eurostat
Methodologies&
Working papers

Manuel des indices des prix de l'immobilier résidentiel

Édition 2013

ILO
INTERNATIONAL MONETARY FUND
OECD
BETTER POLICIES FOR BETTER LIVES
UNITED NATIONS
UNCTAD
THE WORLD BANK
eurostat
EUROPEAN COMMISSION

eurostat
Methodologies &
Working papers

Manuel des indices des prix de l'immobilier résidentiel (IPIR)

Édition 2013

eurostat
EUROPEAN COMMISSION

Des informations complémentaires sur l'Union européenne sont disponibles sur Internet (http://europa.eu).

Les données de catalogage sont disponibles à la fin de la publication.

Luxembourg : Office des publications de l'Union européenne, 2013.

FMI
Traduction des Services linguistiques du Département des services généraux et de la technologie du FMI
Français, ISBN 978-1-49830-156-5 (version imprimée)
Français, ISBN 978-1-49833-908-7 (version PDF)
Français, doi : 10.5089/9781498301565.069 (version imprimée)
Français, doi : 10.5089/9781498339087.069 (version PDF)

Les opinions et les arguments contenus dans ce manuel ne reflètent pas nécessairement la position officielle de la Banque mondiale, de la Commission économique des Nations Unies pour l'Europe, du FMI, de l'OIT ou de l'OCDE, ni celle des gouvernements des États membres de ces organisations, ou d'Eurostat et de la Commission européenne.

Thème : Économie et finance
Série : Methodologies & Working papers

© Banque mondiale, Commission économique des Nations Unies pour l'Europe, Fonds monétaire international, Organisation de coopération et de développement économiques, Organisation internationale du travail, Union européenne, 2013
La reproduction est autorisée sous réserve de la citation des sources.

Photos : © Phovoir
La reproduction des photos est autorisée mais limitée à des fins non commerciales et uniquement dans le contexte de cette publication.

Imprimé en Belgique

IMPRIMÉ SUR PAPIER BLANCHI SANS CHLORE ÉLÉMENTAIRE (ECF)

Également disponible sous le titre *Manuel des indices des prix de l'immobilier résidentiel (IPIR)*

OIT
Anglais, ISBN 978-92-2-127359-2 (version imprimée)
Anglais, ISBN 978-92-2-127360-8 (version PDF)

UNION EUROPÉENNE
ISBN 978-92-79-25984-5
doi : 10.2785/34007
Cat. No KS-RA-12-022-EN-N

OCDE
Anglais, ISBN 978-92-6-419718-3 (version PDF)

Table des matières

Avant-propos 7
Préface 9
1. Introduction 11
2. Utilisations des indices des prix de l'immobilier résidentiel 15
3. Éléments d'un cadre conceptuel 23
4. Méthodes de stratification ou de poststratification 39
5. Méthodes de régression hédonique 51
6. Méthodes des ventes répétées 69
7. Méthodes basées sur l'estimation 77
8. Décomposition d'un IPIR en composantes «terrain» et «bâtiments» 85
9. Sources de données 105
10. Méthodes utilisées actuellement 117
11. Exemples empiriques 141
12. Recommandations 157
Glossaire 165
Bibliographie 171
Index 181

Liste des tableaux

3.1. Estimations en pourcentage des ratios loyer/valeur (taux de capitalisation) 37
4.1. Probabilité de vente dans chaque strate pour l'échantillon 46
4.2. Indices-chaînes de Fisher à produits appariés et à base fixe, et indices des prix moyens, médians et à modèles représentatifs 48
4.3. Indice de Fisher à base fixe pour l'année glissante, indice-chaîne de Fisher à moyenne mobile, et indice de Fisher des prix à base fixe et à moyenne mobile 50
5.1. Indices de prix log-linéaires à variables indicatrices temporelles et indice-chaîne de Fisher des prix moyens à échantillon stratifié 63
5.2. Indices de prix linéaires à variables indicatrices temporelles, indice de prix log-log à variables indicatrices temporelles et indice-chaîne de Fisher des prix moyens à échantillon stratifié 65
5.3. Indices-chaînes des prix à imputation hédonique de Laspeyres, Paasche et Fisher 67
6.1. Indice des prix reposant sur la méthode des ventes répétées, indice-chaîne de Fisher des prix moyens à échantillon stratifié et indice des prix de Fisher avec imputation hédonique 75
7.1. Indices RPVE, indice de Fisher à imputation hédonique et indice des ventes répétées 84
8.1. Prix du terrain (P_{L1}), prix des bâtiments ajustés en fonction de la qualité (P_{S1}), indice du coût total de la production de logements (P_1) et indice de Fischer à imputation hédonique pour les logements 91
8.2. Prix du terrain (P_{L2}), prix des structures (P_{S2}), indice global des prix de l'immobilier utilisant la modélisation par lignes brisées du terrain (P_2) et indice de Fisher à imputation hédonique 93
8.3. Prix du terrain (P_{L3}), prix des bâtiments ajustés en fonction de la qualité (P_{S3}), indice global des prix de l'immobilier avec restrictions de monotonie sur les structures (P_3) et indice global des prix de l'immobilier utilisant la méthodologie des lignes brisées pour le terrain (P_2) 95
8.4. Prix du terrain (P_{L4}), prix des bâtiments ajustés en fonction de la qualité (P_{S4}) et indice global des prix de l'immobilier utilisant des informations exogènes sur le prix des bâtiments (P_4) 96
8.5. Indices des prix de l'immobilier utilisant des informations exogènes (P_4) et des restrictions de monotonie (P_3), indice-chaîne de Fisher à imputation hédonique et indice-chaîne de Fisher à échantillon stratifié 98
8.6. Prix du terrain (P_{L4}), prix des bâtiments ajustés en fonction de la qualité (P_{S4}), indice global des prix de l'immobilier utilisant des informations exogènes sur le prix des bâtiments (P_4) et leurs équivalents à fenêtres mobiles (P_{RWL}) et (P_{RW}) 99
8.7. Indices approximatifs de prix du stock et indices à imputation hédonique (P_{Stock1}) et à stratification (P_{Stock2}), et indice de Fisher des prix de vente à imputation hédonique 101
8.8. Indices approximatifs de prix du stock de maisons (P_{Stock}), du stock de terrains (P_{LStock}), du stock de bâtiments (P_{SStock}) et indices correspondants basés sur les ventes (P_{L4} et P_4) 103
10.1. Indices des prix de l'immobilier publiés au Japon 127
10.2. Indices des prix de l'immobilier résidentiel publiés au Royaume-Uni 131
10.3. Décomposition par statut du stock de logements en Afrique du Sud (selon le recensement de 2001) 137
10.4. Distribution du nombre de pièces dans les habitations informelles 138
10.5. Déterminants des prix 139
10.6. Décomposition en pourcentage des matériaux utilisés dans la construction des habitations informelles et traditionnelles en Afrique du Sud 139
10.7. Évaluation des obstacles 139
11.1. Moyenne, médiane, variation en pourcentage, écart-type et coefficient d'asymétrie 143
11.2. Dépenses régionales, prix et volumes (quantités implicites), utilisant le prix médian comme prix régional 145

11.3. Indice global des prix du logement utilisant la médiane des prix et d'autres formules pour agréger les régions A, B et C 146
11.4. Dépenses régionales, prix et volumes (quantités implicites), utilisant la moyenne pour calculer les prix régionaux 147
11.5. ÍIndice global des prix du logement utilisant la moyenne des prix et d'autres formules pour agréger les régions A, B et C 147
11.6. Résultats du modèle log-linéaire pour un exemple simple 148
11.7. Résultats de la régression groupée pour les années 2006 et 2007 150
11.8. Résultats de la régression groupée pour les années 2006 à 2008 150
11.9. Résultats de la régression groupée pour les années 2007 et 2008 151
11.10. Résultats de la régression pour 2006 152
11.11. Résultats de la régression pour 2007 152
11.12. Valeurs moyennes des caractéristiques pour la période de référence (2006) 153
11.13. Données de ventes répétées 153
11.14. Variables indicatrices des ventes répétées 154
11.15. Régression non pondérée des ventes répétées 155
11.16. Régression pondérée des ventes répétées 155
11.17. Indice des ventes répétées (2002 = 100) 156
11.18. Taux de croissance en pourcentage des différents indices de prix du logement (2007) 156

Liste des graphiques

4.1. Indices-chaînes de Fisher à produits appariés et à base fixe, et indices des prix moyens, médians et à modèles représentatifs 48
4.2. Indice de Fisher à base fixe pour l'année glissante, indice-chaîne de Fisher à moyenne mobile, et indice de Fisher des prix à base fixe et à moyenne mobile 49
5.1. Indices de prix log-linéaires à variables indicatrices temporelles et indice-chaîne de Fisher des prix moyens à échantillon stratifié 62
5.2. Indices de prix linéaires à variables indicatrices temporelles, indice de prix log-log à variables indicatrices temporelles et indice-chaîne de Fisher des prix moyens à échantillon stratifié 64
5.3. Indices-chaînes des prix à imputation hédonique de Laspeyres, Paasche et Fisher 67
5.4. Indice de Fisher à imputation des prix, indice-chaîne de Fisher des prix moyens à échantillon stratifié, indice de prix linéaires à variables indicatrices temporelles et indice de prix log-log à variables indicatrices temporelles 68
6.1. Indice des prix reposant sur la méthode des ventes répétées, indice-chaîne des prix de Fisher à échantillon stratifié et indice des prix de Fisher à imputation hédonique 75
7.1. Indices RPVE, indice de Fisher à imputation hédonique et indice des ventes répétées 84
8.1. Prix du terrain (P_{L1}), prix des bâtiments ajustés en fonction de la qualité (P_{S1}), indice du coût total de la production de logements (P_1) et indice de Fischer à imputation hédonique pour les logements 90
8.2. Prix du terrain (P_{L2}), prix des bâtiments (P_{S2}), indice global des prix de l'immobilier utilisant la modélisation par lignes brisées du terrain (P_2) et indice de Fisher à imputation hédonique 92
8.3. Prix du terrain (P_{L3}), prix des bâtiments ajustés en fonction de la qualité (P_{S3}), indice global des prix de l'immobilier avec restrictions de monotonie sur les structures (P_3) et indice global des prix de l'immobilier utilisant la méthodologie des lignes brisées pour le terrain (P_2) 95
8.4. Prix du terrain (P_{L4}), prix des bâtiments ajustés en fonction de la qualité (P_{S4}) et indice global des prix de l'immobilier utilisant des informations exogènes sur le prix des bâtiments (P_4) 96
8.5. Indices des prix de l'immobilier utilisant des informations exogènes (P_4) et des restrictions de monotonie (P_3), indice-chaîne de Fisher à imputation hédonique et indice-chaîne de Fisher à échantillon stratifié 97
8.6. Indices approximatifs de prix du stock et indices à imputation hédonique (P_{Stock1}) et à stratification (P_{Stock2}), et indice de Fisher des prix de vente à imputation hédonique 101
8.7. Indices approximatifs de prix du stock de maisons (P_{Stock}), du stock de terrains (P_{LStock}), du stock de bâtiments (P_{SStock}) et indices correspondants basés sur les ventes (P_{L4} et P_4) 103
Diagramme : Chronologie de l'achat d'un logement et indices des prix des logements 107
10.1. Quatre indices des prix de l'immobilier résidentiel au Canada 124
10.2. Transmission d'informations immobilières 128
10.3. Quatre indices des prix de l'immobilier résidentiel au Japon (janvier 1999 = 100) 129
10.4. Déroulement de l'achat d'un logement 130
10.5. Indices NHB RESIDEX : indices des prix en zone urbaine en Inde 132
10.6. Indice national trimestriel des prix des logements existants : valeur nominale et réelle 134
10.7. Indice national trimestriel des prix réels des logements existants : variations annuelles en pourcentage 134
10.8. Indice national annuel des prix des logements existants 135
10.9. Indice annuel des prix réels des logements existants : principales zones métropolitaines 135
10.10. Indice annuel des prix réels des logements existants : logements subventionnés (VIS) et non subventionnés (NOVIS) 136
11.1. Distribution des prix du logement en 2008 144

Avant-propos

Ce *Manuel des indices des prix de l'immobilier résidentiel (IPIR)* constitue le premier tour d'horizon complet des questions conceptuelles et pratiques concernant l'établissement des indices des prix de l'immobilier résidentiel.

Ce *Manuel*, dont la rédaction a été coordonnée par le Bureau statistique de l'Union européenne (Eurostat) par le biais d'un groupe de travail intersecrétariat sur les statistiques de prix, est le fruit de la collaboration de six organisations : la Banque mondiale, le Bureau statistique de l'Union européenne (Eurostat), la Commission économique des Nations Unies pour l'Europe, le Fonds monétaire international (FMI), l'Organisation de coopération et de développement économiques (OCDE) et l'Organisation internationale du travail (OIT). L'ouvrage est une publication collective de ces organisations.

Le *Manuel* a pour but d'offrir des orientations pratiques pour l'établissement des indices des prix de l'immobilier et d'en améliorer la comparabilité internationale. Il présente les différentes utilisations qui sont faites de ces indices, fournit des détails sur les données nécessaires à leur élaboration et sur les méthodes applicables en la matière, et formule des recommandations. Son principal objectif est d'aider les statisticiens chargés d'établir les indices des prix de l'immobilier, en particulier dans les pays où ces indices sont en cours de révision ou de création. Le *Manuel* s'appuie sur une somme considérable d'expérience et de connaissances pour présenter des méthodes de mesure pratiques et adaptées. Il devrait aussi aider les pays à construire leurs propres IPIR d'une manière qui rende ces indices plus comparables. Dans la mesure où il met à profit un vaste corpus de connaissances sur le sujet, le *Manuel* peut être utilisé aussi bien pour l'autoenseignement qu'à titre d'outil didactique comme support de cours sur les indices des prix de l'immobilier résidentiel.

