de manière plus générale, la moyenne en valeur absolue est de 23.642 € avec un écart type de 19.368 €. Cette erreur est importante mais si nous comparons ce chiffre au prix moyen de notre échantillon (183.801 €) nous constatons que cela représente 12,86 % du prix. Ce pourcentage est légèrement supérieur à la barre des 10 % (Evans ,1995).

Sur l'histogramme des résidus (graphique 11), nous pouvons voir que les résidus suivent approximativement une distribution normale. Gretl réalise un test de normalité des résidus pour lequel on peut rejeter l'hypothèse nulle ( $p$  critique observé(0,0011) est inférieur au  $p$  critique défini (0,05)). Donc l'erreur n'est significativement pas distribuée comme une loi normale. On peut observer que nous avons un nombre important de résidus élevés. Pour améliorer le modèle, on pourrait penser à supprimer ces observations avec des résidus extrêmes. Nous l'avons déjà réalisé une fois et il n'est pas conseillé de refaire cette opération car on diminuerait trop notre échantillon (Des Rosiers, 2001).

Graphique 11 : Distribution des résidus du modèle 4



Graphique 12 : Diagramme quantile

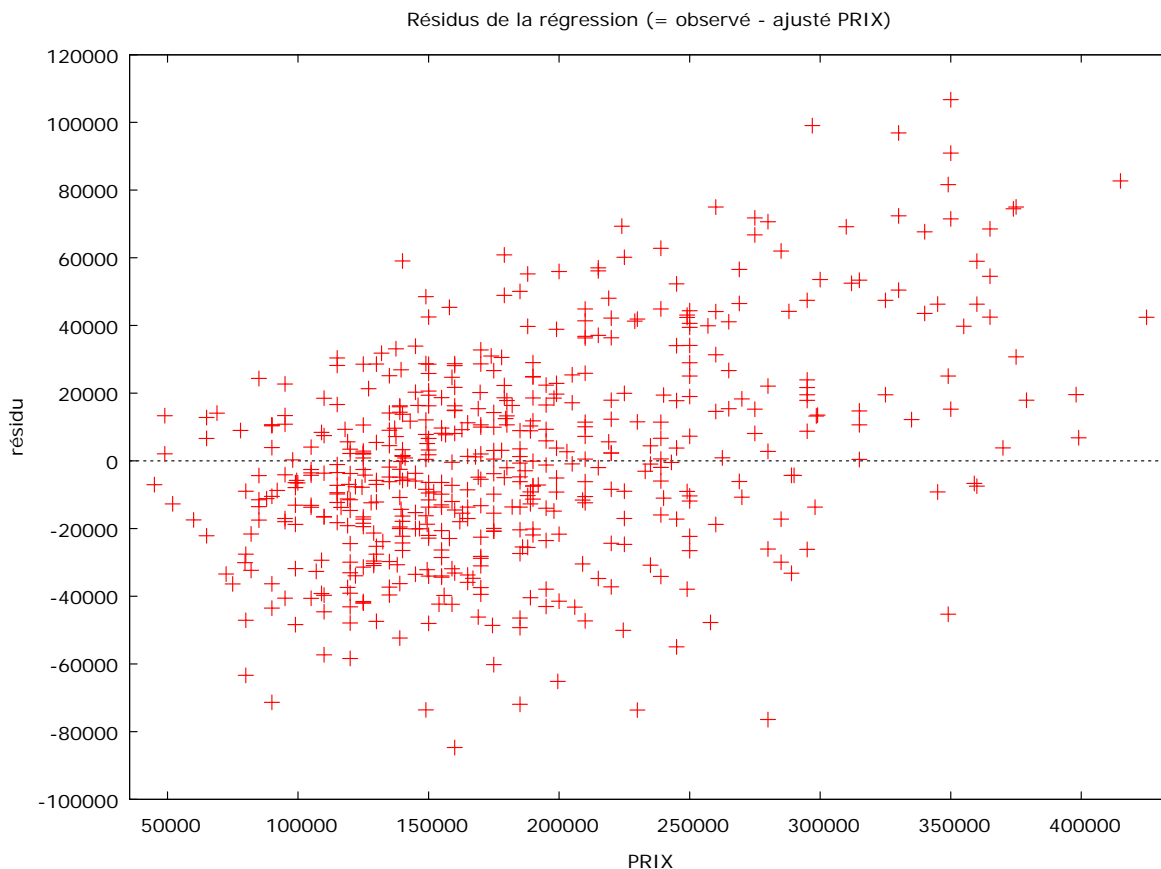


Une autre manière de vérifier la normalité des résidus est de regarder le graphique du QQ-plot aussi appelé diagramme quantile-quantile. Ce graphique représente en abscisse les

quantiles correspondant à la loi théorique et en ordonnée ceux observés. Nos résidus suivent bien la droite de Henry mais s'en écartent aux deux extrémités. On observe que pour les résidus très négatifs, ceux-ci sont moins négatifs que prévu. Tandis que les résidus très positifs sont encore plus importants.

Enfin, si l'on étudie les résidus en fonction du prix de vente (graphique 13), nous observons que les résidus sont plus importants et positifs en moyenne pour les biens supérieurs à 250.000 €. Notre modèle sous-évalue donc ces maisons. Premièrement, parce qu'un modèle hédonique n'est applicable qu'à un seul marché et que dans ce cas nous avons affaire à plusieurs sous-marchés vu l'étendue des prix. Ensuite, nous n'avons peut-être pas pris en considération certaines caractéristiques qui apportent une plus-value importante.

**Graphique 13 : Distribution des résidus par rapport aux prix**



Nous avons dès lors envisagé de diminuer notre échantillon aux biens immobiliers à prix modéré et de réaliser un modèle pour les biens immobiliers inférieurs à 300.000 €, 250.000 € et 200.000 € (annexe 6). Le problème de ces modèles est que nous allons réduire fortement notre échantillon. Les 3 modèles sont moins bien déterminés donc nous n'allons pas les utiliser. Il est quand même intéressant de constater que les variables explicatives sont différentes entre les modèles.

### 5.4.1. Autocorrélation spatiale des résidus

Un dernier test important pour la validation de notre modèle hédonique est la vérification d'absence d'autocorrélation spatiale des résidus. Nous avons réalisé le même test qu'au point 2.3.2. Les résultats sont les suivants :

```
Global Moran's I Summary
Moran's Index:    0,010192
Expected Index:  -0,001828
Variance:        0,000260
Z Score:         0,745001
p-value:         0,456271
```

L'indice de Moran est proche de zéro et en plus le test ne permet pas de rejeter l'hypothèse nulle. Donc, il n'y a pas d'autocorrélation spatiale.

## 5.5. Modèle final

Notre quatrième modèle (MCG, Hétéroscédasticité corrigée, utilisant les observations 1-530, Variable dépendante: PRIX) est celui que nous avons décidé de choisir comme modèle représentant le marché immobilier liégeois des maisons unifamiliales d'un standing normal. Le tableau 10 reprend les différents coefficients estimés pour chaque variable, son erreur standard d'estimation, un test de Student (test de conformité de chaque coefficient) ainsi qu'une p-value associée à la valeur du t de Student.

**Tableau 9 : Statistiques basées sur les données initiales du modèle 4**

Moy. var. dép.	183801,4	Éc. type var. dép.	72659,67
Somme carrés résidus	4,95e+11	Éc. type de régression	31235,67

Tableau 10 : Coefficients du modèle 4

	<i>Coefficient</i>	<i>Erreur Std</i>	<i>t de Student</i>	<i>p. critique</i>	
Constante	130300	17893,6	7,2819	<0,00001	***
A rénover	-20989,5	3921,68	-5,3522	<0,00001	***
Remis à neuf	6491,65	3851,73	1,6854	0,09253	*
Superficie terrain (log m <sup>2</sup> )	25772,9	3913,86	6,5850	<0,00001	***
Superficie habitable (m <sup>2</sup> )	434,176	28,5741	15,1947	<0,00001	***
4 façades	35121,2	4167,24	8,4279	<0,00001	***
0 étage	-22036,6	5354,17	-4,1158	0,00005	***
0 salle de bain	-8468,76	4104,45	-2,0633	0,03959	**
2 salles de bain	10703,8	4104,44	2,6079	0,00938	***
1 chambre	-32443,3	6575,93	-4,9337	<0,00001	***
2 chambres	-6313,53	3000,7	-2,1040	0,03587	**
5 chambres ou plus	22935,5	6216,29	3,6896	0,00025	***
Cuisine équipée	12469,8	2742,34	4,5471	<0,00001	***
Présence d'une terrasse	8744,64	2390,95	3,6574	0,00028	***
Présence d'une piscine	48944,3	11627,9	4,2092	0,00003	***
Age (log année)	-18911,6	1493,94	-12,6589	<0,00001	***
Moyenne des revenus médians	3,5563	0,360516	9,8645	<0,00001	***
Densité (habitant/km <sup>2</sup> )	-0,800927	0,397391	-2,0155	0,04438	**
Distance au centre Liège (min)	-1165,13	223,431	-5,2147	<0,00001	***
Part modale des modes lents	1165,65	402,842	2,8936	0,00397	***
Présence d'un arrêt de bus à 700 m	-44247,2	4847,82	-9,1272	<0,00001	***
Présence d'industrie à 200m	-9621,47	2630,82	-3,6572	0,00028	***
Diversité dans un rayon de 100m	6891,42	3019,74	2,2821	0,02290	**

\*p-value inférieur à 10 %

\*\*p-value inférieur à 5 %

\*\*\*p-value inférieur à 1 %

## 6. ANALYSE DES RÉSULTATS

Sur l'ensemble de nos variables, seules 22 variables sont significatives à un seuil de 10 %.

Une grande part des variables structurelles reste significative. La plupart de celles-ci sont binaires et doivent être analysées par rapport au bien de référence choisi. En effet, nous n'avons pas sélectionné à priori les variables suivantes : la présence de 2 façades, de 1 salle de bain, de 1 étage et de 3 chambres, un bien en état normal et l'absence de cuisine afin d'éviter des problèmes de colinéarité. L'ensemble des coefficients doivent être interprétés en tenant compte de ces caractéristiques. Ainsi, la présence de 2 salles de bains apporte une plus-value de 10.704 €. Donc pour un bien immobilier passant de 1 à 2 salles de bains, la salle de bain supplémentaire coûte 10.700 €.

Parmi les variables absentes de notre modèle, on peut souligner celles-ci : la présence d'un garage, d'un grenier, d'une buanderie, d'une véranda ne semble pas apporter de plus-value au bien immobilier. Grignet (2011) l'avait déjà remarqué dans son modèle. On ne peut que confirmer que ces caractéristiques ne semblent pas importantes dans le choix d'une maison et de ce fait, elles n'influencent pas le prix.

Sur l'état de rénovation du bien, les résultats sont du même ordre de grandeur que ceux de Grignet (2011). Ces variables sont aussi dépendant des variables sur l'équipement de la maison. Nous avons fait un test d'indépendance ( $X^2=77,15$  ;  $dl=6$  ;  $p=0,00$ ) qui montre que ces variables sont dépendantes. Donc souvent les maisons qui sont dans un bon état ou rénovées ont une cuisine équipée donc elles ont une plus-value encore plus importante.

Les maisons 4 façades ont une plus-value de 35.121 €. Ce montant confirme l'envie des Belges d'une maison individuelle 4 façades (3.2.3.1) et donc la plus-value. Par contre, la maison 3 façades ne ressort pas dans le modèle. Elle n'est donc pas considérée comme étant une plus ou une moins-value.

Si la maison possède une ou deux chambres, elle subit une moins-value tandis que si elle est constituée de plus de 5 chambres, une plus-value. Donc les personnes sont à la recherche de maisons avec de nombreuses pièces, ce qui est donc directement corrélé à la superficie. La corrélation entre le nombre de chambres et la superficie habitable est de 0,58. De même, si le bien est de plain-pied, il perd de sa valeur car il est souvent moins grand. Avoir plus de 2 étages n'influence pas le prix d'une maison.



## Partie I : La méthode hédonique

L'analyse du coefficient de la variable « superficie de terrain » nécessite une transformation mathématique pour pouvoir l'appréhender car c'est le logarithme de cette variable qu'est introduit dans le modèle. Nous pouvons déduire le prix du terrain en sachant que :

$$P_{terrain} = Coefficient * \log_{10} Superficie$$

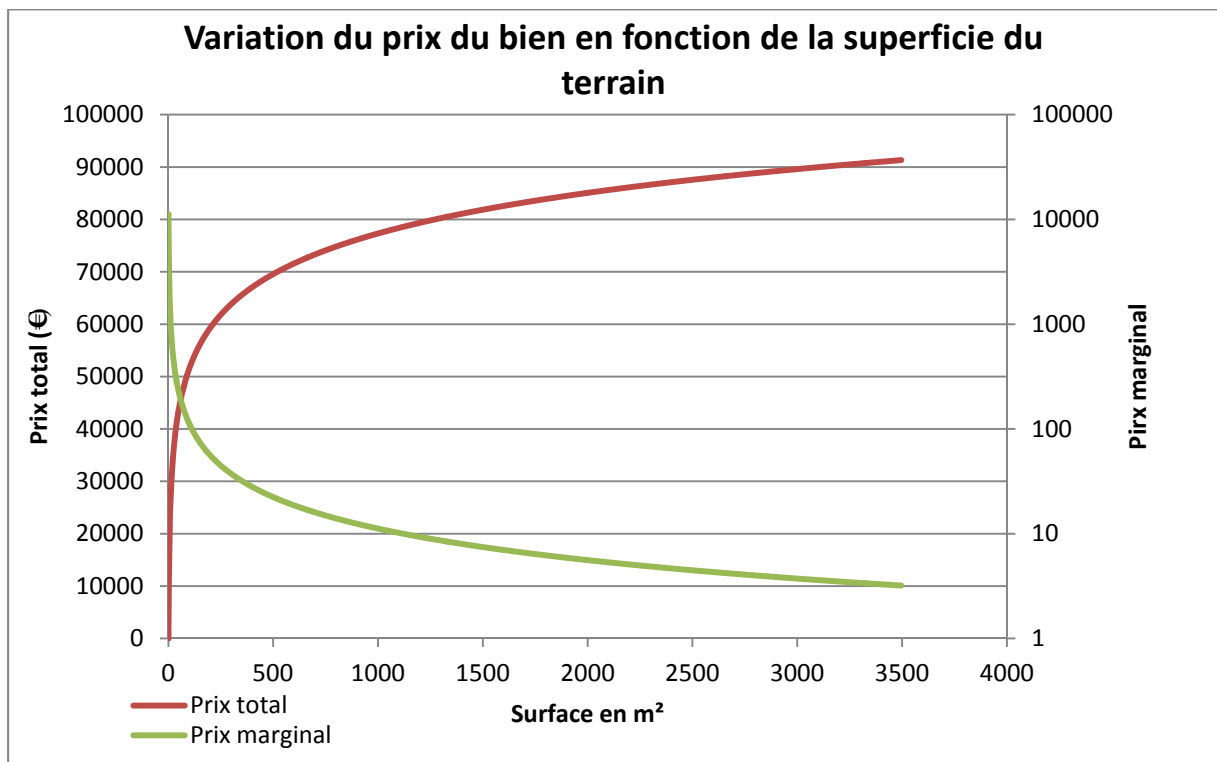
Le prix pour une unité de surface supplémentaire, le prix marginal, se calcule par dérivée partielle du prix par rapport à la superficie :

$$\frac{\partial P_{terrain}}{\partial Superficie} = Coefficient * \frac{\partial \log_{10} Superficie}{\partial Superficie}$$

$$\text{Sachant que } \log'_a(x) = \frac{1}{x * \ln(a)}$$

$$\text{Alors } \frac{\partial P_{terrain}}{\partial Superficie} = Coefficient * \frac{1}{Superficie * \ln(10)}$$

Graphique 14 : Variation du prix du bien en fonction de la superficie du terrain



Grâce à cette formule, nous avons pu représenter graphiquement le prix marginal du terrain. Le graphique 14 représente d'une part le prix total du terrain qui ne cesse d'augmenter avec la superficie (mais qui se stabilise à partir de 1000 m²) et d'autre part le prix marginal qui ne cesse de diminuer. Grâce à ce graphique, on peut comprendre l'intérêt de la transformation de la variable en logarithme. Elle permet de prendre en compte le concept d'utilité marginale

décroissante : la satisfaction de l'acheteur pour une unité supplémentaire décroît plus s'il possède une quantité initiale importante.

Le prix marginal est de l'ordre de 70 €/m<sup>2</sup> pour des surfaces autour de la moyenne (150 m<sup>2</sup>). Si la superficie du terrain augmente, le prix marginal n'est plus que 22 €/m<sup>2</sup> à 500 m<sup>2</sup> et atteint même 10 €/m<sup>2</sup> pour les grands terrains de 1000 m<sup>2</sup>. D'autres analyses sur le prix marginales ont été réalisées (annexe 7).

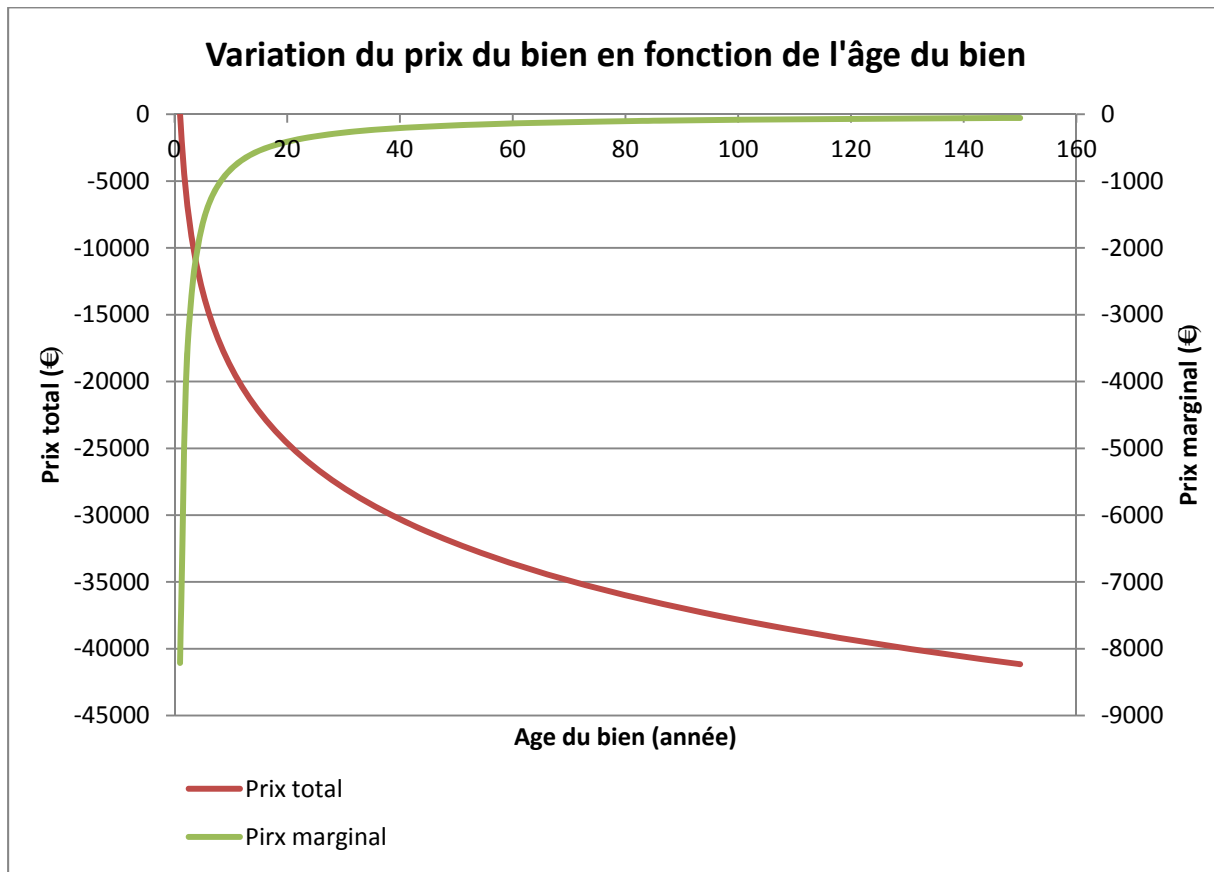
Pour la superficie habitable, nous n'avions pas pris en compte cette variation du prix marginal décroissant. Le prix est constant et est de 434 €/m<sup>2</sup>. Si l'on compare ces valeurs avec celles obtenues par Grignet (2011) (20 €/m<sup>2</sup> pour le terrain et 417 €/m<sup>2</sup> pour l'habitable), les nôtres sont légèrement supérieures mais ne rencontrent pas le coût de production du mètre carré habitable dans le cas d'une construction neuve. Ce coût était de 1.303 €/m<sup>2</sup> en 2004 (CPDT - Thème 3.1 - CAT du 21 septembre 2004) que Grignet (2011) a indexé pour avoir un prix de 1.515 €/m<sup>2</sup> en 2010. Nous expliquons la grande différence de prix par le fait que le coût de production du mètre carré habitable tient compte de l'aménagement des pièces et que, pour nous, cet aménagement est pris en compte par d'autres variables telles que l'état du bien, l'équipement de la cuisine, la présence d'une salle de bain...

L'autre variable qui avait subi une transformation logarithmique était l'âge du bien. Pour calculer le prix marginal, le prix supplémentaire pour avoir un bien d'un an de plus, il faut aussi effectuer la dérivée partielle de la fonction :

$$\frac{\partial P_{age}}{\partial Age} = * \frac{Coefficient}{Age * \ln(10)}$$

Dans ce cas-ci, le bien ne retire pas une plus-value quand la variable augmente mais une moins-value. C'est lors des 20 premières années que le bien perd le plus. Ensuite, la diminution du prix ralentit. Les moins values sont importantes autour de la médiane (80 ans) : pour chaque année supplémentaire, le bien perd encore 100 €/an. Grignet (2011) avait pu mettre en avant que lors des premières années, le prix ne diminuait pas mais qu'au-delà de 15 ans, le bien perdait énormément.

Graphique 15 : Variation du prix du bien en fonction de l'âge du bien



Pour terminer avec les caractéristiques intrinsèques, le prix pour une piscine est moitié moindre que celui obtenu par Grignet (2011). La présence d'une piscine est une plus-value de 48.944,3 €. Ce coût reste supérieur au coût d'une vraie piscine. Selon le site Packtravaux (<http://www.packtravaux.com>), le prix de la construction d'une piscine peut varier de 10.000 € à 40.000 € et même plus. Le prix change énormément selon le type de piscine construite (piscine en kit, en coque, maçonnée...). 49.000 € n'est donc pas un prix extravagant mais il est supérieur à la moyenne des coûts de construction des piscines de l'échantillon. Cette surestimation est due au coût de coordination des travaux que les propriétaires ont dû supporter pour construire la piscine mais aussi au fait que le modèle est moins bien adapté aux maisons de « standing ». Une manière de capitaliser les prix importants des maisons de « standing » est d'augmenter les prix de leurs caractéristiques propres dont fait partie la piscine. En surestimant le prix des piscines, le modèle permet un meilleur ajustement du prix des maisons de « standing ». Les terrasses constituent aussi une plus-value de l'ordre de 8.700 €, ce montant nous semble là réaliste pour une terrasse dont les coûts sont aussi très variables.

Concernant les caractéristiques du secteur statistique, le revenu médian et la densité ont été retenus dans le modèle. Pour chaque euro supplémentaire de revenu médian, le prix d'une

maison augmentera de 3,56 €. Grignet (2011) avait montré que cette valeur était cohérente. Si l'on considère un taux d'imposition de 50 %, l'augmentation du revenu net sera dès lors de 0,5 €. En moyenne, un acheteur emprunte sur 20 ans et n'emprunte pas au-delà d'un tiers de son salaire. Donc avec une augmentation de 1 € du salaire brut, il peut acheter un bien de 3,33 € plus cher ( $1 \times 1/2 \times 1/3 \times 20 = 3,33$ ).

La densité du quartier constitue une moins-value. Le secteur le moins dense pour lequel nous avons une observation est Mehagne (près d'Embourg) où il y a 13 habitants/km<sup>2</sup> tandis que le plus dense, Bressoux, a plus de 16.500 habitants/km<sup>2</sup>. La moins-value qui pouvait sembler faible, puisque qu'elle est inférieure à 1 € provoque de grandes différences de prix puisque la densité est fort variable. Entre les deux extrêmes, nous avons une différence de 13.263 €. Les espaces moins denses sont recherchés car ils permettent l'achat de plus grands terrains et de maisons isolées. Il y a d'ailleurs une faible corrélation entre ces variables.

Pour les variables relatives à l'accessibilité, plusieurs ressortent significativement, notamment la distance à Liège. Un gain de 1 minute au centre produit une plus-value de 1.165 €. Cela entraîne une différence de l'ordre de 46.000 € entre une maison située au centre de Liège et une autre située en périphérie de la région urbaine. Cela représente de l'ordre de 25 % dans le prix d'une maison de 180.000 € (moyenne de l'échantillon). Il faut aussi se rappeler que cette variable a été choisie parmi d'autres. L'accessibilité à l'emploi et aux nodules commerciaux étaient fortement corrélées à cette variable. Au travers de la variable « distance à Liège », ces variables sont aussi prises en compte.

D'autres variables d'accessibilité ont un impact sur le prix du terrain : la part modale des modes lents apporte une plus-value de 1.166 €. Même si ce coefficient est similaire à celui de la distance à Liège, il entraîne une moins grande variabilité des prix car la variance de cette variable est plus faible. Cela peut aller jusqu'à 26.066 €. Pour la présence d'un arrêt de bus à moins de 700 mètres, le bien a une moins-value de 44.247 €. Cette caractéristique est donc considérée comme une nuisance. Il semble pourtant difficile d'imaginer une telle portée pour un arrêt de bus. Dans l'étude de Cervero & Kang (2011) qui étudient l'impact de l'installation d'un bus à haute vitesse à Seoul, ils identifient la zone de nuisances (bruits, vibrations...) un rayon de 100m autour des arrêts. Dans notre travail, cet effet négatif provient sûrement du fait que les arrêts de bus sont positionnés dans des milieux denses ou à proximité de voies rapides et ce sont ces caractéristiques qui entraînent probablement une moins-value.

Par contre, la part modale du bus et du train n'est pas significative. Le train est un moyen de transport nettement moins utilisé à Liège que les autres modes. « *Il faut aussi constater*

*l'usage limité du train par les travailleurs du nord-ouest du pays, la Campine anversoise, le Limbourg et la province de Liège. Sans doute le train n'y représente-t-il pas encore une réelle alternative* » (Direction Mobilité, 2010, p. 13). Pour une grande partie des gens, cela ne figure pas dans leurs critères de sélection d'une maison. Pour le bus, la part modale est un peu plus importante de manière générale mais cela n'influence pas avec certitude le choix d'une maison. La grande majorité des maisons est située à proximité d'arrêts de bus. Dans notre échantillon, seulement 1,5 % des biens étaient situés à plus de 700 mètres d'un arrêt.

Dans nos variables mesurant les nuisances, seule la proximité à une industrie est significative. La nuisance est de courte portée puisque parmi les différents buffers construits, un seul rayon a été retenu, celui de 200 mètres autour des sites industriels. La moins-value est portée à 9.621 €, ce qui équivaut à une baisse de 5 % pour une maison valant 180.000 €. Les sites d'extraction ne semblent pas affecter les prix des biens immobiliers. Nous pensons que cela est expliqué par le manque de biens se situant à proximité immédiate (200 mètres) de tels sites. Seuls 6 biens sur 548 sont dans cette zone. Ils ne sont peut-être pas assez représentatifs des préjudices causés par les carrières. La proximité des sites d'activité économique désaffectés, SAED, n'est pas aussi ressortie dans notre modèle. Les nuisances liés aux SAED sont peut-être très limitées et ne sont répercutées que pour les maisons voisines. Enfin la variable sur la proximité des grandes routes n'est pas non plus significative. Cela peut-être dû au fait que nous n'avons pas pris uniquement les maisons situées sur une grande route mais à proximité de celle-ci. De plus les routes sélectionnées sont celles où la vitesse autorisée excède les 50 km/heure et ce critère ne prend pas en compte l'importance du trafic routier générateur de nuisances.

