



Liberté • Égalité • Fraternité  
RÉPUBLIQUE FRANÇAISE



Ministère de l'Écologie  
et du Développement Durable

# Document de travail

ETUDES – METHODES – SYNTHESSES



## GUIDE DE BONNES PRATIQUES POUR LA MISE EN ŒUVRE DE LA METHODE DES PRIX HEDONIQUES

SERIE METHODE  
05-M01

SEBASTIEN TERRA

Site Internet : <http://www.ecologie.gouv.fr>  
20 avenue de Ségur – 75302 Paris 07 SP

DIRECTION DES ETUDES ECONOMIQUES ET DE L'EVALUATION ENVIRONNEMENTALE

## SOMMAIRE

- I – Introduction
- II – Sources de données disponibles
  1. Les bases de données notariales
  2. Autres bases de données
- III – Quelles sont les variables à inclure dans le modèle ?
  1. Les caractéristiques du logement
  2. Les variables environnementales
  3. L'évolution des prix immobiliers dans le temps
- IV – Comment estimer un modèle hédonique
  1. Une estimation en deux étapes
  2. Mesure des variations de bien-être
- V – Les modèles économétriques
  1. Modèles avec variable expliquée en niveau
  2. Modèle avec variable expliquée en logarithme
  3. Transformation de Box-Cox
- VI – Comment calculer les variations de bien-être ?
  1. Variation de bien-être liée à un changement marginal
  2. Variation de bien-être liée à un changement non marginal
- VII – Choix de la forme fonctionnelle
  1. Brève revue de la littérature
  2. La variable expliquée
  3. Les variables explicatives
- VIII – Problèmes économétriques
  1. Variables explicatives et colinéarité
  2. Autres problèmes économétriques
- IX - Conclusion
- X – Applications à la valorisation des aménités et des dommages
  1. Le Nouvel Urbanisme
  2. Les sites contaminés
  3. Aménités des cours d'eau
- XI – Annexe économétrique
- XII – Références

## RÉSUMÉ

*Ce guide présente les éléments nécessaires pour mettre en œuvre une étude utilisant la méthode des prix hédoniques.*

*La première partie de ce guide est principalement destinée au commanditaire de l'étude. Il présente les éléments nécessaires pour suivre et orienter le déroulement d'une étude utilisant la méthode des prix hédoniques.*

*La seconde partie présente les principales difficultés soulevées par la méthode des prix hédoniques. Elle s'adresse plus spécifiquement au prestataire de l'étude.*

*La méthode des prix hédoniques appliquée à la valorisation des biens environnementaux repose sur l'idée que le prix d'un bien immobilier dépend de ses caractéristiques, parmi lesquelles certaines sont liées à la qualité de l'environnement. Les différences de prix constatées entre des biens présentant par ailleurs des caractéristiques identiques traduisent alors des différences en matière d'environnement et fournissent une information sur le prix implicite du bien environnemental et donc sur le consentement à payer des ménages pour bénéficier de ce bien environnemental.*

*La méthode des prix hédoniques permet de mesurer la valeur d'aménités ou de dommages environnementaux en utilisant des données de marché, en l'occurrence des transactions immobilières (préférences révélées), et non des intentions déclarées par des personnes interrogées lors d'une enquête (méthode d'évaluation contingente). C'est pourquoi elle est considérée par certains comme plus fiable que la méthode d'évaluation contingente, mais catégoriser précisément le type de bénéfice que capte la valeur ainsi mesurée reste délicat.*

*Ce guide technique de mise en œuvre de la méthode des prix hédoniques évoque principalement six points :*

- *quelles sont les sources de données disponibles ?*
- *quelles sont les variables explicatives à inclure dans les modèles hédoniques ?*
- *comment estimer un modèle hédonique ?*
- *quels sont les types de modèles économétriques utilisables pour estimer un modèle hédonique ?*
- *comment calculer les variations de bien-être engendrées par une modification de la qualité de l'environnement ?*
- *comment choisir la forme fonctionnelle la plus adaptée aux données disponibles ?*
- *quels sont les problèmes économétriques spécifiques à la méthode des prix hédoniques ?*

## I – INTRODUCTION

Ce guide présente les éléments nécessaires pour mettre en œuvre une étude utilisant la méthode des prix hédoniques.

La première partie de ce guide (paragraphe II et III) est principalement destinée au commanditaire de l'étude. Il présente les éléments nécessaires pour suivre et orienter le déroulement d'une étude utilisant la méthode des prix hédoniques.

La seconde partie présente les principales difficultés soulevées par la méthode des prix hédoniques. Elle s'adresse plus spécifiquement au prestataire de l'étude.

**La méthode des prix hédoniques appliquée à la valorisation des biens environnementaux repose sur l'idée que le prix d'un bien immobilier dépend de ses caractéristiques, dont certaines sont liées à la qualité de l'environnement. Les différences de prix constatées entre des biens présentant par ailleurs des caractéristiques identiques traduisent alors des différences en matière d'environnement et fournissent une information sur le prix implicite de l'actif qui améliore (ou dégrade) la qualité de l'environnement.**

En économie de l'environnement, Desaigues et Point (1993) et Scherrer (2004) présentent de façon plus détaillée le principe de la méthode des prix hédoniques.

Ce guide de mise en œuvre de la méthode des prix hédoniques évoque principalement **six points** :

- quelles sont les sources de données disponibles ?
- quelles sont les variables explicatives à inclure dans les modèles hédoniques ?
- comment estimer un modèle hédonique ?
- quels sont les types de modèles économétriques utilisables pour estimer un modèle hédonique ?
- comment calculer les variations de bien-être engendrées par une modification de la qualité de l'environnement ?
- comment choisir la forme fonctionnelle la plus adaptée aux données disponibles ?
- quels sont les problèmes économétriques spécifiques à la méthode des prix hédoniques ?

**Le paragraphe X présente trois exemples d'application de la méthode des prix hédoniques.**

**Le paragraphe XI présente de manière formalisée les différents modèles économétriques.**

## II – SOURCES DE DONNEES DISPONIBLES

Les informations relatives aux biens immobiliers peuvent être obtenues à partir de **deux sources de données : les bases notariales et les bases des services fiscaux**. La superficie est une variable parfois manquante pour de nombreuses transactions, en particulier pour les mutations les plus anciennes.

Si l'objectif de l'étude est de mesurer *ex post* la valeur d'un changement de qualité environnementale, il est nécessaire que les données sur les prix des mutations (variable expliquée) permettent d'intégrer à la fois des mutations antérieures et postérieures au changement de qualité (au minimum deux à trois ans de part et d'autre du changement de qualité).

### 1. Les bases de données notariales

#### 1.1. La base BIEN pour l'Ile-de-France

La **base de données notariale BIEN** (Base d'Informations Economiques Notariales) recense chaque mutation depuis 1996 pour les huit départements d'Ile-de-France.

Les principaux champs d'information sont :

- la localisation du bien (commune, quartier, rue, ...) ;
- ses caractéristiques (taille, étage, époque de construction, usage, nombre de salles de bains, présence d'un jardin privatif, ...) ;
- les caractéristiques de la vente (type de mutation, état d'occupation, caractère neuf ou ancien, montant de la vente, crédit, fiscalité, ...) ;
- les profils socio-économiques des vendeurs et acheteurs (qualité, origine géographique, catégories socioprofessionnelles).

#### 1.2. Base de données hors Ile-de-France

La société **Perval**, dont le capital est détenu à 100 % par le Notariat, collecte pour toute la province (hors Ile-de-France) de multiples informations sur les biens immobiliers (type, surface, nombre de pièces, prix, ...) recueillies lors des ventes et transmises par les notaires. Cette base constitue l'une des bases de données les plus complètes en France avec 4 000 000 de références immobilières.

## 2. Autres bases de données

Les **services fiscaux** du cadastre disposent aussi d'une base de données ŒIL (Observatoire des Evaluations Immobilières Locales) qui peut être accessible sous certaines conditions (notamment de confidentialité). L'accès aux données est à étudier en contactant directement les services fiscaux du (ou des) département(s) concerné(s) par l'aire d'étude.

Cette base, qui est à l'échelle du département, associe à chaque section cadastrale une adresse, le prix de la mutation et différentes données relatives à l'habitation (type de logement, surface de la parcelle, surface du logement, année de construction, ...).

### III – QUELLES SONT LES VARIABLES EXPLICATIVES A INCLURE DANS LE MODELE ?

Trois types de variables explicatives influencent le prix d'un logement :

- les caractéristiques du logement ;
- les caractéristiques de l'environnement (au sens large) ;
- l'année au cours de laquelle le logement a été vendu.

#### 1. Les caractéristiques du logement

Parmi les **caractéristiques du logement**, les variables suivantes sont souvent de bons déterminants du prix de vente :

- superficie du logement ;
- superficie du terrain ;
- nombre de pièces ;
- nombre de salles de bain ;
- type de logement : maison, appartement ;
- présence d'un ascenseur (pour les appartements) ;
- âge de l'habitation<sup>1</sup> ;
- présence d'équipements de "luxe" : piscine, tennis ;
- présence d'un garage.

#### 2. Les variables environnementales

##### 2.1. Variables environnementales au sens large

Parmi les **variables environnementales**, figurent les variables correspondant au bien étudié mais aussi les variables liées au quartier (taux de chômage, revenu médian) et différentes variables d'accessibilité (comme la proximité des services publics, la proximité d'un parc urbain, d'une gare, d'un bassin d'emplois ...).