Le *Manuel* pourrait aussi être utile aux autres utilisateurs des IPIR, tels que les entreprises, les décideurs et les chercheurs, comme source d'information, non seulement sur les différentes méthodes employées pour recueillir les données et établir les indices, mais aussi sur les limites de ces méthodes. À cet égard, il pourrait faciliter l'interprétation des résultats.

La rédaction et la révision de cet ouvrage ont nécessité un grand nombre de réunions sur une période de trois ans, auxquelles ont participé des spécialistes des IPIR provenant de bureaux statistiques nationaux, d'organisations internationales et régionales, d'universités et d'instituts de recherche. Sans leurs conseils, empreints d'expérience et de sagesse, la réalisation de ce *Manuel* n'aurait pas été possible. Une version électronique du *Manuel* est disponible à l'adresse Internet http://epp.eurostat.ec.europa.eu. Le groupe de travail intersecrétariat considère ce *Manuel* comme un «document vivant» qui a vocation à être modifié et actualisé pour en préciser certains points.

Le groupe de travail intersecrétariat encourage les lecteurs à lui faire part de leurs observations éventuelles; elles sont à adresser à Eurostat (courriel : ESTAT-hicp-methods@ec.europa.eu) et il en sera tenu compte dans les révisions futures.

Walter Radermacher
Chef statisticien de l'Union européenne
Directeur général
Eurostat – Bureau statistique de l'Union européenne

Alfredo M. Leone
Directeur par intérim
Département des statistiques
Fonds monétaire international

Lidia Bratanova
Directrice de la Division de statistique
Commission économique des Nations Unies pour l'Europe

Rafael Diez de Medina
Chef statisticien
Directeur du Département de statistique
Organisation internationale du travail

Martine Durand
Chef statisticienne
Directrice de la Direction des statistiques
Organisation de coopération et de développement économiques

Shaida Badiee
Directrice, Groupe de gestion des données sur le développement
Banque mondiale

Préface

Introduction

Ce *Manuel* a pour but de faciliter l'établissement d'indices des prix de l'immobilier résidentiel dans les pays où ils n'existent pas encore et, là où ils existent déjà, permettre leur amélioration si le besoin s'en fait sentir. Il est conçu pour donner des orientations pratiques en vue de la construction d'indices des prix de l'immobilier résidentiel, quel que soit le degré de développement du pays considéré, et pour permettre une meilleure comparabilité internationale de ces indices. Il décrit les différentes utilisations qui sont faites de ces indices, fournit des détails sur les données nécessaires à leur élaboration et sur les méthodes applicables en la matière, et formule des recommandations. La production de ce *Manuel* a été financée par Eurostat, qui en a aussi assuré la coordination.

Toile de fond

La nécessité de disposer d'indices des prix de l'immobilier adaptés a été reconnue lors d'une conférence organisée à Washington en octobre 2003 par le Fonds monétaire international (FMI) et la Banque des règlements internationaux (BRI). À la suite de cette conférence, un chapitre sur les indices des prix de l'immobilier résidentiel a été ajouté au *Guide d'établissement des indicateurs de solidité financière*, publié par le FMI. L'idée d'élaborer un manuel plus détaillé remonte à un atelier sur les indices des prix de l'immobilier organisé à Paris en novembre 2006 par l'Organisation de coopération et de développement économiques (OCDE) et le FMI. Le manuel envisagé devait compléter les manuels internationaux existants sur les indices des prix à la consommation, les indices des prix à la production et les indices des prix d'importation et d'exportation, qui sont publiés sous l'égide du groupe de travail intersecrétariat sur les statistiques de prix.

Compte tenu des liens étroits entre cette question et ses propres travaux sur l'inclusion des logements occupés par leurs propriétaires dans l'indice harmonisé des prix à la consommation et de la place des indices des prix de l'immobilier parmi les *Principaux indicateurs économiques européens*, Eurostat a accepté de prendre la tête de cette initiative en en assurant la coordination et en finançant la préparation du *Manuel*.

Le projet de manuel a été discuté lors de la conférence Eurostat–IAOS–SFI sur les indices des prix de l'immobilier résidentiel qui s'est tenue à Bâle les 11 et 12 novembre 2009. Des versions préliminaires du *Manuel* auront été présentées et examinées à plusieurs occasions, en particulier lors de la réunion de la Commission économique des Nations Unies pour l'Europe et de l'OIT consacrée aux indices des prix à la consommation, qui s'est tenue à Genève du 10 au 12 mai 2010, d'un atelier organisé les 10 et 11 février 2011 à La Haye et de la douzième réunion du Groupe d'Ottawa qui a eu lieu du 4 au 6 mai 2011 à Wellington.

Un guide pour les lecteurs

Bien que les chapitres ne soient pas tous autonomes, le *Manuel* n'est pas conçu comme un tout dont les parties seraient indissociables. Par exemple, certains chapitres pourront facilement être laissés de côté sans problème par les statisticiens qui s'intéressent plus particulièrement aux questions méthodologiques. Le chapitre 1 donne des précisions sur le contenu du *Manuel*.

Le *Manuel* ne peut pas être trop prescriptif, et ce pour deux raisons. Premièrement, il n'est pas toujours possible de donner des orientations pratiques, car pour certains problèmes conceptuels la solution n'est pas toujours clairement définie et des choix doivent être faits quant à son application pratique. Deuxièmement, ce sont les données et les ressources dont dispose l'institut national de statistiques (ou tout autre organisme statistique) qui détermineront ce qui est applicable et ce qui peut être accompli dans chaque pays.

Remerciements

Le *Manuel* a été rédigé sous la direction de Statistics Netherlands; Bert M. Balk a coordonné les activités relatives au projet. Jan de Haan et W. Erwin Diewert ont été les principaux rédacteurs. Les auteurs des différents chapitres sont les suivants :

Préface Bert Balk, Jan de Haan et David Fenwick

1. *Introduction* Bert Balk
2. *Utilisation des indices des prix de l'immobilier résidentiel* David Fenwick

3. *Éléments d'un cadre conceptuel* Erwin Diewert
4. *Méthodes de stratification ou de poststratification* Jan de Haan et Erwin Diewert
5. *Méthodes de régression hédonique* Jan de Haan et Erwin Diewert
6. *Méthodes des ventes répétées* Jan de Haan
7. *Méthodes basées sur l'estimation* Jan de Haan
8. *Décomposition d'un IPIR en composante «terrain» et «bâtiments»* Erwin Diewert
9. *Sources de données* David Fenwick
10. *Méthodes utilisées actuellement* David Fenwick
11. *Exemples empiriques* Marc Prud'homme et Erwin Diewert
12. *Recommandations* David Fenwick, Erwin Diewert et Jan de Haan

Glossaire Jan de Haan

La qualité du *Manuel* a été grandement rehaussée par les contributions précieuses d'un grand nombre de personnes et d'organisations, notamment des statisticiens et des utilisateurs d'indices des prix de l'immobilier résidentiel dans différentes parties du monde. Ils sont naturellement bien trop nombreux pour être remerciés nommément ici.

La BRI (et en particulier Paul Van den Bergh) ont été d'excellents hôtes lors de l'atelier qui s'est tenu à Bâle en 2009. Un grand merci aussi à la Commission économique des Nations Unies pour l'Europe (et en particulier à Carsten Boldsen) qui a largement participé à l'organisation de l'atelier de Bâle, et de la séance spéciale consacrée au *Manuel* à l'occasion de la réunion conjointe de 2010 de la Commission et de l'OIT sur l'indice des prix à la consommation.

Des remerciements particuliers sont dus à Irmtraud Beuerlein, Simon Coté, Lee Everts, Gregory Klump, Jose Vicente Romero, Patrick Sabourin, A.P. Saxena et Chihiro Shimizu, qui ont apporté leur contribution aux études de cas par pays, ainsi qu'à Emily Carless, Preechaya Chavalittumrony, Ali Hepşen, Marissa Gonzalez Guzman et Hector Zarate, qui ont fourni d'autres informations de base sur les indices publiés. Carlos Brás, Morris Davis, Martin Eiglsperger, Timothy Erickson, Rui Evangelista, Dennis Fixler, John Greenlees, Brian Graf, Vanda Guerreiro, Ronald Johnson, Marcel van Kints, Andrew Leventis, Bogdan Marola, Daniel Santos, Mick Silver, Leo Sveidkauskas, Randall Verbrugge, David Wasshausen ont fait part d'utiles observations sur les avant-projets de manuel, de même que les participants à l'atelier de La Haye, en particulier Marc Francke et Jan Walschots. Les commentaires d'Eurostat, de la BRI, du FMI et de la BCE ont aussi été fort appréciés. Enfin, le projet a bénéficié des observations de Rens Hendriks et Ning Huang ainsi que de leur aide pour traiter les données. Qu'ils en soient ici tous remerciés.

1

Introduction

1 Introduction

1.1 L'immobilier résidentiel est une source de richesse et, dans la mesure où les propriétaires occupent leur bien foncier, un facteur déterminant du coût de la vie. Le prix d'un logement se distingue du coût des services qu'il fournit, bien que les deux notions soient à l'évidence étroitement liées.

1.2 Le suivi de l'évolution des prix immobiliers est considéré comme important, surtout en période de tourmente économique. Or, la méthode utilisée à cette fin varie selon les pays et il existe parfois deux méthodes concurrentes ou plus dans un pays. Il va sans dire que cette situation n'est pas propice à l'élaboration de mesures gouvernementales cohérentes, fondées sur des comparaisons internationales valables.

1.3 Dans ce contexte, le projet d'élaboration d'un manuel relatif aux indices des prix du logement, ou plus généralement de l'immobilier à usage résidentiel, prend tout son sens[1]. Les objectifs fondamentaux du manuel sont les suivants :

- donner des orientations aux statisticiens désireux d'établir des indices des prix de l'immobilier résidentiel ou de modifier les indices existants dans un souci d'harmonisation internationale;
- fournir une analyse et une comparaison des différents objectifs et des cadres conceptuels correspondants;
- établir un inventaire des pratiques existantes.

Le contenu du manuel est présenté ci-dessous.

1.4 Le *chapitre 2* passe en revue divers domaines dans lesquels les indices des prix de l'immobilier résidentiel (IPIR) jouent un rôle. Il en examine les applications suivantes :

- comme indicateur macroéconomique de l'activité économique;
- aux fins de politique monétaire et de ciblage de l'inflation;
- comme instrument d'évaluation de la valeur d'une composante du patrimoine réel;
- comme indicateur de la stabilité ou de la solidité financière pour mesurer l'exposition au risque;
- comme déflateur dans la comptabilité nationale;
- en tant qu'élément d'information permettant aux citoyens de décider s'ils doivent vendre ou acheter un bien immobilier résidentiel;
- à titre d'intrant de l'indice des prix à la consommation; et
- aux fins d'utilisation dans les comparaisons interrégionales et internationales.

1.5 Le *chapitre 3,* qui porte sur les différentes utilisations d'un IPIR, s'attache à combler les lacunes du Système de comptabilité nationale et de la compilation d'un indice des prix à la consommation. Si l'on parvient à construire des IPIR pour pallier ces déficiences, la famille d'IPIR en résultant satisfera probablement aux besoins de la plupart des utilisateurs.

1 Cette proposition a en fait été formulée à l'issue de l'atelier conjoint OCDE–FMI sur les indices de prix immobiliers qui s'est tenu à Paris les 6 et 7 novembre 2006.

1.6 Globalement, on peut distinguer deux types d'IPIR : un indice des prix à qualité constante du *stock* de logements résidentiels à un moment donné, et un indice des prix à qualité constante des *ventes* de biens immobiliers résidentiels intervenues au cours d'une période donnée. La construction de ces deux types d'indices diffère, notamment en ce qui concerne la pondération associée à chacun d'eux.

1.7 Le chapitre 3 résume ensuite les quatre principales méthodes de construction d'un IPIR. Les dernières sections traitent divers sujets, comme la périodicité de l'IPIR, la compatibilité des estimations mensuelles et trimestrielles et celle des estimations trimestrielles et annuelles, les politiques de révision et les ajustements saisonniers.

1.8 Les *chapitres 4 à 7* analysent en profondeur les principales méthodes d'établissement des IPIR. Les plus simples se fondent sur une mesure de la tendance centrale de la distribution des prix de transaction au cours d'une période donnée, en particulier la moyenne ou la médiane. Les distributions des prix du logement présentant généralement une asymétrie positive (ce qui tient essentiellement au caractère hétérogène du logement, à l'asymétrie positive des distributions des revenus et au prix minimal nul des transactions), on utilise plus souvent la médiane que la moyenne. Comme le calcul de la médiane ne nécessite pas de données sur les caractéristiques du logement, il est facile de construire un indice qui suit l'évolution du prix de vente du logement médian d'une période sur l'autre. Les indices médians présentent en outre l'intérêt d'être faciles à comprendre.

1.9 Les indices fondés sur une médiane simple ont pour inconvénient majeur de fournir des estimations très bruitées de l'évolution des prix. L'ensemble (ou un échantillon) des logements ayant effectivement fait l'objet d'une transaction au cours d'une période est généralement de petite taille et n'est pas nécessairement représentatif du stock total de logements. Les variations dans la composition des biens immobiliers vendus influenceront donc davantage le prix médian de l'échantillon que le prix médian du stock de logements. L'erreur systématique, ou biais, présente un problème peut-être plus important que le bruit à court terme. Un indice médian peut être entaché d'un biais quand la qualité du stock de logements varie au cours du temps. Des biais peuvent également apparaître si certaines catégories de logements sont vendues plus souvent que d'autres et si les prix des deux catégories évoluent différemment au même moment.