Pour l'environnement de proximité, contrairement à l'étude de Geoghegan et al. (1997), notre modèle explique qu'une diversité croissante est une plus-value pour un bien immobilier. Le fait d'obtenir une estimation contraire à ce qui est normalement attendu est expliqué par le fait que, dans notre cas, une faible diversité est liée à l'urbanisation et que, comme pour les bus, l'impact est négatif.

Au final, nous n'avons pas énormément de variables environnementales explicatives du prix de l'immobilier significative. Certaines d'entre elles apportent une valeur contre-intuitive (accessibilité au mode lent, arrêt de bus à 700 mètres et diversité dans un rayon de 100 mètres). On peut y voir trois raisons :

- Les caractéristiques environnementales jouent à la marge du prix. Elles n'expliquent que quelques pourcents du prix immobilier qui est principalement dû aux

caractéristiques intrinsèques. Vu leur faible influence, il est plus difficile de le montrer statistiquement.

- Certaines caractéristiques environnementales sont légèrement corrélées avec des variables intrinsèques notamment avec les 4 façades et la superficie du terrain. Les maisons situées dans un milieu peu bâti sont généralement des maisons 4 façades avec de grands terrains. Une partie des caractéristiques est donc prise en compte par les variables structurelles.
- Enfin, rappelons que notre échantillon est de faible taille. Dans la majorité des études, les échantillons dépassent le millier d'observations.

## 7. COMPARAISON AVEC LE MODÈLE DE GRIGNET (2011)

Il nous semblait important de faire une comparaison plus détaillée de nos résultats avec ceux de Grignet (2011). Nous avons réalisé un modèle hédonique à partir des mêmes données et souhaitons améliorer les résultats obtenus par Grignet. Les deux grands axes que nous avons essayé d'améliorer sont : les caractéristiques d'accessibilité et l'analyse des résidus (notamment la prise en compte de l'hétéroscédasticité).

Le tableau 11 donne une idée de grandeur entre les coefficients de régression obtenus par Grignet (2011) et ceux obtenus dans ce mémoire. Le rapport est obtenu en divisant la valeur de Grignet (2011) par la valeur de Dethier (2012). Cinq caractéristiques ont un coefficient fort proche (différence inférieure à 10 %). Il y a des caractéristiques structurelles qui pour nous présentent une certaine robustesse et des caractéristiques extrinsèques : la moyenne des revenus médians et la distance-temps au centre de Liège. La moyenne des revenus médians est dans les deux cas proche de 3,3, valeur que nous avons estimée comme réaliste. Par contre, malgré de nombreuses variables sur l'accessibilité (accessibilité à l'emploi, aux pôles commerciaux, au centre de Liège, aux gares, aux transports en commun), nous obtenons des résultats fort identiques qui sont contre-intuitifs. De ce point de vue, il nous semble que l'amélioration que nous avons voulu apporter n'a pas été rencontrée.

Tableau 11 : Comparaison entre le modèle de Grignet et de Dethier

	<i>Grignet</i>	<i>Dethier</i>	<i>Rapport</i>
Constante	61515,4	130300	0,47
A rénover	-21832,8	-20989,5	1,04
Remis à neuf	13300,30	6491,65	2,05

Partie I : La méthode hédonique

Superficie habitable (m <sup>2</sup> )	416,8	434,176	0,96
4 façades	36755,6	35121,2	1,05
0 étage	-13317,8	-22036,6	0,60
0 salle de bain	-12989,2	-8468,76	1,53
2 salles de bain	26357,3	10703,8	2,46
1 chambre	-25136	-32443,3	0,77
2 chambres	-9045,5	-6313,53	1,43
5 chambres ou plus	15045,7	22935,5	0,66
Cuisine équipée	16336,7	12469,8	1,31
Présence d'une terrasse	15253,5	8744,64	1,74
Présence d'une piscine	93118,9	48944,3	1,90
Moyenne des revenus médians	3,2	3,5563	0,90
Distance au centre Liège (min)	-1261,4	-1165,13	1,08
Part modale des modes lents	1359,5	1165,65	1,17

Différence inférieur à 10 %

Par contre la prise en compte de l'hétéroscédasticité nous semble avoir permis une certaine amélioration du modèle. Comme la régression par MCG que nous avons réalisée, prend moins en compte les observations extrêmes ayant moins d'informations, l'influence des maisons de « standing » est moins importante. Ainsi les caractéristiques suivantes : présence d'une piscine, présence d'une terrasse, présence de 2 salles de bains ont maintenant des prix de régression plus cohérents avec leur valeur. Ces caractéristiques plus fréquemment rencontrées dans les maisons chères ont des valeurs plus raisonnables.

## **8. CONCLUSION ET PERSPECTIVES**

La comparaison du modèle de Grignet (2011) avec le nôtre a montré une certaine robustesse pour les variables structurelles : nous avons obtenu, malgré des méthodes différentes, des valeurs similaires. Elles pourront donc être confrontées avec les valeurs obtenues par l'approche par contingence pour servir de contrôle.

Par contre, parmi nos variables extrinsèques, le modèle ne semble toujours pas être optimal. Nous avons des valeurs cohérentes pour la moyenne de revenu médian, la densité de la population du quartier et la proximité d'industries. Mais pour les autres variables significatives, ce n'est pas le cas : des variables importantes (notamment la proximité des espaces verts, le cadre bâti ou la proximité de grandes routes) ne sont pas ressorties comme explicatives. En termes d'accessibilité, notre modèle ne s'est pas amélioré par rapport à celui de Grignet (2011). L'approche par contingence sera l'occasion d'approfondir ces faiblesses.

Enfin notre travail a permis de tenir compte de l'hétéroscédasticité, aspect qui avait été peu étudié.



**Partie 2: Les préférences  
déclarées par l'approche par  
contingence**

## **INTRODUCTION**

La deuxième partie de ce mémoire est consacrée à la mise en application de la méthode d'évaluation par contingence dont les résultats doivent être confrontés à notre modèle hédonique. Des enquêtes ont été réalisées en vue de connaître le consentement à payer de la population pour différentes caractéristiques d'un bien immobilier. Ces enquêtes nous ont aussi servi à explorer des caractéristiques qui ne sont pas significatives dans notre modèle hédonique.

Le premier chapitre de cette partie expose les différents choix méthodologiques effectués pour la réalisation des enquêtes. Nous y définissons notre population, le type d'enquête, la méthode d'évaluation et les grandes lignes de notre questionnaire.

Le chapitre 2 est consacré à l'explication des différents biais présents lors d'une enquête et plus précisément lors de ce type d'enquête. Il nous semblait primordial d'y consacrer un chapitre afin d'exposer notre méthodologie visant à réduire ces biais.

Dans le chapitre 3, nous présenterons le questionnaire. Les questions servant à déterminer le consentement à payer des personnes seront exposées et nous expliquerons le but de chacune d'elles. Certaines questions servent en effet de validation vis-à-vis du modèle hédonique et d'autres sont plus exploratoires.

Une description des communes où a lieu l'enquête sera présentée au chapitre 4. Nous y présenterons les communes de Liège, Juprelle et Remicourt.

Au chapitre 5, nous étudierons les 80 enquêtes réalisées. Dans un premier temps, nous analyserons les résultats portant sur la population interrogée et sur le logement. Ensuite, nous effectuerons différents tests statistiques sur les données fournies par les questions d'évaluation immobilière. Nous passerons en revue les 8 situations qui ont été présentées dans notre questionnaire.

Le chapitre 6 clôturera cette partie par une discussion des résultats obtenus.

# 1. PRÉSENTATION DE LA MÉTHODOLOGIE

La mise en application de la méthode d'évaluation des préférences déclarées demande au préalable de définir 4 notions :

- La population interrogée. Il faut définir la population ciblée mais aussi l'échantillon choisi pour prendre en considération cette population.
- Le type d'entretien réalisé. Il faut sélectionner parmi les nombreuses méthodes d'enquêtes existantes : face à face, par téléphone, par mail, par courrier. Chacune présentent des avantages et des inconvénients. De plus, nous devons faire un choix entre une enquête directive, semi-directive ou non directive
- La méthode d'évaluation du consentement à payer, CAP. Il existe différentes manières d'aider le répondant à choisir le montant le plus proche de son CAP. Nous les présenterons et expliquerons le choix que nous avons opéré.
- Le questionnaire. Afin que la personne enquêtée comprenne au mieux la tâche qui lui est demandée, le questionnaire doit présenter certaines caractéristiques. Nous expliquerons les outils mis en œuvre pour faciliter la compréhension de l'interviewé.

## 1.1. Echantillonnage

Comme pour toute enquête, il faut effectuer un échantillonnage de la population interrogée. Il faut donc définir cette population et aussi déterminer la taille de l'échantillon.

Dans notre cas, nous devons interroger une population qui agit sur le marché immobilier de la région liégeoise. Cependant, pour faciliter notre enquête, nous allons limiter cette population à une population qui a été active récemment sur le marché et à la population de certaines communes.

Nous limitons notre population à des individus qui connaissent le marché immobilier. Cette sélection est faite pour éviter d'interroger des personnes ignorant les prix pratiqués sur ce marché et fausser ainsi les résultats de l'enquête. Pour identifier ces individus, nous décidons de sélectionner les personnes qui ont construit récemment. Ainsi nous avons demandé à diverses communes<sup>4</sup> de la région liégeoise une liste des permis d'urbanisme pour des constructions d'habitation ou de maison pour les années 2008, 2009 et 2010. Trois années ont été demandées pour obtenir un nombre suffisamment important de permis, de l'ordre d'une centaine pour les communes dites de banlieue.

---

<sup>4</sup> Une lettre a été envoyée à différentes communes (Liège, Bassenge, Juprelle, Awans, Crisnée, Remicourt, Fexhe-le-Haut Clocher et Donceel) pour demander cette liste. Cette lettre se trouve à l'annexe 8.

Nous avons décidé de limiter notre population à quelques communes : Liège, Juprelle et Remicourt<sup>5</sup>. De cette manière, nous réduisons les coûts de cette enquête en limitant les déplacements. Dans ces communes, nous avons une population d'une commune urbaine, les Liégeois, et une population de communes périurbaines, la population des communes de Juprelle et Remicourt. Nous gardons ainsi une certaine diversité de la population de la région liégeoise. Cette diversité pourra donner lieu à des études comparatives des résultats obtenus pour les deux groupes de populations.

Le deuxième point important concernant la population interrogée est la taille de l'échantillon. Elle est déterminée à la fois par la précision souhaitée de nos résultats mais aussi par le coût. Comme nous sommes dans le cadre d'un mémoire, nous limiterons notre échantillon à une petite taille de l'ordre de 80. Ce nombre est basé sur le mémoire de Strée J. (2010). 40 enquêtes ont été réalisées à Liège et 40 autres dans les communes périurbaines.

Dans la suite du travail, les termes urbain et périurbain seront employés pour nommer les personnes qui ont récemment construit dans ces régions. Nous ne voulons pas par ces noms appelés l'ensemble de la population de ces communes qui ont des caractéristiques socio-économiques et des avis sur l'évaluation de biens différents.

## **1.2. Le type d'enquête**

Dans le cadre d'une enquête, nous devons définir le type d'interaction et de vecteur (téléphone, mail, face à face...).

Pour l'évaluation par contingence, nous avons des questions ciblées. C'est la raison pour laquelle les enquêtes sont systématiquement directives.

Les différents vecteurs présentent chacun des avantages et des inconvénients. Certains ont de faibles coûts mais un taux de réponse faible ; d'autres ont des coûts plus importants mais un taux de réponse aussi plus élevé. Nous avons choisi de faire nos enquêtes face à face car ceci permet un taux de réponse important et aussi un meilleur moyen d'expliquer et de montrer la situation hypothétique que la population va devoir évaluer.

Pour augmenter notre taux de réponses, il nous a été conseillé (Loiseau) de déposer un courrier prévenant de notre passage. Une lettre a été déposée dans les maisons où nous souhaitions interroger les habitants. Les personnes interrogées semblaient beaucoup plus réceptives à notre demande parce qu'elles avaient été prévenues des aboutissements de l'enquête. Cette lettre se trouve en annexe 9.

---

<sup>5</sup> Sur l'ensemble des lettres envoyées, ont accepté de donner la liste : Liège, Juprelle, Remicourt, Fexhe-le-Haut Clocher (sous certaines conditions), Crisnée et Awans

### 1.3. La méthode d'évaluation du CAP

Il existe différents moyens de formuler l'enquête afin de permettre à l'enquêté d'estimer son CAP (Grelot 2004) :

- Méthode de l'enquête ouverte ;
- Méthode de la carte de paiement et de la liste d'intervalles ;
- Méthode d'enchères itératives d'offres à prendre ou à laisser.

L'enquête ouverte pose directement la question du montant maximal de son CAP à l'interviewé. La réponse à cette question ouverte n'est pas simple. C'est une tâche d'autant plus difficile que la personne enquêtée ne s'est jamais posé la question. Cette méthode implique le risque d'une mauvaise estimation du CAP. Elle est donc peu utilisée.

La deuxième méthode pose aussi directement la question du CAP mais propose soit une liste de valeurs, soit une liste d'intervalles de valeurs. Ces valeurs sont reprises dans un tableau et vont d'une valeur nulle à une valeur suffisamment élevée pour éviter que quelqu'un veuille proposer plus. Ces méthodes simplifient la tâche de l'interviewé puisqu'il doit choisir dans cette liste.

Pour la méthode des enchères, l'enquêteur propose un prix et l'interviewé répond s'il accepterait ce prix ou pas. Il doit effectuer un choix binaire (oui ou non). Cette question est répétée à plusieurs reprises avec des prix différents afin d'affiner le CAP. Cette méthode est recommandée (Grelot 2004). Elle simplifie la tâche de la personne enquêtée mais nécessite un grand échantillon (Grelot 2004). C'est pour cette raison que nous ne l'utiliserons pas mais préférons proposer une liste de valeurs.

Nous proposerons à la fin de chaque mise en situation, la question avec cette liste de valeurs présentée à la figure 4.

Figure 4 : Echelle de valeurs croissantes

A

Echelle de valeurs				
<p>Par rapport à la situation de référence, <b>combien accepteriez-vous de payer pour acquérir ce bien ?</b> Je vous demande le <b>montant maximum</b> que vous accepteriez de payer. Faites un choix parmi les propositions ci-dessous.                      En rouge, vous avez le prix identique à la situation de référence. Donc si vous choisissez cette ligne c'est que vous estimez que le nouveau bien doit se vendre au même prix que la situation de référence. Les valeurs en bleu indiquent des prix inférieurs à la maison de référence. Si vous choisissez cette couleur, c'est que vous estimez que le bien que je viens de vous présenter a moins de valeur que la maison de référence. Au contraire, les valeurs vertes sont supérieures au prix de référence. Si vous choisissez cette couleur, c'est que vous estimez que le bien que je viens de vous présenter a plus de valeur que la maison de référence.</p>				
1	- de 40 %	Cette maison vous coûtera moins de	150.000 €	(soit plus de 100.000 € en moins)
2	- 30 à - 40 %	Cette maison vous coûtera entre	150.000 € et 175.000 €	(soit de 75.000 € à 100.000 € en moins)
3	- 25 à - 30 %	Cette maison vous coûtera entre	175.000 € et 187.500 €	(soit de 62.500 € à 75.000 € en moins)
4	- 20 à - 25 %	Cette maison vous coûtera entre	187.500 € et 200.000 €	(soit de 50.000 € à 62.500 € en moins)
5	- 15 à - 20 %	Cette maison vous coûtera entre	200.000 € et 212.500 €	(soit de 37.500 € à 50.000 € en moins)
6	- 10 à - 15 %	Cette maison vous coûtera entre	212.500 € et 225.000 €	(soit de 25.000 € à 37.500 € en moins)
7	- 5 à - 10 %	Cette maison vous coûtera entre	225.000 € et 237.500 €	(soit de 12.500 € à 25.000 € en moins)
8	0 à - 5 %	Cette maison vous coûtera entre	237.500 € et 250.000 €	(soit jusqu'à 12.500 € en moins)
9	0 %	Cette maison vous coûtera	250.000 €	Situation de référence
10	+ 0 à + 5 %	Cette maison vous coûtera entre	250.000 € et 262.500 €	(soit jusqu'à 12.500 € en plus)
11	+ 5 à + 10 %	Cette maison vous coûtera entre	262.500 € et 275.000 €	(soit de 12.500 € à 25.000 € en plus)
12	+ 10 à + 15 %	Cette maison vous coûtera entre	275.000 € et 287.500 €	(soit de 25.000 € à 37.500 € en plus)
13	+ 15 à + 20 %	Cette maison vous coûtera entre	287.500 € et 300.000 €	(soit de 37.500 € à 50.000 € en plus)
14	+ 20 à + 25 %	Cette maison vous coûtera entre	300.000 € et 312.500 €	(soit de 50.000 € à 62.500 € en plus)
15	+ 25 à + 30 %	Cette maison vous coûtera entre	312.500 € et 325.000 €	(soit de 62.500 € à 75.000 € en plus)
16	+ 30 à + 40 %	Cette maison vous coûtera entre	325.000 € et 350.000 €	(soit de 75.000 € à 100.000 € en plus)
17	+ de 40 %	Cette maison vous coûtera plus de	350.000 €	(soit plus de 100.000 € en plus)

## 1.4. Le questionnaire

Le questionnaire doit toujours commencer par une introduction au projet de recherche et une explication claire de la situation hypothétique à évaluer. Les interviewés ne connaissent pas l'étude et ne savent pas la tâche qui les attend. Il faut donc veiller à ce qu'ils comprennent bien le principe de l'enquête, à savoir estimer leur CAP pour un bien. Ensuite, ils doivent bien comprendre ce qu'on leur demande d'évaluer.

Pour faciliter la compréhension des situations à évaluer, nous avons réalisé des planches illustratives (figure 5) qui reprennent une image du bien immobilier à évaluer mais aussi un descriptif de la situation ainsi que les mots clés. Nous avons essayé de réaliser les images les plus neutres possibles.

Figure 5 : Planches illustratives du questionnaire

### Situation de référence

Une maison neuve est à vendre. C'est une maison **4 façades** construite sur un terrain de **700 m<sup>2</sup>**. La maison a une surface habitable de **200 m<sup>2</sup>** avec **quatre chambres**. Elle se situe à **20 minutes** en voiture du centre de Liège dans un quartier neuf. Dans la rue, il y a peu de trafic routier et l'accessibilité en transport en commun n'est pas bonne. Cette maison coûte, tout frais compris, 250.000 €.

**Prix : 250.000 €**

A photograph showing a row of modern, two-story houses with light-colored facades and dark roofs. A red arrow points to the central house in the row.

### Situation 2

La maison, indiquée par la flèche rouge, est identique à la maison de référence. La différence est qu'elle est devenue une maison **2 façades**. De plus, elle est construite sur un terrain de **300 m<sup>2</sup>** et non plus de 700 m<sup>2</sup>.

A photograph of the same row of houses as in Situation 1. A red arrow points to the same central house. A text box in the bottom right corner of the image reads "2 façades 300 m<sup>2</sup> de terrain".

### Situation 3

La maison à vendre, indiquée par la flèche rouge, est, comme pour la situation précédente, de **2 façades** et construite sur un terrain de **300 m<sup>2</sup>**. Mais cette maison a une surface habitable de **150 m<sup>2</sup>** (à la place de 200 m<sup>2</sup>) et elle ne compte plus que **deux chambres** (à la place de quatre chambres).



### Situation 4

La maison est identique à la maison de référence à l'exception de sa localisation. La maison se situe à **5 minutes** du centre de Liège et non plus à 20 minutes.

**Centre de Liège à 5 minutes en voiture**





### Situation 5

La maison à vendre est identique à la maison de la situation de référence mais l'**accessibilité en transport en commun** est bonne (un passage de bus toutes les 15 minutes à proximité).

**Bonne accessibilité en transport en commun**



### Situation 6

La maison est identique à la maison de référence. La différence est le cadre dans lequel se situe la maison. Derrière la maison se trouve un **parc industriel**, un zoning.

**Parc industriel**



### Situation 7

La maison présentée sur l'image est identique à la maison de référence. La différence est le quartier dans lequel se situe la maison. C'est un quartier « ancien ». Elle est toujours à 20 minutes du centre de Liège dans une rue avec peu de trafic routier.

#### Quartier "ancien"



### Situation 8

La maison est identique à la maison de la situation de référence. La différence est la rue dans laquelle se situe la maison. Elle est maintenant sur une route nationale avec un trafic automobile assez important.

#### Route nationale



### Situation 9

La maison présentée sur l'image est identique à la maison de référence. La différence est le quartier dans lequel se situe la maison. Il est plus aéré. La maison est toujours à 20 minutes du centre de Liège dans une rue avec peu de trafic routier.



Les images ont été réalisées à partir de vues de rue de Google Streets dans différents quartiers (allée Saint-Vincent à Jambes, rue Général Leman à Ans, Grande route à Beyne-Heusay et rue de Fêchereux à Ans). La maison a été inspirée d'un modèle « clé sur porte » réalisée par l'entreprise Blavier. Enfin, toutes ces images ont été conçues dans le logiciel GIMP (version 2.8), un logiciel libre de retouche d'images. GIMP (<http://www.gimp.org>) fut conçu par Peter Mattis et Spencer Kimball et est actuellement maintenu par Sven Neumann et Mitch Natterer (<http://www.gimpfr.org>)

## 2. BIAIS DE LA MÉTHODE

Cette méthodologie n'est cependant pas idéale. Nous allons passer en revue les principaux biais présentés dans les travaux de Grelot (2004) et de Hardelin et al. (2010). Certains biais sont dus à la méthode, à la formulation des questions et à l'explication des scénarii, d'autres portent sur la personne enquêtée et l'enquêteur. Nous essayerons de palier à ces problèmes.

### 2.1. Biais liés à la méthode

**Biais d'éventail** : les valeurs proposées dans le tableau servant à évaluer le CAP influencent les réponses. Pour éviter ce problème, nous avons construit des tableaux

proposant des valeurs variant de façon identique dans les négatifs et dans les positifs. De plus nous avons alterné deux tableaux : un tableau croissant et un autre décroissant.

**Biais relationnel** : si l'on présente d'autres biens, l'interviewé sera inspiré par ces éléments de comparaison. Il est donc recommandé de s'en tenir à une explication complète du bien à évaluer et de ne pas donner d'indication sur des points de comparaison.

**Biais d'importance** : le simple fait de poser la question s'il y a une plus-value pour un bien immobilier influence déjà la réponse.

**Biais de position** : l'ordre dans lequel sont posées les questions peut suggérer une gradation dans les plus-values. Dans notre cas, nous interrogeons notamment sur le CAP pour un environnement urbain traditionnel, périurbain dense, périurbain aéré, le long de grandes routes. Pour éviter ce biais, nous avons réalisé différents questionnaires en alternant les questions.

## **2.2. Biais liés aux scénarii**

**Biais théorique de mauvaise spécification** : le scénario présente des erreurs introduites par le chercheur.

**Biais de la mauvaise spécification du bien** : l'interviewé ne perçoit pas convenablement le bien à évaluer. Son évaluation ne sera pas adaptée au projet

**Biais de mauvaise spécification contextuelle** : l'individu interrogé n'intègre pas dans son analyse le contexte proposé lors de l'enquête.

Ce travail et ces enquêtes ont été relues par J.-M. Halleux, le promoteur de ce mémoire, C. Viaud-Mouclier, attachée de recherche du SEGEFA et J. Strée, chercheuse au Lepur, afin d'être les plus explicites possibles. De plus, nous avons testé l'enquête sur 5 personnes de notre échantillon afin de nous assurer que les questions étaient bien comprises.

## **2.3. Biais liés à l'interviewé**

**Biais de l'interrogateur** : l'interviewé modifie son estimation de CAP en vue de satisfaire l'enquêteur. Il est donc important de bien spécifier lors de l'enquête qu'aucune réponse n'est fautive.

**Biais stratégique** : l'interviewé peut répondre de manière à encourager ou à minimiser sa participation financière dans le projet. Il pourrait augmenter son évaluation de CAP en vue de montrer l'importance qu'il accorde aux projets. Par contre, il pourrait minimiser son CAP de

crainte de devoir participer financièrement au projet. Dans le cadre de ce travail, comme nous n'évaluons pas un projet collectif, ce biais sera moins présent.

De plus, Loiseau (2011) a écrit dans son mémoire que cette interrogation ne pouvait pas convenir à toutes les personnes. C'est une tâche logico-mathématique. Une partie de l'échantillon pourrait refuser de faire cette évaluation.

## **2.4. Biais liés à l'enquêteur**

Comme dans toutes enquêtes, l'enquêteur influence le déroulement de l'enquête. Il doit rester le plus neutre possible et se restreindre au protocole d'enquête.

## **3. LE QUESTIONNAIRE**

Le questionnaire que nous avons réalisé peut être divisé en trois grandes parties : une introduction, une deuxième partie concernant les questions d'évaluation immobilière et une troisième concernant des questions d'identification.

L'introduction sert dans un premier temps à se présenter et à expliquer à l'interviewé le cadre dans lequel se déroule cette enquête. Nous pouvons envisager une prise de rendez-vous si la personne n'est pas disponible lors de notre premier passage. Certaines questions concernant la description de la personne enquêtée et de sa maison doivent être aussi complétées par l'intervieweur. Nous y demandons :

- la localité du bien
- le nombre de façades
- Le sexe du répondant

Dans la partie sur l'évaluation par contingence, nous demandons d'évaluer 8 situations par rapport à la situation de référence (situation 1) que nous décrivons au début de l'enquête.

La situation de référence, situation 1, est la suivante : une maison neuve est à vendre. C'est une maison 4 façades construite sur un terrain de 700 m<sup>2</sup>. La maison a une surface habitable de 200 m<sup>2</sup> avec quatre chambres. Elle se situe à 20 minutes en voiture du centre de Liège dans un quartier neuf. Dans la rue, il y a peu de trafic routier et l'accessibilité en transport en commun n'est pas bonne. Cette maison coûte, tous frais compris, 250.000 €.

Nous demanderons à l'individu d'estimer son CAP par rapport à ce bien qui lui a été décrit. Pour l'ensemble des situations, nous demandons ceci : « Par rapport à la situation de référence, combien accepteriez-vous de payer pour acquérir ce bien? Je vous demande le

montant maximum que vous accepteriez de payer. Faites un choix parmi les propositions ci-dessous ». Lors du test de notre enquête, nous nous sommes rendu compte que pour certaines situations, les personnes ne voulaient pas acheter un ou des biens présentés. Pour eux, il n'était donc pas possible de déterminer le prix qu'ils accepteraient de payer. Nous avons donc décidé de permettre à la personne enquêtée de répondre qu'il n'accepterait pas d'acheter une telle maison si celui-ci nous signalait lors de la question qu'il ne l'achèterait pas. Cette solution permet de ne pas trop sous-évaluer la réalité du marché.

Les 8 situations à évaluer sont les suivantes :

**Situation 2 :** La maison est identique à la maison de référence. La différence est qu'elle est devenue une maison 2 façades. De plus, elle est construite sur un terrain de 300 m<sup>2</sup> et non plus de 700 m<sup>2</sup>. Dans ce cas, nous demandons d'évaluer le passage à une maison 2 façades qui est accompagné d'une réduction de terrain pour que la situation soit la plus plausible. En effet, nous avons pu constater l'importance de la difficulté à modifier une et une seule caractéristique d'un bien. Il est rare qu'un bien évolue vers une situation en ne changeant qu'une caractéristique. Nous avons préféré changer plusieurs caractéristiques simultanément pour garder une situation plausible plutôt que de modifier une seule caractéristique. Pourtant, pour simplifier la tâche de l'interviewé et aussi individualiser chacune des caractéristiques afin d'obtenir leur propre prix, il aurait été préférable de ne faire évoluer qu'une caractéristique.

**Situation 3 :** La maison à vendre est, comme pour la situation précédente, de 2 façades et construite sur un terrain de 300 m<sup>2</sup>. Mais cette maison a une surface habitable de 150 m<sup>2</sup> (à la place de 200 m<sup>2</sup>) et elle ne compte plus que deux chambres (à la place de quatre chambres). Nous tentons par cette situation d'identifier le prix de la surface habitable mais cela ne peut être approché que par une réduction de pièces. Cette réduction du nombre de chambres permet une meilleure perception de l'importance de la réduction de la surface habitable par le répondant. Cette question, tout comme la première, sert de contrôle entre la méthode hédonique et l'approche par contingence.