Les résultats du recensement de la population organisé par l'INSEE permettent d'obtenir des informations sur les caractéristiques des quartiers (IRIS) et des communes. La partie communale de ces informations est accessible gratuitement à partir du site de l'INSEE consacré au recensement ([www.insee.fr](http://www.insee.fr), rubrique « Recensement »). En revanche, les informations détaillées au niveau infra-communal (quartier) ne sont pas accessibles librement. Il est alors nécessaire de contacter la direction régionale de l'INSEE compétente pour obtenir ces données (payantes).

Concernant les variables liées à l'accessibilité, l'utilisation d'un système d'information géographique (SIG) est généralement possible. A cet égard, il faut noter que la plupart des bases de données relatives aux transactions immobilières contiennent des informations permettant de localiser (géo-référencer) les habitations.

---

<sup>1</sup> Certains travaux ont mis en évidence que cette variable pouvait être complétée par la variable « âge de l'habitation au carré » dans le modèle explicatif du prix de vente.

## 2.2. Variables correspondant au bien environnemental étudié

Enfin, il faut inclure la (ou les) variable(s) représentant le **bien environnemental** que l'on cherche à étudier. Il est généralement nécessaire de construire cette variable, soit à partir de relevés sur le terrain, soit à partir d'un système d'information géographique si les données sont géo-référencées. La variable retenue doit répondre précisément aux objectifs de l'étude, aussi sa construction doit-elle être l'objet d'une **réflexion** relativement **approfondie**.

**Schématiquement**, on peut distinguer **deux grands types de variables** généralement utilisés pour évaluer les bénéfices ou les dommages liés à la qualité de l'environnement :

- **variable mesurant l'état de l'environnement** (pollution par les nitrates, SO<sub>2</sub>). On cherche alors à mesurer l'impact d'une variation de cette concentration sur le prix d'une habitation. Cela suppose que l'on dispose de stations de mesure en nombre suffisant pour qu'à chaque logement soit attribuée une concentration et pour que l'on dispose d'une variabilité suffisante dans la distribution des concentrations. Ce type de variable peut par exemple être utilisé pour prévoir *ex ante* les bénéfices d'une réduction de la concentration en polluant.
- **variable mesurant une distance**. Si l'on cherche à mesurer les aménités procurées par la présence d'une rivière à proximité du logement, une variable intéressante peut être la distance entre le cours d'eau et chaque logement. L'utilisation de données géo-référencées favorise le calcul de telles distances. Ce type de variables peut aussi être utilisé pour mesurer l'impact d'une amélioration de la qualité du milieu en comparant l'effet de la distance sur le prix avant et après le programme d'amélioration. Il faut souligner que cette distance peut être mesurée de différentes façons : distance en mètres (ou kilomètres), distance en nombre de rues séparatrices (entre le logement et l'aménité étudiée).

## 3. L'évolution des prix de l'immobilier dans le temps

En raison de l'évolution des prix de l'immobilier dans le temps, la valeur d'un bien immobilier est étroitement liée à l'année au cours de laquelle la transaction a eu lieu. Pour procéder à une analyse toutes choses étant égales par ailleurs, il est nécessaire de tenir compte de l'évolution des prix. La façon la plus simple de procéder est d'inclure dans le modèle économétrique une variable binaire pour chaque année (sauf une qui sert de référence) qui vaut 1 si la transaction a eu lieu l'année considérée et 0 sinon.

## IV – COMMENT ESTIMER UN MODELE HEDONIQUE ?

L'estimation d'un modèle hédonique repose, théoriquement, sur une **procédure en deux étapes**. En pratique, seule la première étape est considérée, ce qui est justifiée quand l'externalité (ou l'aménité) valorisée est très localisée.

### 1. Une estimation du modèle en deux étapes

Pour Rosen (1974), l'estimation d'un modèle hédonique se déroule en deux étapes.

- **La fonction** (ou l'équation) **de prix hédonique est estimée**. Les prix implicites marginaux pour chacune des caractéristiques sont calculés à partir des coefficients estimés de l'équation.
- Ces prix marginaux et les caractéristiques socio-économiques des consommateurs sont utilisés pour estimer les paramètres des équations comportementales des consommateurs. Suivant les cas, ces équations correspondent aux équations de demande ou de demande inverse ou à la fonction d'utilité.

**L'estimation de la seconde étape est rarement effectuée en pratique. Cela semble justifié lorsque les externalités sont localisées et affectent une faible proportion du marché immobilier** (voir ci-dessous). Néanmoins, afin d'utiliser les résultats du modèle hédonique à des fins de transfert de valeurs, il faudrait estimer la seconde étape du modèle.

Dans ce document, seuls les éléments relatifs à l'estimation de la première étape sont présentés. Taylor (2003) présente en détail l'estimation de la deuxième étape.

### 2. Mesure des variations de bien-être

**L'équation hédonique fournit des informations sur le consentement à payer marginal pour une amélioration de la qualité de l'environnement** car les consommateurs maximisent leur utilité en égalisant le taux marginal de substitution entre la caractéristique et le numéraire au prix marginal estimé par l'équation de prix hédonique.

On a montré que les bénéfices des aménités (telles qu'une amélioration de la qualité de l'air) pouvaient être estimés à partir de l'équation hédonique dans le cas d'un modèle urbain de petite ville ouverte. Par ailleurs, Palmquist (1992) montre que **le consentement à payer pour un changement environnemental peut être déterminé à partir de l'estimation du modèle hédonique dans le cas d'une externalité « localisée »**. Une externalité localisée affecte seulement un petit nombre de propriétés dans le marché, de sorte que l'équation de prix d'équilibre n'est pas modifiée par le changement.

Si le changement environnemental n'est pas localisé, le prix d'équilibre est modifié suite à la politique mise en œuvre. L'estimation de la seconde étape est alors nécessaire pour déterminer un consentement à payer non-marginal. Pour de nombreux polluants atmosphériques, par exemple, une amélioration de la qualité de l'air affectera les logements dans toute la ville ; par ailleurs, les politiques sont généralement conçues pour réduire de manière significative (et non marginale) la pollution. L'estimation de la demande pour la qualité de l'air est nécessaire pour évaluer de telles politiques.



## V – LES MODELES ECONOMETRIQUES

Pour estimer des modèles hédoniques, les économistes ont généralement le choix entre plusieurs formes fonctionnelles (linéaire, loglinéaire, semi-log, Box-Cox).

**Le paragraphe XI présente de manière plus formalisée les équations correspondant aux différents modèles et complète la présentation plus littéraire de cette section.**

### 1. Modèles avec variable expliquée en niveau

**Le modèle linéaire est la forme fonctionnelle la plus simple** utilisée dans l'estimation des régressions hédoniques. Il relie le prix de vente (en niveau) de l'habitation aux différentes variables explicatives (en niveau).

Le coefficient associé à chaque variable correspond au prix implicite de cette caractéristique. Par ailleurs, une augmentation d'une unité d'une caractéristique donnée entraîne une variation (en euros) du prix de vente égale au coefficient de cette variable.

**Ce modèle linéaire peut aussi prendre une autre forme, qualifiée parfois de modèle « semi-log »** (ou de lin-log). Dans ce cas, le modèle relie le prix de vente en niveau aux variables explicatives dont certaines sont en niveau et d'autres en logarithme.

Une augmentation de 1 % d'une variable en logarithme entraîne un changement (en euros) du prix de vente égal au coefficient de cette variable divisé par 100.

**Cette spécification du modèle linéaire est intéressante car elle permet de modéliser une relation non linéaire entre le prix de vente et certaines variables explicatives et une relation linéaire entre le prix de vente et d'autres caractéristiques du logement.**

### 2. Modèles avec variable expliquée en logarithme

**Le modèle log-linéaire (appelé aussi log-log) relie le logarithme du prix de vente aux logarithmes des différentes variables explicatives.**

**Pour les variables continues, le coefficient d'une variable en logarithme correspond à l'élasticité du prix de vente par rapport à cette caractéristique.** Ainsi, un accroissement de 1 % de cette caractéristique correspond à une variation (en pourcentage) du prix de vente égale au coefficient de cette variable.

Les variables binaires (c'est-à-dire des variables qui prennent une valeur 0 ou 1) figurent toujours dans le modèle sous une forme non transformée.

Comme pour le modèle linéaire, **il existe aussi un modèle semi-log (appelé aussi modèle log-lin) reliant le logarithme du prix de vente aux variables explicatives en niveau.**

Pour une variable continue, un accroissement d'une unité de cette variable entraîne un changement (en pourcentage) du prix de vente égal à 100 fois le coefficient de cette variable.

Pour des variables binaires, l'interprétation des coefficients est différente. L'annexe XI. 2 présente le mode de calcul et l'interprétation de ce type de variables.

**Tous les modèles de régression présentés jusqu'ici peuvent être estimés par la méthode des moindres carrés ordinaires.**

### 3. Transformation de Box-Cox

**La transformation de Box-Cox est généralement considérée comme une forme fonctionnelle flexible bien adaptée pour estimer les modèles hédoniques (voir le paragraphe VII. 1) mais son estimation est plus complexe que celle des modèles présentés précédemment.**

La transformation de Box-Cox permet d'estimer plusieurs types de modèle.