1.10 Une méthode générale de réduction du biais de sélection de l'échantillon est la (post)stratification. Cette technique est analysée au *chapitre 4.*

1.11 Le *chapitre 5* examine la méthode de régression hédonique. Celle-ci reconnaît que des biens hétérogènes peuvent être décrits par leurs attributs ou caractéristiques. Autrement dit, chaque bien est essentiellement un ensemble de caractéristiques de performance. Dans le contexte de l'immobilier, ces caractéristiques peuvent concerner tout autant le bâtiment que l'emplacement des biens concernés. Bien qu'il n'existe

pas de marché pour les caractéristiques, puisque celles-ci ne peuvent être vendues séparément, la demande et l'offre de biens immobiliers déterminent implicitement leur contribution marginale aux prix immobiliers. On peut faire appel à des techniques de régression pour estimer ces contributions marginales ou les prix implicites.

1.12 Le chapitre analyse, sous un angle non technique, les principaux modèles utilisés et les méthodes employées pour établir des IPIR à partir de l'estimation de ces modèles. De manière générale, la méthode de régression hédonique offre probablement le meilleur moyen de construire des IPIR à qualité constante pour différentes catégories de biens immobiliers résidentiels. Néanmoins, c'est aussi celle qui exige le plus grand nombre de données.

1.13 La méthode des ventes répétées, examinée au *chapitre 6,* fait appel aux informations concernant les biens immobiliers qui ont été vendus plus d'une fois. Comme elle utilise uniquement des «paniers fixes», il n'y a pas de variation des paramètres de qualité à neutraliser. Sous sa forme de base, les seules informations requises sont le prix, la date de la vente et l'adresse du bien immobilier. Cette méthode exige donc beaucoup moins de données que les méthodes hédoniques. Elle tiendra en outre automatiquement compte de la micro-localisation (l'adresse), ce que les méthodes hédoniques ne sont pas en mesure de faire.

1.14 La méthode des produits appariés (ou du panier fixe), qui compare les prix du même bien au cours du temps, est le point de départ naturel de la construction d'un indice de prix. Étant donnée la faible fréquence des transactions et l'évolution constante de la qualité des logements, la méthode standard du panier fixe ne peut être appliquée directement. La méthode des ventes répétées s'efforce de résoudre ce problème en tenant uniquement compte des biens immobiliers qui ont été vendus plus d'une fois au cours de la période observée. Cette approche risque néanmoins de produire un nombre relativement faible d'observations et un biais de sélection de l'échantillon. Pour pallier ces problèmes, on peut faire appel aux valeurs estimées des biens immobiliers.

1.15 Il existe dans de nombreux pays des évaluations gouvernementales officielles de l'ensemble des biens immobiliers, car ces données sont nécessaires à des fins fiscales. Si les évaluations concernent une date de référence donnée, on peut construire un IPIR en associant les prix effectifs des ventes aux valeurs estimées. Cette démarche constitue une variante de la méthode de l'appariement de modèles, dont elle se distingue par la prise en compte des variations de composition. Dans ce cas, il est inutile de recourir à des techniques économétriques. Les diverses méthodes d'évaluation, et en particulier celle du ratio prix de vente/prix estimé, sont examinées au *chapitre 7.*

1.16 Les chapitres 4 à 7 présentent tous en conclusion des exemples empiriques testés sur des données réelles afin d'illustrer les méthodes examinées et de fournir des éléments de référence complémentaires. Le jeu de données couvre les ventes de biens immobiliers résidentiels dans 14 quartiers d'une ville relativement petite des Pays-Bas. Comme ces chapitres le montreront, la plupart des méthodes ne sont pas en mesure de décomposer un IPIR en terrains et en bâtiments. Le *chapitre 8* se penche sur la façon dont on peut utiliser les méthodes de régression hédoniques pour obtenir cette décomposition et examine comment construire un IPIR pour le *stock* de biens immobiliers quand des méthodes de régression hédoniques sont employées. Se fondant sur des données réelles, il propose en outre différents moyens de surmonter les problèmes pratiques souvent rencontrés dans les travaux empiriques de cette nature, comme la forte corrélation entre la superficie du bâtiment et celle du terrain.

1.17 Dans la pratique, étant donné le coût élevé lié à la conduite d'études des prix immobiliers à des fins spécifiques, les méthodes adoptées par les organismes statistiques et d'autres organismes pour construire des IPIR ont été essentiellement déterminées par les jeux de données des prix immobiliers résultant de processus, juridiques et autres, intervenant dans le cadre de l'achat d'un logement. Les indices ainsi construits peuvent varier en fonction du moment où le prix est évalué, par exemple selon que le prix définitif de la transaction ou celui précédemment estimé pour obtenir un crédit est retenu. La quantité d'informations détaillées disponibles sur les caractéristiques des biens immobiliers vendus influencera en outre les méthodes de calcul de l'indice et limitera souvent le nombre de techniques utilisables pour mesurer l'effet qualité des modifications de taille ou d'emplacement des logements. Ainsi, le choix de la méthode pour construire un indice a toujours été conditionné par les données disponibles.

1.18 Le *chapitre 9* examine, sous un angle qualitatif, les différentes sources de données utilisables pour construire des IPIR : presse écrite, agences immobilières, sociétés de prêts hypothécaires, cadastres et centres des impôts. Il se penche dans sa dernière partie sur la situation de nombreux pays en développement où les données sont rares et où la question de la propriété foncière est ambiguë.

1.19 Le *chapitre 10* recense les IPIR disponibles dans différents pays et présente quelques études de cas. Il s'appuie sur les métadonnées recueillies par divers organismes, notamment la Banque centrale européenne et la Banque des règlements internationaux, et sur une enquête récemment conduite par Eurostat en relation avec l'intégration des coûts des logements occupés par leurs propriétaires dans l'indice harmonisé des prix à la consommation de l'Union européenne, qui a été élargi de manière à couvrir quelques pays non membres de l'UE.

1.20 Le *chapitre 11* fournit d'autres indications pratiques en expliquant le fonctionnement des méthodes de construction d'un IPIR (à l'exclusion de la méthode rapport prix de vente/évaluation) présentées aux chapitres 4, 5 et 6 au moyen d'exemples simples faisant appel à des jeux restreints de données.

1.21 Pour conclure, le *chapitre 12* formule des recommandations.

Utilisations des indices des prix de l'immobilier résidentiel

2

Introduction

2.1 Dans de nombreux secteurs de la société, des particuliers ou des organismes font appel, directement ou indirectement, à des indices des prix immobiliers résidentiels (IPIR) pour influencer, dans la pratique, le processus de décision, ou éclairer la formulation et la conduite de la politique économique. Les différents usages auxquels un indice est destiné peuvent influencer sensiblement le choix de son champ et la méthode utilisée pour le construire.

2.2 L'immobilier constitue souvent le principal investissement des ménages. Il compte également pour la plus grande part de la richesse dans le compte de patrimoine de la majorité des pays. Les variations des prix du logement peuvent avoir des conséquences majeures pour les particuliers. Celles de la valeur des biens immobiliers et du niveau d'endettement des ménages, par exemple, peuvent se répercuter sur l'ensemble de l'économie. De fait, l'évolution des prix immobiliers joue souvent sur les dépenses de consommation en raison des effets de patrimoine et de leurs retombées sur la confiance des consommateurs. Ces prix influent sur les dépenses d'aménagement et de rénovation qui, dans de nombreux pays, sont supérieures aux dépenses globales de construction de logements neufs. Ils jouent un rôle essentiel pour évaluer l'accessibilité à la propriété, dont certains pays ont fait l'objectif central de leur politique de logement. Leurs fluctuations influencent également la décision de construire des nouveaux logements (l'offre) et celle de devenir propriétaire (la demande)[1]. Les investisseurs font appel aux indices des prix immobiliers pour mesurer le patrimoine, mais aussi pour évaluer les taux de rendement actuels et futurs[2].

2.3 De manière plus générale, les analystes, les responsables politiques et les organismes financiers suivent l'évolution des prix immobiliers pour mieux appréhender la conjoncture sur les marchés de l'immobilier et du crédit, en observer les retombées sur l'activité économique, la stabilité et la viabilité financière[3]. À titre d'exemple, les établissements de crédit hypothécaire se fondent sur les informations relatives à l'inflation des prix immobiliers pour évaluer le risque de défaillance. Les banques centrales s'appuient souvent sur les fluctuations des indices des prix immobiliers pour suivre la capacité d'emprunt des ménages et leur endettement[4] et leurs effets sur la consommation globale[5].

2.4 Dans ce contexte, il convient de souligner que les différentes utilisations des indices des prix de l'immobilier résidentiel peuvent faire appel à des bases conceptuelles et des méthodologies distinctes, même si, dans la pratique, d'autres facteurs peuvent intervenir, comme la disponibilité des données[6]. En général, aucun indicateur de l'évolution des prix immobiliers ne peut satisfaire à lui seul à toutes les finalités. Ainsi, pour estimer la dynamique des prix du marché immobilier en vue d'assurer le suivi de l'inflation immobilière, telle que ressentie par les acheteurs, la meilleure méthode consistera à recueillir des informations sur les prix courants des transactions et à les utiliser pour construire un *indice des prix de vente des logements*. En revanche, pour estimer le stock (réel) de richesse d'une économie, les informations relatives à l'échantillon de logements vendus devront dans l'idéal être complétées par des renseignements sur le stock de logements non vendus pour construire un *indice des prix du stock de logements*. Pour ce faire, on pourra procéder à une repondération de manière à tenir compte de la différence entre la composition du stock de logements et celle des logements vendus, mais, pour que cette méthode soit valable, il faut que les logements ayant effectivement fait l'objet d'une transaction puissent servir d'indicateur pour les autres. Si le prix des logements qui n'ont pas changé de propriétaire n'est pas disponible, et si les informations concernant leur nombre et leurs caractéristiques sont limitées, voire inexistantes, l'utilisateur doit vérifier que la structure des transactions est représentative de l'ensemble du stock de logements. Dans la pratique, cette dernière condition risque de ne pas être pleinement satisfaite dans la mesure où différents secteurs du marché immobilier peuvent être influencés par des facteurs distincts, et où un nombre limité de transactions risque de livrer des données peu fiables, voire aucune, sur les prix dans certains de ces secteurs.

2.5 Les caractéristiques (qui déterminent le prix) des logements individuels évoluent souvent avec le temps. Ces changements tiennent notamment aux améliorations apportées au logement sous forme de rénovation des cuisines et des salles de bains, de remplacement des fenêtres par un double vitrage ou de l'installation de systèmes de chauffage ou de climatisation économes en énergie, ainsi qu'à l'agrandissement des bâtiments découlant de l'évolution vers des logements de plus grande superficie récemment observée dans de nombreux pays. Les améliorations et agrandissements seront en partie neutralisés par la dépréciation des bâtiments. Indépendamment de sa finalité, un IPIR idéal doit être corrigé de manière à tenir compte de toutes ces modifications. Autrement dit, il doit représenter l'évolution des prix de biens immobiliers de qualité comparable dans le temps.

2.6 Les ajustements de la qualité ne se limitent cependant pas à la prise en compte des aménagements et de la dépréciation. La gamme de logements vendus au cours d'une période différera vraisemblablement de celle vendue au cours de la période suivante, par exemple dans le cas où l'échantillon des ventes est constitué de logements plus grands. Les variations de cette nature revêtent parfois un caractère cyclique, car généralement les ventes de logements de grande superficie diminuent quand une économie entre en récession. Les variations de composition de l'échantillon dans le temps, tout comme celles de la qualité des logements, ne doivent pas être interprétées de la même manière que les variations de prix — il faut faire appel à des formules de calcul pour corriger l'évolution des prix en fonction de celle

1 Voir Duffy (2009).
2 L'investissement dans la construction de logements compte pour 5 % du PIB environ dans la zone euro.
3 Voir Case et Wachter (2005).
4 Voir Finocchiaro et von Heideken (2007).
5 Voir Case *et al.* (2001), Phang (2004) et Belsky et Prakken (2004).
6 Voir Fenwick (2006) et le chapitre 9.

des paramètres de qualité. Les différentes méthodes permettant de résoudre le problème des variations (des paramètres) de la qualité sont brièvement présentées au chapitre 3, et analysées en détail aux chapitres 4 à 7.

Tour d'horizon des différentes utilisations des indices des prix de l'immobilier résidentiel

2.7 Les indices des prix de l'immobilier résidentiel ont plusieurs applications importantes :

- comme indicateur macroéconomique de la croissance économique;
- aux fins de politique monétaire et de ciblage de l'inflation;
- comme instrument d'évaluation de la valeur du logement en tant que composante du patrimoine;
- comme indicateur de la stabilité ou de la solidité financière pour mesurer l'exposition au risque;
- comme déflateur dans la comptabilité nationale;
- comme élément d'information permettant aux citoyens de décider d'acheter (ou de vendre) un bien immobilier résidentiel;
- à titre d'intrant pour l'établissement de l'indice des prix à la consommation, lequel est à son tour utilisé à des fins de négociations salariales et d'indexation des salaires[7];
- aux fins d'utilisation dans les comparaisons interrégionales et internationales.

Chacune de ces applications est examinée ci-après.