**Situation 4 :** La maison est identique à la maison de référence à l'exception de sa localisation. La maison se situe à 5 minutes du centre de Liège et non plus à 20 minutes. L'objectif de cette situation est d'identifier l'impact de la proximité du centre de Liège et cela sans modification du milieu. Dans le modèle hédonique, nous avons constaté que les caractéristiques urbaines étaient synonymes de moins-value (ce qui est contre-intuitif), nous voulions par cette question étudier l'impact de la proximité du centre urbain et vérifier si effectivement cette caractéristique apporte une moins-value.

**Situation 5 :** La maison à vendre est identique à la maison de la situation de référence mais l'accessibilité en transport en commun est bonne (un passage de bus toutes les 15 minutes à proximité). Dans cette situation, nous souhaitons voir l'impact de la bonne accessibilité aux transports en commun. Dans notre modèle hédonique, nous avons trouvé une moins-value pour la proximité des transports en commun. Il nous paraît important d'explorer plus ce résultat car il va à l'encontre de nos attentes.

**Situation 6 :** La maison est identique à la maison de référence. La différence est le cadre dans lequel se situe la maison. Derrière la maison se trouve un parc industriel, un « zoning ». Cette situation est exploratoire. Dans notre modèle hédonique, nous avons peu de variables concernant l'environnement. Cette situation, ainsi que les trois suivantes, vise à explorer cette caractéristique de l'environnement. Dans ce cas-ci, nous avons essayé de faire varier le quartier vers un milieu proche d'industries avec une nuisance surtout visuelle.

**Situation 7 :** La maison est identique à la maison de référence. La différence est le quartier dans lequel se situe la maison. C'est un quartier « ancien ». Elle est toujours à 20 minutes du centre de Liège dans une rue avec peu de trafic routier. Nous avons localisé la maison dans un quartier plus urbain, plus dense avec peu d'espace vert.

**Situation 8 :** La maison est identique à la maison de la situation de référence. La différence est la rue dans laquelle se situe la maison. Elle est maintenant sur une route nationale avec un trafic automobile assez important. L'objectif est d'identifier les moins-values engendrées par les nuisances sonores. Le quartier présente aussi peu d'espace vert et est économiquement plus pauvre que dans la situation de référence.

**Situation 9 :** La maison présentée sur l'image est identique à la maison de référence. La différence est le quartier dans lequel se situe la maison. Il est plus aéré. La maison est toujours à 20 minutes du centre de Liège dans une rue avec peu de trafic routier. Cette dernière situation doit permettre d'identifier les plus-values d'un milieu peu dense. La rue est typique du milieu périurbain belge, grand consommateur d'espace. Cette situation représente le modèle « 4 façades » que beaucoup de personnes recherchent.

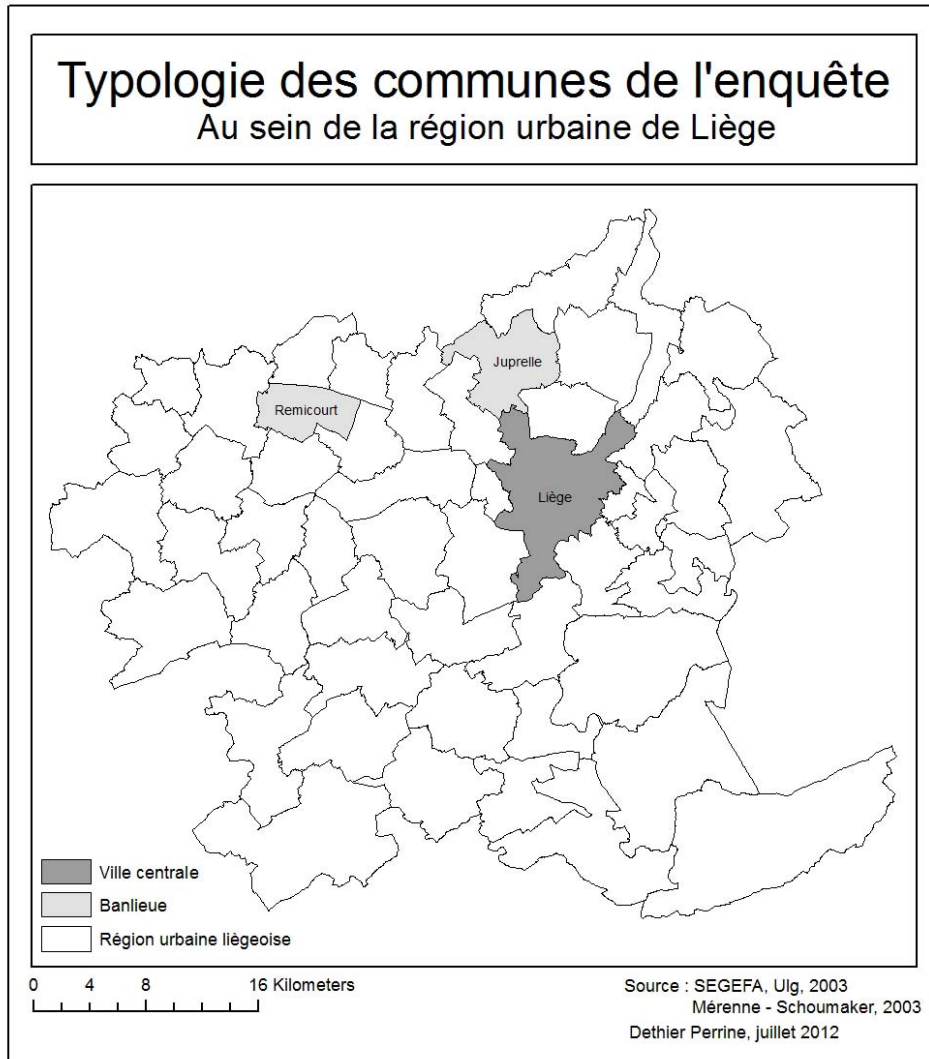
Pour éviter un biais lié à l'ordre des questions nous avons décidé d'intervertir les situations 7, 8 et 9.

La troisième partie correspond à des questions d'identification. Nous y retrouvons des questions portant sur le ménage, l'emploi ou l'âge. Le questionnaire se trouve à l'annexe 10.

## 4. DESCRIPTION DES TROIS COMMUNES

Les enquêtes ont été réalisées au des communes de Liège, Juprelle et Remicourt. Liège est la ville centrale et les deux autres des communes de la banlieue périurbaine. Elles sont situées dans la région liégeoise qui a été notre zone d'étude de la méthode hédonique (carte 4).

Carte 4 : Typologie des communes de l'enquête



Les communes connaissent toutes les trois une croissance de leurs habitants même si, pendant de nombreuses années, la commune de Liège en a perdu.



**Tableau 12 : Populations des communes**

	2001	2008	Taux de croissance entre 2001 et 2008
Juprelle	8.149	8.619	5,77 %
Remicourt	4.891	5.191	6,13%
Liège	184.604	190.102	2,98 %

Source : INS 2001 et 2008

En ce qui concerne le revenu total (transferts, patrimoine, travail) (tableau 13 Tableau 13), les deux communes de banlieues ont en moyenne un revenu plus important par habitant que la commune de Liège. Les habitants de ces communes gagnent presque de 2.000 € de plus par an que les Liégeois. Pour pouvoir comparer ces chiffres à notre question de l'enquête sur le revenu, nous décidons de calculer le revenu par mois pour un ménage avec 2 adultes. Il faut cependant faire attention que ces montants datent déjà de 2001 et qu'il faudrait les actualiser pour pouvoir les comparer.

**Tableau 13 : Revenus totaux des communes**

Commune	Estimation des revenus totaux (transferts, patrimoine, travail) en 2001 (en €)	Montant moyen des revenus totaux (transferts, patrimoine, travail) par habitant (en €)	Revenu moyen par mois pour 2 habitants (en €)
JUPRELLE	102.541.000	12.580	2.097
REMICOURT	62.663.000	12.810	2.135
LIEGE	2.013.087.000	10.900	1.817

Source : CPDT, 2008

D'un point de vue de l'urbanisation, les communes ont des profils différents. Premièrement, la commune de Liège a un territoire fortement urbanisé par de l'habitat puisqu'un quart de la commune est construit par du résidentiel. Les communes de banlieue ne le sont qu'à 7-8 %. Par contre, si l'on regarde l'évolution entre 2001 et 2006, les communes périurbaines connaissent une croissance relative de l'habitat plus importante.

**Tableau 14 : Evolution des terrains résidentiels par commune entre 2001 et 2006**

	2001		2006		Evolution	
	Ha	%	Ha	%	Ha	%
JUPRELLE	257	7,2	276	7,8	20	7,7
REMICOURT	167	7,3	178	7,8	11	6,5
LIEGE	1.655	24,2	1.682	24,6	27	1,6

Source : CPDT, 2008

## 5. ANALYSE DES ENQUÊTES

Dans ce chapitre, nous allons présenter les résultats de nos enquêtes. Dans un premier temps, nous analyserons la population interrogée à l'aide des caractéristiques socio-économiques et de logement. Cette section permettra de mettre en lumière les caractéristiques des personnes construisant (ou achetant) des habitations neuves dans la région urbaine liégeoise.

La seconde section concerne le traitement des questions portant sur le CAP. Nous essayerons d'identifier des liens entre ces réponses et les caractéristiques économiques mais aussi des différences entre les populations « urbaine » et « périurbaine ». Mais le point primordial sera de dégager des tendances à comparer avec notre modèle hédonique.

Pour l'analyse des résultats de l'enquête, nous effectuerons des tests statistiques. Les principales données à analyser sont les CAP qui sont des variables catégorisées en 17 classes. Nous décidons d'utiliser le centre de chaque catégorie comme donnée et de la traiter comme une variable quantitative. Nous aurions pu faire le choix de les traiter comme des variables qualitatives. Mais vu le nombre de catégories et le nombre d'enquêtes réalisées, les tests statistiques, notamment le test d'indépendance du  $X^2$ , effectués n'auraient pas été efficaces car le nombre d'individus pour certaines catégories n'aurait pas été suffisant.

Trois tests statistiques seront utilisés : tests de corrélation, tests d'analyse de variance, tests t de Student (test de conformité). Les corrélations mesurent l'association entre deux variables quantitatives. Ensuite, nous effectuerons des analyses de la variance à un critère (ANOVA-1). Ce test permet d'observer si des populations distinctes selon une variable qualitative à k modalités ont des moyennes hétérogènes. L'hypothèse nulle d'une ANOVA-1 est que les moyennes sont homogènes. Enfin, si nous voulons tester l'égalité des moyennes entre seulement deux populations, un test t de Student suffit, une ANOVA-1 ne doit pas être réalisée.

Il faut cependant tenir compte du fait que la taille de notre échantillon est relativement limitée. Les résultats présents dans cette section ne peuvent pas être extrapolés à l'ensemble de la population de la région urbaine liégeoise. Nous pouvons cependant y dégager des tendances qui pourraient être affinées par l'augmentation de notre échantillon.

## 5.1. Population interrogée

Nous distinguons dans cette partie les caractéristiques socio-économiques d'une part et les caractéristiques portant sur l'habitation de 80 personnes interrogées. Cette partie permettra de mieux connaître la population ciblée par nos enquêtes.

### 5.1.1. Caractéristiques socio-économiques

La majorité des personnes enquêtées sont des femmes, principalement dans la commune de Liège (67,5 %). Nous expliquons ce pourcentage important par l'heure de nos enquêtes, généralement en après-midi et aussi par le fait que les femmes répondaient plus volontiers (taux de réponses relativement faible car une bonne partie des personnes n'étaient pas à leur domicile et que certaines habitations présentent dans notre liste n'étaient encore pas construites).

**Tableau 15 : Sexe du répondant**

	Pourcentage Femme	Pourcentage homme	n
Liège	67,5 %	32,5 %	40
Communes périurbaines	52,5 %	47,5 %	40
<b>TOTAL</b>	<b>60 %</b>	<b>40 %</b>	<b>80</b>

La moyenne d'âge des personnes enquêtées est de 42 ans. Il y a une légère différence entre la population urbaine et population périurbaine mais cela ne conduit pas à une hétérogénéité des populations. Le test de Student ( $t=0,73$ ,  $df=78$ ,  $p=0,46$ ) ne permet pas de rejeter l'hypothèse nulle, les échantillons sont donc homogènes du point vu de l'âge.

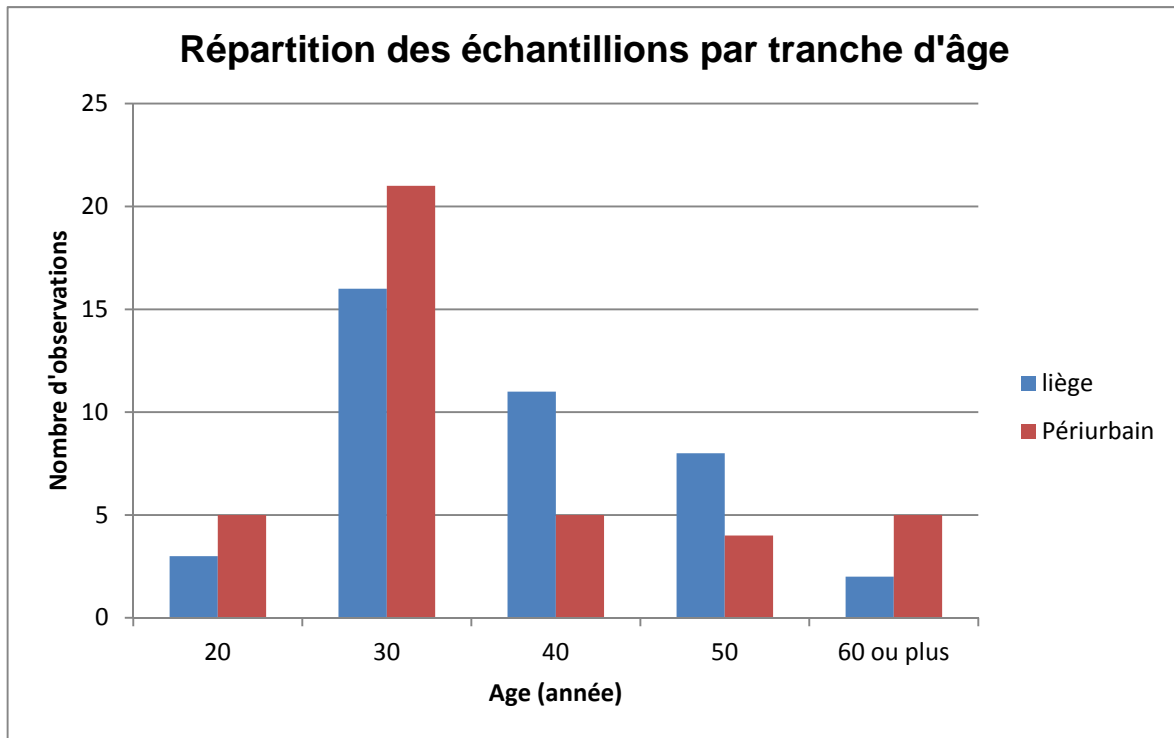
**Tableau 16 : Statistiques descriptives de l'âge du répondant (en année)**

n=80	Moyenne	Médiane	Ecart-type	Minimum	Maximum
Liège	43	41,5	9,23	27	64
Communes périurbaines	41	35	12,75	26	73
<b>TOTAL</b>	<b>42</b>	<b>38,5</b>	<b>11,10</b>	<b>26</b>	<b>73</b>

La population interrogée se situe principalement dans la tranche des 30 - 40 ans mais l'ensemble des tranches actives a été interrogé, de même que quelques personnes retraitées. Cette répartition montre que les jeunes (20-30 ans) construisent moins. Cela est sans doute lié à l'importance des besoins financiers initiaux pour la construction d'une

maison. De plus, les personnes retraitées sont très peu représentées car elles ont certainement moins envie d'investir dans l'immobilier : cela demande une charge de travail importante et, par le passé, elles ont pu construire leur maison.

**Graphique 16 : Répartition des échantillons par tranche d'âge**

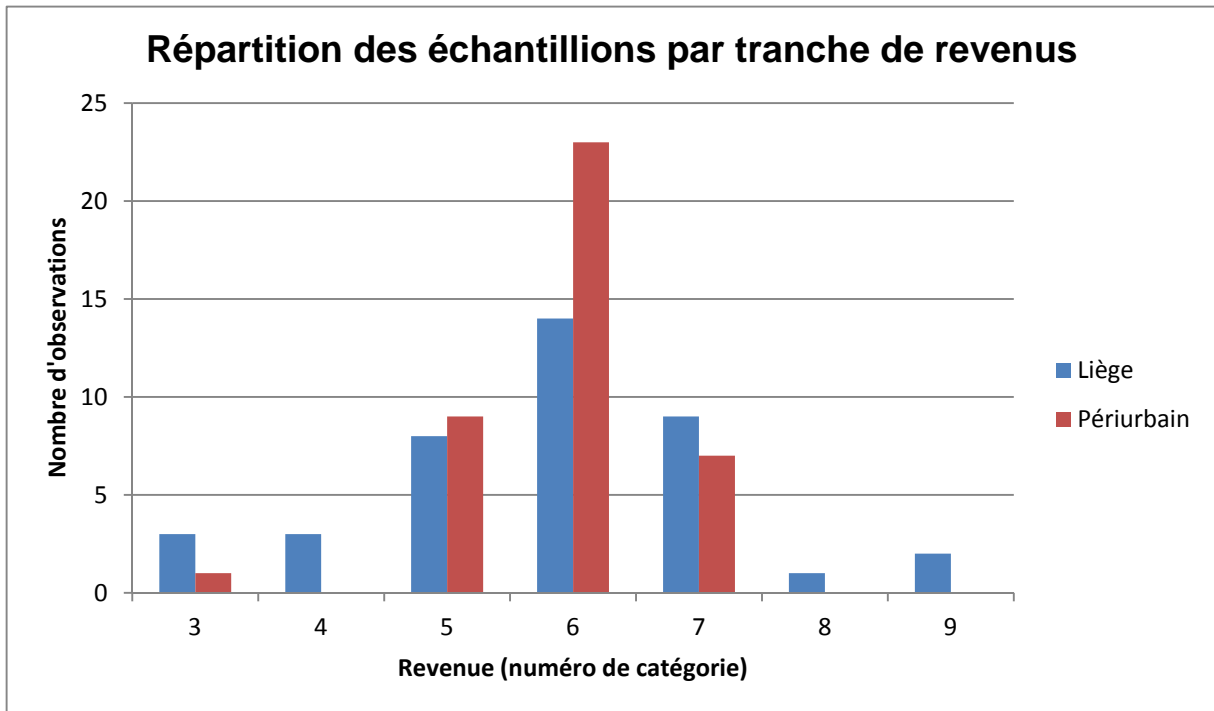


Le revenu net du foyer est une variable catégorisée (tableau 17) puisque les personnes interrogées devaient choisir dans une liste de propositions de revenus nets. Les moyennes pour les deux échantillons se situent dans la catégorie 5 (entre 2.500 € et 3.000 €) et les modes dans la catégorie 6 (entre 3.000 € et 4.000 €). Les deux populations, « urbaine » et « périurbaine », sont donc homogènes (t de Student avec  $t=0,20$ ,  $dl=78$ ,  $p=0,84$ ).

**Tableau 17 : Catégorie de revenus**

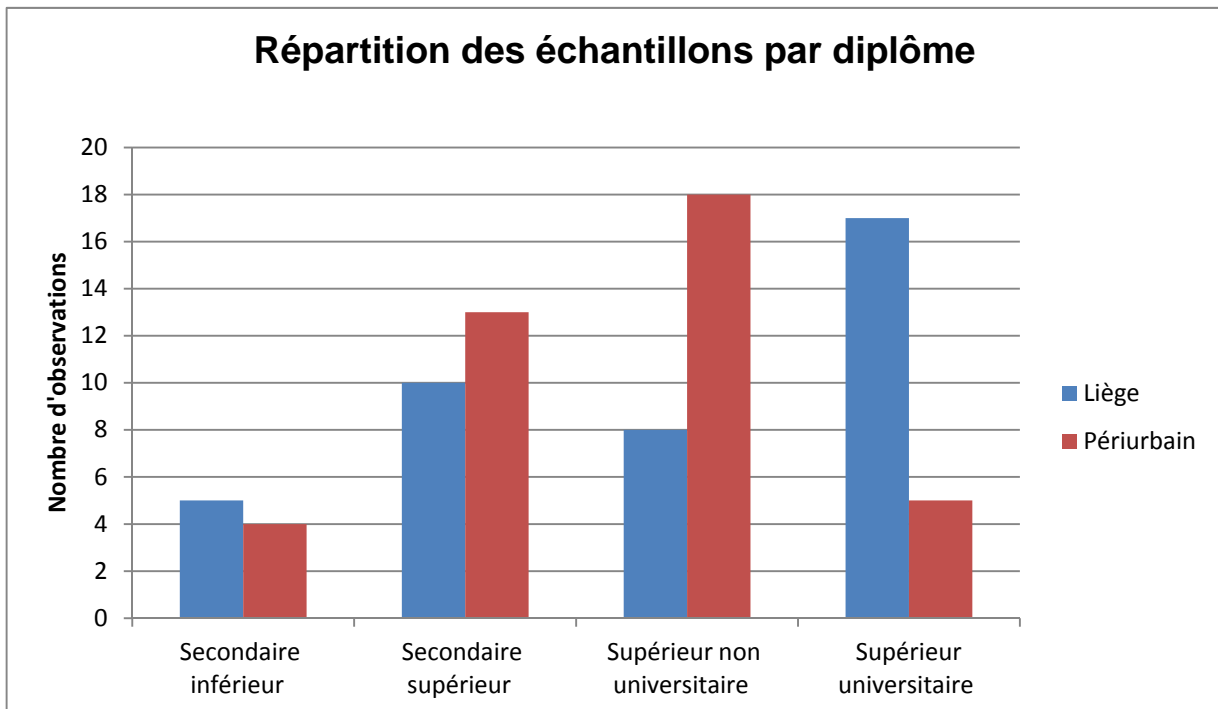
1	< 1.000 €
2	1.000 € à 1.500 €
3	1.500 € à 2.000 €
4	2.000 € à 2.500 €
5	2.500 € à 3.000 €
6	3.000 € à 4.000 €
7	4.000 € à 5.000 €
8	5.000 € à 6.000 €
9	6.000 € à 7.000 €
10	> 7.000 €

Graphique 17 : Répartition des échantillons par tranche de revenus



Sur l'ensemble des personnes interrogées, seulement 6 personnes ne travaillaient pas : 2 étaient sans-emploi (dont un pour invalidité) et 4 personnes étaient retraitées. De plus ces personnes ont pour la grande majorité leur diplôme du secondaire supérieur.

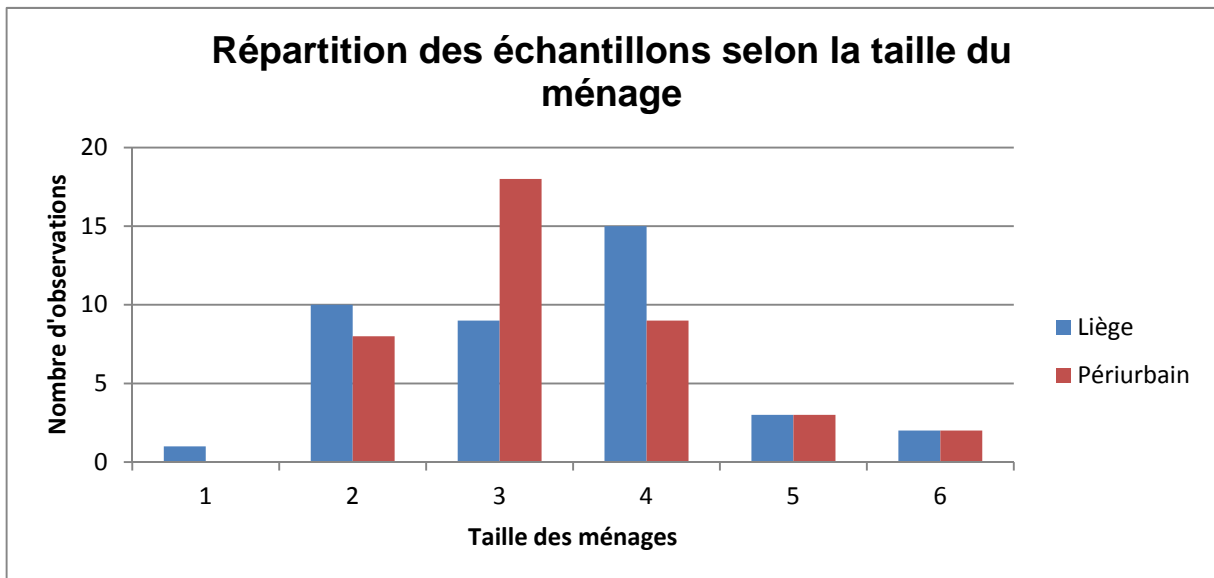
Graphique 18 : Répartition des échantillons par diplôme



Le niveau de formation est significativement dépendant de la commune de l'interviewé ( $X^2=10,89$ ,  $dl=3$ ,  $p=0,01$ ). En effet, sur le graphique 18, nous constatons qu'à Liège, il y a plus de diplômés universitaires et dans les communes périurbaines, on retrouve plus de diplômés de l'enseignement supérieur non universitaire.

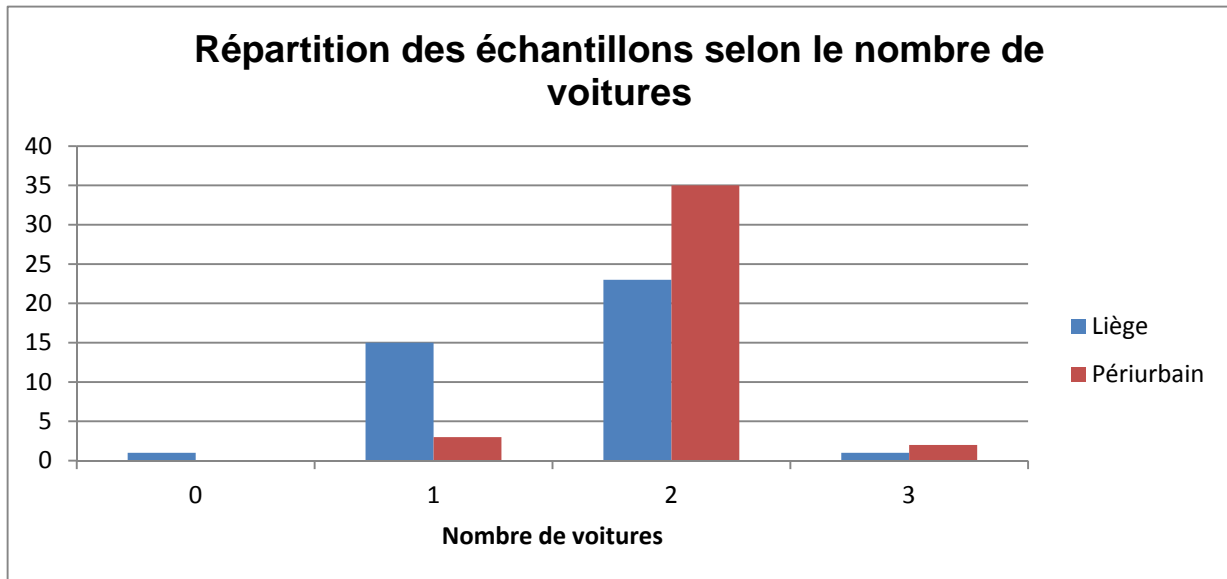
La taille des ménages est similaire entre Liège et les communes périurbaines ( $t$  de Student :  $t=0,201399$ ,  $dl=78$ ,  $p=0,84$ ). La moyenne tourne autour de 3,3 personnes ce qui est supérieur à la moyenne liégeoise (1,87), de Remicourt (2,52) et de Juprelle (2,47). Le graphique 19 nous montre que la majorité des ménages est constituée de 3 à 4 personnes. En effet, lors de nos enquêtes, nous avons rencontré beaucoup de foyers constitués de deux adultes et de un ou deux enfants. D'ailleurs, le nombre d'enfants de moins de 15 ans dans le foyer est fortement dépendant à la taille des ménages ( $X^2=130,12$ ,  $dl=20$ ,  $p=0,00$ ).