- Dans le premier modèle (Équation 7, p. 25), seule la variable expliquée (le prix de vente) est transformée. Ce modèle admet deux cas particulier : le modèle log-linéaire (log-lin) et le modèle linéaire.
- Dans le deuxième modèle (Équation 8, p. 25), la variable expliquée (le prix de vente) et les variables explicatives sont transformées. Les deux transformations peuvent être paramétrées par des coefficients différents. Ce modèle est parfois appelé modèle de Box-Cox linéaire.

Quand une variable explicative est une variable binaire, une transformation de Box-Cox de cette variable n'a pas de sens.

## VI – COMMENT CALCULER LES VARIATIONS DE BIEN-ETRE ?

Pour calculer les variations de bien-être qui résultent d'une amélioration de la qualité de l'environnement, il est utile de distinguer deux cas<sup>2</sup> :

- effet d'une variation marginale de la qualité ;
- effet d'une variation non marginale de la qualité.

### 1. Variation de bien-être liée à un changement marginal de la qualité

Dans ce cas, la variation de bien-être, c'est-à-dire le consentement à payer pour bénéficier d'une amélioration marginale de la qualité de l'environnement, se calcule à partir des coefficients estimés par la régression hédonique.

Le mode de calcul dépend de la forme fonctionnelle retenue. Les détails des calculs sont présentés au paragraphe XI. 4.

Le consentement à payer total (pour la collectivité) est la somme des prix implicites pour toutes les propriétés qui bénéficient de l'amélioration de la qualité de l'environnement<sup>3</sup>.

### 2. Variation de bien-être liée à un changement non marginal de la qualité

Dans ce cas, on considère que l'acheteur d'une maison est prêt à payer un montant égal à  $P^1 - P^0$  pour bénéficier de l'amélioration de la qualité de l'environnement, avec  $P^0$  le prix de vente associé au niveau initial de qualité et  $P^1$  le prix associé au nouveau niveau de qualité.

Par exemple, pour calculer le consentement à payer moyen pour une amélioration de la qualité de l'eau (réduction de la pollution par les nitrates d'une concentration  $q_0$  à une concentration  $q_1$ ), on calcule le prix de vente d'une habitation ayant les caractéristiques moyennes des habitations de l'échantillon et on fait varier la variable de qualité de l'eau. Supposons que, au niveau de qualité  $q_0$ , cette habitation moyenne se vende au prix de 100 000 euros, prix estimé à partir du modèle économétrique. Si, au niveau de qualité  $q_1$ , cette habitation se vend 110 000 euros, alors cela signifie que le consentement à payer moyen des acheteurs pour bénéficier d'une réduction de la concentration en nitrates de l'eau de  $q_0$  à  $q_1$  est de 10 000 euros.

Le consentement à payer total est la somme de tous les changements de prix pour toutes les habitations qui bénéficient du changement de qualité environnementale.

<sup>2</sup> Dans les deux cas, il s'agit de dévaluer un changement localisé (voir le paragraphe IV. 2).

<sup>3</sup> Il faut noter que l'identification de toutes les propriétés affectées par le changement de qualité environnementale n'est pas toujours immédiate.

## VII – CHOIX DE LA FORME FONCTIONNELLE

### 1. Brève revue de la littérature

Dans l'une des premières applications du modèle de Box-Cox dans le contexte des prix immobiliers, les auteurs montrent qu'une forme fonctionnelle linéaire est particulièrement restrictive et qu'elle est généralement rejetée en faveur d'une transformation de Box-Cox. D'autres auteurs discutent également la spécification appropriée pour une fonction de prix hédonique et recommandent l'utilisation de la transformation de Box-Cox.

Cropper et al. (1988) étudient la performance de six formes fonctionnelles différentes. Quand tous les attributs sont parfaitement observés, ils montrent que le modèle Box-Cox linéaire (Équation 8, p. 25) et le modèle Box-Cox quadratique (non présenté dans ce guide) fournissent la plus petite erreur dans l'estimation des prix marginaux des caractéristiques. En revanche, quand certains attributs ne sont pas observés ou mesurés avec erreur, alors les formes fonctionnelles linéaire et Box-Cox linéaire fournissent l'erreur la plus faible.

Taylor (2003) souligne que la **forme linéaire n'est généralement pas adaptée** car elle suppose que le changement de prix engendré par une augmentation d'une unité du niveau d'une caractéristique est identique quel que soit le niveau de cette caractéristique. En d'autres termes, la forme linéaire suppose, par exemple, que l'effet sur le prix de vente d'une habitation de la proximité d'une rivière est le même que cette distance passe de 10 à 11 mètres ou de 100 à 101 mètres. En réalité, il est très vraisemblable que les prix implicites ne soient pas constants.

Enfin, d'autres auteurs, à partir de tests de spécification, prônent l'utilisation du logarithme du prix de vente et d'une forme quadratique des variables explicatives.

En d'autres termes, **la littérature sur le choix d'une forme fonctionnelle appropriée pour estimer un modèle hédonique ne fournit pas, à l'heure actuelle, de réponse définitive.**

**On peut néanmoins souligner les trois points suivants :**

- **la forme linéaire est la plus simple à estimer et à interpréter, mais elle n'est généralement pas adaptée ;**
- **la forme log-linéaire est également facile à estimer mais son interprétation est plus complexe.** En effet, certains paramètres, notamment pour les variables binaires, ne sont pas directement interprétables. Mais cette difficulté peut être aisément surmontée par une présentation adéquate.
- **les modèles de Box-Cox sont plus difficiles à estimer** (surtout le modèle Équation 8, p. 25) **et à interpréter.** Leur avantage réside surtout dans leur plus grande généralité : les autres formes évoquées en sont des cas particuliers.

**En pratique, les modèles estimés présentent souvent une variable expliquée en logarithme.**

## 2. La variable expliquée

Sous quelle forme la variable expliquée doit-elle figurer dans le modèle ? Comment choisir entre une spécification en niveau, en logarithme ou une transformation de type Box-Cox ?

La première idée pourrait être de comparer, à nombre de variables égal, le  $R^2$  des modèles linéaire et log-linéaire. Malheureusement, le  $R^2$  mesure la part de la variation de la variable expliquée qui est expliquée par le modèle. Par conséquent, il correspond, dans un cas, à la variation du prix et, dans l'autre, à la variation du logarithme du prix. **Les  $R^2$  de ces deux modèles ne sont donc pas directement comparables.** Néanmoins, Wooldridge (2000) a proposé une méthode pour calculer un  $R^2$  dans le cas d'un modèle où la variable expliquée est en logarithme. L'annexe XI. 6 présente cette méthode.

De façon plus formelle, il est possible de tester différentes hypothèses relatives à la spécification de la variable expliquée dans les modèles hédoniques. Parmi les différents tests disponibles, deux seront présentés, dans l'annexe XI. 7, en raison de la simplicité de leur mise en œuvre.

## 3. Les variables explicatives

Un modèle économétrique peut inclure toutes les variables explicatives appropriées, et pourtant être mal spécifié du point de vue de la forme fonctionnelle (le modèle ne rend pas correctement compte de la relation entre la variable expliquée et les variables explicatives). **Une approche utile pour tester, de façon générale, la mauvaise spécification de la forme fonctionnelle est l'utilisation du test RESET (*REgression Specification Error Test*).** L'idée qui sous-tend ce test est simple : si le modèle est correctement spécifié, alors aucune fonction non linéaire des variables explicatives ne devrait être significative quand elle est ajoutée au modèle. Puisque les valeurs prédites par le modèle estimé sont une combinaison linéaire des variables explicatives, il faut ajouter au modèle le carré et/ou le cube de ces valeurs prédites. Le test RESET consiste à réestimer le modèle de départ en ajoutant ces nouvelles variables et à effectuer un test F de Fisher dont l'hypothèse nulle est que ces variables n'ont aucun pouvoir explicatif.

**Faut-il inclure les variables explicatives en niveau ou en logarithme ?** Les procédures d'inférence classiques ne permettent pas de répondre à cette question car les hypothèses (ou alternatives) ne sont pas emboîtées. Proche, dans la démarche, du test RESET, le « **test J** » permet de répondre à la question sur le choix entre logarithme et niveau pour la variable expliquée (voir l'annexe XI. 5). Il peut permettre d'indiquer que l'un des modèles est meilleur que l'autre. Néanmoins, dans cette stratégie de test, quatre situations peuvent se produire : rejeter les deux modèles, n'en rejeter aucun, en rejeter un.

Enfin, concernant la sélection des variables explicatives pertinentes, les **procédures usuelles** comme la sélection *backward* (général vers particulier), *forward* (particulier vers général) et *stepwise* sont toujours disponibles. De même, les critères d'information de Akaike (AIC) et de Schwartz (BIC) peuvent toujours être utilisés. Enfin, le  $R^2$  ajusté doit être utilisé pour comparer deux modèles (avec la même variable expliquée) présentant un nombre différent de variables explicatives.

## VIII – QUELQUES PROBLEMES ECONOMETRIQUES

### 1. Variables explicatives et colinéarité

Les modèles hédoniques souffrent généralement de **problèmes de colinéarité entre les variables explicatives**. Par exemple, plus les maisons ont une superficie élevée, plus elles sont susceptibles d'avoir de nombreuses pièces et plusieurs salles de bain.