Comme indicateur macroéconomique de la croissance économique

2.8 La hausse des prix immobiliers est souvent associée aux périodes d'expansion économique, tandis que leur baisse correspond souvent à une décélération de l'économie. Goodhart et Hofmann (2006) démontrent pour 16 pays industrialisés une forte corrélation entre les prix de l'immobilier et l'activité économique. En réalité, les six grandes crises bancaires qu'ont connues les pays avancés depuis le milieu des années 70 ont été associées à l'explosion d'une bulle immobilière (Reinhart et Rogoff, 2009)[8]. Généralement, les prix de l'immobilier sont traités comme un indicateur avancé, même si la question de savoir si leur évolution est un indicateur avancé, retardé ou simultané ne fait pas l'unanimité.

2.9 Il apparaît en revanche clairement que, souvent, trois facteurs au moins expliquent le lien entre l'augmentation des prix immobiliers et la croissance économique :

- La hausse (relative) des prix immobiliers a généralement pour effet de *stimuler l'activité du bâtiment*, ce qui se traduit par un accroissement de l'emploi et des revenus pour un large éventail d'intervenants du marché de l'immobilier, comme les agents immobiliers, les ouvriers du bâtiment et les professionnels de la sphère financière et juridique. L'augmentation escomptée des rendements des placements immobiliers amène les constructeurs à lancer de nouveaux projets, ce qui s'accompagne d'une poussée de la demande des propriétaires occupants et des investisseurs dans les secteurs connexes[9]. Par ailleurs, l'activité de construction se développe sous l'effet du nombre croissant de travaux de rénovation.
- Le renchérissement des prix immobiliers entraîne généralement une *expansion des ventes de logements existants*, laquelle se traduit par une hausse des recettes fiscales sous forme des droits de mutation dérivant de l'augmentation du volume et du montant des ventes de biens immobiliers. Cette *croissance des recettes fiscales* peut donner lieu à une poussée des dépenses publiques, qui a elle-même pour effet de donner une nouvelle impulsion à l'économie.
- Le relèvement des prix immobiliers amène une amélioration de la situation patrimoniale des ménages (*l'effet de richesse*), laquelle se traduit généralement par une hausse de leurs dépenses de consommation et d'investissement[10]. Selon un rapport de l'Office du budget du congrès américain (2007), durant la flambée des prix immobiliers qu'ont connue les années 90 et la première décennie du siècle, les dépenses de consommation ont progressé plus vite que les revenus. Cet effet sur le patrimoine des ménages entraîne généralement, outre une hausse des dépenses allouées aux autres biens et services, une augmentation de celles consacrées aux travaux de rénovation et de réparation de leur logement.

2.10 À l'évidence, les effets stimulants de cette hausse s'inversent lorsque les prix des logements (réels) chutent. Il importe donc que les responsables publics et ceux de la politique économique disposent d'informations précises et régulières sur l'évolution des prix immobiliers.

7 L'intégration d'un indice des prix immobiliers dans le calcul de l'IPC est fonction des objectifs de l'IPC et, en particulier, de la méthode retenue — à savoir le concept d'acquisition, de paiement ou du coût d'usage. Ces questions font l'objet d'une analyse approfondie dans le Manuel de l'indice des prix à la consommation (BIT *et al.*, 2004) et le Guide pratique pour l'établissement d'indices des prix à la consommation (Nations Unies, 2009).

8 Claessens, Kose et Terrones (2008; 25) observent que «... les récessions associées à un effondrement des prix immobiliers durent en moyenne un trimestre de plus que les autres. Les baisses de la production (et les pertes cumulées correspondantes) liées aux premières sont en outre généralement beaucoup plus importantes, à savoir 2,2 (3,7) % contre 1,5 (2,3) % pour les secondes. Ces écarts appréciables concernent également d'autres variables macroéconomiques, dont la consommation, l'investissement et le taux de chômage.»

9 Voir Zhu (2005).

10 Voir Campbell et Cocco (2007).

2.11 Les prix des actifs, prix immobiliers compris, sont un indicateur indispensable à une meilleure compréhension de la dynamique économique[11]. Selon Plosser (2007), ils renferment des informations importantes quant à l'état actuel et futur de l'économie, et peuvent jouer un rôle majeur dans les délibérations des banques centrales désireuses d'atteindre leurs objectifs de stabilité des prix et de croissance durable de la production.

Aux fins de politique monétaire et de ciblage de l'inflation

2.12 Outre l'intérêt général que présente le suivi des prix immobiliers, de nombreuses banques centrales se sont fixé des *cibles d'inflation* qui peuvent faire intervenir directement des indices des prix immobiliers. Certaines utilisent ainsi un indice des conditions monétaires (ICM) à titre de cible fonctionnelle courante pour conduire la politique monétaire. Une version élargie de cet indice, comme celle proposée par Jarociński et Smets (2008) et Goodhart et Hofmann (2007), comporterait une mesure des prix immobiliers en raison du rôle capital que cette variable joue dans le processus inflationniste et dans le fonctionnement de l'économie. D'autres banques centrales ont établi des cibles d'inflation fondées sur l'indice des prix à la consommation (IPC) et tiendront indirectement compte des fluctuations des prix immobiliers pour fixer les taux d'intérêt, en partie selon le traitement des logements occupés par leur propriétaire dans l'IPC national. Cette question est analysée en détail au chapitre 3.

2.13 On peut avancer que les prix immobiliers résidentiels vont probablement jouer un rôle grandissant dans la conduite de la politique monétaire. Ces dernières années, un nombre croissant de pays ont recouru à un dispositif de cible d'inflation pour définir et appliquer leur cadre de politique monétaire. Le FMI (2007) établit une liste de 28 pays ayant fixé une cible d'inflation en fonction de leur «régime de change» (sans préciser la cible ou l'indicateur d'inflation). Carare et Stone (2003) ont élargi cette analyse en classant les pays qui utilisent le ciblage de l'inflation aux fins de politique monétaire en trois catégories : ceux qui ont adopté un régime de ciblage strict, ceux qui appliquent des dispositifs «éclectiques» et ceux qui ont opté pour un ciblage souple; ce classement est établi en fonction de la clarté et de la crédibilité[12] de l'engagement de ces pays à tenir leur cible. Les auteurs recensent ensuite 42 banques centrales de pays moyens et grands qui ont mis en place, sous une forme ou une autre, un régime de change flottant (autrement dit, qui n'ont pas adopté de taux de change fixe) et ont fait de leur détermination à respecter leur cible d'inflation l'objectif monétaire déterminant. Ils ont estimé qu'en 2001, 7 pays industriels et 11 pays émergents appliquaient un régime de ciblage strict de l'inflation; autrement dit, «ils affichent un degré de crédibilité moyen à élevé, sont manifestement déterminés à respecter leur cible d'inflation et institutionnalisent cet engagement sous forme d'un cadre monétaire transparent qui favorise la responsabilité de la banque centrale à cet égard». Le nombre de pays appliquant un régime de ciblage strict de l'inflation a progressé au fil des ans.

11 Voir Turvey (1989) et Goodhart (2001).

12 La clarté est évaluée par l'annonce publique d'une cible d'inflation et par la présence de dispositifs institutionnels d'obligation de rendre compte. La crédibilité est mesurée indirectement, en prenant pour indicateurs l'inflation réelle et l'évaluation par le marché de la dette publique à long terme en monnaie locale.

Comme instrument d'évaluation de la valeur du logement en tant que composante du patrimoine

2.14 Les prix immobiliers sont un outil d'évaluation de la richesse globale d'un pays. Les logements existants font partie des comptes de patrimoine du Système de comptabilité nationale (SCN). Il est donc nécessaire de disposer d'un indice des prix de cette catégorie d'actifs pour établir des estimations du patrimoine réel des ménages. Comme indiqué dans l'introduction à ce chapitre, la hausse des prix immobiliers produit un effet de richesse qui peut entraîner une augmentation de la consommation et des emprunts des ménages.

2.15 Plus généralement, les particuliers ont un intérêt indirect dans les prix des actifs immobiliers, biens résidentiels compris, au travers des fonds de pension et d'autres investissements directs dans l'immobilier.

Comme indicateur de la stabilité ou de la solidité financière pour mesurer l'exposition au risque

2.16 Les indicateurs de solidité financière (ISF) sont des indicateurs de l'état de santé et de solidité du système et des établissements financiers d'un pays et des entreprises et ménages qui les composent. Ils sont constitués de données agrégées concernant les différents établissements financiers et d'indicateurs représentatifs des marchés sur lesquels ceux-ci opèrent, y compris des statistiques des prix immobiliers. L'établissement et la diffusion des ISF visent à soutenir la surveillance nationale et internationale des systèmes financiers. Le FMI les a mis au point dans l'objectif de surveiller et de consolider le système financier mondial et de renforcer la stabilité après les crises qui ont secoué les marchés financiers à la fin des années 90, et pour combattre les crises bancaires qui se sont ensuite multipliées dans le monde. Le *Guide d'établissement des indicateurs de solidité financière* fournit des conseils pour l'élaboration d'un indice des prix immobiliers, tout en reconnaissant le manque relatif de directives et d'antécédents internationaux et l'absence d'un cadre exhaustif pour construire ce type d'indice. Plus récemment, le rapport sur la crise financière et les lacunes en matière d'information établi en octobre 2009 à l'intention des ministres des finances et des gouverneurs des banques

centrales des pays membres du G-20[13] indique que les informations relatives aux logements et à l'évolution de leur prix sont des éléments indispensables à une analyse des politiques de stabilité financière.

2.17 La chute brutale des prix immobiliers a un effet préjudiciable sur la santé et la solidité du secteur financier ainsi que sur la situation financière des particuliers et des ménages, dans la mesure où elle altère les cotes de solvabilité, la valeur des garanties et les ratios dette/fonds propres.

2.18 Il n'y a rien d'étonnant à ce que la relation entre les cycles immobiliers et les cycles économiques soit bien établie et à ce que le rôle des prix immobiliers dans les crises de financement de la dette et les crises financières ait été reconnu de longue date. Cela a conduit à utiliser les indices des prix de l'immobilier résidentiel comme indicateurs de la stabilité financière, surtout dans les pays où l'immobilier représente une part substantielle du patrimoine national et de celui des ménages, et où la propension à la propriété immobilière est relativement élevée.

2.19 Des études analytiques approfondies ont préconisé l'utilisation de l'évolution des prix de l'immobilier résidentiel, et des prix de l'immobilier en général, comme indicateur de la solidité financière. On citera parmi les nombreuses publications portant sur ce sujet une étude de Nabarro et Key (2003), qui présentent un modèle applicable aux cycles de l'immobilier et du crédit, étayé par des études de cas. Leur étude suit l'évolution du cycle, depuis les premiers indicateurs fournis par le marché locatif jusqu'aux prix immobiliers, en passant par les bilans des emprunteurs et des prêteurs, et attire l'attention sur plusieurs indicateurs pertinents du marché immobilier. Elle décrit ce que les auteurs appellent «la dangereuse interdépendance entre les cycles immobiliers et les systèmes financiers». Sans contester le caractère fortement imprévisible des cycles immobiliers et leurs caractéristiques et particularités distinctes d'un cycle à l'autre, ils analysent les synergies entre ces cycles et le financement de la dette pour définir dans quels domaines l'amélioration des informations favoriserait l'élaboration de stratégies et politiques correctives efficaces. Ils expliquent comment un système fiable et rentable d'évaluation et de suivi des résultats peut être établi et mis en œuvre, et indiquent en quoi il peut constituer un mécanisme de prise de décision analytique, conçu pour influer sur le comportement du secteur immobilier.

2.20 Les informations relatives aux prix de l'immobilier résidentiel et à d'autres catégories de biens immobiliers doivent être complétées par des analyses détaillées sérieuses et actualisées et par d'autres renseignements, comme le pourcentage de logements achetés au comptant, et non financés par un crédit. Le rapport moyen entre le montant du crédit et le prix du bien immobilier — et sa distribution — donne une idée de l'exposition de l'emprunteur et du prêteur; il en va de même du rapport prix/bénéfices et, dans une certaine mesure, du volume de transactions[14]. Pareillement, une analyse plus détaillée des catégories de logements vendues par région indiquera si l'activité du marché immobilier est concentrée sur des segments particuliers, comme les biens immobiliers haut de gamme, ou dans certaines zones, comme la capitale ou les grandes agglomérations urbaines.

13 Disponible à l'adresse suivante : http://www.imf.org/external/np/seminars/eng/2010/infogaps/index.htm.

Comme déflateur dans la comptabilité nationale

2.21 Les organismes statistiques nationaux utilisent de deux façons au moins les indices des prix immobiliers. En premier lieu, ils utilisent souvent la composante «bâtiments» d'un indice des prix des logements neufs pour déflater les prix courants des constructions résidentielles dans les comptes nationaux; voir Bover et Izquierdo (2003). Ensuite, les indices immobiliers peuvent être intégrés dans la construction d'un IPC, selon la base conceptuelle retenue. Cette seconde utilisation est examinée ci-après, et de manière plus approfondie au chapitre 3.

2.22 Les indices des prix et les déflateurs sont apparemment des entités distinctes au sein d'un groupe plus large de statistiques associées aux prix[15]. Dans ce contexte, il convient de noter que deux des ouvrages de référence les plus récents et largement disponibles concernant l'établissement et l'utilisation des déflateurs des comptes nationaux, le SCN (1993) et le *Manuel de la mesure des prix et des volumes dans les comptes nationaux* d'Eurostat (2001), sont antérieurs au *Manuel de l'indice des prix à la consommation* (Manuel de l'IPC, 2004) et au *Manuel de l'indice des prix à la production* (Manuel de l'IPP, 2004).