Graphique 19 : Répartition des échantillons selon la taille du ménage



Pour ce qui est des voitures, la plupart des foyers ont deux voitures. Mais on constate une différence entre les habitants de la commune de Liège et ceux de Juprelle et Remicourt. Le  $t$  de Student permet de rejeter l'hypothèse nulle ( $t=-3,44$ ,  $dl=78$ ,  $p=0,00$ ), ce qui implique une hétérogénéité des deux échantillons. En effet, à Liège, nous avons une moyenne de 1,6 voitures par ménage (avec une personne ne disposant d'aucune voiture) tandis que dans les communes périurbaines, la moyenne est plus élevée (2 voitures).

Graphique 20 : Répartition des échantillons selon le nombre de voitures



Enfin, sur l'ensemble des personnes interrogées, seulement deux personnes n'étaient pas propriétaires mais locataires.

### 5.1.2. Caractéristiques de l'habitation

Lors de l'enquête, des informations sur le logement ont été demandées : la superficie habitable de la nouvelle maison, la superficie du terrain et le nombre de façades.

Les superficies habitables sont homogènes dans les deux échantillons ( $t=-1,52$ ,  $dl=78$ ,  $p=0,13$ ). La médiane de 200 m<sup>2</sup> est identique. Nous retrouvons quelques observations plus extrêmes dont notamment un maximum à 450 m<sup>2</sup> dans les communes périurbaines. Lors de nos enquêtes, nous avons déjà relevé cet extrême et il nous semblait aberrant au vu de la maison lors de l'enquête.

Tableau 18 : Statistiques descriptives de la surface habitable (m<sup>2</sup>)

n=80	Moyenne	Médiane	Ecart-type	Minimum	Maximum
Liège	191,7	200	46,4	99	330
Communes périurbaines	209,5	200	57,7	100	450
<b>TOTAL</b>	<b>200,6</b>	<b>200</b>	<b>52,7</b>	<b>99</b>	<b>450</b>

Pour la superficie des terrains, le test du t de Student nous indique que les deux échantillons sont significativement hétérogènes ( $t= -4,31$ ,  $dl=78$ ,  $p=0,00$ ). En effet, les terrains sont plus grands dans les communes périurbaines (autour de 870 m<sup>2</sup>) tandis que la surface des

terrains se situe plutôt autour de 550 m<sup>2</sup> à Liège. On peut cependant trouver de grands terrains à Liège comme le montre le maximum de 1900 m<sup>2</sup>.

**Tableau 19 : Statistiques descriptives de la surface du terrain (m<sup>2</sup>)**

n=80	Moyenne	Médiane	Ecart-type	Minimum	Maximum
Liège	546,5	552	343,9	50	1900
Communes périurbaines	871,6	825	330,8	200	1950
<b>TOTAL</b>	<b>709,1</b>	<b>700</b>	<b>373,1</b>	<b>50</b>	<b>1950</b>

Enfin, pour les façades, les maisons liégeoises sont souvent de 2 ou 3 façades. Dans le périurbain, on trouve plus régulièrement des maisons 4 façades. Le test du t de Student a d'ailleurs montré que les deux échantillons étaient hétérogènes ( $t = -4,94$ ,  $dl = 78$ ,  $p = 0,00$ )

**Tableau 20 : Statistiques descriptives sur le nombre de façades**

n=80	Moyenne	Médiane
Liège	2,7	3
Communes périurbaines	3,5	4
<b>TOTAL</b>	<b>3,1</b>	<b>3</b>

## 5.2. Les consentements à payer

Dans cette partie nous étudions les résultats sur les questions d'évaluation immobilière. Nous avons repris, dans le tableau 21, les différents tests que nous allons effectuer en fonction des variables socio-économiques et des CAP. Nous testerons aussi l'indépendance entre les réponses des CAP des différentes situations.

**Tableau 21 : Classification des variables selon le test statistique effectué**

T de Student	ANOVA-1	Corrélation
Sexe	Nombre d'enfants	Age
Planche d'échelle de valeurs	Nombres de façades	Revenu (moyenne de la catégorie)
	Niveau d'étude	Taille du ménage
	Nombres de voitures	Superficie habitable
	Numéro du questionnaire	Superficie du terrain



Comme il a été admis que les personnes puissent refuser de proposer un prix, nous devons en tenir compte dans nos analyses. Nous avons réalisé deux analyses : une n'incluant pas les réponses nulles et l'autre considérant les refus comme le CAP minimum. Les deux méthodes donnant des résultats forts similaires, nous ne présentons dans notre travail les deux résultats uniquement pour les statistiques descriptives des CAP de la population totale.

### 5.2.1. Situation 2 (petit terrain et 2 façades)

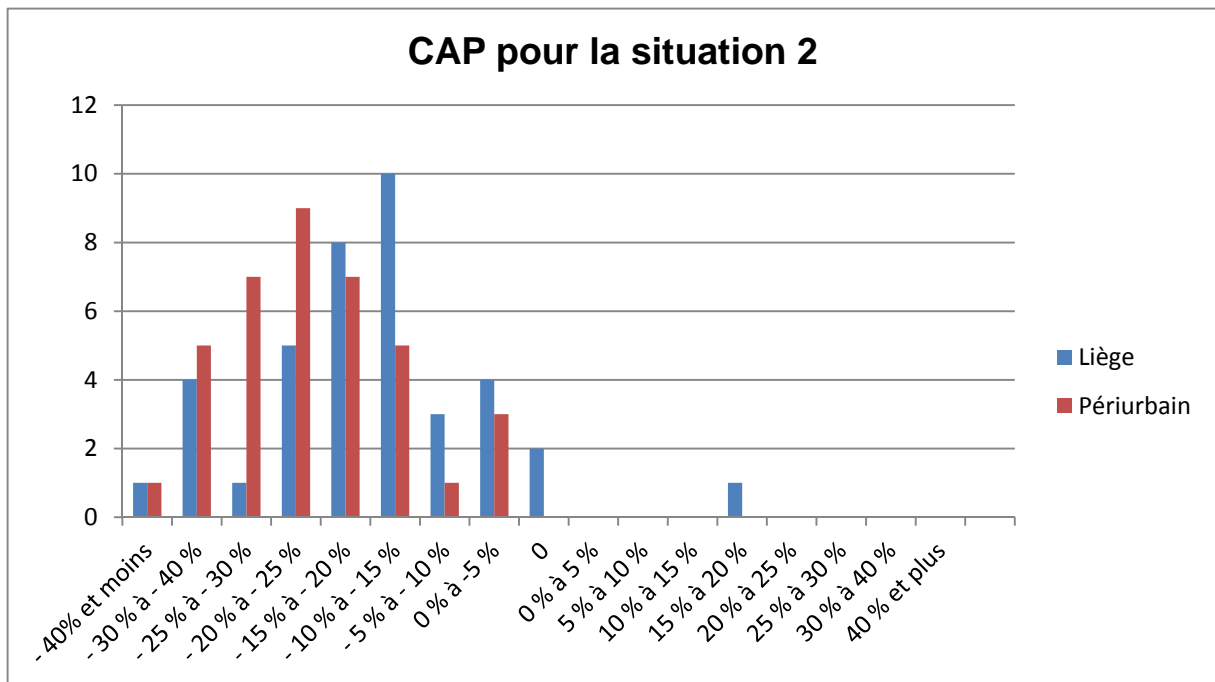
La situation 2 fait référence à une maison possédant un terrain plus petit (300 m<sup>2</sup> au lieu de 700) et une maison 2 façades au lieu de 4 façades.

Pour la grande majorité des personnes interrogées, cette situation présente une moins-value. Une personne trouve cette situation plus avantageuse évoquant un gain d'isolation comme l'indique le tableau 22. Trois personnes ont indiqué le refus de l'achat d'une telle maison. Dans le graphique 21 nous remarquons que pour l'échantillon périurbain, la majorité évoque une moins-value de l'ordre de 20 à 25 % tandis qu'à Liège cette moins-value n'atteint que 10 à 15 %. Le t de Student, réalisé sur ces données, montre que les deux échantillons sont hétérogènes ( $t=2,13$ ,  $t=78$ ,  $p= 0,03$ ). Les urbains accordent donc moins d'importance à la grandeur du terrain et à l'envie d'avoir une maison 4 façades. A l'inverse les périurbains évaluent plus négativement la perte de jardin et la mitoyenneté de leur maison.

Tableau 22 : Statistiques descriptives sur le CAP pour la situation 2

n=80	Moyenne	Médiane	Minimum	Maximum
Liège	- 15 % à - 20 %	- 10 % à - 15 %	- 40% et moins	15 % à 20 %
Communes périurbaines	- 20 % à - 25 %	- 20 % à - 25 %	- 40% et moins	0 % à -5 %
<b>TOTAL</b>	<b>- 15 % à - 20 %</b>	<b>- 15 % à - 20 %</b>	<b>- 40% et moins</b>	<b>15 % à 20 %</b>
<b>TOTAL (incluant refus)</b>	<b>- 20 % à - 25 %</b>	<b>- 15 % à - 20 %</b>	<b>- 40% et moins</b>	<b>15 % à 20 %</b>

Graphique 21 : CAP pour la situation 2



Les différents tests de corrélation montrent que la corrélation est non nulle entre les CAP pour la situation 2 et la superficie du terrain ( $p=0,02$ ) mais aussi entre les CAP pour la situation 2 et l'âge du répondant ( $p=0,00$ ).

Le coefficient de corrélation entre les CAP pour la situation 2 et la superficie de terrain est de  $-0,25$ . La relation est négative : plus les personnes possèdent un grand terrain moins elles seraient prêtes à payer pour acheter une maison avec un petit terrain. Cette corrélation peut expliquer l'hétérogénéité entre nos deux échantillons puisque les urbains possèdent des terrains plus petits que les populations périurbaines. D'ailleurs, il n'y a pas de corrélation significative pour des tests de corrélation uniquement avec l'échantillon périurbain ou urbain.

En ce qui concerne la corrélation entre les CAP pour la situation 2 et l'âge, le coefficient est positif ( $0,30$ ) : plus les personnes sont âgées moins elles diminuent le prix du bien immobiliers. Certains retraités nous ont parlé de la pénibilité de l'entretien de grands terrains qui est moins envisagée chez les plus jeunes. Cette perspective peut expliquer cette corrélation. Cette corrélation est toujours observée pour l'échantillon uniquement urbain.

### 5.2.2. Situation 3 (petite maison)

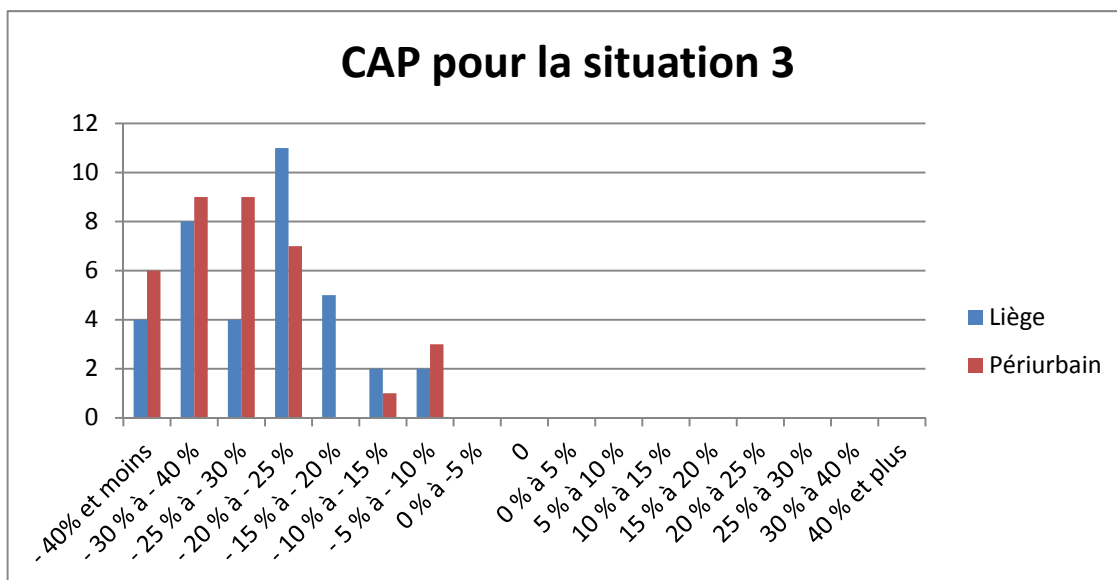
La situation 3 illustre une maison sur un petit terrain et 2 façades avec une superficie habitable plus petite (150 m<sup>2</sup> au lieu de 200 m<sup>2</sup>) et ne possédant plus que 2 chambres.

Sur l'ensemble des personnes interrogées, 9 personnes ne voudraient pas acheter cette maison : principalement parce que le nombre de chambres n'est pas suffisant. Les résultats obtenus pour les deux populations sont fort identiques. D'ailleurs le t de Student ( $t= 0,63$ ,  $df=78$ ,  $p= 0,53$ ) ne rejette pas l'hypothèse d'homogénéité. La majorité des personnes donnent une valeur inférieure à au moins 25 - 30 % (tableau 23) par rapport au prix initial. Il est intéressant de noter que, contrairement à la situation précédente, les choix de valeurs sont plus homogènes entre les deux échantillons. L'espace habitable est donc perçu de la même manière pour des personnes vivant à Liège ou en dehors.

Tableau 23 : Statistiques descriptives sur le CAP pour la situation 3

n=80	Moyenne	Médiane	Minimum	Maximum
Liège	- 25 % à - 30 %	- 20 % à - 25 %	- 40% et moins	- 5 % à - 10 %
Communes périurbaines	- 25 % à - 30 %	- 25 % à - 30 %	- 40% et moins	- 5 % à - 10 %
<b>TOTAL</b>	<b>- 25 % à - 30 %</b>	<b>- 25 % à - 30 %</b>	<b>- 40% et moins</b>	<b>- 5 % à - 10 %</b>
<b>TOTAL (incluant refus)</b>	<b>- 25 % à - 30 %</b>	<b>- 25 % à - 30 %</b>	<b>- 40% et moins</b>	<b>- 5 % à - 10 %</b>

Graphique 22 : CAP pour la situation 3



Par rapport à la situation 2, nous pouvons observer sur le graphique 22 que l'ensemble des valeurs a diminué. La majorité des personnes choisissaient souvent une catégorie en-

dessous de ce qu'ils avaient répondu précédemment. Cette manière de procéder explique que les deux situations sont corrélées avec un  $r$  de 0,51 ( $p=0,00$ ).

### 5.2.3. Situation 4 (proche de Liège)

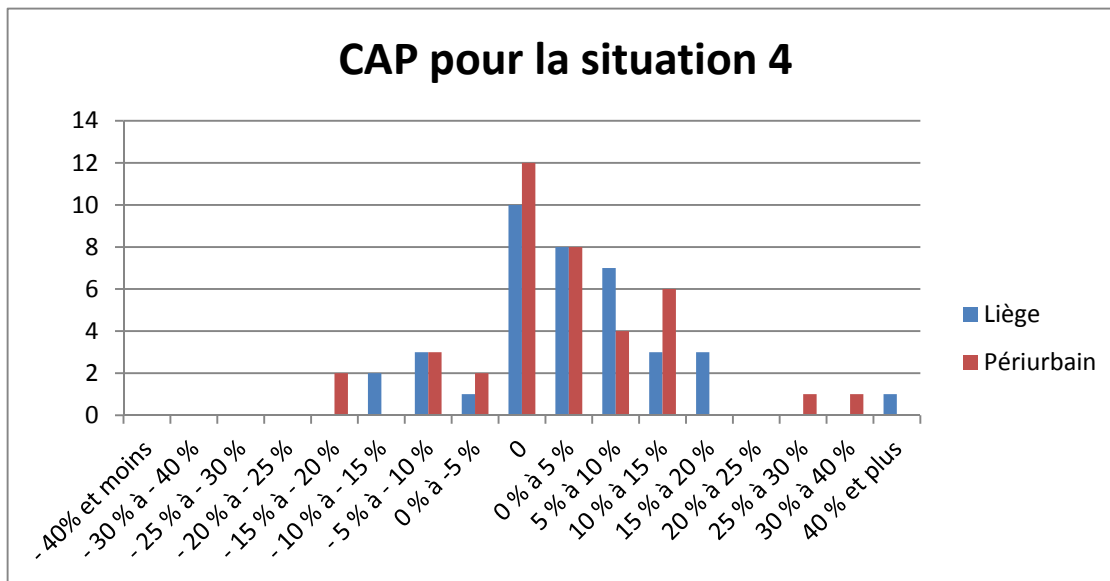
La situation 4 représentait la proximité de Liège. L'habitation passait de 20 à 5 minutes en voiture du centre de Liège.

Trois personnes refuseraient d'acheter un tel bien. Pour les autres, les avis sont plus partagés que pour les deux premières situations : 13 personnes considèrent que la proximité de Liège est une moins-value tandis que 45 personnes considèrent cela comme une plus-value pour l'habitation. Les deux échantillons sont identiques puisque le  $t$  de Student affichait un  $p$ -value de 0,42 ( $t= 0,81$ ,  $dl=33$ ,  $p=0,42$ ). La moyenne indique un prix stable par rapport à la maison de référence mais la médiane montre une légère plus-value. Cette caractéristique d'accessibilité est donc perçue de manière très variable au sein de notre échantillon. La proximité de Liège n'est pas toujours perçue comme un avantage. Le graphique 23 l'illustre bien puisque les observations sont de part et d'autre du prix de référence.

Tableau 24 : Statistiques descriptives sur le CAP pour la situation 4

<b>n=80</b>	<b>Moyenne</b>	<b>Médiane</b>	<b>Minimum</b>	<b>Maximum</b>
Liège	0 % à 5 %	0 % à 5 %	- 10 % à - 15 %	40 % et plus
Communes périurbaines	0 %	0 % à 5 %	- 15 % à - 20 %	30 % à 40 %
<b>TOTAL</b>	<b>0 %</b>	<b>0 % à 5 %</b>	<b>- 15 % à - 20 %</b>	<b>30 % à 40 %</b>
<b>TOTAL (incluant refus)</b>	<b>0 %</b>	<b>0 % à 5 %</b>	<b>- 40% et moins</b>	<b>30 % à 40 %</b>

Graphique 23 : CAP pour la situation 4



Une des corrélations est positive : les CAP pour la situation 4 et la taille des ménages. Elle est de  $-0,24$  et  $p=0,03$  : plus la taille des ménages est grande, plus les personnes diminuent le prix de la maison. Il serait intéressant d'approfondir cette relation. On pourrait l'expliquer par le fait que les quartiers périurbains ont un trafic routier généralement plus calme ce qui permet aux enfants de jouer dans la rue ou par un sentiment de plus grande sécurité pour les enfants à se promener en rue. Les parents ne voudraient pas élever leurs enfants à Liège.

Il y a aussi une corrélation négative de  $-0,22$  ( $p=0,05$ ) entre les CAP pour la situation 4 et ceux de la situation 3 : plus l'interrogé a évalué de manière positive la situation 4 moins il voit positivement la situation 3 (petite superficie habitable). Les personnes qui considèrent qu'habiter près de la ville est un avantage accordent un grand intérêt à l'importance de la taille de la maison.

#### 5.2.4. Situation 5 (accessibilité en transport en commun)

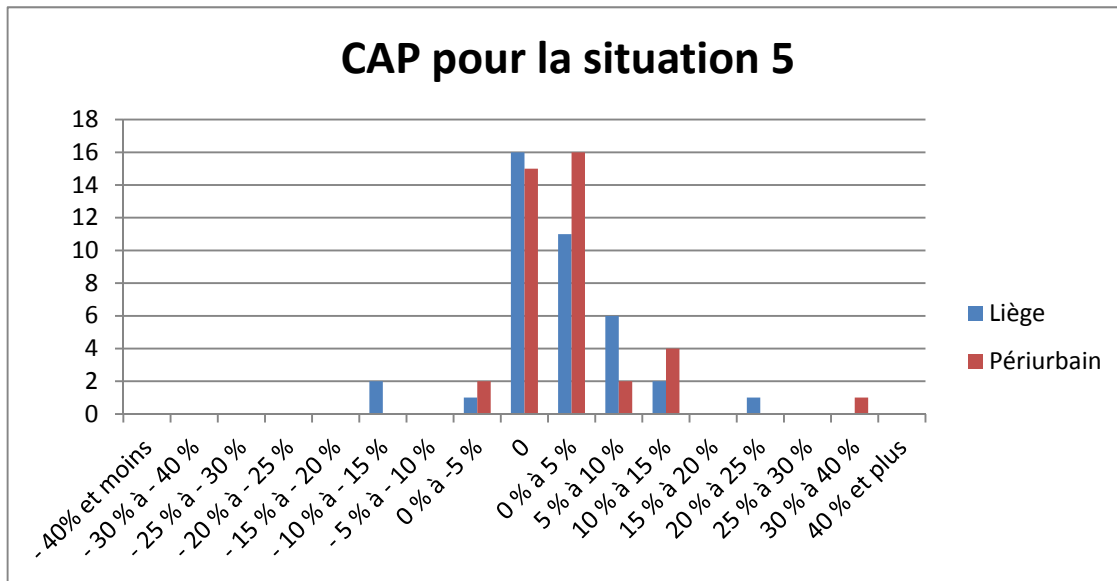
La situation 5 présentait un bien immobilier dont l'accessibilité en transport en commun était bonne.

Seule une personne ne souhaiterait pas acheter cette maison. Les moyennes sont fort similaires à la situation 4 même si on peut noter une légère augmentation. En effet, il n'y a que 5 personnes qui considèrent la proximité et la haute fréquence des bus comme une moins-value au contraire des 13 personnes pour la situation 4. Les échantillons ne sont pas hétérogènes (test de t de Student :  $t= -0,57$ ,  $dl=78$ ,  $p=0,57$ ) et la moyenne se situe entre une plus-value de 0 à 5 %.

Tableau 25 : Statistiques descriptives sur le CAP pour la situation 5

	Moyenne	Médiane	Minimum	Maximum
Liège	0 %	0 % à 5 %	- 10 % à - 15 %	20 % à 25 %
Communes périurbaines	0 % à 5 %	0 % à 5 %	0 % à -5 %	30 % à 40 %
<b>TOTAL</b>	<b>0 % à 5 %</b>	<b>0 % à 5 %</b>	<b>- 10 % à - 15 %</b>	<b>30 % à 40 %</b>
<b>TOTAL (incluant refus)</b>	<b>0 %</b>	<b>0 % à 5 %</b>	<b>- 40% et moins</b>	<b>30 % à 40 %</b>

Graphique 24 : CAP pour la situation 5



Parmi les corrélations, deux sont significatives. Il y a une corrélation ( $r=0,46$  et  $p=0,00$ ) entre la situation 4 et 5. Les choix de CAP entre la situation 4 et 5 sont similaires. Pour 35 personnes, la situation 4 est identique à la situation 5. Une autre corrélation est présente entre les situations 5 et 3. Celle-ci est négative ( $r=-0,30$  et  $p=0,01$ ) : plus les gens évaluent positivement une habitation avec une bonne accessibilité en transport en commun moins ils mettent un prix élevé pour les petites maisons. Il en était de même entre les situations 3 et 4.

### 5.2.5. Situation 6 (parc industriel)

La situation 6 illustre une maison identique à la situation de référence à l'exception qu'au fond du jardin se situait un parc industriel.

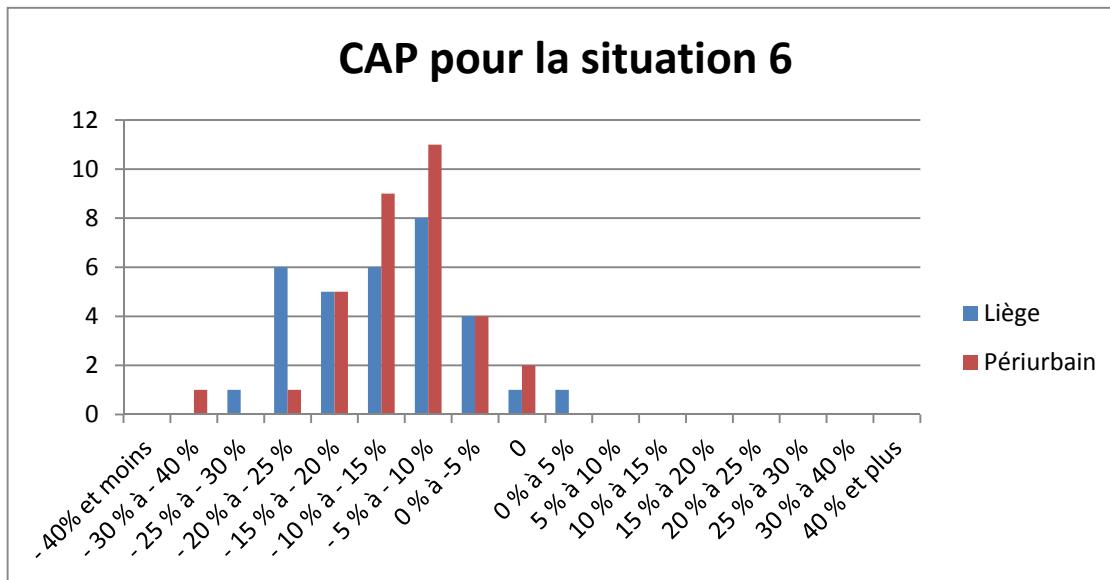
Cette situation a essuyé 15 refus. Pour ces personnes, il ne serait pas envisageable d'avoir un parc industriel visible de leur jardin. Pour le reste des personnes enquêtées, la situation représente une moins-value d'en moyenne 10 à 15 % du prix original. Les deux échantillons ont des avis homogènes (t de Student :  $t= -0,58$ ,  $dl=78$ ,  $p=0,56$ ). Nous pouvons remarquer

sur le graphique 25 qu'une seule personne estime cette situation comme intéressante. Pour elle, la proximité d'un parc industriel permet d'être aussi à proximité d'emplois ou de services.

**Tableau 26 : Statistiques descriptives sur le CAP pour la situation 6**

	Moyenne	Médiane	Minimum	Maximum
Liège	- 10 % à - 15 %	- 10 % à - 15 %	- 25 % à - 30 %	0 % à 5 %
Communes périurbaines	- 10 % à - 15 %	- 5 % à - 10 %	- 30 % à - 40 %	0 %
<b>TOTAL</b>	<b>- 10 % à - 15 %</b>	<b>- 10 % à - 15 %</b>	<b>- 30 % à - 40 %</b>	<b>0 % à 5 %</b>
<b>TOTAL (incluant refus)</b>	<b>- 15 % à - 20 %</b>	<b>- 10 % à - 15 %</b>	<b>- 40% et moins</b>	<b>0 % à 5 %</b>

**Graphique 25 : CAP pour la situation 6**



Les réponses à cette situation sont corrélées avec la surface habitable. Le coefficient vaut - 0,22 ( $p=0,05$ ) : plus les personnes ont une grande habitation plus leur CAP diminue.

Il existe aussi une corrélation positive entre les CAP pour la situation 6 et la situation 3 égale à 0,31 ( $p=0,00$ ).

### 5.2.6. Situation 7 (quartier « ancien »)

Pour cette situation, le bien immobilier était situé dans un quartier dit « ancien ». La densité de population était plus importante et il y avait moins d'espace vert.