Dans sa version la plus extrême<sup>4</sup> (rarement rencontrée en pratique), la colinéarité se traduit par l'impossibilité d'estimer les paramètres de la régression par la méthode des moindres carrés. Dans une version plus courante, **la colinéarité se traduit par des estimations imprécises des paramètres et par des écarts-types élevés**.

Greene (2003, chapitre 4) propose plusieurs indicateurs pour déceler la colinéarité.

**Il n'existe pas de « bonne façon » de traiter les problèmes de colinéarité.**

Une **première solution** consiste à **ne pas** essayer de **corriger la colinéarité**. Ceci pourrait être justifié si la variable de qualité environnementale, qui est au cœur de l'analyse, n'est pas affectée par les problèmes de colinéarité. Dans ce cas, les paramètres des autres caractéristiques risquent d'être estimés de façon imprécise, mais cela n'affectera vraisemblablement pas l'estimation du paramètre d'intérêt et donc du prix implicite.

D'autres solutions sont proposées dans les manuels d'économétrie (par exemple Greene (2003, chapitre 4)), mais aucune solution ne semble parfaite. Une de ces solutions est d'utiliser un petit nombre de composantes principales issues d'une analyse en composantes principales sur les variables explicatives. La colinéarité disparaît puisque, par construction, les composantes principales sont orthogonales les unes aux autres. En revanche, l'interprétation économique des paramètres estimés s'avère très délicate ; en particulier, le calcul des prix implicites des différentes caractéristiques est difficile à réaliser.

Par ailleurs, lors du choix d'une forme fonctionnelle, il faut toujours garder à l'esprit les problèmes de colinéarité des variables explicatives. En effet, plus la forme fonctionnelle est flexible, plus les problèmes de colinéarité sont importants.

---

<sup>4</sup> Dans ce cas, une ou plusieurs variables s'exprime(nt) comme combinaison linéaire exacte d'autres variables.

## 2. Autres problèmes économétriques

D'autres problèmes économétriques sont fréquents : l'hétéroscédasticité et l'auto-corrélation spatiale.

Le premier problème n'est pas spécifique à la méthode des prix hédoniques. En règle générale, il convient de tester la présence d'hétéroscédasticité, ou, a minima, d'utiliser, dans les inférences sur les coefficients estimés, la matrice de White robuste à l'hétéroscédasticité (par exemple).

La détection et la correction du second problème nécessitent le recours à des techniques économétriques plus avancées et ne peuvent s'effectuer que si les données disponibles sont géo-référencées. Taylor (2003) présente quelques éléments sur les aspects spatiaux dans la méthode des prix hédoniques.

## IX – CONCLUSION

La méthode des prix hédoniques permet de mesurer la valeur d'aménités ou de dommages environnementaux en utilisant des données de marché, en l'occurrence des transactions immobilières (préférences révélées), et non des intentions déclarées par des personnes interrogées lors d'une enquête (méthode d'évaluation contingente). C'est pourquoi elle est considérée par certains comme plus fiable que la méthode d'évaluation contingente, mais catégoriser précisément le type de bénéfice que capte la valeur ainsi mesurée reste délicat. En effet, la valeur fournie par la méthode renvoie à une valeur d'usage et probablement, pour une faible part, à une valeur de legs

Pour estimer les modèles hédoniques, plusieurs formes fonctionnelles sont disponibles. Une brève revue de la littérature montre qu'aucun consensus n'émerge véritablement quant à la meilleure spécification du modèle. Les modèles les plus flexibles (modèles de Box-Cox) présentent souvent les problèmes de colinéarité les plus importants. Le modèle de Box-Cox est utile car il admet comme cas particuliers les modèles linéaire et log-linéaire. En revanche, son estimation est plus difficile.

En l'absence de consensus sur la meilleure forme fonctionnelle, chacun est donc réduit à déterminer, en fonction des données dont il dispose, la forme fonctionnelle la plus adaptée. Plusieurs tests ont été présentés afin de guider la recherche de cette forme fonctionnelle.

**On peut néanmoins souligner les trois points suivants :**

- **la forme linéaire est la plus simple à estimer et à interpréter, mais n'est généralement pas adaptée ;**
- **la forme log-linéaire est également facile à estimer mais son interprétation est plus complexe.** En effet, certains paramètres, notamment pour les variables binaires, ne sont pas directement interprétables. Mais cette difficulté peut être aisément surmontée par une présentation adéquate.
- **les modèles de Box-Cox sont plus difficiles à estimer et à interpréter.** Leur avantage réside surtout dans leur plus grande généralité : les autres formes évoquées en sont des cas particuliers.

En pratique, les modèles estimés présentent souvent une variable expliquée en logarithme.



## X – APPLICATIONS A LA VALORISATION DES AMENITES ET DES DOMMAGES

Parmi les travaux récents d'estimation des aménités et des dommages environnementaux, trois exemples ont été choisis : les aménités procurées par le Nouvel Urbanisme aux Etats-Unis ; les externalités négatives causées par les sites contaminés aux Etats-Unis ; les aménités procurées par la rivière la Scarpe, dans le nord de la France.

### 1. Le Nouvel Urbanisme aux Etats-Unis

L'article de Song et Knapp (2003) a pour but d'évaluer l'effet d'une nouvelle forme d'urbanisme, le Nouvel Urbanisme, sur la valeur des habitations dans la région de Portland (Oregon) aux Etats-Unis. La méthode des prix hédoniques est appliquée à un échantillon de ventes immobilières réalisées entre 1990 et 2000 afin de déterminer les prix implicites des caractéristiques de ce nouvel urbanisme.

#### 1.1. Contexte de l'étude

Le Nouvel Urbanisme est un mouvement architectural créé aux Etats-Unis en 1993. Ses promoteurs espèrent « créer des immeubles, quartiers et régions qui permettent une grande qualité de vie à leurs habitants, tout en protégeant l'environnement naturel. » Ils souhaitent, en outre, limiter l'étalement urbain.

La création de quartier à activités mixtes (résidentielles, commerciales, industrielles), la mise en place de réseaux de transport public, de pistes cyclables et de rues piétonnes adaptés sont parmi les principes directeurs de ce Nouvel Urbanisme. Il faut aussi noter la volonté des architectes du Nouvel Urbanisme de doter l'espace urbain de zones ouvertes stratégiquement placées et de développer une architecture destinée à encourager les interactions sociales.

L'étude présentée dans cet article a été réalisée en observant les transactions immobilières réalisées dans la région de Portland aux Etats-Unis entre janvier 1990 et décembre 2000. Un système d'information géographique (SIG) est utilisé afin de disposer de mesures quantitatives de la forme urbaine, mesures qui seront ensuite utilisées dans une analyse des prix hédoniques.

Le choix de la ville de Portland s'explique par sa réputation en matière de lutte contre l'étalement urbain.

En 1991, les autorités locales ont mis au point un plan d'urbanisme, approuvé en 1997. Ce plan a pour objectifs, selon les principes du Nouvel Urbanisme, la transformation de l'espace urbain traditionnel en une forme urbaine polynucléaire, le développement d'un système de transport multimodal et la création de centres-villes regroupant des activités diverses (résidentielles, commerciales, industrielles).

#### 1.2. Les prix implicites des caractéristiques des habitations

Le modèle économétrique retenu par Song et Knapp (2003) est un modèle semi-log, reliant le logarithme du prix vente à un ensemble de caractéristiques.

La superficie du terrain et la superficie de l'habitation ont un effet positif sur la valeur de celle-ci. De plus, la qualité des services publics dans le quartier ajoute de la valeur à la maison, en particulier en ce qui concerne les services éducatifs. En revanche, le niveau de la taxe d'habitation réduit la valeur immobilière.

La localisation de l'habitation a aussi un impact puisque sa valeur diminue avec la distance qui la sépare du *Central Business District* (CBD), qui est notamment un bassin d'emplois.

Les auteurs intègrent aussi des indicateurs d'aménités et de nuisances. L'existence d'un terrain de golf accroît la valeur des habitations proches. De même, la vue sur un espace naturel (par exemple, un massif montagneux) augmente le prix de vente. En revanche, la proximité immédiate d'un axe de circulation majeur (route, voie ferrée) diminue la valeur immobilière.

Les auteurs retiennent différents groupes d'indicateurs pour étudier la forme urbaine des quartiers.

- *Forme des rues et systèmes de circulation* - Les acheteurs préfèrent des quartiers plutôt isolés par rapport à l'extérieur, desservis par un réseau de rues nombreuses. Ils ont aussi une préférence pour les impasses, de longueur modérée.
- *Densité* - Les auteurs observent une préférence pour des quartiers dont la densité de population est faible.
- *Répartition des activités* - Les acheteurs accordent plus de valeur aux habitations individuelles. Ils préfèrent une répartition uniforme des activités (résidentielles, commerciales, industrielles) sur le territoire.
- *Accessibilité* - La valeur d'une habitation augmente quand elle se trouve à proximité des parcs publics. En revanche, cette valeur diminue quand la maison est proche des zones commerciales.
- *Choix du mode de transport* - La mise en place d'un transport ferroviaire urbain augmente la valeur des habitations situées à proximité des gares.
- *Accès piétonniers* - Les consommateurs préfèrent avoir une meilleure accessibilité piétonne aux zones commerciales. En revanche, selon l'étude, ils préféreraient ne pas être trop proches des arrêts de bus.