2.23 Ces deux derniers ouvrages ont été élaborés en parallèle et ont bénéficié des travaux de recherche les plus récents sur la théorie et la pratique des indices, qui ne sont pas pleinement pris en compte dans les ouvrages officiels sur les comptes nationaux[16]; ils s'appuient en substance sur les mêmes principes économiques et la même théorie statistique. Ils donnent une présentation complète et cohérente des questions conceptuelles et théoriques associées aux indices des prix à la consommation et à la production, et les traduisent en formules valides aux fins de mesures pratiques. Le Manuel de l'IPC a également servi de catalyseur à la nouvelle résolution de l'OIT concernant les indices des prix à la consommation, qui a été adoptée en 2003.

14 Des observations antérieures indiquent que, lorsque les ratios prix/bénéfices atteignent un niveau intenable, l'ajustement intervient dans un premier temps sous forme d'une réduction du volume des ventes, et non d'une baisse des prix de transaction.

15 La théorie qui sous-tend les déflateurs et les indices des prix (directs) est néanmoins identique; voir le chapitre 16 du SCN (1993). Samuelson et Swamy (1974) notent ce qui suit : «Bien que les ouvrages spécialisés s'intéressent surtout aux indices des prix… Une fois estimés, d'une manière ou d'une autre, les indices des prix sont en fait utilisés, s'ils le sont, essentiellement pour déflater les totaux nominaux ou monétaires de manière à obtenir des estimations des grandeurs «réelles» sous-jacentes.»

16 Le Manuel de l'IPC et le Manuel de l'IPP cadrent avec le texte du chapitre 16 du SCN (1993) et avec le *Système de comptabilité nationale* de 2008, mais analysent plus en profondeur les problèmes associés à la construction des indices des prix, en particulier aux niveaux inférieurs d'agrégation.

Comme élément d'information permettant aux citoyens de décider d'acheter (ou de vendre) un bien immobilier résidentiel

2.24 L'achat ou la vente d'un logement est généralement l'opération financière la plus importante à laquelle un ménage procédera au cours de sa vie. Les variations des prix immobiliers devraient donc sensiblement influencer sa décision d'acquérir un logement, de même que les prévisions budgétaires et les décisions d'épargne des acheteurs et vendeurs éventuels. De nombreux propriétaires occupants considèrent cet achat à la fois comme un moyen d'obtenir des services de logement et comme un investissement en capital, ce dernier leur offrant la possibilité de réaliser des plus-values substantielles à plus long terme. Le niveau et l'évolution courants des prix, conjugués aux mouvements ultérieurs prévus des prix immobiliers et des taux d'intérêt hypothécaires[17], influeront sur la décision d'un particulier de procéder immédiatement à l'achat ou de le reporter. Le coût d'opportunité associé aux sommes concernées jouera également dans la mesure où les acheteurs potentiels évaluent les choix qui s'offrent à eux. Ils prendront ainsi souvent en considération les effets des variations des prix immobiliers sur la valeur locative de marché.

2.25 De manière plus générale, les particuliers ont également un intérêt indirect dans le prix des actifs immobiliers au travers des fonds de pension et d'autres investissements sur lesquels les prix immobiliers sont susceptibles d'avoir des retombées. À titre d'exemple, les portefeuilles de certains fonds de pension comportent des immeubles d'habitation dont les loyers assurent des revenus et dont l'augmentation de la valeur immobilière est censée produire une plus-value.

À titre d'intrant pour l'établissement de l'indice des prix à la consommation (IPC)

2.26 Les prix immobiliers influent directement sur la mesure de l'inflation dans les cas où l'IPC comprend le coût des logements occupés par leur propriétaire et où la méthode de calcul utilise ces prix comme intrant. Ils l'influencent indirectement s'ils agissent sur la valeur locative de marché, qui constitue un autre élément de l'IPC, et si les loyers fictifs servent en outre d'indicateur du coût des logements occupés par leur propriétaire. La location et l'achat peuvent se substituer mutuellement, et les prix immobiliers influent sur le taux de rendement obtenu par un propriétaire sur son investissement, de même que sur le loyer appliqué.

2.27 Le traitement des logements occupés par leur propriétaire dans les indices des prix à la consommation (IPC) est l'une des questions les plus complexes pour le statisticien. Il existe plusieurs façons théoriques de les traiter, et la méthode retenue peut avoir des retombées appréciables sur l'indice global, en modifiant à la fois la pondération attribuée à ces logements (et, par voie de conséquence, à un IPIR) et la mesure du taux d'inflation. Il existe en substance quatre grandes méthodes pour intégrer les logements occupés par leur propriétaire dans un IPC : le concept d'acquisition, le concept de paiement ou de déboursement, le concept de coût d'usage et le concept d'équivalent-loyer. Les trois premières exigent la construction d'un indice des prix immobiliers. Ces différentes méthodes sont examinées plus en détail au chapitre 3.

Aux fins d'utilisation dans les comparaisons interrégionales et internationales

2.28 Les indices des prix immobiliers sont également utilisés en association avec des données de référence (comparables) relatives aux prix immobiliers dans les différents pays ou régions afin de produire des comparaisons interrégionales ou internationales des écarts de coût de la vie. Les problèmes liés à l'évaluation du prix des services fournis par les logements occupés par leur propriétaire dans un contexte national interviennent également dans celui des comparaisons interrégionales et internationales. Dans ce dernier cas, néanmoins, les problèmes sont un peu plus complexes que, par exemple, dans celui de comparaisons nationales dans le temps, car les comparaisons interrégionales et internationales doivent confronter des catégories comparables de logements dans les différents pays ou régions (ou des informations comparables sur les caractéristiques des logements dans les différentes régions si une technique de régression hédonique est utilisée) pour construire un indice des prix à qualité constante.

2.29 La Banque centrale européenne (BCE), en coopération avec les banques centrales des différents pays de la zone euro et de l'Union européenne, s'intéresse aux mesures comparatives des variations des prix de l'immobilier résidentiel dans les différents pays de la zone euro et dans l'ensemble de la zone. Les données brutes utilisées ici proviennent de différentes sources nationales et ont essentiellement été recueillies et établies par la Banque des règlements internationaux (BRI)[18]. Depuis 2001, la BCE a établi pour la zone euro un indice d'agrégat en pondérant les variations des prix des logements dans les différents pays de la zone euro[19]. Les méthodes utilisées pour obtenir les chiffres dans chaque pays et l'agrégat de la zone euro ont été perfectionnées ces dernières années, mais ne sont peut-être pas à la hauteur des normes appliquées à d'autres statistiques économiques ou indicateurs de prix pour la zone euro[20]. La BRI a également réuni des statistiques des prix de l'immobilier

17 La politique de taux d'intérêt jouera aussi bien sur l'inflation que sur le revenu disponible net après paiement des intérêts.

18 Le jeu de données utilisé par la BRI pour établir les statistiques des prix de l'immobilier résidentiel est disponible à l'adresse suivante : http://www.bis.org/statistics/pp.htm.

19 Voir l'encadré intitulé «Indications provisoires concernant les évolutions des prix de l'immobilier résidentiel dans la zone euro» dans le Bulletin mensuel d'octobre 2001 de la BCE.

20 Voir Eiglsperger (2010), page 233.

résidentiel pour les pays de l'Union européenne non membres de la zone euro et s'est heurtée dans de nombreux cas à des problèmes encore plus aigus en termes de comparabilité et de qualité des données.

2.30 Ces comparaisons peuvent être faussées par des différences de méthodologie et de champ, et par des disparités en termes de périodicité et d'actualisation des données. Cette hétérogénéité tient en partie aux différentes sources de données utilisées pour établir les indices nationaux. Le chapitre 9 examine ces sources plus en détail et le chapitre 10 dresse un inventaire des différentes méthodes utilisées par les pays pour établir leurs indices des prix de l'immobilier résidentiel. On constatera qu'un pourcentage notable de pays, y compris certains pays développés, ne disposent pas d'indices fiables des prix de l'immobilier résidentiel.

Éléments d'un cadre conceptuel

3

Introduction

3.1 Pourquoi est-il si difficile de construire un indice des prix de l'immobilier résidentiel (IPIR)? Cette question a été traitée au chapitre 1 du présent manuel, mais il est utile de rappeler au lecteur les principaux problèmes :

- L'établissement d'un indice des prix se fonde généralement sur l'*appariement* des prix de produits identiques dans le temps. Or, dans le contexte du logement, chaque bien immobilier occupe un emplacement spécifique et présente d'ordinaire un ensemble particulier de caractéristiques structurelles. Il est donc difficile, voire impossible, d'appliquer la méthode de l'appariement de modèles.
- Les transactions sont sporadiques.
- Le concept d'indice souhaité n'est pas toujours clair; autrement dit, *divers objectifs* président à l'établissement d'un IPIR et, de manière générale, des objectifs différents requièrent des indices distincts.
- Certains objectifs, notamment la construction de comptes de patrimoine nationaux et l'estimation des coûts d'usage des logements occupés par leur propriétaire, appellent une décomposition du prix du bien immobilier en deux éléments, «terrain» et «bâtiments», mais on ignore comment procéder au mieux à cette décomposition. Ce problème sera analysé plus précisément au chapitre 8.

3.2 Comme l'indiquent les citations suivantes, les ouvrages consacrés à l'évaluation des prix immobiliers sont bien conscients des deux premières difficultés :

«Les prix de l'immobilier sont plus difficiles à évaluer que ceux des autres biens et actifs, car ils présentent trois caractéristiques essentielles. D'abord, et surtout, les logements sont hétérogènes. Il n'y a pas deux logements identiques, ne serait-ce que parce qu'ils ne peuvent se situer exactement au même endroit. Autrement dit, les prix relevés risquent de constituer un mauvais indicateur de l'ensemble des prix immobiliers parce qu'on ne peut pas toujours prédire de manière fiable le prix de vente d'un logement donné à partir du prix d'un autre logement.» Robert Wood (2005; 213).

«Le *problème fondamental* auquel les statisticiens sont confrontés lorsqu'ils s'efforcent de construire un indice des prix immobiliers est qu'il est *impossible d'apparier exactement les biens immobiliers dans le temps*, et ce pour deux raisons : i) le bien immobilier se déprécie au fil du temps (*le problème de la dépréciation*) et ii) il peut faire l'objet de réfections, d'agrandissements ou d'aménagements d'importance entre les deux périodes examinées (*le problème des rénovations*). Il est en conséquence impossible de construire directement des indices des prix de l'immobilier à qualité constante; des imputations ou estimations indirectes s'imposeront sous une forme ou une autre.» Erwin Diewert (2009b; 92).

Il ressort de ces constatations que la construction d'un IPIR sera nettement plus complexe que celle d'un indice des prix «normal», fondé sur un échantillon constant. Il convient d'admettre dès le départ que, compte tenu des difficultés dérivant de la spécificité de chaque logement, il sera impossible de construire un IPIR «parfait»; il faudra se contenter d'une approximation de l'indice théoriquement idéal pour chaque objectif.

3.3 La question de la finalité d'un IPIR a été abordée au chapitre 2, qui s'est penché sur les nombreux usages d'un tel indice. Le présent chapitre s'intéresse aux utilisations des IPIR pour combler les lacunes du Système de comptabilité nationale et de la construction d'un IPC. Si des IPIR appropriés peuvent être construits pour remédier à ces insuffisances, la famille d'IPIR en résultant satisfera vraisemblablement aux besoins de la plupart des utilisateurs.

3.4 De manière générale, on peut distinguer deux types d'IPIR : 1) un indice des prix à qualité constante pour le *stock* de logements résidentiels à un moment donné; et 2) un indice des prix à qualité constante pour les *ventes* de biens immobiliers résidentiels intervenues au cours d'une période donnée. La méthode de construction de ces deux catégories d'indice diffère, par exemple en ce qui concerne la pondération associée à chacune d'elle. Ce chapitre examine brièvement les principales méthodes de construction d'un IPIR, qui seront analysées en détail aux chapitres 4 à 7.

3.5 Les quatre dernières parties du chapitre traitent de sujets divers : la périodicité de l'IPIR et des besoins des usagers; la cohérence des estimations mensuelles et trimestrielles et des estimations trimestrielles et annuelles; les politiques de révision; et l'ajustement saisonnier.

Les indices des prix de l'immobilier résidentiel et le Système de comptabilité nationale

3.6 Le *Système de comptabilité nationale (SCN) 1993* et sa récente mise à jour, le *Système de comptabilité nationale 2008*[1], constituent pour un pays un cadre comptable complet. Le SCN répartit les flux de valeur de l'économie en plusieurs catégories pertinentes et procède au rapprochement des comptes de flux et des comptes de stocks correspondants. Il est en outre recommandé de décomposer les valeurs figurant dans les cellules des comptes en prix et en volumes (ou quantités).

3.7 Le SCN comporte trois rubriques dans lesquelles les indices des prix de l'immobilier résidentiel doivent convertir des valeurs nominales en volumes ou en valeurs réelles :

- le *stock* de biens immobiliers résidentiels existant en un lieu précis et à un moment donné dans le pays;

1 Voir Banque mondiale, Eurostat, FMI, Nations Unies et OCDE, (1993) et (2009).

- les *ventes* de biens immobiliers résidentiels intervenues en un endroit précis et pendant une période donnée dans le pays; et
- la composante «bâtiments» des *ventes* de biens immobiliers résidentiels *neufs* intervenues en un endroit précis et pendant une période donnée dans le pays.

3.8 Le stock immobilier résidentiel d'un pays est une composante de son patrimoine national. Un indice des prix de l'immobilier résidentiel est donc nécessaire pour établir des estimations du *patrimoine réel par composante* du compte de patrimoine[2]. Les estimations de la richesse nationale réelle établissent généralement une distinction entre la composante «bâtiments» des biens immobiliers résidentiels et la composante «terrains». Si des estimations du stock réel de bâtiments résidentiels et du stock réel de terrains résidentiels d'un pays doivent être produites, il conviendra de décomposer la valeur des biens immobiliers résidentiels en ces deux éléments distincts et de construire des indices des prix pour chacun d'eux.