Sur l'ensemble de notre échantillon, 12 personnes n'achèteraient pas cette maison. Pour les autres, la moyenne tourne autour de - 5 à - 10 % tandis que la médiane est légèrement plus

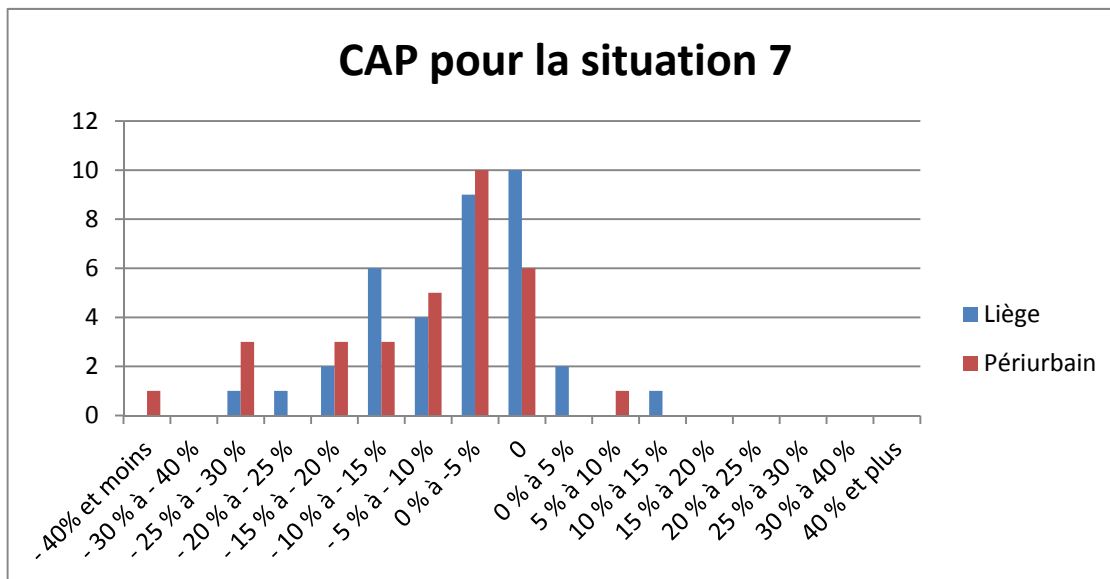
élevée. Les personnes interrogées dévaluent moins la maison dans un quartier « ancien » que la maison proche d'un parc industriel.

**Tableau 27 : Statistiques descriptives sur le CAP pour la situation 7**

	Moyenne	Médiane	Minimum	Maximum
Liège	- 5 % à - 10 %	0 % à -5 %	- 25 % à - 30 %	10 % à 15 %
Communes périurbaines	- 5 % à - 10 %	0 % à -5 %	- 40% et moins	5 % à 10 %
<b>TOTAL</b>	<b>- 5 % à - 10 %</b>	<b>0 % à -5 %</b>	<b>- 40% et moins</b>	<b>10 % à 15 %</b>
<b>TOTAL (incluant refus)</b>	<b>- 10 % à - 15 %</b>	<b>- 5 % à - 10 %</b>	<b>- 40% et moins</b>	<b>10 % à 15 %</b>

Le graphique 26 montre, que pour certaines personnes, cet environnement apporte une plus-value immobilière. Pour une majorité de Liégeois, ce quartier n'influencerait pas leur consentement à payer pour l'habitation tandis que pour les périurbains cette situation est légèrement dévalorisante. Cela n'empêche que les deux échantillons sont homogènes (t de Student :  $t=0,77$ ,  $dl=78$ ,  $p=0,44$ ).

**Graphique 26 : CAP pour la situation 7**



Cette situation est corrélée négativement à la fois avec la superficie habitable ( $r=-0,40$  et  $p=0,00$ ) et avec la superficie du terrain ( $r=-0,27$  et  $p=0,01$ ). On pourrait expliquer cette corrélation par le fait que cet environnement urbain donne la perception que les habitations sont plus petites et que les terrains sont moins grands. Malgré la signalisation dans notre enquête que le bien est identique à la première situation, les interviewés ont des difficultés à imaginer cette situation et à se rendre compte que malgré le quartier, il est possible d'avoir une grande maison avec un grand terrain.



Enfin, les résultats de la situation 7 sont aussi corrélés avec ceux de la situation 2. Le  $r$  vaut 0,2 ( $p=0,05$ ). Les personnes qui donnent les plus hautes valeurs pour une maison avec un petit terrain et 2 façades donnent aussi le plus haut prix pour une maison construite dans un quartier dense.

Au vu de ces corrélations, cela donne à penser que vivre dans un quartier dense équivaut à vivre dans une maison et un jardin plus petits.

### 5.2.7. Situation 8 (route nationale)

Dans cette situation, la maison était présentée le long d'une route nationale avec un trafic automobile assez important.

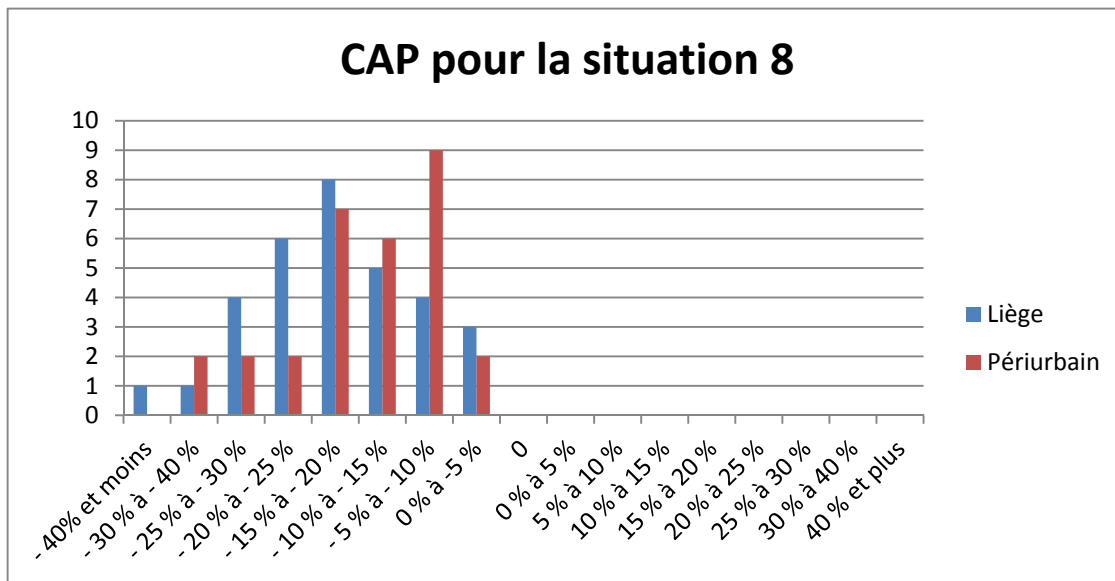
Pour presque le quart des personnes interrogées, 18 personnes au total, cette situation n'est pas acceptable. Elle est la situation qui a essuyé le plus de refus. Pour les personnes qui ont donné un prix, celui-ci est en moyenne de -15 à -20 % par rapport à la situation de référence (tableau 28). Personne n'a trouvé cette situation équivalente ou avantageuse par rapport à la situation de référence.

Tableau 28 : Statistiques descriptives sur le CAP pour la situation 8

	Moyenne	Médiane	Minimum	Maximum
Liège	- 15 % à - 20 %	- 15 % à - 20 %	- 40% et moins	0 % à -5 %
Communes périurbaines	- 10 % à - 15 %	- 10 % à - 15 %	- 30 % à - 40 %	0 % à -5 %
<b>TOTAL</b>	<b>- 15 % à - 20 %</b>	<b>- 15 % à - 20 %</b>	<b>- 40% et moins</b>	<b>0 % à -5 %</b>
<b>TOTAL (incluant refus)</b>	<b>- 20 % à - 25 %</b>	<b>- 15 % à - 20 %</b>	<b>- 40% et moins</b>	<b>0 % à -5 %</b>

Sur le graphique 27, il est observable que pour les « urbains », une moins-value de l'ordre de 15 à 20 % a été donnée par le plus de personnes. Par contre, pour les périurbains, la majorité a évalué le bien avec une baisse de 5 à 10 %. Il n'y a cependant pas d'hétérogénéité statistiquement significative entre les deux groupes ( $t$  de Student :  $t=-11,30$ ,  $df=78$ ,  $p=0,20$ ).

Graphique 27 : CAP pour la situation 8



Parmi les corrélations, trois sont significatives. Premièrement, il ressort une corrélation négative de -0,32 ( $p=0,00$ ) avec la surface habitable : plus les personnes interrogées ont une grande habitation plus elles diminuent le prix de la maison le long de la grand route.

Les autres corrélations concernent les CAP de la situation 8 avec la situation 6 (zoning) et la situation 7 (quartier ancien). Les deux corrélations sont positives et valent respectivement 0,38 ( $p=0,00$ ) et 0,43 ( $p=0,00$ ). Les personnes enquêtées ont tendance à rester sur leur position. S'ils avaient fortement dévalué une situation, ils dévalueront fortement les autres situations. Une explication à ce phénomène est que les individus interrogés comparent les situations entre elles et se servent de leurs premières évaluations pour déterminer leurs CAP suivants.

### 5.2.8. Situation 9 (périurbain)

La dernière situation représentait une maison située dans un quartier appelé plus « aéré ». Ce quartier est moins densément construit que dans les autres situations, les habitations sont plus espacées. Il symbolise l'image des quartiers périurbains actuels.

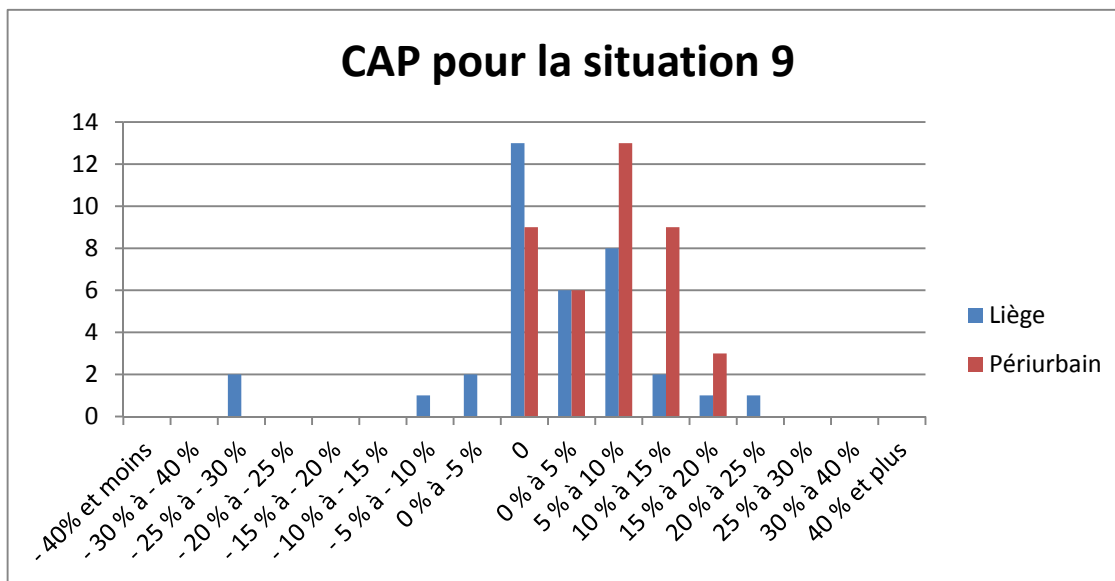
Cette situation a été rejetée par 4 personnes. Pour les autres, il y a une distinction entre l'échantillon urbain et périurbain. La moyenne et la médiane du tableau 29 pour Liège sont inférieures aux communes périurbaines. Le t de Student a d'ailleurs montré une hétérogénéité puisque le p value vaut : 0,01 ( $t=-2,62$ ,  $dl=78$ ,  $p=0,01$ ). Pour les Liégeois, cette situation est en moyenne stable par rapport à la situation de référence tandis que, pour les périurbains, elle est plus plaisante et donc plus valorisée.

Tableau 29 : Statistiques descriptives sur le CAP pour la situation 9

	Moyenne	Médiane	Minimum	Maximum
Liège	0 %	0 %	- 25 % à - 30 %	20 % à 25 %
Communes périurbaines	5 % à 10 %	5 % à 10 %	0 %	15 % à 20 %
<b>TOTAL</b>	<b>0 % à 5 %</b>	<b>0 % à 5 %</b>	<b>- 25 % à - 30 %</b>	<b>20 % à 25 %</b>
<b>TOTAL (incluant refus)</b>	<b>0 %</b>	<b>0 % à 5 %</b>	<b>- 40% et moins</b>	<b>15 % à 20 %</b>

Sur le graphique 28, nous pouvons constater que, pour 5 Liégeois, cette situation procure une moins-value. Cette maison ne correspond pas du tout à leurs critères de recherche. Tandis que pour une majorité de périurbains, cette maison a une plus-value de 5 à 10 %. Cette différence entre les deux échantillons est tout à fait compréhensible. Une grande partie de périurbains ont fait le choix de s'installer en dehors de l'agglomération liégeoise pour pouvoir acquérir une maison de ce type : 4 façades, grand terrain, environnement plus campagnard. La maison présentée correspond donc à leurs envies plus que la première, ce qui explique cette plus-value.

Graphique 28 : CAP pour la situation 9



Le test d'ANOVA-1 a montré une hétérogénéité des réponses en fonction du niveau d'étude. Le p-value équivaut à 0,049 ( $F=2,73$ ,  $p=0,049$ ). Si l'on approfondit les recherches en comparant les moyennes deux à deux à l'aide d'un test de Newman-Keuls, nous pouvons constater que deux moyennes sont différentes : la moyenne des personnes avec un diplôme universitaire avec la moyenne des personnes avec un diplôme secondaire inférieur. Il est important de réaliser ce test de comparaisons car il permet d'identifier quels groupes diffèrent entre eux. Via ce test, il est apparu que deux moyennes, diplômés universitaires et

diplômés du secondaire inférieur, étaient hétérogènes avec un p-value de 0,017. Pour les diplômés universitaires, cette situation n'apporte pas de plus-value tandis que pour les diplômés du secondaire inférieur la plus value est en moyenne de 18750 € (catégorie de 5 à 10 %).

Cependant, le niveau d'étude est dépendant de la commune de résidence. La majorité des universitaires se situent à Liège. Donc l'hétérogénéité montrée par l'ANOVA-1 est peut-être due non pas au diplôme mais à la commune de résidence qui elle aussi engendre une différence dans le CAP pour la situation 9. Nous avons réalisé une ANOVA-1 pour les deux échantillons, urbain et périurbain, afin de vérifier si, au sein de ces populations, il existait une différence en fonction de niveau d'étude. En faisant cette analyse, nous limitons encore la taille de notre échantillon. Les résultats sont des indications mais ne peuvent pas être extrapolés à la population.

Pour le périurbain, l'ANOVA-1 ( $F=0,54$ ,  $p=0,65$ ) n'indique pas de différence en fonction du niveau d'étude. Les personnes interrogées ne donnent pas un CAP différent pour une maison périurbaine. Dans l'échantillon liégeois, l'ANOVA-1 ( $F=2,66$ ,  $p=0,062$ ) ne permet pas de rejeter l'hypothèse nulle, donc il n'y a pas d'hétérogénéité des CAP. Cependant, la p-value est à la limite du seuil de 0,05. Il serait donc intéressant d'approfondir cette hétérogénéité par des enquêtes supplémentaires.

## **6. DISCUSSION GÉNÉRALE DES RÉSULTATS**

Sur l'ensemble des situations, c'est la 8 (maison le long d'une route nationale) qui est évaluée la plus négativement. Elle atteint en moyenne un minimum de -15 à - 20 %. De plus, presque un quart des personnes interrogées refuseraient d'acheter un tel bien. Les personnes évoquent comme principale problème des nuisances sonores et ne trouvent pas que la proximité immédiate de la grande route soit réellement un avantage. Notre modèle hédonique n'avait pas mis en avant cette moins-value. Certains nous ont signalé que l'on pouvait trouver un endroit plus calme tout en étant proche d'une grande route. La deuxième situation la plus négative est la présence d'un parc industriel mais une partie des personnes ne trouve pas cette situation dévalorisante. Le principal problème évoqué est une nuisance paysagère.

Les situations les plus valorisées sont la situation 5 (bonne desserte en transport en commun) et la situation 9 (quartier aéré). Dans ces situations, les plus-values envisagées sont en moyenne entre 0 à 5 %. Cependant, si on regarde par échantillon, nous pouvons noter que pour les communes périurbaines, un avantage est donné à la situation 9 et pour

## *Partie 2 : L'approche par contingence*

Liège, les situations 4 et 5 sont fort équivalentes. La situation 4, proximité du centre de Liège, est la plus ambivalente. Nous y rencontrons à la fois des personnes qui y trouvent un avantage et d'autres un désavantage. Les avantages évoqués sont surtout des économies en temps et en frais de transport et les désavantages cités sont les nuisances sonores et la qualité environnementale du quartier.

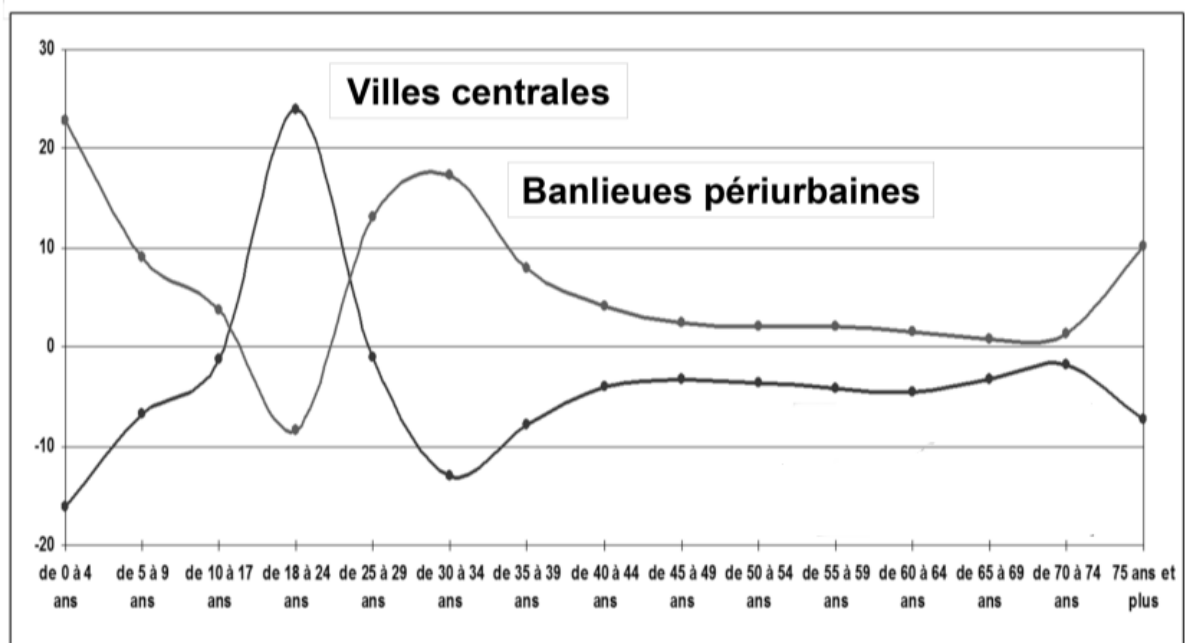
Le deuxième constat de l'analyse de ces résultats est que pour seulement deux situations (2 et 9), les CAP sont différents entre l'échantillon urbain et périurbain. Dans le premier cas, les périurbains dévalorisent le plus la maison 2 façades, construite sur un plus petit terrain. Dans le deuxième cas, ils valorisent le plus la maison construite dans un environnement périurbain typique. Ces résultats semblent donc cohérents entre le choix que les personnes ont fait il y a quelques années et leur prix d'acceptation actuel. C'est en dehors de l'agglomération qu'elles ont pu rencontrer leurs envies, construire sur un grand terrain, dans un espace aéré. Ces deux situations permettent de distinguer ces deux échantillons et de mettre en évidence les divergences de critères dans la construction d'une maison. Pour ce qui est des autres situations, les avis entre les deux échantillons sont plus homogènes.

Pour l'ensemble des situations, il existe un bon nombre de corrélations entre les CAP. Ceci peut être expliqué : les interviewés se servent de leurs évaluations antérieures pour estimer leurs CAP. Effectivement, lors des enquêtes, c'est lors de la première question que le temps de réponse était le plus long. Tandis que, plus on avançait dans l'enquête, plus les réponses semblaient venir rapidement. Comme il avait été dit dans la littérature, l'exercice demandé à l'interviewé l'expose à une situation nouvelle dans laquelle il n'a pas l'habitude de se placer. Il lui faut un certain temps d'adaptation et au fur et à mesure que l'enquête passe, l'exercice devient plus facile. Au vu de ces résultats, l'alternance des questions semble être judicieux même si nous n'avons pas constaté de différence statistiquement significative entre les réponses. Une ANOVA-1 a été réalisé qui ne permet pas de conclure à une hétérogénéité significative des réponses données selon le questionnaire utilisé.

Parmi les autres corrélations, nous en avons trouvé une intéressante entre la superficie habitable et la diminution du prix pour une maison sur un plus petit terrain. Elle montre la cohérence entre le choix qu'ont fait les individus et leurs réponses. Pour les personnes qui ont un grand terrain, ce critère est essentiel et donc un terrain plus petit est nettement moins attractif. Il existe aussi une relation entre l'augmentation de l'âge et la diminution du prix pour une maison sur un plus petit terrain. Avec l'âge, les personnes sont moins intéressées par de grands terrains. Ceux-ci demandent de l'énergie en entretien qui n'est peut être toujours pas proportionnelle à l'utilité que les personnes en retirent.

La situation 4 est corrélée à la taille des ménages. Cette constatation peut être mise en parallèle avec le graphique 29. Sur ce graphique, on constate que les banlieues sont attractives pour les jeunes familles tandis que les jeunes adultes et les personnes âgées (la taille des ménages est réduite) préfèrent la ville. Les jeunes parents quittent la ville pour des causes d'agrandissement du ménage et une insatisfaction du logement urbain. Ils désirent s'installer en banlieue où le modèle culturel du « 4 façades » est réalisable.

**Graphique 29 : Solde migratoire (%) par tranche d'âges pour les régions urbaines belges (1988-1996)**



Source : Halleux, 2005

Il existe une dépendance entre le niveau d'étude et la situation 9 représentant le modèle « 4 façades ». Il a été constaté que les diplômés universitaires ne valorisaient pas spécialement cette habitation tandis que les diplômés de secondaire inférieur valorisaient fortement ce logement. Parmi les universitaires interrogés, certains ont évoqué la limite de ce modèle. Pour eux, il est impossible de continuer à développer de telles habitations coûteuses en espace et en énergie (mauvaise isolation de l'habitation, éloignement des villes). Ce type de maison ne correspond pas à leur critères et donc ils n'accepteraient pas une plus-value sur ce bien et certains ont même dévalué le bien ou signalé qu'ils ne l'achèteraient pas. Peut-être que leur éducation leur inculque cette vision de durabilité. Les diplômés du secondaire inférieur ont, une vision plus traditionnelle de l'habitation et rêvent toujours au modèle culturel « 4 façades ».

Enfin nous n'avons constaté aucune différence de réponses selon le sexe et le revenu des personnes interrogées.

## **Partie 3: Comparaison des deux méthodes**

## INTRODUCTION

La mise en œuvre de deux méthodes d'évaluation immobilière au sein de la même région permet de comparer nos résultats en vue de répondre à notre question de recherche. La comparaison va donc être réalisée dans cette dernière partie de notre travail.

Cette partie est divisée en quatre chapitres. Le premier concerne un état de l'art sur les études comparatives existantes. Dans le deuxième chapitre, nous réalisons les statistiques de comparaisons des deux méthodes. Au troisième chapitre, nous tirons les conclusions de ces comparaisons. Pour terminer, nous exposerons les conclusions de cette partie.

## 1. DES ÉTUDES COMPARATIVES

La majorité des études n'exploite qu'une méthode mais certains articles effectuent des comparaisons. Carson et al. (1996) ont repéré 83 études qui fournissent 616 comparaisons entre les méthodes qui estiment les préférences révélées et l'évaluation par contingence. Ces études évaluent une grande variété de biens publics. Parmi ces études, 13 concernent des comparaisons (62 au total) entre la méthode hédonique et l'approche par contingence.

**Tableau 30 : Etudes comparatives entre méthode hédonique et approche par contingence reprises dans Carson et al. (1996)**

Auteurs	Biens évalués	Nombre de comparaisons
Darling (1973)	<i>Amenities at three urban lakes in California</i>	6
Loehman, Boldt, & Chaikin (1981)	<i>Changes in air quality in Los Angeles and the San Francisco Bay</i>	7
Brookshire et al. (1982)	<i>Improvements in Los Angeles air quality</i>	11
Blomquist (1984)	<i>Lake and high-rise views, Chicago</i>	14
Gegax (1984)	<i>Job-related risk reduction</i>	1
Brookshire et al. (1985)	<i>Housing locations inside and outside Los Angeles County's special earthquake study zones</i>	1
Gegax, Gerking, & Schulze (1985)	<i>Job-related risk reduction</i>	2
Blomquist (1988)	<i>Lake and high-rise views, Chicago</i>	4
IADB (1988)	<i>Three types of housing structures</i>	6



Pommerehne (1988)	<i>Road and aircraft noise in Basle, Switzerland</i>	2
d'Arge & Shogren (1989)	<i>Water quality in the Okoboji Lakes region of Iowa</i>	3
Handall & Kriesel (1990)	<i>25 percent reductions in both air and water pollution in the United States</i>	1
Shechter (1992)	<i>Air pollution in the Haifa area, Israel</i>	4

Source: Carson et al., 1996

Pour l'ensemble des comparaisons, Carson et al. (1996) ont trouvé que, de manière générale, l'approche par contingence sous-estimait les valeurs par rapport aux préférences révélées. Les valeurs obtenues par la méthode hédonique sont, en moyenne supérieures de 40 % par rapport à aux valeurs de l'approche par contingence. Cependant, certaines études montrent des résultats forts différents et Carson et al. (1996) affirment « *qu'il ne faut pas conclure que les estimations par l'approche par contingence sont toujours plus petites que les estimations des préférences révélées* ».

D'un point de vue plus théorique, l'article de Hardelin et al. (2010) pointe une grande différence entre les deux méthodes. Quand on réalise des enquêtes, on prend en considération l'ensemble des avis d'une population. Par contre, dans la méthode hédonique, la valeur est fixée par le prix du marché. Ce prix reflète la valeur maximale du consentement à payer de la population. Si le prix est inférieur au consentement à payer d'une personne, elle pourra proposer un prix plus élevé puisqu'elle considère que cette caractéristique à une plus haute valeur.

Ces différentes lectures nous ont permis de poser comme hypothèse que les valeurs obtenues par l'évaluation par contingence sont inférieurs aux valeurs obtenues par la méthode hédonique

## 2. LES COMPARAISONS

Avant de réaliser les comparaisons, nous allons expliquer les choix méthodologiques effectués pour les réaliser.

### 2.1. La méthodologie

Si l'on veut vérifier que deux échantillons indépendants sont hétérogènes, le test statistique adéquat est le t de Student. Pour ce faire, il nous faut, pour les deux échantillons, la moyenne et l'écart-type. Or, nous ne disposons pas de l'écart type pour le modèle hédonique car les variables n'ont pas été calculées pour les différentes situations.

Ne disposant dès lors que des moyennes, nous adopter la solution de Carson et al. (1996) : étudier le rapport entre la moyenne du modèle d'évaluation par contingence (CV) et la valeur fournie par le modèle hédonique (HD).

$$R = \frac{CV}{HD}$$

Nous avons aussi décidé de comparer le coefficient de régression de notre modèle hédonique avec les médianes de l'évaluation par contingence. Les médianes sont, peut être, plus proches des valeurs hédoniques car elles ne sont pas influencées par les valeurs extrêmes qui sont parfois aberrantes

## 2.2. Les résultats

Nous allons réaliser dans un premier temps les rapports pour la situation 2 et 3. Ces deux scénarii ont été conçus pour servir de points de comparaison entre les deux méthodes. C'est sur ces deux points de comparaison que nous pensons que les résultats seront presque identiques.

Pour les situations exploratoires, de la situation 4 à 9, nous avons estimé que le modèle hédonique n'était pas assez précis sur ces caractéristiques. Au point 2.2.2, nous réalisons les rapports pour les situations pour lesquelles il est possible de trouver les caractéristiques identiques dans notre modèle hédonique et dont la valeur est fiable. Les résultats pour les situations 4, 5 et 6 sont effectivement comparables aux résultats de la méthode hédonique même si, pour certaines, il n'y pas d'équivalence parfaite et que les valeurs du modèle hédonique sont moins certaines. Par contre, au point 2.2.3, nous n'avons pas de variable similaire dans le modèle hédonique pour comparer aux situations 7, 8 et 9 donc aucune comparaison n'est possible.