### 1.3. Estimation de la "valeur ajoutée" apportée par le Nouvel Urbanisme

A partir des résultats obtenus par le modèle hédonique, les auteurs estiment la « prime » que paient les acheteurs pour habiter dans des quartiers aménagés selon les principes du Nouvel Urbanisme. Cette "prime" mesure donc la valeur ajoutée apportée par cette nouvelle forme d'organisation urbaine.

Pour déterminer cette « prime », Song et Knapp (2003) comparent deux quartiers distincts : le premier, purement fictif, est un quartier traditionnel ; le second est un exemple d'application des principes du Nouvel Urbanisme.

La comparaison des deux quartiers suggère que les indicateurs de forme urbaine des quartiers semblent rendre compte des différences significatives observées entre les quartiers périphériques classiques et les quartiers dont l'organisation est inspirée par le Nouvel

Urbanisme. Par ailleurs, ces différences se traduisent par une différence dans la valeur immobilière dans ces deux quartiers.

Pour mesurer cet écart, les auteurs estiment le prix de vente des maisons de chacun des quartiers en fonction de leurs caractéristiques. En outre, ils calculent le changement dans le prix de vente qui résulte d'une modification d'une unité dans la quantité d'une des variables explicatives, toutes choses étant égales par ailleurs.

La valeur de la « prime » est alors déterminée en considérant un prix de vente originel de 150 000 dollars. La valeur d'une maison dans le quartier traditionnel est estimée à 132 731 dollars, alors que, dans un quartier "Nouvel Urbanisme", cette valeur est de 156 986 dollars. Les acheteurs paient donc une "prime" de 24 255 dollars pour habiter dans un quartier dont l'organisation est inspirée par les principes du Nouvel Urbanisme.

Ce surcroît de valeur provient des caractéristiques suivantes : meilleurs connections internes des rues, rues plus nombreuses, meilleur accès piétonnier aux zones commerciales, proximité d'une gare et multiplicité des modes de transport.

## 2. Les externalités négatives des sites contaminés aux Etats-Unis

L'article de Ihlanfeldt et Taylor (2004), propose de mesurer les nuisances causées par les sites contenant des déchets toxiques et d'étudier la viabilité d'un mode de financement particulier pour assurer le nettoyage de ces sites.

### 2.1. La méthode et les données

Pour mesurer les effets externes des sites à risque, les auteurs utilisent la méthode des prix hédoniques.

Dans cette étude, c'est la proximité d'un site contaminé qui constitue l'indicateur de la qualité environnementale dont on cherche à mesurer l'effet. L'objectif de l'étude est en effet d'estimer la relation entre le prix d'une propriété et la distance qui la sépare du site contaminé le plus proche (cette relation est aussi appelée gradient).

Parmi les autres éléments pouvant influencer le prix de la propriété, on trouve trois autres catégories de variables : les caractéristiques de la propriété (superficie, âge, qualité du bâtiment, ...), des variables d'accessibilité (proximité de l'aéroport, distance jusqu'à la sortie d'autoroutes ou la station de métro la plus proche, ...) et des caractéristiques socio-démographiques du quartier (densité de population et d'emploi, revenu réel médian, ...).

La zone d'étude est le Comté de Fulton (qui comprend notamment la ville d'Atlanta), Géorgie, aux Etats-Unis. Les auteurs étudient les transactions immobilières effectuées entre 1981 et 1998. Cinq types de propriétés sont inclus dans l'analyse : les logements, les immeubles de bureaux, les locaux à usage commercial, les locaux à usage industriel et les terrains vagues.

Les auteurs distinguent deux catégories de sites à risque. Les sites de ce qu'ils appellent la « Liste 1 » sont des sites présentant un risque de pollution, selon l'*Environmental Protection*

*Agency* (EPA) ou l'*Environmental Protection Division* (EPD) de l'Etat de Géorgie. Le classement sur cette liste est fondé sur un score établi en tenant compte de la toxicité, de la quantité de la substance rejetée, ainsi que de la proximité des habitations et des sources d'approvisionnement en eau potable. Dans la zone étudiée, la corrélation est très forte entre la distance des propriétés vendues et les sites listés les plus proche.

Les sites de la « Liste 2 » sont des sites qui figuraient auparavant sur la liste 1 mais qui en ont été retirés. L'une des trois raisons suivantes peut expliquer ce retrait : (1) aucune contamination n'a été découverte ; (2) une contamination a été découverte mais elle a été rapidement éliminée ; (3) une contamination a été découverte mais elle a été jugée insuffisamment grave pour justifier une intervention fédérale.

## 2.2. Les résultats

Pour les appartements, les bureaux et les locaux à usage commercial, le modèle descriptif des déterminants des prix des transactions explique plus de 49 % de la variation des prix de vente au sein de leur catégorie respective. Pour les catégories "Industrie" et "Terrains vagues", la portée explicative du modèle est moindre. Les auteurs expliquent que cela est dû à l'omission d'une variable explicative importante : pour l'industrie, il s'agit de la valeur des machines vendues avec la propriété ; pour les terrains vagues, de la forme et de la pente du terrain.

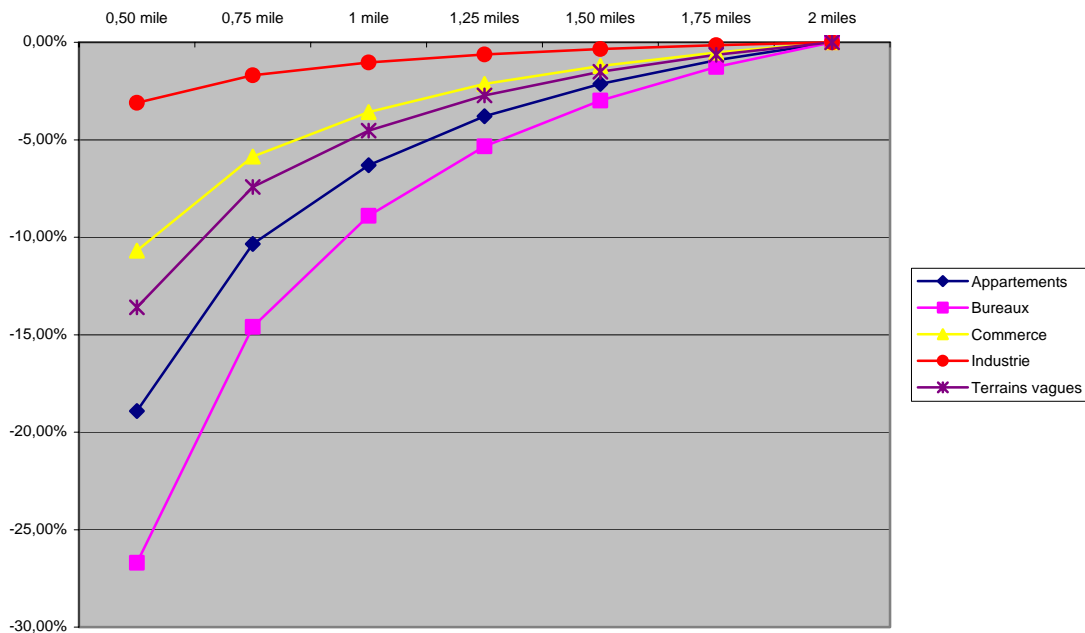
Les transactions peuvent avoir lieu avant le classement du site ou après son classement.

On n'observe pas d'influence significative de la distance entre une propriété et le site contaminé de la liste 1 le plus proche lorsque le site n'a pas encore été répertorié sur la liste 1.

En revanche, après le classement, la proximité d'un site classé sur la liste 1 influence à la baisse le prix de vente des locaux.

Pour illustrer l'ampleur de ces effets, le prix d'une propriété "moyenne" de chaque catégorie est calculé à partir du modèle estimé précédemment. Seule la distance varie, tous les autres paramètres sont constants (fixés à leur valeur moyenne sur l'échantillon). Pour un immeuble de bureau "moyen", l'augmentation de la distance qui le sépare du site "Liste 1" le plus proche de 0,5 à 2 miles entraîne une hausse du prix de vente de cet immeuble de 36 %, comme l'indique la figure suivante.

**Perte de valeur par rapport à une propriété située à 2 miles du site contaminé le plus proche**



A partir de ces résultats, il est possible de calculer la perte de valeur de chacune des propriétés situées à proximité d'un site "Liste 1". Le tableau suivant présente les pertes totales de valeur des propriétés (qu'elles aient été vendues ou non) situées à moins de 1,5 miles du site contaminé le plus proche pour chacune des cinq catégories de propriété. Les pertes totales pour l'ensemble des cinq usages pour le Comté de Fulton représentent plus d'un milliard de dollars, soit environ 10 % de la valeur totale estimée des propriétés situées à moins de 1,5 miles du site contaminé le plus proche.

	Appartements	Bureaux	Commerce	Industrie	Terrains vagues	Ens.
Nombre de propriétés	2823	703	2167	1868	2709	10270
Perte valeur (US \$ millions)	377	347	63.5	54.3	231	1073
Perte valeur / valeur totale	18 %	13 %	7 %	5 %	19 %	10 %

Ces différents résultats montrent que les sites contaminés réduisent bien la valeur des propriétés qui sont situées à proximité.