3.9 Les raisons pour lesquelles un indice des prix de vente des biens immobiliers résidentiels s'impose aux fins de comptabilisation du revenu national n'apparaissent pas toujours d'emblée. Cet indice sert à estimer la *production réelle* du secteur des services immobiliers résidentiels, autrement dit, le secteur qui fournit les services nécessaires aux transactions immobilières. Quelques calculs permettront de comprendre pourquoi un indice des prix de vente des biens immobiliers résidentiels est indispensable à cet égard.

3.10 Supposons que le montant des commissions d'un agent immobilier pour une catégorie donnée de transactions immobilières est de V_C^t au cours de la période t, et supposons que le montant correspondant des ventes pour le même groupe de biens immobiliers (commissions comprises) est V_S^t. Supposons en outre qu'un indice des prix à qualité constante a été construit pour cette catégorie de ventes et que la valeur de cet indice pour la période t est P_S^t[3]. Une estimation du volume des ventes pour cette catégorie d'opérations immobilières au cours de la période t, disons Q_S^t, peut être calculée selon la relation suivante :

$$Q_S^t \equiv V_S^t / P_S^t \tag{3.1}$$

2 Un indice des prix du stock de biens immobiliers résidentiels présente également une certaine utilité pour les banques centrales désireuses de surveiller les prix immobiliers en vue de déceler l'éventuelle formation de bulles dans le pays; voir le chapitre 2.

3 On peut également utiliser un indice des prix de vente plutôt qu'un indice des prix d'achat. Lorsque l'on construit un indice des prix du logement à qualité constante, doit-on procéder à l'ajustement au titre de la qualité au moyen des caractéristiques déterminant le prix du vendeur ou de celles déterminant le prix de l'acheteur? On pourrait préconiser l'utilisation des secondes pour corriger le prix des biens immobiliers résidentiels dans la mesure où, si l'acheteur ne confère pas suffisamment de valeur à un bien immobilier, celui-ci ne sera pas acheté. La construction d'un indice des prix d'achat à qualité constante semblerait ainsi s'imposer, par opposition à un indice des prix de vente à qualité constante. On peut cependant aussi faire valoir que, si le prix de vente d'un bien immobilier (considéré comme une fonction des caractéristiques de ce bien) n'est pas assez élevé, les producteurs de logements neufs n'en construiront pas, et ce sont donc les caractéristiques déterminantes du prix du vendeur qui devraient entrer en ligne de compte, tout au moins dans le contexte de l'évaluation des logements neufs. Rosen (1974) s'est penché sur ces questions. Selon son analyse des facteurs déterminants de la surface hédonique, dans le cas de logements neufs, son analyse du cas n° 1 est vraisemblablement pertinente, la situation des différentes entreprises étant identique sur le plan des coûts, et la surface hédonique étant donc déterminée par l'offre sur le marché; voir Rosen (1974; 50–51).

3.11 Le secteur de l'immobilier peut être traité comme un secteur de détail ou de gros; il s'agit d'un secteur à marge dont on peut considérer qu'il achète un bien à un prix avant commission et le vend à un prix après commission. La valeur du service est égale au revenu de la commission, V_C^t, et la quantité de services est proportionnelle au volume des ventes, Q_S^t. Ainsi, le volume des services immobiliers, Q_C^t, est égal à Q_S^t :

$$Q_C^t \equiv Q_S^t \tag{3.2}$$

3.12 Enfin, l'indice des prix de la période t pour le sous-secteur de l'immobilier associé aux ventes, de valeur V_S^t au cours de la période t, est fixé à un niveau égal au montant des commissions correspondantes, V_C^t, divisé par le volume correspondant, Q_C^t :

$$P_C^t \equiv V_C^t / Q_C^t \tag{3.3}$$

$$= V_C^t / [V_S^t / P_S^t] \qquad \text{en utilisant (3.1) et (3.2)}$$

$$= [V_C^t / V_S^t] P_S^t$$

$$= m_C^t P_S^t$$

où $m_C^t = V_C^t / V_S^t$ est le *taux de marge de la période t* pour cette catégorie d'opérations immobilières; autrement dit, m_C^t est le rapport des commissions de la période t à la valeur d'achat totale correspondante des opérations immobilières. Ainsi l'indice des prix de la production de ce segment du secteur immobilier pour la période t est le produit du taux de marge m_C^t et de l'indice des prix à qualité constante des biens immobiliers vendus au cours de la période t, P_S^t. Cette démonstration explique pourquoi les indices des prix à qualité constante des ventes immobilières résidentielles sont utiles pour établir les comptes du revenu national.

3.13 Le troisième élément des comptes nationaux pour lequel un déflateur des prix du logement est nécessaire est la *valeur des logements neufs construits* en différents endroits du pays au cours d'une période de référence. Cette valeur fait partie de la *formation brute de capital* du pays. Quand un bien immobilier neuf est produit durant la période de référence, et qu'aucune amélioration n'a été apportée au terrain sur lequel il est construit, le pourcentage du prix de vente applicable au terrain doit être déduit du prix de vente; le reliquat fait alors partie de la formation brute de capital et de la production du secteur de la construction. Ainsi, un IPIR de la *composante «bâtiments» des ventes de biens immobiliers résidentiels neufs* est nécessaire dans les comptes nationaux. Il faudra décomposer les ventes de ces biens neufs en deux éléments — *terrains et bâtiments* — et construire un indice des prix à qualité constante de la composante «bâtiments» pour répondre aux besoins des comptes nationaux.

3.14 Reprenons l'analyse de la modélisation de la production du secteur immobilier. Les divers coûts de transaction associés à la vente de biens immobiliers neufs (les commissions des agents immobiliers, par exemple) créent

dans le système de comptabilité nationale des difficultés qui n'ont pas encore été définitivement résolues. Du point de vue du secteur de la construction, ces coûts ne font pas partie des recettes qui lui reviennent; ils ne doivent donc pas être intégrés dans la valeur de sa production. Pour le secteur qui achète les logements neufs, en revanche, ils constituent un coût réel et doivent être comptabilisés. Il existe plusieurs moyens de traiter les *coûts de transaction* associés à l'achat d'un logement neuf (du point de vue de l'acquéreur) :

- simplement imputer la totalité des coûts à la période d'achat et traiter les coûts de transaction sous forme de dépenses de l'acheteur[4] (application du concept d'acquisition);
- intégrer les coûts de transaction à la composante «bâtiments» de la valeur de l'achat de manière à ce qu'ils soient amortis au même taux que celui appliqué au bâtiment;
- amortir séparément les coûts de transaction en fonction de la durée moyenne de détention d'un bien immobilier de la catégorie considérée avant sa revente.

Sur un plan conceptuel, la dernière méthode paraît préférable[5], mais la première et la deuxième aboutiront à un jeu de comptes plus simple. Il faut que les comptables nationaux approfondissent l'étude de ces questions au moyen d'informations communiquées par l'ensemble de la communauté économique.

Les indices des prix de l'immobilier résidentiel et l'indice des prix à la consommation

3.15 Le suivi des prix des services fournis par les logements occupés par leur propriétaire dans un indice des prix à la consommation (IPC) est traité en profondeur dans le *Manuel de l'indice des prix à la consommation*[6]. Il n'existe pas de consensus universel à ce sujet, mais le manuel propose quatre méthodes[7]. Celles-ci traitent différemment la spécificité des logements occupés par leur propriétaire, qui suppose à la fois l'achat d'un logement et la consommation au fil du temps du flux de services associés à ce dernier.

- Le *concept de paiement ou de déboursement*. Cette approche consiste simplement à totaliser les dépenses directes liées à la propriété d'un logement. Il s'agit des dépenses consacrées à l'entretien et aux réparations, des intérêts des prêts immobiliers, des primes d'assurance, des taxes foncières et des charges de copropriété (s'il s'agit d'une copropriété). Deux catégories importantes de coûts implicites et un avantage implicite majeur ne sont pas pris en compte. Il s'agit pour les premiers du coût d'amortissement et du coût d'opportunité lié au capital financier immobilisé par l'achat du logement, et pour le second des éventuelles plus-values (nettes) dégagées par le propriétaire au cours de la période considérée[8]. Cette méthode est utile si l'analyste s'intéresse essentiellement au revenu disponible des ménages, mais pas particulièrement en tant que mesure des services de consommation des ménages (en raison de l'omission des coûts et de l'avantage mentionnés plus haut).
- Le *concept des acquisitions (nettes)*[9]. Dans le cadre de cette approche, l'IPC ne tient pas compte des services fournis par les logements occupés par leur propriétaire, si ce n'est lorsqu'un logement neuf est mis sur le marché. Le prix d'achat de ce dernier est imputé à la période d'acquisition de sorte que l'achat d'un logement neuf est traité de la même manière que celui d'un bien ou d'un service non durable — autrement dit, comme celui d'autres biens durables. Une variante de cette méthode consiste à décomposer le prix de vente du bien immobilier résidentiel neuf en terrains et en bâtiments, et à intégrer uniquement le prix des bâtiments dans l'IPC.
- Le *concept d'équivalent-loyer*. Ici, un prix est imputé aux services fournis par les logements (neufs et anciens) occupés par leur propriétaire, qui est égal au prix auquel le logement pourrait être loué[10].
- Le *concept du coût d'usage*. Cette méthode calcule le coût d'opportunité financier associé à la propriété d'un logement et à l'utilisation de ses services pendant la période de référence.

Depuis la rédaction du manuel de l'IPC (2004), un cinquième concept a été proposé pour assurer le suivi des prix des services fournis par les logements occupés par leur propriétaire[11] :

- Le *concept du coût d'opportunité*. Dans le cadre de cette approche, le prix des services de logements occupés par leur propriétaire est fixé à un niveau égal *aux prix maximaux* établis par les méthodes de l'équivalent-loyer et des coûts d'usage.

4 L'indice des prix susceptible d'être utilisé pour convertir la valeur nominale des frais de transaction en un montant (ou un volume) réel est un indice composite des prix d'achat des biens immobiliers concernés couvrant les deux composantes : «terrain» et «bâtiments».

5 C'est celle qu'utilise le Bureau des statistiques australien. Un problème non résolu est celui du choix du déflateur des prix pour établir les charges réelles d'amortissement. Autrement dit, faut-il utiliser un indice des prix des bâtiments ou un indice composite des bâtiments et des terrains? Dans le cas de l'immobilier, les commissions sont généralement proportionnelles au prix total du bien (la somme des bâtiments et des terrains); il serait donc approprié d'utiliser un indice des prix composite pour déflater cette composante des coûts de transaction. Comme les terrains et les bâtiments peuvent être assujettis à des taxes de mutation ou des droits de timbre de taux différents, l'établissement d'un prix réel approprié pour cette composante des coûts de transaction peut s'avérer assez complexe. Là encore, pour éviter tous ces problèmes, on pourra simplement utiliser un indice composite des prix d'achat pour procéder à la déflation.

6 Voir OIT, FMI, OCDE, Eurostat, Nations Unies et Banque mondiale (2004), chapitre 23.

7 Diewert (2002) (2009a) (2009b) présente une analyse détaillée d'autres méthodes.

8 Le concept de décaissement est expliqué en détail dans Baldwin, Nakamura et Prud'homme (2010).

9 Pour un traitement pratique complet du concept des acquisitions nettes, voir le *Technical Manual on Owner Occupied Housing* d'Eurostat (2012).

10 Cette méthode cadre avec le traitement des logements occupés par leur propriétaire dans les comptes nationaux. Dans le SCN, ces logements sont considérés comme des actifs immobilisés, contrairement à d'autres biens durables (lave-linge, meubles, automobiles, etc.). L'achat d'un logement est considéré comme un investissement; il est intégré à la formation brute du capital fixe et donc exclu des dépenses de consommation finale des ménages; il en va de même de la réfection et de l'agrandissement du logement. La propriété d'un logement offre néanmoins un service qui est consommé par le propriétaire au fil du temps; la valeur de ce service est intégrée aux dépenses de consommation finale des ménages.

11 Voir Diewert (2009b), Diewert et Nakamura (2009) et Diewert, Nakamura et Nakamura (2009).

3.16 Il convient de souligner les différences conceptuelles entre ces méthodes. La *méthode de l'équivalent-loyer* et la *méthode du coût d'usage* évaluent les prix des services d'un logement occupé par son propriétaire. Le *concept de paiement* mesure les dépenses à la charge du propriétaire du logement. La méthode fondée sur le *concept d'acquisitions nettes* adopte un point de vue complètement différent et impute implicitement l'ensemble des services fournis par le logement neuf à la période d'acquisition.

3.17 Ces méthodes, excepté celles reposant sur le *concept de paiement* et le *concept de l'équivalent-loyer*, font appel à des indices des prix à qualité constante pour les logements neufs ou pour le stock des logements existants. Le recours aux méthodes du *coût d'usage* et du *coût d'opportunité* pour évaluer le prix des services fournis par un logement ne va pas sans poser de problème. L'appendice au présent chapitre décrit leur fonctionnement.