### 2.2.1. Situations 2 et 3

Ces situations sont représentatives de caractéristiques robustes qui sont intrinsèques aux biens immobiliers.

Le rapport entre les résultats de l'évaluation par contingence et le modèle hédonique (tableau 31) est de 1,026 avec la moyenne et 0,981 si on utilise la médiane. La moyenne de la moins-value de l'évaluation par contingence est donc de 2,5 % plus élevée que les pronostics du modèle hédonique. La médiane est quant à elle 2 % plus faible.

**Tableau 31 : Résultats finaux des deux modèles pour la situation 2**

	Modèle hédonique	Modèle contingence	
Passer de 700 m <sup>2</sup> de superficie de terrain à 300 m <sup>2</sup>	-9.484 €		
Passer de 4 façades à 2 façades	-35.121 €		
TOTAL	-44.605 €	<b>Moyenne</b>	<b>Médiane</b>
		-45.779 €	-43.750 €
Rapport (CV/HD)		1,026	0,981

**Tableau 32 : Résultats finaux des deux modèles pour la situation 3**

	Modèle hédonique	Modèle contingence	
Partant de la situation 2	-44.605 €		
Passer de 200 m <sup>2</sup> de superficie habitable à 150 m <sup>2</sup>	-21.709 €		
Passer de 4 chambres à 2 chambres	-6.313 €		
TOTAL	-72.627 €	<b>Moyenne</b>	<b>Médiane</b>
		-67.606€	-68.750 €
Rapport (CV/HD)		0,931	0,946

Dans cette situation, tableau 32, le rapport avec la moyenne et la médiane valent respectivement 0,931 et 0,946. L'évaluation par contingence sous-estime les moins-values d'environ 5 % à 7 % par rapport au modèle hédonique.

### **2.2.2. Situations 4, 5 et 6**

Dans le cas de la situation 4 (proximité de Liège), la valeur du modèle hédonique pour la distance à Liège apportait une importante plus-value au bien. Les préférences déclarées donnent une plus-value inférieure à 10.000 €. Les deux valeurs sont différentes. La plus-value est deux fois plus importante pour le modèle hédonique.

**Tableau 33 : Résultats finaux des deux modèles pour la situation 4**

	Modèle hédonique	Modèle contingence	
Passé de 20 minutes à 5 minutes du centre de Liège	17.476,95 €		
TOTAL	17.476,95 €	Moyenne	Médiane
		9.009 €	6.250 €
Rapport (CV/HD)		0,515	0,357

Pour les situations 5 et 6, les caractéristiques étudiées dans les 2 méthodes ne sont pas équivalentes. Une comparaison n'est donc guère pertinente. Nous allons cependant l'effectuer, en sachant que des conclusions sur la similitude des méthodes ne pourront en être tirées.

Pour la situation 5 (bonne accessibilité en transport en commun), nous constatons que l'évaluation par contingence conduit à la bonne desserte en transport en commun comme une plus-value. Le modèle hédonique, lui, suggère que la proximité d'un arrêt de bus, une des deux composantes de la bonne accessibilité en transport en commun (l'autre étant la fréquence), est une importante moins-value. Il y a donc une contradiction entre les deux modèles

**Tableau 34 : Résultats finaux des deux modèles pour la situation 5**

	Modèle hédonique	Modèle contingence	
		Moyenne	Médiane
Bonne déserte en transport en commun	-44.247 €	7.199 €	6.250 €

La présence d'un parc industriel en fond de jardin apporte, selon les enquêtes, une moins-value de l'ordre de 30.000 €. Le modèle hédonique estimait la caractéristique « proximité à 200 mètres d'industries ». Il n'y a donc pas forcément une vue sur les industries. Non seulement les parcs industriels sont pris en compte mais aussi d'autres industries isolées. Le montant de cette moins-value est 3 fois moindre que les résultats de l'évaluation par contingence.

**Tableau 35 : Résultats finaux des deux modèles pour la situation 6**

	Modèle hédonique	Modèle contingence	
		Moyenne	Médiane
Industrie situé à proximité	-9.621 €	-28.558 €	-31.250 €

### **2.2.3. Situations 7, 8 et 9**

Pour les situations 7 (quartier ancien), 8 (route nationale) et 9 (quartier périurbain), il n'y a pas de valeurs de comparaison présentes dans notre modèle hédonique dans lequel la densité de population, la part de bâti et la proximité d'une grande route n'étaient pas significatives dans notre modèle final. Ces analyses avaient d'ailleurs été créées pour explorer ces caractéristiques. Les seules valeurs disponibles sont celles produites par les enquêtes.

Nous pouvons quand même constater que si les caractéristiques ne sont pas significatives dans ce modèle, c'est que nous n'avons pas assez d'observations et qu'elles ont peu une influence sur le prix.

## **3. DISCUSSIONS DES COMPARAISONS**

### **3.1. Pour les variables de contrôle**

Nous pouvons constater que, dans les 2 cas, les estimations sont très identiques. On est bien en-dessous d'une variation de 10 %, ordre de grandeur de la moyenne d'erreur lors d'une évaluation immobilière. Ces résultats nous permettent de dire que les deux méthodes semblent donner des résultats similaires.

Nous ne pouvons pas nous avancer plus dans le constat car nous n'avons pas pu estimer un risque d'erreur puisqu'aucun test d'hypothèse n'a été réalisé. De plus, les deux méthodes ont été appliquées avec des échantillons assez restreints avec une erreur standard assez importante. Les résultats ne sont donc pas précis mais ils sont encourageants dans la validation de notre hypothèse de départ (expliquée au point 2 de l'introduction).

### **3.2. Pour les variables exploratoires**

Pour les variables où il a été possible de faire un point de comparaison, nous avons constaté de grandes différences entre les approches.

En ce qui concerne la distance à Liège, le modèle hédonique aboutit à une forte plus-value. Tandis que lors des enquêtes, de nombreuses personnes semblaient intéressées à vivre proche du centre-ville pour diminuer leurs déplacements, mais voyaient aussi des désavantages liés à la proximité de Liège : qualité du voisinage, nuisances sonores, qualité du paysage, coût plus élevé pour le terrain... Le prix de cette caractéristique est donc très complexe à déterminer. Nous pouvons dire que c'est une plus-value de l'ordre de quelques pourcents pour la maison mais tout dépend de l'environnement dans lequel se situera la

maison plus proche de Liège : en s'approchant de Liège, il y a un risque de dégradation de l'environnement.

La desserte en transport en commun ressort bien comme une plus-value lors de nos enquêtes et non, comme l'avait estimé le modèle hédonique, comme une moins-value. Le modèle hédonique ne permet de différencier les différentes évolutions qu'entraîne la proximité d'un arrêt de bus. Les valeurs de l'évaluation par contingence nous semblent plus en adéquation avec notre attente.

En ce qui concerne la proximité d'industries, il est plus difficile de trancher entre les deux modèles. Nous n'avons pas d'étude comparative pour avoir une idée de l'ordre de grandeur des moins-values causées par la proximité d'industrie. Nous pensons cependant que le prix auquel va se vendre un bien ayant un parc industriel dans le fond du jardin sera moins élevé que s'il n'y en a pas. Mais, comme nous l'avons vu dans les enquêtes, pour certaines personnes, ce n'est pas un inconvénient si les nuisances sont uniquement visuelles. Elles seraient prêtes à accepter des prix proches de la situation de référence. Dans ce cas-ci, nous ne pensons pas que le modèle hédonique permette de mieux tenir compte de la réalité du marché immobilier, toutes les localisations sont spécifiques.

Pour le quartier ancien et pour la route nationale, nous n'avons pas de point de comparaison. Nous pensons que nous pouvons conclure, comme pour la proximité de l'industrie, que les valeurs moyennes estimées par l'évaluation par contingence sont un minimum pour un prix de vente comme le dit l'étude de Hardelin et al. (2010). Sur le marché, il y a des personnes pour qui vivre dans un quartier d'habitat dense n'apporte pas de moins-value. Donc, la maison pourrait être vendue à un prix identique. Par contre, pour la situation le long d'une grande route, il y aura certainement une moins-value car aucune personne interrogée ne voudrait donner le même prix. Cependant, les nuisances sonores et l'insécurité liées au trafic gênent peu certaines personnes qui iraient jusqu'à accepter un prix correspondant à 95 % à la valeur de la maison.

Pour la situation d'une maison de type périurbain nous avons pu constater que c'est la situation la plus valorisée : elle aurait le prix le plus élevé sur le marché et, selon l'évaluation par contingence, de l'ordre de 10 %. Certaines personnes pourraient augmenter de 20 % par rapport au prix de la maison initiale. Mais encore faut-il qu'elles passent à l'acte !

## **4. CONCLUSIONS**

La confrontation des résultats des deux méthodes a montré une grande ressemblance pour les variables intrinsèques. De ce point de vue, les méthodes donnent des résultats

### *Partie 3 : Comparaison des deux méthodes*

identiques. Pour les autres caractéristiques, la comparaison des résultats entre les deux modèles a été moins satisfaisante. Une partie des caractéristiques n'étaient pas présentes dans les deux modèles. Nos conclusions sont que pour certaines variables, la méthode hédonique prend en compte d'autres caractéristiques et donc donne une fausse impression de l'impact de la variable. Dans l'approche par contingence, certains prix obtenus sont plus faible que ceux du modèle hédonique. Les prix de biens immobiliers peuvent être plus élevés que nos moyennes obtenues puisqu'il suffit de quelques amateurs pour que ceux-ci augmentent.

# Conclusions générales

Les résultats de la méthode hédonique ont montré une certaine robustesse en ce qui concerne les caractéristiques intrinsèques. Par rapport au modèle de Grignet (2011), il y a eu une amélioration du modèle par la prise en compte de l'hétéroscédasticité grâce à la création d'un modèle s'ajustant moins aux données extrêmes. Un dernier aspect positif de notre modèle est qu'il y a absence d'autocorrélation spatiale des résidus.

Un certain nombre de caractéristiques extrinsèques sont ressorties significatives dans notre modèle. Cependant, la valeur qui leur a été attribuée ne semble pas toujours correspondre aux hypothèses attendues. Les variables d'accessibilité, qui était le deuxième objectif de la réalisation d'un modèle hédonique avec des données identiques à celle de Grignet (2011), n'ont pas apporté l'amélioration escomptée. De plus, les résidus finaux montrent une certaine croissance par rapport aux prix de mise en vente.

Pour améliorer le modèle, il serait intéressant de mettre en application la méthode hédonique pour des sous-marchés plus restrictifs que celui que l'on a utilisé. Il faut limiter l'échantillon à un espace géographique plus homogène et à des prix moins étendus. Cet échantillon n'est possible que si la source de données en fournit plus sur les biens mis en vente (adresse, caractéristiques intrinsèques).

Au vu des résultats de notre modèle hédonique, l'enquête a été conçue pour valider les variables les plus fortes et explorer des caractéristiques environnementales. Sa réalisation s'est faite à Liège et dans les communes de Juprelle et Remicourt pour avoir une certaine diversité de la population de la région liégeoise.

Un questionnaire illustré a été conçu pour permettre d'être le plus explicite. Nous avons aussi tenté de limiter les biais en inversant l'ordre des questions et l'ordre des prix dans l'échelle des valeurs proposées à l'interviewé. Un protocole d'enquête a été réalisé afin de limiter les différences entre les enquêteurs et leur influence sur les personnes interrogées.

Les principaux résultats de l'évaluation par contingence sont les suivants :

- Hétérogénéité des CAP entre urbains et périurbains pour les petits terrains et les maisons 4 façades ;
- Les personnes âgées dévaluent moins les maisons construites sur un petit terrain ;
- La situation la plus contraignante est le long de la route nationale ;
- Les résultats sont assez divergents pour la proximité à Liège ;



## *Conclusions générales*

- Les maisons éloignées du centre-ville sont plus valorisées par les familles ;
- Les répondants se servent de leurs réponses antérieures pour évaluer les nouvelles situations. Ceci entraîne de nombreuses corrélations entre les prix des différentes situations.

L'enquête n'a cependant été réalisée qu'auprès de 80 personnes et dans 3 communes différentes. Les résultats devraient être approfondis par une augmentation de la taille de l'échantillon.

Les deux méthodes présentent des avantages et des inconvénients mais on a pu remarquer que le désavantage principal de ces méthodes porte sur la difficulté de mesurer les caractéristiques environnementales. Il est souvent impossible d'isoler une dimension, que ce soit dans la méthode hédonique ou dans l'approche par contingence. La caractéristique est difficilement appréhendable en faisant totalement abstraction du contexte.

La comparaison réalisée entre les deux méthodes donne des résultats distincts selon les caractéristiques. Pour les variables structurelles, les valeurs sont identiques tandis que pour les variables extrinsèques, les préférences déclarées donnent une estimation inférieure aux préférences révélées. Ce constat rejoint notre hypothèse qui était que les valeurs obtenues par l'évaluation par contingence sont inférieures aux valeurs obtenues par la méthode hédonique. Les préférences révélées reflètent le prix pratiqué sur le marché tandis que les préférences déclarées sont le résultat de l'avis de l'ensemble d'une population qui n'est pas forcément intéressé par le bien immobilier présenté. Le prix de vente est l'aboutissement de la manifestation de certains amateurs.

# Bibliographie

- Albert A. (2005). *Biostatistique* (5<sup>e</sup> éd). Liège : Les Editions de l'Université de Liège.
- Bureau B. & Glachant, M. (2010). Évaluation de l'impact des politiques. *Economie & prévision*, (1), 27–44.
- Carson R. T., Flores N. E., Martin, K. M. & Wright, J. L. (1996). Contingent valuation and revealed preference methodologies: comparing the estimates for quasi-public goods. *Land Economics*, 80–99.
- Cervero R. & Kang, C. D. (2011). Bus rapid transit impacts on land uses and land values in Seoul, Korea. *Transport Policy*, 18(1), 102–116.
- Choumert J. & Travers, M. (2010). La capitalisation immobilière des espaces verts dans la ville d'Angers. *Revue économique*, 61(5), 821. doi:10.3917/reco.615.0821
- Ciriacy-Wantrup S. V. (1947). Capital returns from soil-conservation practices. *Journal of Farm Economics*, 29(4), 1181–1196.
- Cornet Y., Daxhelet D., Halleux J.-M., Klinkenberg A.-C. & Lambotte J.-M. (2005). Cartographie de l'accessibilité par les alternatives à la voiture. In F. Witlox & V. Van Acker (Eds.), *Les Journées Géographiques Belges. Mobilité, Société et Environnement en Cartes* (pp. 355-364). Zelzate : DCL Print & Sign . URL : <http://orbi.ulg.ac.be/handle/2268/3620>
- Cottrell A. & Lucchetti R. (2012). *Gretl User's Guide*. En ligne : <http://gretl.sourceforge.net/#man>
- Davis R. (1963). The value of outdoor recreation: an economic study of the Maine Woods. *Doctoral Dissertation in Economics*, Harvard University, Cambridge, MA.
- del Saz Salazar S. & Garcia Menendez L. (2007). Estimating the non-market benefits of an urban park: Does proximity matter? *Land Use Policy*, 24(1), 296–305.
- Direction Mobilité (2010). Diagnostique des déplacements domicile - travail 2008, rapport final. *Service public fédéral Mobilité et Transports*, 46.
- Donnay J.-P. (2010). *Analyse spatiale*. Notes de cours, Université de Liège, Faculté des Sciences, Département de Géographie, Liège, inédit.
- Evans A. W. (1995). The property market: ninety per cent efficient? *Urban Studies*, 32(1), 5.

## Bibliographie

- Des Rosiers F. (2001). *La modélisation statistique en analyse et évaluation immobilières: Guide méthodologique*. Université Laval, Québec.
- Gao X. & Asami Y. (2007). Effect of urban landscapes on land prices in two Japanese cities. *Landscape and urban planning*, 81(1), 155–166.
- Geoghegan J., Wainger L. A. & Bockstael N. E. (1997). Spatial landscape indices in a hedonic framework: an ecological economics analysis using GIS. *Ecological economics*, 23(3), 251–264.
- Gibbons S. & Machin S. (2005). Valuing rail access using transport innovations. *Journal of Urban Economics*, 57(1), 148–169.
- Grelot F. (2004). *Gestion collective des inondations. Peut-on tenir compte de l'avis de la population dans la phase d'évaluation économique a priori?* Thèse de doctorat en Sciences Économiques, Ecole doctorale de l'ENSAM, Paris. URL : <http://tel.archives-ouvertes.fr/docs/00/42/03/88/PDF/GrelotF2004b.pdf>
- Grignet Y. (2011). *Approche hédonique du marché immobilier sur base d'annonces internet: les maisons d'habitation au sein de la région urbaine de Liège*. Mémoire de master en Sciences Géographiques. Université de Liège, Liège, inédit.
- Halleux J.M. (2005). *Structuration spatiale des marchés fonciers et production de l'urbanisation. Application à la Belgique et à ses nouveaux espaces résidentiels*. Thèse de doctorat en Sciences Géographiques. Université de Liège, Liège. URL : <http://orbi.ulg.ac.be/handle/2268/63714>
- Halleux J.M. (2005). Mobilité résidentielle, mobilité voiture et logiques familiales en Belgique : entre étalement de la périurbanisation et intraurbanisation. *NetCom-Réseaux-Communications-Territoires*, 19(3), 157-177
- Halleux J.M. (2010). *Développement et aménagement des territoires* (Note de cours). Notes de cours, Université de Liège, Faculté des Sciences, Département de Géographie, Liège, inédit.
- Halleux J.M. (2010). *Marchés fonciers et immobiliers*. Notes de cours, Université de Liège, Faculté des Sciences, Département de Géographie, Liège, inédit.
- Halleux J.-M. & Strée, J. (2012). Les éco-quartiers : vers la généralisation d'une alternative à l'étalement urbain ? In P. Harismendy & S. Gaudin (Eds.), *Habitat durable, habitat social ?* Saint-Brieuc : Ville de Saint-Brieuc. URL : <http://orbi.ulg.ac.be/handle/2268/113555>

## Bibliographie

- Hardelin J., Katosky A. & Marical F. (2010). Pourquoi et comment monétariser l'Environnement ? *La Revue du CGDD*, 13.
- Kong F., Yin, H. & Nakagoshi, N. (2007). Using GIS and landscape metrics in the hedonic price modeling of the amenity value of urban green space: A case study in Jinan City, China. *Landscape and Urban Planning*, 79(3-4), 240-252. doi:10.1016/j.landurbplan.2006.02.013
- Lancelot R. & Lesnoff M. (2005). Sélection de modèles avec l'AIC et critères d'information dérivés. *CIRAD, Montpellier*.
- Lejeune B. (2011). *Econométrie I*. Notes de cours, Université de Liège, HEC, Liège.
- Loiseau V. (2011). *Toits verts et rénovation urbaine : Monétarisation et internalisation de l'externalité paysagère*. Mémoire de master en Sciences Géographiques. Université de Liège, Liège, inédit.
- Pérée F.-P. (2005). *Problèmes statistiques et utilisation de logiciels*. Notes de cours, Université de Liège, Faculté de Psychologie et de Sciences de l'Education, Liège, inédit.
- Rodrigue J.-I. (1998). *Site Web Géographie des Transports*. Hofstra University: Department of Economics and Geography.
- Rosen S. (1974). Hedonic prices and implicit markets: product differentiation in pure competition. *The Journal of Political Economy*, 82(1), 34–55.
- Schmitz S. (2011). Analyse et gestion des paysages. Notes de cours, Université de Liège, Faculté des Sciences, Département de Géographie, Liège, inédit.
- Allin S. & Henneberry J. (2010). *Action 1.2 Baseline Analysis of Existing Economic Valuation Tools for Application to Green Infrastructure Investments ( No. Final Report) (p. 87)*. Department of Town and Regional Planning, University of Sheffield.
- Strée J. (2010). *Les écoquartiers en Wallonie: étude de l'offre et approche de la demande*. Mémoire de master en Sciences Géographiques. Université de Liège, Liège, inédit.
- Terra S. (2005). Guide de bonnes pratiques pour la mise en œuvre de la méthode des prix hédoniques. Document de travail, MEDD, D4E.
- Theebe M. A. J. (2004). Planes, trains, and automobiles: the impact of traffic noise on house prices. *The Journal of Real Estate Finance and Economics*, 28(2), 209–234.
- Vandermotten C., Istaz D. & Marissal P. (1996). Marchés fonciers et immobiliers acquisitifs et locatifs : 1994-1995. *Etudes et documents logements*, 2.

## Bibliographie

Williams R. (2011). *Heteroscedasticity*. University of Notre Dame,

Yusuf, A. A., & Resosudarmo, B. P. (2009). Does clean air matter in developing countries' megacities? A hedonic price analysis of the Jakarta housing market, Indonesia. *Ecological Economics*, 68(5), 1398–1407.

Zabel, J. E., & Guignet, D. (2012). A hedonic analysis of the impact of LUST sites on house prices. *Resource and Energy Economics*, 34(4), 549-564. doi:10.1016/j.reseneeco.2012.05.006

### Sites internet:

ArcGIS Desktop Help : <http://webhelp.esri.com/arcgisdesktop/9.2/index.cfm?TopicName=welcome>, consulté entre avril et juillet 2012

Immoweb : <http://www.immoweb.be/fr/>, consulté en avril et mai 2012

Packtravaux : <http://www.packtravaux.com>, consulté en mai 2012

Statistica : <http://www.statsoft.fr/index.php>, consulté de mai à juillet 2012

# Annexes

## ANNEXE 1 : CODES DES PROGRAMMES PYTHON

### Répétition de la fonction « *cost distance* » d'ArcGIS

```
import sys, string, os, arcgisscripting

gp = arcgisscripting.create()
gp.CheckOutExtension("spatial")
gp.AddToolbox("C:\Program Files\ArcGIS\ArcToolBox\Toolboxes\Spatial Analyst Tools.tbx")

i = 0

while (i<=547):
    InSourceRaster =
        "C:\Users\Perrine\Documents\unif\master\memoire\Projet_ArcGis\Accessibilité-
        potentiel\BIENS\Observations"+str(i)+".shp"
    InCostRaster =
        "C:\Users\Perrine\Documents\unif\master\memoire\Projet_ArcGis\Accessibilité-
        potentiel\STREET\cosstreets1.img"
    OutDistanceRaster =
        "C:\Users\Perrine\Documents\unif\master\memoire\Projet_ArcGis\Accessibilité-
        potentiel\DIST_BIENS\cost_dist"+str(i)
    InMaxDistance = "36000000"

    gp.CostDistance_sa(InSourceRaster, InCostRaster, OutDistanceRaster,
        InMaxDistance)

    i = i + 1
```

## Répétition de la fonction « *extract values to points* » d'ArcGIS

```
import sys, string, os, arcgisscripting

gp = arcgisscripting.create()
gp.CheckOutExtension("spatial")
gp.AddToolbox("C:\Program Files\ArcGIS\ArcToolBox\Toolboxes\Spatial Analyst Tools.tbx")

i = 0

while (i<=547):
    InPointsFC =
        "C:\Users\Perrine\Documents\unif\master\memoire\Projet_ArcGis\Accessibilité-
        potentiel\Emploi_centre.shp"
    InRaster =
        "C:\Users\Perrine\Documents\unif\master\memoire\Projet_ArcGis\Accessibilité-
        potentiel\DIST_BIENS\cost_dist"+str(i)
    OutPointsFC =
        "C:\Users\Perrine\Documents\unif\master\memoire\Projet_ArcGis\Accessibilité-
        potentiel\DIST_POINTS\Dist_P_O"+str(i)+".shp"

    gp.ExtractValuesToPoints_sa(InPointsFC, InRaster, OutPointsFC, "NONE",
        "VALUE_ONLY")

    i = i + 1
```

## Répétition de la fonction « *tabualte area* » d'ArcGIS

```
import sys, string, os, arcgisscripting

gp = arcgisscripting.create()
gp.CheckOutExtension("spatial")
gp.AddToolbox("C:\Program Files\ArcGIS\ArcToolBox\Toolboxes\Spatial Analyst Tools.tbx")

i = 0
# changer dans la condition while le i en fct du nombre de fichier shp à traiter

while (i<=547):
    BUFFER =
    "C:\Users\Perrine\Documents\unif\master\memoire\Projet_ArcGis\SHANNON-
    COSW\Buffer\Buffer"+str(i)+".shp"
    Landcover_LB72 =
    "C:\Users\Perrine\Documents\unif\master\memoire\Projet_ArcGis\Donnees\Occup_so
    I\GRID\OCCSOL"
    Tab_BUFFER =
    "C:\Users\Perrine\Documents\unif\master\memoire\Projet_ArcGis\SHANNON-
    COSW\TABULATE_AREA\TAB"+str(i)
    cellSize = 20
    inField1 = "CLE"
    inField2 = "VALUE"

    gp.TabulateArea_sa(BUFFER, inField1, Landcover_LB72, inField2, Tab_BUFFER,
    cellSize)

    i = i + 1
```



## ANNEXE 2 : MATRICES DE CORRÉLATION

Coefficients de corrélation de Spearman

Variable	Corrélations (RESULTATS) Corrélations significatives marquées à $p < ,05000$ N=548 (Observations à VM ignorées)								
	PRIX	SUP_TERR	SUP_HAB	AGE	MOY_R_MED	Const_81__	DENSITE_SS	NIVEAU_INSTRUC UC	ACC_EMPLOI
PRIX	1,000000	0,577611	0,648148	-0,474023	0,424879	0,415260	-0,108268	0,444778	-0,213004
SUP_TERR	0,577611	1,000000	0,307967	-0,181406	0,310855	0,304826	-0,142155	0,257038	-0,382139
SUP_HAB	0,648148	0,307967	1,000000	-0,248421	0,187641	0,203826	-0,021738	0,239782	-0,051293
AGE	-0,474023	-0,181406	-0,248421	1,000000	-0,273901	-0,270310	0,063004	-0,209041	0,107339
MOY_R_MED	0,424879	0,310855	0,187641	-0,273901	1,000000	0,646508	-0,167540	0,789023	-0,427715
Const_81__	0,415260	0,304826	0,203826	-0,270310	0,646508	1,000000	-0,121513	0,456488	-0,335623
DENSITE_SS	-0,108268	-0,142155	-0,021738	0,063004	-0,167540	-0,121513	1,000000	-0,114131	0,218760
NIVEAU_INSTRUC UC	0,444778	0,257038	0,239782	-0,209041	0,789023	0,456488	-0,114131	1,000000	-0,151869
ACC_EMPLOI	-0,213004	-0,382139	-0,051293	0,107339	-0,427715	-0,335623	0,218760	-0,151869	1,000000
ACC_EMPL2	-0,155788	-0,231288	-0,044351	0,104710	-0,311930	-0,263540	0,083824	-0,070234	0,677216
DIST_LIEGE	0,170664	0,381389	0,044445	-0,056104	0,344839	0,211981	-0,217085	0,193188	-0,835882
ACC_BUS	-0,065057	-0,079774	0,048653	0,138207	-0,159341	-0,103102	0,070461	0,017233	0,314123
ACC_ML	-0,285148	-0,347675	-0,053860	0,203383	-0,519140	-0,420536	0,128928	-0,334647	0,364841
ACC_TRAI	-0,065687	-0,053083	-0,001880	0,069867	-0,041622	-0,068530	-0,104338	0,053638	0,086586
DIST_GUILL	0,194513	0,380812	0,069348	-0,041055	0,320580	0,274021	-0,167371	0,159633	-0,821134
ACC_NODULE	-0,150046	-0,351667	-0,028548	0,018101	-0,317818	-0,160665	0,164437	-0,138836	0,853029
SHANNON_10	0,197700	0,227149	0,128744	-0,033861	0,287435	0,273999	-0,110316	0,254163	-0,234750
SHANNON_50	0,161501	0,052200	0,119179	-0,141431	0,149384	0,019734	0,072567	0,169803	-0,038605
Part_batit_10	-0,363143	-0,460417	-0,153323	0,135675	-0,504209	-0,495552	0,211413	-0,369284	0,482987
PAYSAGE	0,057586	0,089806	0,020716	0,023955	0,056700	-0,074874	0,049258	0,120751	-0,103855