### 3. Les aménités d'un cours d'eau dans le Nord de la France

La méthode des prix hédoniques peut également être employée pour évaluer l'impact d'un cours d'eau et de sa qualité sur les prix immobiliers. C'est dans cet objectif qu'une étude a été réalisée par Bertrand Zuideau et Virginie Fromon en 1999 à Douai – ville traversée par la Scarpe – pour estimer les externalités que la rivière génère dans ce milieu urbain.

Si la Scarpe était source de nuisances dans les années 70, de gros investissements ont été consentis (aménagement des espaces publics sur les quais, ravalements des façades, réseaux d'égout et stations d'épuration). Ces efforts ont permis d'améliorer de façon notable la qualité de l'eau et ont été vecteur de la réappropriation de la rivière et de ses abords par les habitants de la ville.

La variable environnementale qui a été retenue dans cette étude est la distance entre l'habitation et la rivière, exprimée en nombre de rues (ou carrefours) à traverser pour se rendre sur les quais. Les auteurs expliquent ce choix car ils estiment que c'est en comptant le nombre de ruptures dans le parcours entre la maison et la rivière que les individus appréhendent la proximité aux aménités. Une seconde modélisation a par ailleurs été réalisée avec la variable environnementale « vue sur la rivière » (variable binaire indiquant s'il est possible de voir la Scarpe depuis la maison en question).

Les variables physiques classiques de description des logements ont été prises en compte, mais aucune variable de voisinage (sur le quartier) n'apparaît dans le modèle. Les auteurs ont en effet jugé que le centre de la ville de Douai (dont sont issues toutes les données) présentait un caractère compact avec des services uniformément répartis.

Un échantillon de 189 mutations immobilières, effectuées entre 1988 et 1998, a servi de fondement à cette étude. Plusieurs modes de présentation de la variable ont été testés (linéaire, logarithmique, quadratique, inverse, etc.), la forme linéaire ayant finalement été retenue. Les variables significatives sont l'année de la transaction, la catégorie cadastrale (appréciation de confort donnée par les impôts), le nombre de pièces principales et la présence de garage(s) – avec en plus le nombre de salles de bains pour le modèle intégrant la variable « vue sur la rivière ». Les  $R^2$  de ces deux modèles sont respectivement de 0,57 et de 0,54.

Les résultats révèlent un effet non négligeable de la Scarpe sur la formation des prix des habitations à proximité de celle-ci. En effet, l'éloignement d'une rue supplémentaire par rapport à la rivière diminue en moyenne la valeur du logement de 7,5% (soit 6 700 €). Par ailleurs, la vue sur le cours d'eau augmente le prix d'une maison de 21,5% (soit en moyenne 19 400 €).

## XI – ANNEXE ECONOMETRIQUE

### 1. Modèles avec variable expliquée en niveau

**Le modèle linéaire est la forme fonctionnelle la plus simple** utilisée dans l'estimation des régressions hédoniques. Il relie le prix de vente  $p_i$  (en niveau) de l'habitation  $i$  aux  $J$  variables explicatives (en niveau)  $\mathbf{x}_i = (x_{i1}, \dots, x_{iJ})$  par l'équation (Équation 1) :

$$\text{Équation 1} \quad p_i = \sum_{j=1}^J x_{ij} \beta_j + \varepsilon_i$$

Chaque  $\beta_j$  correspond au prix implicite de la caractéristique  $j$ . Par ailleurs, une augmentation d'une unité de la caractéristique  $x_j$  entraîne une augmentation de  $\beta_j$  euros du prix de vente.

Ce modèle linéaire peut aussi prendre une autre forme, qualifiée parfois de modèle « semi-log » (ou de lin-log). Dans ce cas, le modèle relie le prix de vente ( $p_i$ ) en niveau aux variables explicatives dont certaines ( $x_j$ ) sont en niveau et d'autres ( $z_j$ ) en logarithme :

$$\text{Équation 2} \quad p_i = \sum_{j=1}^J x_{ij} \beta_j + \sum_{j=1}^J \ln(z_{ij}) \gamma_j + \varepsilon_i$$

où  $\beta$  et  $\gamma$  sont les vecteurs de paramètres à estimer. Une augmentation de 1 % de  $z_j$  entraîne un changement de  $(\gamma_j / 100)$  euros du prix de vente.

## 2. Modèles avec variable expliquée en logarithme

Le modèle log-linéaire (appelé aussi loglog) relie le logarithme du prix de vente aux logarithmes des différentes variables explicatives :

$$\text{Équation 3} \quad \ln(p_i) = \sum_{j=1}^J x_{ij} \beta_j + \sum_{j=1}^J \ln(z_{ij}) \gamma_j + \varepsilon_i$$

Pour les variables continues, le coefficient  $\gamma_j$  correspond à l'élasticité du prix de vente par rapport à la caractéristique  $j$ . Ainsi, un accroissement de 1 % de la caractéristique  $j$  correspond à une augmentation de  $\gamma_j$  % du prix de vente.

Les variables binaires (c'est-à-dire des variables qui prennent une valeur 0 ou 1) figurent toujours dans le modèle (Équation 3) sous une forme non transformée (par exemple, les  $x_j$ ).

Comme pour le modèle linéaire, il existe aussi un modèle semi-log reliant le logarithme du prix de vente aux variables explicatives en niveau :

$$\text{Équation 4} \quad \ln(p_i) = \sum_{j=1}^J x_{ij} \beta_j + \varepsilon_i$$

Pour une variable  $x_j$  continue, un accroissement d'une unité de  $x_j$  entraîne un changement de  $100 \beta_j$  % du prix de vente.

Pour des variables binaires, l'interprétation des coefficients est différente. Supposons, par exemple, que l'on cherche à étudier l'impact de la présence d'un garage (0 signifie absence de garage et 1 présence d'un garage) sur le prix d'une vente d'un logement à partir des modèles (Équation 3) ou (Équation 4). Une estimation  $g$  en pourcentage de l'impact de cette variable sur la variable expliquée (prix de vente) est donnée par la formule (Équation 5).

$$\text{Équation 5} \quad g = 100(e^{\beta_{\text{garage}}} - 1)$$

où  $\beta_{\text{garage}}$  est le paramètre relatif à la présence d'un garage.



### 3. Transformation de Box-Cox

La transformation de Box-Cox est généralement considérée comme une forme fonctionnelle flexible bien adaptée pour estimer les modèles hédoniques (voir VII. 1) mais son estimation est plus complexe que celle des modèles présentés précédemment. D'une façon générale, la transformation de Box-Cox d'une variable  $x$  est notée  $x(\lambda)$  et est donnée par l'équation (Équation 6).

$$\text{Équation 6} \quad x(\lambda) = \begin{cases} \frac{x^\lambda - 1}{\lambda} & \text{si } \lambda \neq 0 \\ \ln(x) & \text{si } \lambda = 0 \end{cases}$$

La transformation de Box-Cox permet d'estimer plusieurs types de modèle.

- Dans le premier modèle (Équation 7), seule la variable expliquée (le prix de vente) est transformée. Si  $\lambda = 0$ , on retrouve un modèle log-linéaire. Si  $\lambda = 1$ , on retrouve un modèle linéaire.

$$\text{Équation 7} \quad p_i(\lambda) = \sum_{j=1}^J x_{ij} \beta_j + \varepsilon_i$$

- Dans le deuxième modèle (Équation 8), la variable expliquée (le prix de vente) et les variables explicatives sont transformées. Les deux transformations peuvent être paramétrées par des coefficients différents ( $\lambda$  et  $\mu$ ). Ce modèle est parfois appelé modèle de Box-Cox linéaire.

$$\text{Équation 8} \quad p_i(\lambda) = \sum_{j=1}^J x_{ij}(\mu) \beta_j + \varepsilon_i$$

Quand une variable explicative est une variable binaire, une transformation de Box-Cox de cette variable n'a pas de sens. Les variables binaires sont donc incluses dans le modèle sous une forme non transformée.

Dans ce cas des variables binaires, l'interprétation des coefficients pour le modèle de Box-Cox défini par l'équation (Équation 7) est la suivante. Supposons, par exemple, que l'on cherche à étudier l'impact de la présence d'un garage sur le prix d'une vente d'un logement à partir de ce modèle. Une estimation  $g$  en pourcentage de l'impact de cette variable sur la variable expliquée (prix de vente) est donnée par la formule suivante.

$$g = 100 \left[ \left( 1 + \frac{\lambda \beta_{garage}}{p_0^\lambda} \right)^{\frac{1}{\lambda}} - 1 \right]$$

où  $\beta_{garage}$  est le paramètre relatif à la présence d'un garage et  $p_0$  est le prix moyen d'une habitation sans garage.

#### 4. Calcul des prix implicites pour les différentes formes fonctionnelles

Le prix marginal implicite (c'est-à-dire le consentement à payer marginal pour bénéficier de la variation du niveau d'une caractéristique) se calcule différemment pour chacune des formes fonctionnelles précédentes.