3.18 En résumé, on a besoin d'IPIR pour construire un IPC et pour déflater plusieurs flux de valeur et stocks dans les comptes nationaux. Il sera utile ou nécessaire, aux fins d'établissement d'un IPC et des comptes nationaux, de décomposer les indices des prix en deux éléments — les bâtiments et les terrains. Il serait en particulier utile de pouvoir construire le jeu d'IPIR suivant[12] :

- un indice des prix pour la totalité du *stock* immobilier résidentiel à un moment donné, nécessaire pour estimer l'évolution réelle du stock immobilier résidentiel national, une composante de la richesse réelle du pays;
- un indice des prix du *stock de logements occupés par leur propriétaire* (un sous-indice de l'indice précédent), nécessaire pour établir des estimations de la valeur des services fournis par les logements occupés par leur propriétaire en fonction des principes du coût d'usage ou du coût d'opportunité;
- un indice du prix des *ventes* de biens immobiliers résidentiels (logements neufs et existants) intervenues au cours d'une période donnée, nécessaire pour estimer la production réelle du secteur des services d'immobilier résidentiel;
- un indice du prix des *ventes de biens immobiliers résidentiels neufs* au cours d'une période donnée, nécessaire si une méthode d'acquisitions nettes largement définie est utilisée en vertu de laquelle l'achat couvre à la fois les bâtiments et les terrains;
- un indice des prix de la *composante «bâtiments» des biens immobiliers résidentiels neufs* vendus au cours d'une période donnée, nécessaire dans le cadre d'une méthode d'acquisitions nettes étroitement définie en vertu de laquelle l'achat couvre uniquement les bâtiments.

12 Fenwick (2005) (2006) a fait valoir qu'il serait utile d'élaborer un cadre conceptuel cohérent pour une famille d'indices des prix immobiliers. «On constate que les besoins des usagers varient et que, dans certains cas, plus d'une mesure des prix du logement ou de l'inflation des prix immobiliers est nécessaire. On constate également qu'il convient d'assurer la cohérence entre les différentes mesures et d'autres statistiques économiques, ce qui est particulièrement difficile étant donné que les statisticiens ne disposent probablement pas d'un jeu idéal d'indicateurs des prix.» David Fenwick (2006; 8).

Principales méthodes

3.19 Pour mesurer les *variations pures de prix*, les prix immobiliers doivent être corrigés des changements de qualité. En d'autres termes, pour établir un IPIR à *qualité constante*, il faudra neutraliser, d'une manière ou d'une autre, toute variation des caractéristiques déterminantes du prix des biens immobiliers. Les caractéristiques majeures sont les suivantes :

- la *superficie du bâtiment* (en pieds carrés ou en mètres carrés);
- la *superficie du terrain* sur lequel le bâtiment est construit (en pieds carrés ou en mètres carrés);
- l'*emplacement* du bien immobilier;
- l'*ancienneté* du bâtiment;
- le *type de bâtiment*; ce dernier peut occuper la parcelle entière, sans mur commun avec d'autres bâtiments (logements individuels isolés), ou être accolé à un bâtiment voisin (logements individuels jumelés); il peut également s'agir d'un appartement ou logement isolé dans une résidence multifamiliale (immeuble d'habitation ou copropriété);
- les *matériaux utilisés* dans la construction du logement (essentiellement le bois, la brique, le béton, ou les matériaux traditionnels dans le cas de baraques et bidonvilles); et
- les *autres caractéristiques déterminantes du prix* comme le nombre de chambres, le nombre de salles de bains, la présence d'un garage, d'une piscine, de la climatisation, la proximité des services collectifs, etc.

3.20 Les ouvrages spécialisés ont proposé quatre grandes méthodes pour neutraliser les variations du nombre de caractéristiques des biens immobiliers : la stratification (ou poststratification), les méthodes des ventes répétées, les méthodes de régression hédonique et le recours aux estimations des biens immobiliers. On en trouvera une brève description ci-dessous, et une analyse détaillée aux chapitres 4 à 7.

3.21 La *stratification* des transactions en fonction de certaines des caractéristiques déterminantes de leur prix est une méthode de calcul simple qui permet de procéder à des ajustements pour tenir compte des variations de qualité des échantillons à des périodes différentes. En définissant un nombre de strates ou cellules raisonnablement homogènes, on peut utiliser le prix de vente moyen de chaque cellule comme (indicateur d'un) prix à qualité constante de la catégorie de bien immobilier correspondante. On peut ensuite appliquer la théorie courante des indices pour réunir les prix moyens des différentes cellules en un indice global. Ces méthodes de stratification sont également connues sous le nom de *méthodes des cellules homogènes*. Wood (2005) les décrit de la manière suivante :

«Les prix immobiliers relevés sont réunis en groupes ou «cellules» d'observations concernant des logements dont l'emplacement et les attributs physiques sont similaires. […] Les prix moyens de chaque cellule sont pondérés pour obtenir un prix «à structure constante». Un changement dans la composition de l'échantillon modifiera

le nombre d'observations dans chaque cellule. Mais si les cellules sont définies de manière suffisamment précise, de sorte que tous leurs éléments affichent des prix et des évolutions similaires, les variations de la composition de l'échantillon n'influeront pas systématiquement sur le prix des logements à structure constante.» Robert Wood (2005; 214).

3.22 La *méthode des ventes répétées* aborde le problème des paramètres de qualité en comparant les biens immobiliers qui ont été vendus plus d'une fois au cours de la période d'observation. Limiter la comparaison aux logements qui ont été vendus plusieurs fois garantit que les rapports de prix comparent des biens analogues, sous réserve que la qualité du logement n'ait pas changé. La méthode courante fait appel à un modèle de régression qui regroupe les données relatives aux ventes répétées se rapportant à toutes les périodes. Elle peut poser un problème en termes de «révisions» : quand on ajoute de nouvelles périodes et que l'on estime à nouveau le modèle, les indices des prix précédemment estimés changent. Elle a pour avantage que, les biens immobiliers étant appariés au niveau de l'adresse, leur emplacement, un déterminant majeur des prix immobiliers, reste constante.

3.23 Un autre inconvénient potentiel de la méthode des ventes répétées est qu'elle ne tient pas compte de l'évolution de la qualité des logements échantillonnés. Au fil du temps, un logement peut être rénové et subir une dépréciation; la qualité du bien immobilier peut donc varier dans le temps. Les *méthodes de régression hédonique* permettent en principe de procéder à des ajustements visant à prendre en compte, outre le changement des paramètres de qualité des échantillons, ces variations de la qualité. Elles utilisent des informations concernant les caractéristiques pertinentes des biens immobiliers pour estimer les indices des prix corrigés de la qualité au moyen de techniques de régression, mais il est parfois difficile de faire la part de l'emplacement du bien. Il existe plusieurs manières d'estimer les indices de prix hédoniques. Les études sur l'immobilier privilégient la *méthode des variables indicatrices temporelles*. Celle-ci modélise le prix d'un bien immobilier en fonction de ses caractéristiques et d'un ensemble de variables indicatrices temporelles. Comme les données de toutes les périodes sont regroupées, les indices ainsi produits sont susceptibles de révisions, comme dans le cas de la méthode des ventes répétées. Un autre défaut de cette approche est qu'elle impose des restrictions peut-être injustifiées sur les variations des prix des terrains et des bâtiments dans le temps. On peut éviter les difficultés liées à cette variante de la méthode de régression hédonique par l'utilisation d'une autre version connue sous le nom de *méthode à imputation hédonique*.

3.24 De nombreux pays imposent les biens immobiliers et disposent généralement d'un service officiel compétent qui fournit à intervalles réguliers des évaluations de l'ensemble des biens immobiliers imposables. Les *méthodes fondées sur les évaluations* associent les prix de vente et les évaluations pour calculer les rapports de prix (ratio prix de vente/évaluation) et tiennent compte de l'évolution des paramètres de qualité. La méthode du ratio prix de vente/évaluation se fonde sur la méthode du suivi de modèles identiques. Contrairement à la méthode des ventes répétées, elle fait appel à l'ensemble des données sur les ventes (ventes uniques et ventes répétées), et les indices précédemment estimés ne sont pas révisés. Il va de soi que cette méthode n'est applicable que dans les pays où des estimations fiables des biens immobiliers sont disponibles.

3.25 Si la période de référence est une année, toutes les méthodes produiront des estimations similaires de l'évolution tendancielle des prix de l'immobilier résidentiel pour l'ensemble du pays. Néanmoins, comme nous le verrons dans les exemples présentés aux chapitres 4 à 7 et au chapitre 11, des méthodes différentes génèrent des tendances légèrement, mais significativement divergentes; sur des périodes plus courtes, elles peuvent aboutir à des estimations assez dissemblables des variations de prix. Elles pourraient également indiquer des points de retournement distincts.

3.26 Comme les méthodes hédoniques supposent que les informations sur les caractéristiques des biens immobiliers vendus sont connues, les échantillons peuvent être stratifiés et, si l'on dispose d'un nombre suffisant d'observations, on peut estimer un indice pour chaque strate. Autrement dit, les méthodes de régression hédonique peuvent fournir un ensemble d'indices de prix à qualité constante pour différentes catégories de biens immobiliers. Il va sans dire que, si l'on dispose de données portant sur des caractéristiques déterminantes des prix, les méthodes des ventes répétées et les méthodes fondées sur les évaluations peuvent également être combinées à celle de la stratification.

3.27 On peut aussi utiliser la stratification pour estimer un IPIR fondé sur les stocks. Dans ce cas, les pondérations des strates se basent sur les données de recensement concernant la valeur du stock de logements occupés par leur propriétaire. Les indices des prix des différentes strates restent fondés sur les données d'échantillon des biens immobiliers vendus. À l'intérieur de chaque strate, les biens négociés sont maintenant considérés comme un échantillon (aléatoire) du stock. Comme de longs délais séparent habituellement deux recensements, les pondérations de la valeur des stocks ne peuvent être actualisées que très rarement.

3.28 Comme déjà signalé, il convient, pour des raisons diverses, de décomposer le prix global d'un bien immobilier en éléments (additifs) correspondant au prix du bâtiment et au prix du terrain sur lequel il se situe. Le chapitre 8 explique comment les techniques de régression hédonique peuvent être utilisées pour procéder à cette décomposition.

La périodicité de l'IPIR et les besoins des utilisateurs

3.29 Pour le contrôle de l'inflation, la plupart des banques centrales préfèrent disposer d'un IPIR mensuel ou trimestriel. Pour l'établissement des comptes nationaux, des indices

trimestriels suffisent, tandis que celui de l'IPC requiert généralement des indices mensuels. Comme le nombre d'observations nécessaires à un indice des prix mensuel représente à peu près le tiers seulement de celui nécessaire à un indice trimestriel, les organismes statistiques doivent s'efforcer de concilier périodicité de la publication, actualité et exactitude. L'utilisation de données mensuelles peut aboutir à des chiffres assez bruités, quelle que soit la méthode utilisée pour établir l'IPIR. Pour atténuer le bruit, on peut calculer une moyenne mobile, mais, comme expliqué plus loin, cela crée de nouveaux problèmes[13].

3.30 Il est utile de décrire sommairement certains des arbitrages auxquels les organismes statistiques sont confrontés lorsqu'ils s'efforcent de construire des indices des prix immobiliers répondant aux besoins des utilisateurs. Avant de procéder aux arbitrages, ils doivent analyser les besoins des utilisateurs d'une *famille d'indices des prix de l'immobilier résidentiel*. La liste présentée ci-dessous s'inspire de celle établie par Emily Carless (2011) du National Statistician's Office de l'autorité de la statistique du Royaume-Uni. La famille d'IPIR doit[14] :

- être basée sur le prix payé pour les biens immobiliers ayant fait l'objet d'une transaction;
- être stratifiée par région;
- être stratifiée par type de logement (logement individuel, en rangée, tour, type de construction, etc.);
- être établie sur une base mensuelle;
- constituer globalement un indice national cohérent;
- être précise et établie dans les délais prévus avec un minimum de révisions.

Le cinquième critère, à savoir que l'ensemble des sous-indices constitue un indice national cohérent, n'est pas trop difficile à satisfaire. La possibilité de satisfaire le premier (que les indices de prix soient fondés sur les prix des transactions) est fonction de la disponibilité des données. Dans de nombreux pays, les prix de vente réels sont utilisés pour construire les IPIR, mais tous les organismes statistiques n'ont pas accès à ces chiffres. Même s'ils sont disponibles, c'est parfois avec un certain délai (voir le chapitre 9), de sorte que le premier critère pourrait, dans la pratique, être incompatible avec le sixième (la production des indices dans les délais prévus).

3.31 Certains objectifs sont également incompatibles avec d'autres critères : en présence de plusieurs strates, l'établissement d'indices mensuels risque de créer une situation dans laquelle certaines strates ne comporteraient que peu de transactions, ce qui produirait des sous-indices assez instables et imprécis. Le recours aux moyennes mobiles des indices mensuels peut certes atténuer cette volatilité[15], mais cette méthode ne fournira pas d'informations sur les variations de prix en temps voulu. Autrement dit, l'indice moyen en résultant sera centré sur le milieu de la période de la moyenne mobile et ne sera pas disponible avant plusieurs mois[16]. Cette méthode pourrait notamment donner une image trompeuse des phases d'expansion et de ralentissement du marché immobilier. De manière générale, un indice unique ne permettra donc pas de satisfaire l'ensemble des besoins énumérés plus haut. Les organismes statistiques devront donc faire quelques compromis pour répondre aux besoins des différents usagers.

Compatibilité des estimations mensuelles et trimestrielles

3.32 Comment assurer la compatibilité des estimations mensuelles et trimestrielles des variations des prix immobiliers? Il est relativement facile de répondre à cette question si la méthode du prix moyen ou de la valeur unitaire appliquée aux données mensuelles l'est également aux données trimestrielles. Supposons qu'un IPIR mensuel des ventes soit construit au moyen de la méthode de stratification (ou de partition). Comme on le verra plus en détail au chapitre 4, le prix mensuel d'une cellule donnée est le prix de transaction moyen, ou valeur unitaire moyenne, et la quantité correspondante est le nombre total de biens immobiliers négociés. Pour établir l'IPIR trimestriel de cette cellule, on calculera dans un premier temps la valeur unitaire trimestrielle, la quantité correspondante étant le nombre total de transactions dans la strate pour le trimestre. Quelques calculs feront apparaître plus clairement le lien entre les prix et les quantités trimestriels d'une cellule avec les données mensuelles correspondantes[17].