Annexes

Variable	Corrélations (RESULTATS)										
	Corrélations significatives marquées à $p < ,05000$ N=548 (Observations à VM ignorées)										
	ACC_EMPL2	DIST_LIEGE	ACC_BUS	ACC_ML	ACC_TRAI	DIST_GUILL	ACC_NODULE	SHANNON_10	SHANNON_50	Part batit_10	PAYSAGE
PRIX	-0,155788	0,170664	-0,065057	-0,285148	-0,065687	0,194513	-0,150046	0,197700	0,161501	-0,363143	0,057586
SUP_TERR	-0,231288	0,381389	-0,079774	-0,347675	-0,053083	0,380812	-0,351667	0,227149	0,052200	-0,460417	0,089806
SUP_HAB	-0,044351	0,044445	0,048653	-0,053860	-0,001880	0,069348	-0,028548	0,128744	0,119179	-0,153323	0,020716
AGE	0,104710	-0,056104	0,138207	0,203383	0,069867	-0,041055	0,018101	-0,033861	-0,141431	0,135675	0,023955
MOY_R_MED	-0,311930	0,344839	-0,159341	-0,519140	-0,041622	0,320580	-0,317818	0,287435	0,149384	-0,504209	0,056700
Const_81__	-0,263540	0,211981	-0,103102	-0,420536	-0,068530	0,274021	-0,160665	0,273999	0,019734	-0,495552	-0,074874
DENSITE_SS	0,083824	-0,217085	0,070461	0,128928	-0,104338	-0,167371	0,164437	-0,110316	0,072567	0,211413	0,049258
NIVEAU_INSTRUC	-0,070234	0,193188	0,017233	-0,334647	0,053638	0,159633	-0,138836	0,254163	0,169803	-0,369284	0,120751
ACC_EMPLOI	0,677216	-0,835882	0,314123	0,364841	0,086586	-0,821134	0,853029	-0,234750	-0,038605	0,482987	-0,103855
ACC_EMPL2	1,000000	-0,423360	0,296120	0,319153	0,144434	-0,424692	0,489719	-0,091306	-0,004407	0,285481	-0,018182
DIST_LIEGE	-0,423360	1,000000	-0,183605	-0,208371	0,085925	0,886065	-0,825243	0,252387	0,066700	-0,452412	0,142773
ACC_BUS	0,296120	-0,183605	1,000000	0,211301	0,000065	-0,133350	0,264499	-0,077064	-0,062979	0,118891	-0,090308
ACC_ML	0,319153	-0,208371	0,211301	1,000000	0,220919	-0,186483	0,268392	-0,380496	-0,121364	0,622134	-0,172894
ACC_TRAI	0,144434	0,085925	0,000065	0,220919	1,000000	0,050571	0,068063	0,105871	-0,034607	0,055948	-0,107176
DIST_GUILL	-0,424692	0,886065	-0,133350	-0,186483	0,050571	1,000000	-0,780326	0,267789	0,110685	-0,470109	0,083597
ACC_NODULE	0,489719	-0,825243	0,264499	0,268392	0,068063	-0,780326	1,000000	-0,247311	-0,110093	0,437019	-0,215188
SHANNON_10	-0,091306	0,252387	-0,077064	-0,380496	0,105871	0,267789	-0,247311	1,000000	0,103564	-0,624779	0,063334
SHANNON_50	-0,004407	0,066700	-0,062979	-0,121364	-0,034607	0,110685	-0,110093	0,103564	1,000000	-0,087755	0,325665
Part batit_10	0,285481	-0,452412	0,118891	0,622134	0,055948	-0,470109	0,437019	-0,624779	-0,087755	1,000000	-0,067240
PAYSAGE	-0,018182	0,142773	-0,090308	-0,172894	-0,107176	0,083597	-0,215188	0,063334	0,325665	-0,067240	1,000000

Annexes

Coefficients de corrélation bisérial de point

	A_RENOVE	BON_OU_EXC	R_A_NEUF	N3FAC	N4FAC	NOETA	N2_ETA	N1_GAR	NOSDB	N2_SDB	N1_SDD
PRIX	-0,271	0,224	-0,057	-0,055	0,565	0,055	-0,182	0,437	-0,175	0,223	0,099
SUP_TERRAIN	-0,112	0,165	-0,047	-0,108	0,569	0,092	-0,217	0,329	-0,060	0,095	0,071
SUP_HAB	-0,152	0,111	-0,089	0,040	0,259	-0,057	0,061	0,291	-0,153	0,221	0,136
AGE	0,118	-0,184	0,174	0,037	-0,374	-0,204	0,259	-0,437	0,133	-0,065	0,021
MOY_R_MED	-0,047	0,156	-0,118	-0,046	0,417	0,250	-0,267	0,321	-0,011	-0,003	-0,010
Const_81__	-0,132	0,150	-0,099	-0,074	0,368	0,194	-0,213	0,221	-0,038	0,062	0,021
DENSITE_SS	-0,025	-0,021	0,069	-0,062	-0,091	-0,056	0,068	-0,109	0,064	0,021	0,003
NIVEAU_INSTRUC	-0,048	0,110	-0,079	-0,091	0,344	0,125	-0,093	0,224	-0,064	0,038	0,002
ACC_EMPLOI	-0,020	-0,090	0,046	-0,080	-0,423	-0,140	0,342	-0,284	0,006	0,063	0,076
ACC_EMPL2	-0,013	-0,055	0,008	-0,092	-0,209	0,000	0,289	-0,202	0,010	0,034	0,031
DIST_LIEGE	-0,017	0,097	-0,042	0,025	0,409	0,111	-0,213	0,215	-0,035	-0,071	-0,063
ACC_BUS	-0,007	-0,063	0,033	-0,058	-0,094	-0,037	0,180	-0,127	-0,001	-0,034	0,060
ACC_ML	0,042	-0,077	0,012	0,049	-0,378	-0,176	0,259	-0,257	-0,026	-0,026	-0,021
ACC_TRAI	-0,045	0,076	-0,051	0,076	-0,102	-0,115	0,148	-0,055	-0,052	-0,027	-0,034
DIST_GUILL	-0,030	0,075	-0,047	0,035	0,409	0,109	-0,219	0,234	-0,053	-0,053	-0,068
ACC_NODULE	-0,051	-0,064	0,027	-0,052	-0,399	-0,125	0,242	-0,201	-0,021	0,036	0,065
SHANNON_10	-0,062	0,085	-0,054	-0,019	0,258	0,083	-0,011	0,162	-0,025	-0,011	0,014
SHANNON_50	0,003	0,039	-0,006	-0,026	0,152	0,073	0,000	0,153	-0,024	0,012	0,030
Part_batit_10	0,079	-0,141	0,048	0,062	-0,477	-0,180	0,172	-0,282	0,043	-0,003	-0,042
PAYSAGE	0,060	0,028	-0,014	-0,035	0,079	0,022	-0,010	0,037	0,015	0,002	-0,028

Annexes

	N1CHB	N2CHB	N4_CHB	N5_CHB	S_EQ	EQ	CAVE	GREN	TERR	JARD	VERANDA
PRIX	-0,173	-0,288	0,256	0,240	-0,109	0,299	-0,063	0,007	0,341	0,318	0,199
SUP_TERRAIN	-0,046	-0,104	0,126	0,178	-0,078	0,140	-0,114	0,021	0,182	0,250	0,125
SUP_HAB	-0,151	-0,386	0,308	0,359	-0,033	0,157	0,054	0,050	0,229	0,202	0,144
AGE	0,111	0,099	-0,061	-0,007	0,050	-0,147	0,113	0,075	-0,184	-0,233	-0,115
MOY_R_MED	-0,067	-0,060	0,083	-0,044	-0,079	0,142	-0,153	0,002	0,093	0,255	0,112
Const_81__	-0,030	-0,067	0,088	0,011	-0,073	0,171	-0,190	0,004	0,103	0,196	0,161
DENSITE_SS	-0,025	0,035	0,026	0,060	0,040	-0,043	0,048	-0,075	-0,041	-0,067	-0,006
NIVEAU_INSTRUC	-0,077	-0,087	0,142	0,063	-0,031	0,063	-0,098	0,004	0,031	0,164	0,093
ACC_EMPLOI	0,012	0,006	0,019	0,060	0,060	-0,129	0,138	-0,097	-0,108	-0,227	-0,076
ACC_EMPL2	-0,009	0,021	0,024	0,075	0,033	-0,153	0,081	-0,079	-0,122	-0,154	-0,068
DIST_LIEGE	-0,026	-0,009	-0,015	-0,002	-0,031	0,069	-0,152	0,093	0,063	0,158	0,042
ACC_BUS	-0,022	-0,024	0,107	-0,014	0,149	-0,103	0,027	-0,034	-0,084	-0,167	-0,019
ACC_ML	-0,032	0,036	-0,003	0,017	0,095	-0,099	0,137	-0,074	-0,085	-0,194	-0,120
ACC_TRAI	0,023	0,005	-0,047	0,008	0,052	-0,042	0,070	0,059	0,007	0,027	0,013
DIST_GUILL	-0,024	-0,055	0,022	-0,029	-0,045	0,112	-0,153	0,083	0,094	0,172	0,061
ACC_NODULE	-0,011	0,000	0,002	0,020	0,062	-0,077	0,104	-0,083	-0,073	-0,165	-0,064
SHANNON_10	0,045	-0,079	0,046	-0,001	-0,076	0,039	-0,114	0,089	0,062	0,090	0,099
SHANNON_50	-0,070	-0,069	0,034	0,157	-0,112	0,123	-0,060	0,050	0,088	0,032	0,046
Part_batit_10	0,002	0,063	-0,008	-0,037	0,056	-0,079	0,221	-0,056	-0,112	-0,193	-0,128
PAYSAGE	-0,014	0,059	0,026	0,065	0,017	-0,004	-0,048	0,002	0,006	0,015	0,033

Annexes

	PISCINE	BUANDERI	B1_BUS	B5_BUS	B7_BUS	B1_COM	B5_COM	B7_COM	B1_EXTRACT	B2_EXTRACT	B5_EXTRACT
PRIX	0,291	0,218	-0,148	-0,156	-0,160	-0,026	-0,160	-0,169	-0,019	0,078	-0,046
SUP_TERRAIN	0,185	0,114	-0,087	-0,204	-0,155	-0,052	-0,225	-0,288	0,089	0,171	0,053
SUP_HAB	0,073	0,211	-0,014	-0,013	-0,046	-0,004	-0,013	0,017	-0,023	0,047	-0,059
AGE	-0,082	-0,181	0,128	0,055	0,016	0,079	0,052	0,026	-0,045	-0,080	0,025
MOY_R_MED	0,032	0,098	-0,156	-0,196	-0,078	-0,108	-0,334	-0,360	0,005	0,005	-0,062
Const_81__	0,050	0,130	-0,158	-0,233	-0,140	-0,041	-0,283	-0,327	0,026	0,039	-0,031
DENSITE_SS	-0,026	-0,085	0,053	0,031	0,045	0,133	0,069	0,066	0,008	-0,016	-0,060
NIVEAU_INSTRUC	0,065	0,071	-0,072	-0,073	-0,033	-0,022	-0,155	-0,206	-0,040	-0,051	-0,138
ACC_EMPLOI	-0,037	-0,098	0,101	0,164	0,104	0,144	0,316	0,387	-0,028	-0,041	-0,104
ACC_EMPL2	-0,037	-0,082	0,066	0,114	0,064	0,142	0,305	0,310	-0,029	-0,039	-0,079
DIST_LIEGE	0,029	0,053	-0,049	-0,137	-0,071	-0,092	-0,207	-0,292	0,014	0,021	0,126
ACC_BUS	-0,017	-0,084	0,106	0,036	0,018	0,217	0,158	0,137	-0,011	-0,016	-0,039
ACC_ML	-0,109	-0,072	0,155	0,242	0,168	0,136	0,478	0,575	-0,012	-0,050	-0,017
ACC_TRAI	-0,063	0,014	-0,030	0,100	0,061	-0,067	0,187	0,197	-0,055	-0,056	-0,073
DIST_GUILL	0,026	0,087	-0,065	-0,161	-0,073	-0,055	-0,215	-0,320	-0,014	-0,023	0,049
ACC_NODULE	0,006	-0,072	0,023	0,134	0,108	0,085	0,275	0,363	-0,035	-0,051	-0,129
SHANNON_10	0,021	0,110	-0,060	-0,084	-0,060	-0,046	-0,260	-0,317	0,086	0,098	-0,030
SHANNON_50	0,019	0,087	-0,008	-0,020	-0,036	0,020	-0,094	-0,094	-0,013	-0,022	-0,073
Part_batit_10	-0,102	-0,109	0,141	0,210	0,136	0,069	0,359	0,467	-0,041	-0,085	0,039
PAYSAGE	0,027	-0,028	0,049	-0,084	-0,067	-0,003	-0,108	-0,161	0,039	0,062	0,066

Annexes

	B1_INDU	B2_INDU	B5_INDU	B1_SAED	B2_SAED	B5_SAED	B1_VERT2	B5_VERT2	B7_VERT2
PRIX	-0,134	-0,127	-0,140	-0,106	-0,162	-0,238	0,151	-0,016	-0,060
SUP_TERRAIN	-0,093	-0,125	-0,144	-0,051	-0,094	-0,230	0,176	0,009	-0,028
SUP_HAB	-0,050	0,006	-0,036	-0,094	-0,145	-0,122	0,063	-0,023	-0,042
AGE	0,020	0,024	0,046	0,024	0,083	0,109	-0,054	-0,038	0,025
MOY_R_MED	-0,150	-0,139	-0,201	-0,094	-0,203	-0,393	0,155	0,055	0,009
Const_81__	-0,075	-0,081	-0,187	-0,041	-0,106	-0,209	0,076	-0,045	-0,051
DENSITE_SS	-0,002	-0,025	-0,137	0,099	0,124	0,092	0,001	-0,009	0,063
NIVEAU_INSTRUC	-0,174	-0,164	-0,196	-0,160	-0,234	-0,353	0,138	0,023	-0,057
ACC_EMPLOI	0,089	0,152	0,210	0,019	0,099	0,300	-0,211	-0,075	-0,039
ACC_EMPL2	0,079	0,114	0,194	-0,032	0,024	0,182	-0,161	-0,115	-0,107
DIST_LIEGE	-0,065	-0,103	-0,142	-0,067	-0,145	-0,289	0,165	0,033	0,027
ACC_BUS	0,012	0,005	-0,022	-0,024	0,041	0,122	-0,124	-0,133	-0,189
ACC_ML	0,016	0,041	0,147	-0,013	0,045	0,238	-0,189	-0,083	-0,003
ACC_TRAI	0,176	0,207	0,274	-0,053	-0,025	0,053	-0,032	0,017	-0,008
DIST_GUILL	-0,093	-0,138	-0,165	-0,056	-0,129	-0,278	0,163	-0,014	-0,005
ACC_NODULE	0,089	0,144	0,214	0,042	0,114	0,289	-0,195	-0,054	-0,068
SHANNON_10	0,131	0,087	-0,030	0,040	-0,086	-0,122	0,413	-0,017	-0,039
SHANNON_50	-0,079	-0,107	-0,111	-0,090	-0,131	-0,146	0,006	0,014	-0,020
Part_batit_10	0,010	0,110	0,194	-0,017	0,103	0,273	-0,271	0,040	0,078
PAYSAGE	-0,100	-0,116	-0,113	0,057	0,067	-0,063	0,186	0,252	0,206

## ANNEXE 3 : LES MODÈLES HÉDONIQUES

### Modèle 1

Modèle 1: MCO, utilisant les observations 1-548  
Variable dépendante: PRIX

	<i>Coefficient</i>	<i>Erreur Std</i>	<i>t de Student</i>	<i>p. critique</i>	
const	57885,6	28150,7	2,0563	0,04025	**
A_RENOVE	-18178,7	6819,66	-2,6656	0,00792	***
BON_OU_EXC	7864,67	4094,35	1,9209	0,05529	*
R_A_NEUF	12795	6170,23	2,0737	0,03860	**
log_terrain	41406,2	5601,88	7,3915	<0,00001	***
SUP_HAB	426,07	36,2502	11,7536	<0,00001	***
N4FAC	29070,3	5112,88	5,6857	<0,00001	***
N0ETA	-23448	7489,44	-3,1308	0,00184	***
N0SDB	-13187,2	7047,1	-1,8713	0,06186	*
N2_SDB	26882,5	6358,25	4,2280	0,00003	***
N1CHB	-22002,5	11042,7	-1,9925	0,04683	**
N4_CHB	7926,52	4405,34	1,7993	0,07255	*
N5_CHB	14526,5	7099,42	2,0462	0,04124	**
S_EQ	12276,3	5850,11	2,0985	0,03634	**
EQ	22098,7	5467,15	4,0421	0,00006	***
TERR	10892,2	3639,11	2,9931	0,00289	***
PISCINE	104131	13812	7,5392	<0,00001	***
LOG_AGE	-17951	1974,63	-9,0908	<0,00001	***
MOY_R_MED	3,27931	0,606557	5,4064	<0,00001	***
DENSITÉ_SS	-1,21194	0,646985	-1,8732	0,06160	*
ACC_EMPLOI	0,372079	0,0820039	4,5373	<0,00001	***
B7_BUS	-47681,5	13931	-3,4227	0,00067	***
B2_INDU	-10190,1	4515,93	-2,2565	0,02445	**
Moy. var. dép.	188802,5	Éc. type var. dép.		81675,64	
Somme carrés résidus	7,46e+11	Éc. type de régression		37685,87	
R2	0,795664	R2 ajusté		0,787102	
F(22, 525)	92,92290	p. critique (F)		5,0e-165	
Log de vraisemblance	-6540,128	Critère d'Akaike		13126,26	
Critère de Schwarz	13225,30	Hannan-Quinn		13164,97	

**Modèle 2**

Modèle 2: MCO, utilisant les observations 1-548  
Variable dépendante: PRIX

	<i>Coefficient</i>	<i>Erreur Std</i>	<i>t de Student</i>	<i>p. critique</i>	
const	83906,5	26976,5	3,1104	0,00197	***
A_RENOVE	-20998,8	6778,94	-3,0977	0,00206	***
BON_OU_EXC	7696,61	4093,74	1,8801	0,06065	*
R_A_NEUF	13347	6183,97	2,1583	0,03136	**
log_terrain	41553,8	5615,65	7,3996	<0,00001	***
SUP_HAB	417,264	36,5855	11,4052	<0,00001	***
N4FAC	30913,6	5179,17	5,9688	<0,00001	***
N0ETA	-23378,9	7490,83	-3,1210	0,00190	***
N0SDB	-13783,2	7056,93	-1,9531	0,05134	*
N2_SDB	27292	6372,54	4,2828	0,00002	***
N1CHB	-22933,5	11064,4	-2,0727	0,03869	**
N4_CHB	8054,03	4403,47	1,8290	0,06796	*
N5_CHB	16843,9	7059,95	2,3858	0,01739	**
S_EQ	10700,7	5846,13	1,8304	0,06776	*
EQ	19870,3	5443,83	3,6500	0,00029	***
TERR	10376,7	3643,75	2,8478	0,00457	***
PISCINE	105919	13828,8	7,6594	<0,00001	***
LOG_AGE	-17818,7	1995,57	-8,9292	<0,00001	***
MOY_R_MED	3,59271	0,649862	5,5284	<0,00001	***
DENSITE_SS	-1,27055	0,649306	-1,9568	0,05090	*
DIST_LIEGE	-1264,27	294,604	-4,2914	0,00002	***
ACC_ML	1091,58	520,342	2,0978	0,03640	**
B7_BUS	-49602,2	14018,3	-3,5384	0,00044	***
B2_INDU	-8290,63	4533,1	-1,8289	0,06798	*
Moy. var. dép.	188802,5	Éc. type var. dép.	81675,64		
Somme carrés résidus	7,44e+11	Éc. type de régression	37692,88		
R2	0,795978	R2 ajusté	0,787023		
F(23, 524)	88,88465	p. critique (F)	3,3e-164		
Log de vraisemblance	-6539,708	Critère d'Akaike	13127,42		
Critère de Schwarz	13230,77	Hannan-Quinn	13167,81		



**Modèle 3**

Modèle 3: Hétéroscédasticité corrigée, utilisant les observations 1-548  
Variable dépendante: PRIX

	<i>Coefficient</i>	<i>Erreur Std</i>	<i>t de Student</i>	<i>p. critique</i>	
const	109259	20544,5	5,3182	<0,00001	***
A_RENOVE	-23157,9	4459,12	-5,1934	<0,00001	***
log_terrain	31707,9	4733,77	6,6982	<0,00001	***
SUP_HAB	475,589	34,1145	13,9410	<0,00001	***
N3FAC	6838,07	3135,24	2,1810	0,02962	**
N4FAC	34440	5181,21	6,6471	<0,00001	***
N0ETA	-21883,5	6709,61	-3,2615	0,00118	***
N2_SDB	15947,5	5422,21	2,9411	0,00341	***
N1CHB	-27712	8809,54	-3,1457	0,00175	***
N5_CHB	18088,5	7580,85	2,3861	0,01738	**
EQ	12244,2	3083,8	3,9705	0,00008	***
TERR	7475,43	2781,08	2,6880	0,00742	***
PISCINE	81709,5	38994,5	2,0954	0,03661	**
LOG_AGE	-18749,8	1545,74	-12,1300	<0,00001	***
MOY_R_MED	3,48351	0,454913	7,6575	<0,00001	***
DIST_LIÈGE	-1238,8	231,422	-5,3530	<0,00001	***
ACC_ML	1553,14	453,997	3,4210	0,00067	***
B7_BUS	-50519,5	7582,82	-6,6624	<0,00001	***
B2_INDU	-9968,09	3014,92	-3,3063	0,00101	***
SHANNON_10	8972,54	3563,52	2,5179	0,01210	**

## Statistiques basées sur les données pondérées:

Somme carrés résidus	2083,433	Éc. type de régression	1,986428
R2	0,812130	R2 ajusté	0,805370
F(19, 528)	120,1293	p. critique (F)	1,3e-177
Log de vraisemblance	-1143,504	Critère d'Akaike	2327,009
Critère de Schwarz	2413,134	Hannan-Quinn	2360,671

## Statistiques basées sur les données initiales:

Moy. var. dép.	188802,5	Éc. type var. dép.	81675,64
Somme carrés résidus	7,99e+11	Éc. type de régression	38909,48

## ANNEXE 4 : COLINÉARITÉ

Facteurs d'inflation de variance

Valeur minimale possible = 1.0

Valeurs > 10.0 peut indiquer un problème de colinéarité

A_RENOVE	1,511
BON_OU_EXC	1,581
R_A_NEUF	1,395
log_terrain	2,584
SUP_HAB	1,601
N4FAC	2,144
N0ETA	1,171
N0SDB	1,056
N2_SDB	1,104
N1CHB	1,090
N5_CHB	1,252
S_EQ	2,171
EQ	2,576
TERR	1,213
PISCINE	1,061
LOG_AGE	1,448
MOY_R_MED	1,750
DENSITE_SS	1,092
DIST_LIEGE	1,499
ACC_ML	1,634
B7_BUS	1,091
PAYSAGE	1,078

$VIF(j) = 1/(1 - R(j)^2)$ , où  $R(j)$  est un coefficient de corrélation multiple entre la variable  $j$  et les autres variables indépendantes

Propriétés de la matrice  $X'X$ :

1-norm = 2,3065461e+011

Déterminant = 4,7731833e+067

Numéro de la condition réciproque = 5,111552e-012

## ANNEXE 5 : TEST DE BREUSCH-PAGAN

Test de Breusch-Pagan pour l'hétéroscédasticité  
MCO, utilisant les observations 1-548  
Variable dépendante: à l'échelle uhat^2

	coefficient	erreur std.	t de Student	p. critique	
const	-3,28059	1,39845	-2,346	0,0194	**
A_RENOVE	-0,103538	0,344748	-0,3003	0,7640	
BON_OU_EXC	0,105131	0,209051	0,5029	0,6152	
R_A_NEUF	0,132685	0,315570	0,4205	0,6743	
log_terrain	0,736770	0,282134	2,611	0,0093	***
SUP_HAB	0,00819603	0,00176808	4,636	4,50e-06	***
N4FAC	0,259979	0,263769	0,9856	0,3248	
N0ETA	0,316380	0,379411	0,8339	0,4047	
N0SDB	0,0291102	0,360222	0,08081	0,9356	
N2_SDB	0,204030	0,324714	0,6283	0,5301	
N1CHB	-0,271676	0,563808	-0,4819	0,6301	
N5_CHB	-0,534882	0,342072	-1,564	0,1185	
S_EQ	-0,442299	0,298411	-1,482	0,1389	
EQ	-0,276283	0,277894	-0,9942	0,3206	
TERR	-0,150177	0,185558	-0,8093	0,4187	
PISCINE	2,37032	0,705951	3,358	0,0008	***
LOG_AGE	0,149689	0,101798	1,470	0,1420	
MOY_R_MED	-2,53651e-06	3,30799e-05	-0,07668	0,9389	
DENSITE_SS	-1,44610e-05	3,32077e-05	-0,4355	0,6634	
DIST_LIEGE	-0,0168807	0,0151354	-1,115	0,2652	
ACC_ML	0,0306464	0,0268358	1,142	0,2540	
B7_BUS	-0,0711204	0,715729	-0,09937	0,9209	
PAYSAGE	0,00712342	0,00392440	1,815	0,0701	*

Somme des carrés expliquée = 312,871

Statistique de test: LM = 156,435641,  
avec p. critique = P(Chi-deux(22) > 156,435641) = 0,000000

## ANNEXE 6 : MODÈLES HÉDONIQUES POUR DIFFÉRENTES CATÉGORIES DE PRIX

### Prix inférieur à 300.000 €

Modèle 1: Hétéroscédasticité corrigée, utilisant les observations 1-489  
Variable dépendante: PRIX

	<i>Coefficient</i>	<i>Erreur Std</i>	<i>t de Student</i>	<i>p. critique</i>	
const	95948,6	20453,1	4,6912	<0,00001	***
A_RENOVE	-26758,3	3975,94	-6,7301	<0,00001	***
log_terrain	25449,3	3919,64	6,4928	<0,00001	***
SUP_HAB	342,78	31,342	10,9368	<0,00001	***
N3FAC	7810,14	2776,5	2,8129	0,00512	***
N4FAC	29353,2	4347,86	6,7512	<0,00001	***
N0ETA	-10600,9	4692,91	-2,2589	0,02436	**
N0SDB	-7064,04	4058,66	-1,7405	0,08244	*
N2_SDB	18820,5	5215,07	3,6089	0,00034	***
N1CHB	-31335	5317,62	-5,8927	<0,00001	***
N2CHB	-8473,47	2910,28	-2,9116	0,00377	***
N5_CHB	21818,6	4929,71	4,4259	0,00001	***
S_EQ	6625,6	4050,43	1,6358	0,10257	
EQ	15211,6	3767,62	4,0375	0,00006	***
TERR	7943,85	2354,27	3,3742	0,00080	***
PISCINE	13707,2	4745,71	2,8883	0,00406	***
BUANDERI	285,807	2425,1	0,1179	0,90623	
LOG_AGE	-15218,7	1689,07	-9,0101	<0,00001	***
MOY_R_MED	3,65667	0,419189	8,7232	<0,00001	***
DENSITE_SS	-1,01511	0,328743	-3,0879	0,00214	***
DIST_LIEGE	-917,49	229,081	-4,0051	0,00007	***
ACC_BUS	0,0108276	0,0107609	1,0062	0,31485	
ACC_ML	1830,21	436,165	4,1961	0,00003	***
B7_BUS	-36915,4	6854,58	-5,3855	<0,00001	***
B2_EXTRACT	3737,32	10898,3	0,3429	0,73181	
B2_INDU	-6885,71	2963,03	-2,3239	0,02057	**
B2_SAED	6461,04	3215,9	2,0091	0,04511	**
SHANNON_50	9032,32	3173,46	2,8462	0,00462	***
Part_batit_10	-8648,54	5130,04	-1,6859	0,09250	*

Statistiques basées sur les données pondérées:

Somme carrés résidus	1677,998	Éc. type de régression	1,909927
R2	0,834919	R2 ajusté	0,824871
F(28, 460)	83,08996	p. critique (F)	1,4e-160
Log de vraisemblance	-995,3281	Critère d'Akaike	2048,656
Critère de Schwarz	2170,235	Hannan-Quinn	2096,409

Statistiques basées sur les données initiales:

Moy. var. dép.	169819,5	Éc. type var. dép.	55901,70
Somme carrés résidus	3,44e+11	Éc. type de régression	27337,34

**Prix inférieur à 250.000 €**

Modèle 1: Hétéroscédasticité corrigée, utilisant les observations 1-434  
Variable dépendante: PRIX

	<i>Coefficient</i>	<i>Erreur Std</i>	<i>t de Student</i>	<i>p. critique</i>	
const	98645	18888,7	5,2224	<0,00001	***
A_RENOVE	-31110,2	3873,02	-8,0325	<0,00001	***
R_A_NEUF	9278,23	2836,23	3,2713	0,00116	***
log_terrain	23443,2	3911,58	5,9933	<0,00001	***
SUP_HAB	254,702	28,9359	8,8023	<0,00001	***
N3FAC	6150,99	2654,35	2,3173	0,02098	**
N4FAC	15141,9	4320,64	3,5045	0,00051	***
N1_GAR	5931,34	2807,56	2,1126	0,03524	**
N0SDB	-7332,18	4599,34	-1,5942	0,11167	
N2_SDB	10751,9	5057,3	2,1260	0,03410	**
N1CHB	-36410,7	5470,53	-6,6558	<0,00001	***
N2CHB	-12801,7	2752,04	-4,6517	<0,00001	***
N5_CHB	21031	5894,32	3,5680	0,00040	***
TERR	8251,74	2364,85	3,4893	0,00054	***
PISCINE	27156,6	3881,8	6,9959	<0,00001	***
LOG_AGE	-11624,6	1753,51	-6,6293	<0,00001	***
MOY_R_MED	3,81607	0,471665	8,0906	<0,00001	***
DENSITE_SS	-0,473073	0,44529	-1,0624	0,28868	
DIST_LIEGE	-884,445	230,662	-3,8344	0,00015	***
ACC_BUS	0,0165291	0,00901029	1,8345	0,06731	*
ACC_ML	1317,47	372,248	3,5392	0,00045	***
B7_BUS	-42878,2	6225,38	-6,8876	<0,00001	***
B1_INDU	-8372,78	3506,19	-2,3880	0,01739	**
SHANNON_50	7586,42	3339,08	2,2720	0,02360	**

Statistiques basées sur les données pondérées:

Somme carrés résidus	1429,625	Éc. type de régression	1,867321
R2	0,786174	R2 ajusté	0,774179
F(23, 410)	65,54115	p. critique (F)	7,9e-122
Log de vraisemblance	-874,5099	Critère d'Akaike	1797,020
Critère de Schwarz	1894,773	Hannan-Quinn	1835,605

Statistiques basées sur les données initiales:

Moy. var. dép.	156805,9	Éc. type var. dép.	44458,60
Somme carrés résidus	2,44e+11	Éc. type de régression	24387,48

**Prix inférieur à 200.000 €**

Modèle 1: Hétéroscédasticité corrigée, utilisant les observations 1-358  
Variable dépendante: PRIX

	<i>Coefficient</i>	<i>Erreur Std</i>	<i>t de Student</i>	<i>p. critique</i>	
const	139478	18505,9	7,5369	<0,00001	***
A_RENOVE	-20376,2	3463,24	-5,8836	<0,00001	***
log_terrain	10343	4011,45	2,5784	0,01035	**
SUP_HAB	248,737	28,0512	8,8673	<0,00001	***
N3FAC	7541,59	2556,68	2,9498	0,00340	***
N4FAC	16662,8	3796,83	4,3886	0,00002	***
N1_GAR	4578,31	2539,26	1,8030	0,07227	*
N0SDB	-6752,67	3498,43	-1,9302	0,05441	*
N1CHB	-33393,7	4625,91	-7,2188	<0,00001	***
N2CHB	-11692,4	2524,29	-4,6319	<0,00001	***
N5_CHB	18169,2	7427,91	2,4461	0,01495	**
EQ	11947,6	2257,78	5,2918	<0,00001	***
TERR	8855,47	2231,2	3,9689	0,00009	***
LOG_AGE	-7309,89	1884,56	-3,8788	0,00013	***
MOY_R_MED	2,40424	0,392495	6,1255	<0,00001	***
DIST_LIEGE	-929,226	200,301	-4,6392	<0,00001	***
B7_BUS	-47090,3	7716,82	-6,1023	<0,00001	***
B2_INDU	-8822,92	2815,73	-3,1334	0,00188	***

Statistiques basées sur les données pondérées:

Somme carrés résidus	1157,019	Éc. type de régression	1,844722
R2	0,696752	R2 ajusté	0,681590
F(17, 340)	45,95270	p. critique (F)	2,50e-77
Log de vraisemblance	-717,9594	Critère d'Akaike	1471,919
Critère de Schwarz	1541,768	Hannan-Quinn	1499,698

Statistiques basées sur les données initiales:

Moy. var. dép.	142508,8	Éc. type var. dép.	34388,37
Somme carrés résidus	1,43e+11	Éc. type de régression	20504,94

## ANNEXE 7 : ANALYSE DE L'UTILITÉ MARGINALE DÉCROISSANTE

Notre modèle hédonique final considérait l'utilité marginale de nos variables quantitatives de manière constante. Or une des caractéristiques importante du marché immobilier, c'est la décroissance du prix marginal du terrain et des superficies habitables.

Notre modèle prenait déjà en compte cette décroissance du prix marginal pour la superficie du terrain en transformant la variable par son logarithme mais pour les autres variables, notamment la superficie habitable et la distance à Liège.

Pour approcher ce concept, des sous-modèles ont été réalisés. Pour chacune des variables testées, notre échantillon initial a été subdivisé, sur base des quartiles, en 4 groupes. L'approche hédonique reprenant les mêmes variables que notre modèle final a été mise en application pour ces différents groupes. Nous présentons les coefficients pour les différentes catégories. Ce genre de traitements implique une réduction importante de données à régresser. L'ajustement des variables est moins bon que pour notre modèle final et le  $p$  peut-être élevé.

Pour les superficies habitables, le coefficient pour le 3<sup>e</sup> quartile n'est pas significatif. Nous ne tiendrons donc pas compte de cette valeur qui semble, en effet, ne pas correspondre à la réalité (55 €/m<sup>2</sup>). Le principe d'utilité marginale décroissante n'est pas visible pour cette variable. Nous constatons que le prix le plus élevé au mètre carré est au 2<sup>e</sup> quartile, entre 110 et 140 m<sup>2</sup>. Le fait que le prix au mètre carré soit plus important dans cette catégorie qu'au 1<sup>er</sup> quartile peut-être expliqué par une demande plus importante pour ce type de bien. Une maison dont la superficie habitable est inférieure à 110 m<sup>2</sup> est très petite et ne doit pas satisfaire les ménages comportant plusieurs enfants. Le prix au mètre carré se réduit pour les catégories suivantes malgré la non significativité de 3<sup>e</sup> quartile mais cela laisse supposer que l'utilité marginale décroît au delà de 140 m<sup>2</sup>.

**Tableau 36 : Coefficient pour les catégories de superficie habitable**

Catégories des superficies habitables		Coefficient	Erreur Std	t de Student	p. critique
1e quartile	0 - 110	508,806	102,192	4,9789	<0,00001
2e quartile	110 - 140	918,779	238,011	3,8602	0,00021
3e quartile	140 - 180	54,8768	190,256	0,2884	0,77358
4e quartile	180 - 400	292,64	64,6116	4,5292	0,00002

Nous avons testé aussi la superficie du terrain malgré le fait que nous avons transformé la variable dans notre modèle. Les coefficients diminuent de catégorie en catégorie, le prix marginal est donc décroissant puisque l'on passe pour les petits terrains à 133 €/m<sup>2</sup> à moins de 10 €/m<sup>2</sup> pour les grands terrains. La plus forte diminution se situe quand l'on passe le cap des 300 m<sup>2</sup> de terrain. La baisse dépasse les 80 %. L'utilité marginale devient donc fortement réduite au delà de 300 m<sup>2</sup> de terrain.

**Tableau 37 : Coefficient pour les catégories de superficie de terrain**

Catégories des superficies du terrain		Coefficient	Erreur Std	t de Student	p. critique
1e quartile	0 - 157,5	132,508	53,1365	2,4937	0,01409
2e quartile	157,5 -300	75,6827	56,4565	1,3405	0,18291
3e quartile	300 - 798,5	12,2553	16,6762	0,7349	0,46403
4e quartile	798,5 - 5800	9,11006	4,39334	2,0736	0,04049

Tous nos coefficients pour la distance-temps à Liège ne sont pas significatifs. Les résultats laissent penser que l'utilité marginale est croissante excepté quand on dépasse la limite des 15 minutes. L'optimum semble être situé entre 12 et 15 minutes du centre de Liège, une situation où l'on est suffisamment éloigné de Liège pour ne pas souffrir des désagréments du centre urbain mais pas trop éloigné quand même pour profiter de l'accessibilité à l'emploi et aux loisirs qu'offre un centre urbain. Le centre reste quand même une moins-value pour le bien immobilier et toutes les situations sont préférables à cette situation centrale.

**Tableau 38 : Coefficient pour les catégories de distance-temps à Liège**

Catégories du temps à liège		Coefficient	Erreur Std	t de Student	p. critique
1er quartile	0 - 8 min	-1858,64	1543,25	-1,2044	0,23099
2e quartile	8 - 12 min	-1402,76	1627,02	-0,8622	0,39011
3e quartile	12 - 15 min	5290,68	2281,13	2,3193	0,02278
4e quartile	15 - 40 min	-789,213	601,85	-1,3113	0,19248

Nous aurions pu faire ces tests pour d'autres variables mais les résultats nous semblaient moins intéressant ou étaient très peu significatifs.



## ANNEXE 8 : LETTRE ÉCRITE AUX COMMUNES



Université de Liège  
Faculté des Sciences  
Département de Géographie  
ECOGEO  
Jean-Marie Halleux  
Chargé de cours



**ECOGEO**  
economic geography unit  
service de géographie économique

A l'attention de M. le Bourgmestre  
Commune de MACHIN  
Rue de MACHIN  
4XXXX MACHIN

*Concerne : Mémoire de fin d'études traitant des choix résidentiels des ménages réalisé par Mademoiselle Perrine DETHIER*

Liège, le 15 juin 2012

Monsieur le Bourgmestre,

Je supervise actuellement un mémoire de Master en Sciences Géographiques traitant des choix résidentiels des ménages. Ce mémoire est réalisé par Mademoiselle Perrine DETHIER.

Le travail de Mademoiselle DETHIER devant inclure des enquêtes auprès de ménages installés dans des constructions neuves, je me permets de vous solliciter afin d'obtenir les adresses liées à des permis délivrés par votre administration en vue de la construction de maisons unifamiliales neuves. Dans la mesure du possible, nous souhaiterions recevoir les informations pour les permis délivrés entre 2008 et 2010.

Les informations peuvent être directement transmises à Mademoiselle DETHIER. Il est possible de la joindre au 04/..... et à l'adresse postale suivante : XXXX, 4000 Liège. Elle est également joignable à l'adresse électronique suivante : XXXXX.

Dans l'espoir que vous accédiez favorablement à cette requête, je vous prie d'agréer, Monsieur le Bourgmestre, l'expression de mes sentiments très distingués.

Jean-Marie Halleux

# ANNEXE 9 : LETTRE DISTRIBUÉE AUX HABITANTS DE NOUVELLES MAISONS



Liège, le ..... 2012

Madame, Monsieur,

Je m'appelle Perrine Dethier, je suis étudiante en dernière année en géographie à l'Université de Liège. Pour mon mémoire de fin d'étude, je réalise un travail sur les localisations résidentielles, c'est-à-dire les endroits où les personnes décident d'habiter. Dans ce cadre, je souhaite réaliser des enquêtes et votre avis m'intéresse. Je désirerais vous poser quelques questions à ce sujet.

Je passerai dans le courant de cette semaine pour m'entretenir avec vous quelques minutes. Si vous avez la moindre question concernant cette rencontre, vous pouvez me contacter par mail ou sur mon téléphone portable.

Je vous remercie d'avance,

Perrine Dethier  
Etudiante en 2<sup>ème</sup> master à l'ULG  
[perrine.dethier@student.ulg.ac.be](mailto:perrine.dethier@student.ulg.ac.be)  
0497/75.63.43

## ANNEXE 10 : L'ENQUÊTE



### Questionnaire d'enquête

Numéro d'enquête : .....

Adresse : .....

.....

Date : .....

Numéro de la planche de valeurs utilisée :      A       B

Questionnaire utilisé :    1       2       3

Nombre de façades :      2       3       4

Sexe du répondant :    Femme  Homme

Bonjour, je m'appelle ..... et je suis étudiant(e) en dernière année en géographie à l'Université de Liège. Pour mon travail de fin d'étude, je réalise une enquête sur les localisations résidentielles, c'est-à-dire les endroits où les personnes décident d'habiter. Dans ce cadre, je souhaite vous soumettre un questionnaire assez bref et totalement anonyme qui ne prendra que quelques minutes de votre temps.

- *Si la personne répond qu'elle n'a pas le temps :*

Pourrais-je revenir à un autre moment de la journée ou un autre jour qui vous conviendrait mieux ?

Si la personne accepte, inscrire la date et l'heure du rendez-vous :

Date : .....Heure : .....

- *Si la personne accepte de répondre aux questionnaires :*

Avant de commencer le questionnaire, je tiens à préciser que la plupart des questions que je vais vous poser concernent votre opinion et vos préférences. Il n'y a donc pas de bonnes ou de mauvaises réponses.

Si vous n'avez pas de questions concernant le déroulement de cet entretien, je vais commencer.



**Je vais vous demander de bien vouloir considérer la situation de référence suivante.**

*(L'enquêteur présente et lit la planche « situation référence » avec le répondant)*

Une maison neuve est à vendre. C'est une maison 4 façades construite sur un terrain de 700 m<sup>2</sup>. La maison a une surface habitable de 200 m<sup>2</sup> avec quatre chambres. Elle se situe à 20 minutes en voiture du centre de Liège dans un quartier neuf. Dans la rue, il y a peu de trafic routier et l'accessibilité en transport en commun n'est pas bonne. Cette maison coûte, tout frais compris, 250.000 €.

**1. À présent, je vais vous présenter une seconde situation.**

*(L'enquêteur présente et lit la planche « situation 2 » avec le répondant avec en parallèle la planche « situation référence », une fois que la planche « situation 2 » a été présentée, passer à la planche d'échelle de valeurs et puis revenir au questionnaire)*

La maison, indiquée par la flèche rouge, est identique à la maison de référence. La différence est qu'elle est devenue une maison 2 façades. De plus, elle est construite sur un terrain de 300 m<sup>2</sup> et non plus de 700 m<sup>2</sup>.

N° de la catégorie de prix évoquée par le répondant : .....

**2. Imaginez-vous maintenant que cette situation évolue vers une maison plus petite.**

*(L'enquêteur présente et lit la planche « situation 3 » avec le répondant avec en parallèle la planche « situation référence », une fois que la planche « situation 3 » a été présentée, passer à la planche d'échelle de valeurs et puis revenir au questionnaire)*

La maison à vendre, indiquée par la flèche rouge, est, comme pour la situation précédente, de 2 façades et construite sur un terrain de 300 m<sup>2</sup>. Mais cette maison a une surface habitable de 150 m<sup>2</sup> (à la place de 200 m<sup>2</sup>) et elle ne compte plus que deux chambres (à la place de quatre chambres).

N° de la catégorie de prix évoquée par le répondant : .....

**3. Revenons à la situation de référence et imaginez-vous cette nouvelle situation.**

*(L'enquêteur présente et lit la planche « situation 4 » avec le répondant avec en parallèle la planche « situation référence », une fois que la planche « situation 4 » a été présentée, passer à la planche d'échelle de valeurs et puis revenir au questionnaire)*

La maison est identique à la maison de référence à l'exception de sa localisation. La maison se situe à 5 minutes du centre de Liège et non plus à 20 minutes.

N° de la catégorie de prix évoquée par le répondant : .....



**4. Imaginez-vous maintenant cette situation.**

*(L'enquêteur présente et lit la planche « situation 5 » avec le répondant avec en parallèle la planche « situation référence », une fois que la planche « situation 5 » a été présentée, passer à la planche d'échelle de valeurs et puis revenir au questionnaire)*

La maison à vendre est identique à la maison de la situation de référence mais l'accessibilité en transport en commun est bonne (un passage de bus tout les 15 minutes).

N° de la catégorie de prix évoquée par le répondant : .....

**5. Imaginez-vous maintenant que le paysage change.**

*(L'enquêteur présente et lit la planche « situation 6 » avec le répondant avec en parallèle la planche « situation référence », une fois que la planche « situation 6 » a été présentée, passer à la planche d'échelle de valeurs et puis revenir au questionnaire)*

La maison est identique à la maison de référence. La différence est le cadre dans lequel se situe la maison. Derrière la maison se trouve un parc industriel, un zoning.

N° de la catégorie de prix évoquée par le répondant : .....

**6. A présent, la maison change de quartier.**

*(L'enquêteur présente et lit les planches « situation 7 » avec le répondant avec en parallèle la planche « situation référence », une fois que les planches « situation 7 » ont été présentées, passer à la planche d'échelle de valeurs et puis revenir au questionnaire).*

La maison présentée sur l'image est identique à la maison de référence. La différence est le quartier dans lequel se situe la maison. Le quartier est « ancien ». Elle est toujours à 20 minutes du centre de Liège dans une rue avec peu de trafic routier.

N° de la catégorie de prix évoquée par le répondant : .....

**7. Imaginez-vous que la maison se situe le long d'une grande route**

*(L'enquêteur présente et lit la planche « situation 8 » avec le répondant avec en parallèle la planche « situation référence », une fois que la planche « situation 8 » a été présentée, passer à la planche d'échelle de valeurs et puis revenir au questionnaire)*

La maison est identique à la maison de la situation de référence. La différence est la rue dans laquelle se situe la maison. Elle est maintenant sur une route nationale avec un trafic automobile assez important.

N° de la catégorie de prix évoquée par le répondant : .....



**8. Je vous présente maintenant une nouvelle situation.**

*(L'enquêteur présente et lit les planches « situation 9 » avec le répondant avec en parallèle la planche « situation référence », une fois que les planches « situation 9 » ont été présentées, passer à la planche d'échelle de valeurs et puis revenir au questionnaire)*

La maison présentée sur l'image est identique à la maison de référence. La différence est le quartier dans lequel se situe la maison. Il est plus aéré. La maison est toujours à 20 minutes du centre de Liège dans une rue avec peu de trafic routier.

N° de la catégorie de prix évoquée par le répondant : .....

Nous en avons terminé avec les questions d'évaluation immobilière. Nous désirons analyser les résultats de cette étude en fonction de certains critères socio-économique. Il me reste quelques questions à vous poser.

**9. Combien de voitures disposez-vous dans le ménage ?**

0  1  2  3  4

**10. Êtes-vous propriétaires ou locataires ?**

Propriétaire  Locataire

**11. De combien de personnes se compose votre foyer y compris vous-même ?**

1  2  3  4  5  6  7  8  9

**12. Combien y a-t-il d'enfant de moins de 15 ans dans votre foyer ?**

0  1  2  3  4  5  6  7

**13. Quelle est la superficie habitable de votre maison ? .....m<sup>2</sup>**

**14. Quelle est la superficie de votre terrain ? .....m<sup>2</sup>**

**15. Quel est votre situation sociale ?**

Étudiant  Allocataires social  Travailleur  Retraité

**16. Quel est votre niveau d'étude ?**

Primaire  Secondaire inférieur  Secondaire supérieur   
Supérieur non universitaire  Supérieur universitaire

**17. Puis-je vous demander dans quelle catégorie de revenu mensuel se situez votre foyer ?**

*(Proposer les catégories de la planche « revenus mensuels »)*

Catégorie n°= .....

**18. Pour terminer, puis-je vous demander votre année de naissance ? 19.....**

## Annexes

A

<b>Echelle de valeurs</b>				
Par rapport à la situation de référence, <b>combien accepteriez-vous de payer pour acquérir ce bien ?</b> Je vous demande le <b>montant maximum</b> que vous accepteriez de payer. Faites un choix parmi les propositions ci-dessous. En rouge, vous avez le prix identique à la situation de référence. Donc si vous choisissez cette ligne c'est que vous estimez que le nouveau bien doit se vendre au même prix que la situation de référence. Les valeurs en bleu indiquent des prix inférieurs à la maison de référence. Si vous choisissez cette couleur, c'est que vous estimez que le bien que je viens de vous présenter a moins de valeur que la maison de référence. Au contraire, les valeurs vertes sont supérieures au prix de référence. Si vous choisissez cette couleur, c'est que vous estimez que le bien que je viens de vous présenter a plus de valeur que la maison de référence.				
1	<b>- de 40 %</b>	Cette maison vous coûtera moins de	<b>150.000 €</b>	(soit plus de 100.000 € en moins)
2	<b>- 30 à - 40 %</b>	Cette maison vous coûtera entre	<b>150.000 € et 175.000 €</b>	(soit de 75.000 € à 100.000 € en moins)
3	<b>- 25 à - 30 %</b>	Cette maison vous coûtera entre	<b>175.000 € et 187.500 €</b>	(soit de 62.500 € à 75.000 € en moins)
4	<b>- 20 à - 25 %</b>	Cette maison vous coûtera entre	<b>187.500 € et 200.000 €</b>	(soit de 50.000 € à 62.500 € en moins)
5	<b>- 15 à - 20 %</b>	Cette maison vous coûtera entre	<b>200.000 € et 212.500 €</b>	(soit de 37.500 € à 50.000 € en moins)
6	<b>- 10 à - 15 %</b>	Cette maison vous coûtera entre	<b>212.500 € et 225.000 €</b>	(soit de 25.000 € à 37.500 € en moins)
7	<b>- 5 à - 10 %</b>	Cette maison vous coûtera entre	<b>225.000 € et 237.500 €</b>	(soit de 12.500 € à 25.000 € en moins)
8	<b>0 à - 5 %</b>	Cette maison vous coûtera entre	<b>237.500 € et 250.000 €</b>	(soit jusqu'à 12.500 € en moins)
9	<b>0 %</b>	Cette maison vous coûtera	<b>250.000 €</b>	<b>Situation de référence</b>
10	<b>+ 0 à + 5 %</b>	Cette maison vous coûtera entre	<b>250.000 € et 262.500 €</b>	(soit jusqu'à 12.500 € en plus)
11	<b>+ 5 à + 10 %</b>	Cette maison vous coûtera entre	<b>262.500 € et 275.000 €</b>	(soit de 12.500 € à 25.000 € en plus)
12	<b>+ 10 à + 15 %</b>	Cette maison vous coûtera entre	<b>275.000 € et 287.500 €</b>	(soit de 25.000 € à 37.500 € en plus)
13	<b>+ 15 à + 20 %</b>	Cette maison vous coûtera entre	<b>287.500 € et 300.000 €</b>	(soit de 37.500 € à 50.000 € en plus)
14	<b>+ 20 à + 25 %</b>	Cette maison vous coûtera entre	<b>300.000 € et 312.500 €</b>	(soit de 50.000 € à 62.500 € en plus)
15	<b>+ 25 à + 30 %</b>	Cette maison vous coûtera entre	<b>312.500 € et 325.000 €</b>	(soit de 62.500 € à 75.000 € en plus)
16	<b>+ 30 à + 40 %</b>	Cette maison vous coûtera entre	<b>325.000 € et 350.000 €</b>	(soit de 75.000 € à 100.000 € en plus)
17	<b>+ de 40 %</b>	Cette maison vous coûtera plus de	<b>350.000 €</b>	(soit plus de 100.000 € en plus)

B

<b>Echelle de valeurs</b>				
Par rapport à la situation de référence, <b>combien accepteriez-vous de payer pour acquérir ce bien ?</b> Je vous demande le <b>montant maximum</b> que vous accepteriez de payer. Faites un choix parmi les propositions ci-dessous. En rouge, vous avez le prix identique à la situation de référence. Donc si vous choisissez cette ligne c'est que vous estimez que le nouveau bien doit se vendre au même prix que la situation de référence. Les valeurs vertes sont supérieures au prix de référence. Si vous choisissez cette couleur, c'est que vous estimez que le bien que je viens de vous présenter a plus de valeur que la maison de référence. Au contraire, les valeurs en bleu indiquent des prix inférieurs à la maison de référence. Si vous choisissez cette couleur, c'est que vous estimez que le bien que je viens de vous présenter a moins de valeur que la maison de référence.				
1	<b>+ de 40 %</b>	Cette maison vous coûtera plus de	<b>350.000 €</b>	(soit plus de 100.000 € en plus)
2	<b>+ 30 à + 40 %</b>	Cette maison vous coûtera entre	<b>325.000 € et 350.000 €</b>	(soit de 75.000 € à 100.000 € en plus)
3	<b>+ 25 à + 30 %</b>	Cette maison vous coûtera entre	<b>312.500 € et 325.000 €</b>	(soit de 62.500 € à 75.000 € en plus)
4	<b>+ 20 à + 25 %</b>	Cette maison vous coûtera entre	<b>300.000 € et 312.500 €</b>	(soit de 50.000 € à 62.500 € en plus)
5	<b>+ 15 à + 20 %</b>	Cette maison vous coûtera entre	<b>287.500 € et 300.000 €</b>	(soit de 37.500 € à 50.000 € en plus)
6	<b>+ 10 à + 15 %</b>	Cette maison vous coûtera entre	<b>275.000 € et 287.500 €</b>	(soit de 25.000 € à 37.500 € en plus)
7	<b>+ 5 à + 10 %</b>	Cette maison vous coûtera entre	<b>262.500 € et 275.000 €</b>	(soit de 12.500 € à 25.000 € en plus)
8	<b>+ 0 à + 5 %</b>	Cette maison vous coûtera entre	<b>250.000 € et 262.500 €</b>	(soit jusqu'à 12.500 € en plus)
9	<b>0 %</b>	Cette maison vous coûtera	<b>250.000 €</b>	<b>Situation de référence</b>
10	<b>0 à - 5 %</b>	Cette maison vous coûtera entre	<b>237.500 € et 250.000 €</b>	(soit jusqu'à 12.500 € en moins)
11	<b>- 5 à - 10 %</b>	Cette maison vous coûtera entre	<b>225.000 € et 237.500 €</b>	(soit de 12.500 € à 25.000 € en moins)
12	<b>- 10 à - 15 %</b>	Cette maison vous coûtera entre	<b>212.500 € et 225.000 €</b>	(soit de 25.000 € à 37.500 € en moins)
13	<b>- 15 à - 20 %</b>	Cette maison vous coûtera entre	<b>200.000 € et 212.500 €</b>	(soit de 37.500 € à 50.000 € en moins)
14	<b>- 20 à - 25 %</b>	Cette maison vous coûtera entre	<b>187.500 € et 200.000 €</b>	(soit de 50.000 € à 62.500 € en moins)
15	<b>- 25 à - 30 %</b>	Cette maison vous coûtera entre	<b>175.000 € et 187.500 €</b>	(soit de 62.500 € à 75.000 € en moins)
16	<b>- 30 à - 40 %</b>	Cette maison vous coûtera entre	<b>150.000 € et 175.000 €</b>	(soit de 75.000 € à 100.000 € en moins)
17	<b>- de 40 %</b>	Cette maison vous coûtera moins de	<b>150.000 €</b>	(soit plus de 100.000 € en moins)

### Revenus mensuels

Puis-je vous demander dans quelle catégorie de revenu mensuel situez-vous votre foyer ?

1	< 1.000 €
2	1.000 € à 1.500 €
3	1.500 € à 2.000 €
4	2.000 € à 2.500 €
5	2.500 € à 3.000 €
6	3.000 € à 4.000 €
7	4.000 € à 5.000 €
8	5.000 € à 6.000 €
9	6.000 € à 7.000 €
10	> 7.000 €