Dans le cas du modèle linéaire (Équation 1), le prix implicite d'un changement dans la caractéristique  $j$  sur le prix de l'habitation  $i$  est le suivant :

$$\text{Équation 9} \quad \frac{\partial p_i}{\partial x_{ij}} = \beta_j$$

Dans le cas du modèle semi-log (Équation 4), le prix implicite est le suivant :

$$\text{Équation 10} \quad \frac{\partial p_i}{\partial x_{ij}} = \beta_j P$$

Dans le cas du modèle log-log (Équation 3), le prix implicite est le suivant :

$$\text{Équation 11} \quad \frac{\partial p_i}{\partial z_{ij}} = \beta_j \frac{P}{z_{ij}}$$

Dans le cas du modèle de Box-Cox le plus général (Équation 8), le prix implicite d'un changement dans la caractéristique  $j$  sur le prix de l'habitation  $i$  est le suivant :

$$\text{Équation 12} \quad \frac{\partial p_i}{\partial x_{ij}} = p_i^{1-\lambda} (x_{ij}^{\mu-1} \beta_j)$$

Il y a **deux approches possibles pour calculer les prix implicites** quand les formules incluent le niveau d'une variable.

**L'approche la plus courante est d'utiliser le prix moyen (ou la valeur moyenne de la caractéristique) sur l'échantillon.**

Une approche alternative est de calculer le prix implicite pour chaque habitation de l'échantillon et ensuite de calculer la valeur moyenne de ces prix implicites.

## 5. Faut-il inclure les variables explicatives en niveau ou en logarithme ?

Les procédures d'inférence classiques ne permettent pas de répondre à cette question car les hypothèses (ou alternatives) ne sont pas emboîtées.

Par exemple, les deux modèles suivants sont des modèles non emboîtés :

$$\begin{aligned} \text{Équation 13} \quad p_i &= \beta_0 + x_{i1}\beta_1 + x_{i2}\beta_2 + u_{i1} \\ p_i &= \alpha_0 + \ln(x_{i1})\alpha_1 + \ln(x_{i2})\alpha_2 + u_{i2} \end{aligned}$$

En utilisant la même démarche que pour le test RESET, on peut estimer chacun des modèles de (Équation 13), calculer les valeurs prédites et les inclure dans l'autre modèle. Sous l'hypothèse nulle que la première forme est correctement spécifiée, alors une combinaison linéaire des logarithmes des  $x_i$  ne devrait pas améliorer le modèle et le coefficient associé ne devrait pas être significatif. De même, on peut réestimer le second modèle en incluant les valeurs prédites du premier modèle. C'est le principe du « test J ». Ce test peut permettre d'indiquer que l'un des modèles est meilleur que l'autre. Néanmoins, dans cette stratégie de test, quatre situations peuvent se produire : rejeter les deux modèles, n'en rejeter aucun, en rejeter un.

## 6. Mesurer la qualité de l'ajustement dans les modèles log-linéaires

Un indicateur simple généralement utilisé pour mesurer la qualité de l'ajustement d'un modèle de régression log-linéaire est le  $R^2$ . Afin qu'il soit comparable à celui provenant d'un modèle linéaire, des ajustements sont nécessaires.

Dans le cas d'un modèle où la variable expliquée est en logarithme, Wooldridge (2000) a proposé une méthode pour calculer un  $R^2$  comparable avec celui d'un modèle linéaire. La démarche permettant de calculer ce  $R^2$  modifié est la suivante :

1. Estimer le modèle de régression log-linéaire :  $\ln p_i = (\ln x_i)' \beta$ .
2. Calculer les valeurs prédites du modèle :  $\ln \hat{p}_i = (\ln x_i)' \hat{\beta}$ .
3. Calculer l'exponentielle de ces valeurs prédites :  $\tilde{p}_i = e^{\ln \hat{p}_i}$ .
4. Régresser  $p$  sur  $\tilde{p}$  sans constante :  $p_i = \tilde{p}_i' a$ .
5. Pour estimer  $p$  à partir du modèle de départ, calculer :  $\hat{p}_i = \tilde{p}_i' \hat{a}$ .
6. Calculer le coefficient de corrélation linéaire entre ces  $\hat{p}_i$  et les vrais  $p_i$ . Le carré de ce coefficient de corrélation est une sorte de  $R^2$  qui peut être comparé au  $R^2$  d'un modèle linéaire.

## 7. Variable expliquée en logarithme ou en niveau ?

### 7.1. Tester l'hypothèse de linéarité contre l'hypothèse de log-linéarité

Le test  $P_E$  permet de tester l'hypothèse de linéarité contre l'hypothèse de log-linéarité. Par exemple, supposons que les deux modèles en concurrence soient :

$$\begin{aligned} \text{Équation 14} \quad H_0 : y &= X\beta + \varepsilon \\ H_1 : \ln(y) &= (\ln X)\gamma + \varepsilon \end{aligned}$$

Notons  $\hat{\beta}$  et  $\hat{\gamma}$  les estimations des moindres carrés des paramètres des deux modèles. Le test  $P_E$  dont l'hypothèse nulle est  $H_0$  revient à tester la significativité du coefficient  $\hat{\alpha}$  dans le modèle (Équation 15).

$$\text{Équation 15} \quad y = X\beta + \alpha [\ln y - \ln(X\hat{\beta})] + \varepsilon_1$$

Le second terme (à droite) correspond à la différence entre les prédictions de  $\ln y$  obtenues à partir du modèle log-linéaire et celles obtenues à partir du logarithme des prédictions du modèle linéaire. Il est possible d'inverser les rôles et de tester  $H_0$  comme hypothèse alternative. La régression devient alors (Équation 16) :

$$\text{Équation 16} \quad \ln y = (\ln X)\beta + \alpha [\hat{\gamma} - e^{\ln(X)\hat{\gamma}}] + \varepsilon_2$$

### 7.2. Linéaire et log-linéaire vs. Box-Cox

Néanmoins, il est possible qu'aucun de ces deux modèles ne soit pertinent. Par exemple, un modèle plus général de type Box-Cox pourrait être plus adapté. Les modèles linéaire et log-linéaire peuvent être vus comme des cas particuliers d'un modèle de Box-Cox. Le test de Wald peut être utilisé pour tester si le paramètre  $\lambda$  (voir Équation 7 et Équation 8) est égal à 1 (modèle linéaire) ou à 0 (modèle log-linéaire).

Des tests du rapport de vraisemblance (*Likelihood Ratio test*) peuvent être employés pour tester si un modèle linéaire ou log-linéaire est plus adapté qu'un modèle de Box-Cox du type (Équation 8) (Haab et McConnell, 2002).

## XII – REFERENCES

- Cropper, M., Deck, L. et McConnell, K. (1988), 'On the choice of functional form for hedonic price functions', *The Review of Economics and Statistics* **70**(4), 668–675.
- Desaigues, B. et Point, P. (1993), *Economie du patrimoine naturel*, Economica, Paris.
- Greene, W. (2003), *Econometric Analysis*, Prentice Hall, Upper Saddle River, New Jersey.
- Haab, T. et McConnell, K. (2002), *Valuing Environmental and Natural Resources : the Econometrics of Non-Market Valuation*, Edward Elgar, Cheltenham, UK.
- Ihlanfeldt, K. et Taylor, L. (2004), 'Externality effects of small-scale hazardous waste sites : evidence from urban commercial property markets', *Journal of Environmental Economics and Management* **47**(1), 117–139.
- Palmquist, R. (1992), 'Valuing localized externalities', *Journal of Urban Economics* **32**, 40–44.
- Rosen, S. (1974), 'Hedonic prices and implicit markets : Product differentiation in pure competition', **82**(1), 34–55.
- Scherrer, S. (2004), *Comment évaluer les biens et services environnementaux ?*, La documentation Française, Paris.
- Song, Y. et Knapp, G. (2003), 'New urbanism and housing values : a disaggregate assessment', *Journal of Urban Economics* **51**, 218–238.
- Taylor, L.O. (2003), 'The Hedonic Method', in Champ, P.A., Boyle, K.J. et Brown, T.C. (eds), *A Primer on Nonmarket Valuation*, Kluwer Academic Publishers.
- Wooldridge, J. (2000), *Introductory Econometrics : A Modern Approach*, South-Western College Publishing, New York.

**I - Etudes :**

- 05 – E07  
Guillemette BUISSON
- 05 – E06  
Guillemette BUISSON Nouveautés
- 05 – E05  
Place de l'environnement dans le système juridique de l'OMC  
Ruth GABBAY
- 05 - E04  
Comment les politiques publiques peuvent-elles accélérer le progrès sur les technologies de lutte contre le changement climatique ?  
Aurélie VIEILLEFOSSE
- 05 - E03  
Modélisation du découplage des aides et environnement en agriculture  
Elsa LAVAL
- 05 - E02  
Efficacité de la filière piles et accumulateurs  
Olivier ARNOLD
- 05 - E01  
Les régulations environnementales ont-elles un effet sur le commerce extérieur de l'industrie française ?  
Sébastien RASPILLER, Nicolas RIEDINGER, Céline BONNET
- 04 - E10  
Les politiques environnementales ont-elles un impact sur la croissance ?  
Nicolas RIEDINGER
- 04 - E09  
Estimation des nuisances pour la collectivité générées par les éoliennes de Sigean  
Sébastien TERRA
- 04 - E08  
Stratégies d'échantillonnage et modèles de comptage dans la méthode des coûts de transport  
Sébastien TERRA
- 04 - E07  
Bien public global et instruments des politiques nationales unilatérales  
Christine CROS, Sylviane GASTALDO
- 04 - E06  
Principe de précaution et décision médicale  
Dominique BUREAU, Emmanuel MASSE
- 04 - E05  
Préservation des ressources globales et développement économique  
Dominique BUREAU
- 04 - E04  
Evaluation du coût subi par EDF suite à une mesure en faveur de la vie piscicole sur la Dordogne  
Franck FREDEFON
- 04 - E03  
Valorisation économique d'une amélioration de la qualité de l'eau de l'étang de Berre  
Franck FREDEFON
- 04 - E02  
La prise en compte du changement technique endogène affecte-telle l'équivalence entre taxes et permis ?  
Gilles SAINT-PAUL

- 04 - E01 Les différences de sévérité environnementale entre pays influencent-elles les comportements de localisation des groupes français ?  
Sébastien RASPILLER, Nicolas RIEDINGER
- 03 - E09 Evaluation économique des aménités récréatives d'un parc urbain : le cas du parc de Sceaux  
Sylvie SCHERRER
- 03 - E08 Analyse économique de la rentabilité des filtres à particules sur les véhicules diesels neufs  
Emmanuel MASSE
- 03 - E07 Note sur l'évaluation des infrastructures de transport et l'étalement urbain  
Dominique BUREAU, Nicolas THOUVEREZ
- 03 - E06 Evaluation des bénéfices pour le public de la protection des espaces littoraux remarquables  
Sylvie SCHERRER
- 03 - E05 Evaluation économique des aménités récréatives d'une zone humide intérieure : le cas du lac de Der  
Sylvie SCHERRER
- 03 - E04 Exploration des engagements futurs en matière de changement climatique  
Vincent VAN STEENBERGHE
- 03 - E03 Quels instruments pour une politique environnementale ?  
Gilles SAINT-PAUL
- 03 - E02 Couverture des charges d'infrastructure et tarification de l'usage de la route  
Isabelle ROVIRA, Martine PERBET
- 03 - E01 Les dommages visuels et sonores causés par les éoliennes : une évaluation par le consentement à payer des ménages dans le cas des éoliennes de Sigean  
Sylvie SCHERRER
- 02 - E07 Pollutions atmosphériques transfrontières : mise en œuvre du protocole de Goteborg et de la directive plafonds  
Daniel DELALANDE
- 02 - E06 Régulation du bruit à Roissy : efficacité et instruments économiques  
Dominique BUREAU
- 02 - E05 Gisement d'énergie éolienne par région : quelques éléments d'éclairage économique  
Sabine GUILLAUME
- 02 - E04 Les accords de Bonn et Marrakech : analyse quantitative et mise en perspective  
Sandrine ROCARD, Eve ROUMIGUIERES
- 02 - E03 Typologie des modes de gestion des déchets ménagers par les collectivités locales  
Anne-Claire BOITEL, Christine LAGARENNE

- 02 - E02 Evaluation économique des pertes d'usage dues aux tempêtes Lothar et Martin de décembre 1999 : le cas de la forêt de Fontainebleau  
Sylvie SCHERRER
- 02 - E01 Régulation de la durée des contrats dans le secteur de l'eau  
Patrick DERONZIER
- 01 - E07 Effet de serre document de base de la maquette SAGESSE  
Eve ROUMIGUIERES
- 01 - E06 Déterminants de la consommation en produits de l'agriculture biologique  
Sylvie SCHERRER
- 01 - E05 Effet de serre : quantification de l'effort économique par les parties du protocole de Kyoto  
Eve ROUMIGUIERES
- 01 - E04 Déterminants des comportements de tri des ménages  
Christine LAGARENNE, Séverine WILTGEN
- 01 - E03 Combinaison des instruments prix et quantités dans le cas de l'effet de serre  
Boris COURNEDE, Sylviane GASTALDO
- 01 - E02 Politiques nationales de lutte contre le changement climatique et réglementation de la concurrence : le cas de la fiscalité  
Jérôme RIEU
- 01 - E01 Effets économiques du Protocole de Kyoto : une maquette internationale  
Jean-Pierre BERTHIER, Martin GUESPEREAU, Eve ROUMIGUIERES

## II - Méthodes :

- 05 - M06 La monétarisation de l'indice pollution population pour l'analyse coût-bénéfice des projets de transport  
Pierre BARBERA
- 05 - M05 Guide de bonnes pratiques pour la mise en œuvre de la méthode des coûts de transport  
Sébastien TERRA
- 05 - M04 Guide de bonnes pratiques pour la mise en œuvre de la méthode d'évaluation contingente  
Sébastien TERRA
- 05 - M03 Options réelles environnementales Nouveautés  
Emmanuel MASSE, Stéphane GALLON
- 05 - M02 Guide pour l'élaboration de cahiers des charges pour des études de valorisation des dommages et aménités environnementales en 5 questions/réponses  
Sébastien TERRA
- 05 - M01 Guide pour la mise en œuvre de la méthode des prix hédoniques  
Sébastien TERRA



- 04 - M07 Maquette ECHEANCES : Epuisement des Combustibles selon Hotelling et Application Naturelle au Contingentement de l'Effet de Serre  
Hélène OLLIVIER
- 04 - M06 Articulation entre quotas échangeables et mesures de gestion des ressources halieutiques : éléments pour l'évaluation économique d'aires marines protégées  
Dominique BUREAU
- 04 - M05 Qu'est-ce qu'un marché de permis ? Adaptation du jeu de simulation de l'ENSAE à un marché de crédits « Azote »
- 04 - M04 Tourisme, loi littoral et économie de l'environnement  
Dominique BUREAU
- 04 - M03 Fiches DPSEEA élaborées à partir du rapport final de la commission d'orientation pour le plan santé Environnement  
Camille FEVRIER
- 04 - M02 Arbitrages intertemporels, risque et actualisation  
Stéphane GALLON, Emmanuel MASSE
- 04 - M01 Le cycle de la prévention et de l'information sur les risques  
Patrick MOMAL
- 03 - M03 La culture du risque et de la sûreté  
Patrick MOMAL
- 03 - M02 Rapport du groupe de réflexion environnement et applications de l'espace  
Bertrand GALTIER, Michel LEBLANC
- 03 - M01 Le système d'information environnementale français  
Armelle GIRY
- 02 - M02 Santé environnement : problèmes et méthodes  
Benoît VERGRIETTE
- 02 - M01 Intérêts et limites des variables biologiques en écotoxicologie aquatique  
Patrick FLAMMARION
- 01 - M02 Indicateurs environnementaux : méthodes et utilisation pour l'évaluation des politiques publiques  
Xavier DELACHE
- 01 - M01 Méthodologie de valorisation des biens environnementaux  
Sylvie SCHERRER

### III - Synthèses :

- 05 - S03 Les études de monétarisation des externalités associées à la gestion des déchets  
Benoît CHEZE, Olivier ARNOLD
- 05 - S02 Plan National d'Affectation des Quotas : retour d'expérience  
Sébastien MERCERON



- 05 - S01 Les différentes gestions du dossier de l'amiante  
Grégoire LAGNY
- 04 - S07 Mécanismes économiques à l'œuvre sur la biodiversité dans les secteurs de l'agriculture, la forêt, l'eau, la pêche, le tourisme et les transports  
Christine CROS
- 04 - S06 Evolution du régime d'indemnisation des catastrophes naturelles  
Annie ERHARD-CASSEGRAIN, Emmanuel MASSE, Patrick MOMAL
- 04 - S05 Développement durable et aménagement routier : le cas de la RN88  
Stéphanie ANTOINE
- 04 - S04 L'économie de l'effet de serre : point sur les engagements internationaux de lutte contre le changement climatique  
Aurélie VIEILLEFOSSE
- 04 - S03 Entreprises et développement durable  
Irène CABY
- 04 - S02 Références méthodologiques pour la prise en compte de l'environnement dans les projets routiers  
Stéphanie ANTOINE
- 04 - S01 Déchets ménagers en France. Financement du service et recyclage : Exemples de travaux d'évaluation économiques utiles à la décision publique  
Patrick DERONZIER, Olivier ARNOLD
- 03 - S06 L'évaluation des aménités et des dommages environnementaux  
Sylvie SCHERRER
- 03 - S05 Les enseignements pour la France des régimes de responsabilité environnementale en vigueur à l'étranger : l'exemple des Etats-Unis et du Brésil  
Catherine SCHLEGEL, Laurent VERDIER
- 03 - S04 Les engagements futurs dans les négociations sur le changement climatique  
Séminaire D4E
- 03 - S03 Economie de l'environnement et décision publique  
Dominique BUREAU
- 03 - S02 Biens publics mondiaux et négociations internationales  
Hélène FRANCES, François NASS
- 03 - S01 Axes pour la recherche en environnement et en développement durable dans le sixième programme cadre de recherche et développement de l'union européenne  
Groupe thématique national français « recherche européenne, environnement et développement durable »
- 02 - S02 Marchés de droits : expériences passées et débuts pour l'effet de serre  
Christine CROS, Sylviane GASTALDO
- 02 - S01 Microéconomie du développement durable : une introduction  
Dominique BUREAU

- 01 - S05 L'impact économique des tempêtes de décembre 1999  
Annie ERHARD-CASSEGRAIN
- 01 - S04 Ouverture des marchés de l'électricité et environnement  
Dominique BUREAU, Sylvie SCHERRER
- 01 - S03 La responsabilité environnementale  
Patrick MOMAL
- 01 - S02 Gouvernance mondiale et environnement  
Dominique BUREAU, Marie-Claire DAVEU, Sylviane GASTALDO
- 01 - S01 Les rapports environnementaux des entreprises  
Christine LAGARENNE, Marc AVIAM