3.33 Supposons que l'on dispose de T trimestres de données mensuelles. Soit V^t la valeur des transactions trimestrielles d'une cellule donnée du modèle de stratification pour $t = 1,...,T$. Pour chaque trimestre t, la valeur des transactions du premier mois est V_1^t, celle des transactions du deuxième mois, V_2^t, et celle des transactions du troisième mois V_3^t. Les *prix unitaires mensuels* du trimestre t sont représentés par P_1^t, P_2^t et P_3^t, et le *nombre mensuel de transactions* correspondant par Q_1^t, Q_2^t et Q_3^t. À noter que

13 Les moyennes mobiles sont cependant utilisées, par exemple en Islande. Dans le contexte d'un IPC mensuel, il est parfois nécessaire d'utiliser des informations qui datent quelque peu; voir Gudnason et Jónsdóttir (2006; 4).

14 Outre les critères énumérés, M^me^ Carless a fait observer que les utilisateurs souhaitent une explication précise des méthodes utilisées pour construire les statistiques et les indicateurs de la qualité des mesures. Certains souhaitent en outre disposer, en sus des séries non ajustées, des séries corrigées des variations saisonnières.

15 On peut également atténuer la volatilité en combinant certaines strates, mais au détriment de certains détails géographiques ou de la couverture des catégories de logements dont les usagers souhaitaient disposer. Par ailleurs, les nouvelles strates ainsi constituées pourraient ne pas être soumises à la même évolution des prix; l'agrégation des strates risque donc de conduire à un biais de valeur unitaire.

16 Ce nombre est égal à la moitié du nombre de périodes de calcul de la moyenne mobile.

17 Le même type d'analyse peut s'appliquer à la relation entre un IPIR des ventes annuelles (calculé par poststratification) et les estimations trimestrielles correspondantes.

V_m^t est égal à $P_m^t Q_m^t$ pour $m = 1,2,3$ et $t = 1,...,T$. La valeur des transactions pour le trimestre t, V^t, est égale à la somme des transactions mensuelles du trimestre :

$$V^t = V_1^t + V_2^t + V_3^t = P_1^t Q_1^t + P_2^t Q_2^t + P_3^t Q_3^t \quad (3.4)$$
$$t = 1,...,T$$

La série trimestrielle de quantités, Q^t, est la somme des transactions mensuelles du trimestre et la série trimestrielle de prix, P^t, est la valeur unitaire trimestrielle pour la cellule considérée; autrement dit :

$$Q^t = Q_1^t + Q_2^t + Q_3^t \quad (3.5)$$
$$t = 1,...,T$$

$$P^t = V^t / Q^t \quad (3.6)$$
$$= [P_1^t Q_1^t + P_2^t Q_2^t + P_3^t Q_3^t]/[Q_1^t + Q_2^t + Q_3^t]$$
$$= s_1^t P_1^t + s_2^t P_2^t + s_3^t P_3^t$$
$$t = 1,...,T$$

où la *part des transactions du mois m dans le trimestre t*, s_m^t, est définie comme suit :

$$s_m^t = Q_m^t / Q^t \quad (3.7)$$
$$m = 1,2,3;\ t = 1,...,T$$

En conséquence, le prix trimestriel pour la cellule considérée, P^t, est égal à la moyenne, pondérée par les parts des transactions, des prix mensuels P_m^t des mois m du trimestre t.

3.34 Pour les méthodes de construction d'IPIR autres que la stratification (régression hédonique, ventes répétées, utilisation des données d'évaluation), la relation entre les estimations trimestrielles de la variation des prix et les estimations mensuelles correspondantes est plus complexe. Néanmoins, ces méthodes produisent au bout du compte un indice des prix, disons P^t pour la période t, qui est associé à un groupe donné de transactions (ou de stocks). En général, la valeur correspondante de la période t associée à ces stocks, disons V^t, est disponible, ce qui permet de définir un volume correspondant pour la période t, $Q^t = V^t/P^t$, et donc de procéder au calcul ci-dessus.

Politique de révision

3.35 Il semblerait possible de construire un IPIR des ventes immobilières sans que des révisions s'imposent; or, il n'est pas toujours facile de recueillir les données correspondantes dans les délais requis. La construction d'un IPIR du stock de logements est tributaire des données de recensement, souvent disponibles tardivement. Par ailleurs, lorsqu'un nouveau recensement est publié, il est généralement préférable d'utiliser les informations qu'il fournit pour corriger rétrospectivement l'IPIR du stock de logements à la date du recensement précédent. De manière générale, il vaut donc mieux permettre la révision de cette catégorie d'IPIR. Cela ne devrait pas poser de problème majeur pour l'établissement des comptes nationaux, ces derniers faisant régulièrement l'objet de révisions.

3.36 Les révisions créent néanmoins des problèmes dans le contexte de statistiques non révisables comme l'IPC. Le traitement des logements occupés par leur propriétaire dans un IPC requiert un IPIR du stock de logements si les méthodes du coût d'usage ou du coût d'opportunité sont employées[18]. Dans ce cas, il faudra éventuellement utiliser des informations préliminaires pour établir l'IPIR. Une fois que de nouvelles données seront disponibles, une version révisée de l'IPC pourra être publiée sous forme de *série analytique* pour permettre aux analystes d'établir des estimations approximatives du biais éventuellement lié à l'utilisation de l'IPC non corrigé, fondé sur une estimation préliminaire de l'IPIR pour les logements occupés par leur propriétaire.

Correction des variations saisonnières

3.37 Bien que l'emplacement puisse quelque peu varier selon les pays, des fluctuations saisonnières sensibles des *quantités* de biens immobiliers négociés au cours d'une année sont généralement observables. Dans le cadre de la construction d'un IPIR, il s'agit de déterminer si cette saisonnalité des quantités se traduit par une saisonnalité des *prix*. Les données empiriques sont assez contrastées. Meese et Wallace (1991) observent dans leur étude économétrique une saisonnalité limitée des prix. Prasad et Richards (2008) signalent que les prix médians dans les villes australiennes sont saisonniers, mais que cette saisonnalité disparaît une fois que la variation de la composition est neutralisée par la méthode de la stratification. À des niveaux agrégés, et notamment au niveau national, il semble donc peu probable qu'une série d'IPIR affiche de fortes fluctuations saisonnières. Néanmoins, à des niveaux inférieurs d'agrégation, il est utile de vérifier la présence d'une saisonnalité et de la corriger si des séries corrigées des variations saisonnières sont requises. Certains usagers tiennent à disposer de séries corrigées des variations saisonnières (en sus des séries non corrigées) si une saisonnalité des prix est constatée.

3.38 Le chapitre 4 présente un exemple chiffré qui montre comment la saisonnalité peut être traitée au moyen de méthodes indicielles simples. On peut également faire appel aux méthodes courantes d'ajustement des variations saisonnières.

18 La méthode des acquisitions nécessite un nouvel indice des prix du logement qui devrait probablement exclure l'élément «terrains» du prix de vente d'un logement neuf. Un nouvel indice approprié du coût de la construction permettrait d'approcher convenablement ce nouvel indice.

Appendice : Le rôle des indices des prix du logement dans la construction des coûts d'usage

3.39 Le présent appendice explique comment construire les coûts d'usage et les coûts d'opportunité. Il analyse dans un premier temps la construction des coûts d'usage pour les biens durables en général, il se penche ensuite sur les difficultés dérivant du fait que les biens immobiliers sont uniques et qu'ils sont composés de terrains et de bâtiments. Il examine pour finir l'évaluation des prix des services fournis par les logements occupés par leur propriétaire par la méthode du coût d'opportunité.

La construction des coûts d'usage pour les biens durables en général

3.40 Nous présentons ici les éléments d'une théorie du coût d'usage pour un bien de consommation durable. La durabilité consiste en essence à fournir un service à l'acheteur pendant de nombreuses périodes. Dans bien des cas (y compris l'évaluation des dépenses de consommation des ménages allouées aux services fournis par les logements occupés par leur propriétaire), il n'est pas conseillé d'imputer la totalité du prix d'achat d'un bien durable à la période d'achat initiale, mais plutôt de le répartir sur l'ensemble de sa vie utile. Se pose alors la question suivante : comment répartir ce prix intertemporel dans le temps?

3.41 Deux grandes méthodes permettent de déterminer le prix des *services* fournis par un logement occupé par son propriétaire[19] : la *méthode de l'équivalent-loyer* et la *méthode du coût d'usage*. La méthode du coût d'usage est importante en soi — quand le nombre de logements loués dans un pays est faible, il n'est pas réaliste d'évaluer les services fournis par les logements occupés par leur propriétaire par la méthode de l'équivalent-loyer —, mais aussi pour expliquer de quelle manière les bailleurs établissent le montant des loyers. Il est cependant plus difficile d'évaluer les prix des services de logement que, par exemple, ceux des services rendus par un modèle courant d'automobile, parce que les services de logements sont plus complexes[20]. Nous traiterons donc d'abord dans cette partie les problèmes liés à la détermination des prix des services fournis par un *bien de consommation durable ordinaire* (disponible sous la même forme pendant de nombreuses périodes) avant d'aborder les difficultés associées aux logements.

3.42 Le traitement des biens durables par la méthode du coût d'usage est en un certain sens très simple : il consiste à calculer le prix d'achat des biens durables au début de la période considérée et le coût de l'utilisation des services fournis par ces biens durant cette période, puis à déduire de ces coûts le profit qui peut être dégagé de la vente des biens à la fin de la période, en tenant compte des intérêts non perçus du fait que le capital a été investi dans l'achat du bien durable. Plusieurs détails de cette procédure prêtent cependant à controverse, comme le traitement de l'amortissement, des intérêts et des plus-values ou gains de détention.

3.43 Le concept du coût d'usage pose un autre problème en ce qu'il établit une distinction entre les achats effectués durant la période considérée et les stocks physiques du bien durable détenus au début et à la fin de l'exercice comptable. Normalement, dans le Système de comptabilité nationale, on suppose que tous les achats interviennent à un moment précis, par exemple au milieu de la période considérée, et que la consommation a également lieu durant cette période. Ainsi, le bien étant entièrement consommé pendant la période d'achat, il est inutile de tenir compte de la valeur du stock de biens durables de consommation dont disposent les ménages. Comme bon nombre de statisticiens sont peu au fait des problèmes assez complexes que pose la comptabilité des stocks et des flux, il peut être utile d'en donner ici une description détaillée.

3.44 Pour déterminer le coût net de l'utilisation d'un bien durable durant, par exemple, la période 0, on suppose que l'on achète une unité de ce bien durable au début de la période 0 au prix P^0. On peut vendre ce bien durable «usagé» ou «d'occasion» à la fin de la période 0 au prix P_S^1. Il semble donc raisonnable d'établir que le coût net de l'utilisation d'une unité du bien durable pendant la période 0 correspond à son prix d'achat initial P^0 moins la «valeur de rebut» P_S^1 à la fin de la période 0. L'argent reçu à la fin de la période n'a cependant pas autant de valeur que celui reçu au début de la période. Pour convertir la valeur à la fin de la période en valeur équivalente au début de la période, il faut donc *actualiser* le terme P_S^1 par le terme $1+r^0$ où r^0 est le taux d'intérêt nominal appliqué au consommateur au début de la période 0. Par conséquent, le *coût d'usage* u^0 *au cours de la période 0* du bien de consommation durable[21] se définit comme suit :

$$u^0 \equiv P^0 - P_S^1/(1+r^0) \qquad (3.A1)$$

3.45 On peut considérer la formule du coût d'usage (3.A1) autrement : le consommateur achète le bien durable

19 La méthode fondée sur le concept d'acquisition impute implicitement l'ensemble des services rendus par un logement récemment acquis à la période d'achat, mais le Système de comptabilité nationale ne la juge pas valable pour déterminer le prix des services rendus par les logements occupés par leur propriétaire. Il la valide en revanche pour évaluer le prix des services fournis par d'autres biens durables.

20 En particulier, les services de logement recouvrent à la fois les services fournis par les bâtiments et ceux fournis par le terrain qui les porte, et les logements sont généralement des biens uniques.

21 Cette méthode de calcul de la formule du coût d'usage a été utilisée par Diewert (1974), qui s'est lui-même inspiré d'une méthode que l'on doit à Hicks (1946; 326). À noter que ce coût d'usage sera ultérieurement interprété comme le coût d'usage en début de période, puisque tous les coûts d'usage sont exprimés en termes de prix actualisés au début de la période.

au début de la période 0 au prix P^0 et se paye à lui-même le loyer u^0. Le reste du prix d'achat, I^0, défini comme étant :

$$I^0 \equiv P^0 - u^0 \qquad (3.A2)$$

peut être considéré comme un *investissement*, qui doit rapporter le coût d'opportunité du capital approprié r^0 qui s'applique au consommateur. À la fin de la période 0, on peut obtenir ce taux de rendement à condition que I^0, r^0 et le prix de vente du bien durable à la fin de la période P_S^1 vérifient l'équation suivante :

$$I^0(1+r^0) = P_S^1 \qquad (3.A3)$$

Connaissant P_S^1 et r^0, l'équation (3.A3) détermine I^0, qui, P^0 étant donné, détermine à son tour le coût d'usage u^0 via l'équation (3.A2)[22].

3.46 Il ressort clairement de ce qui précède que la méthode du coût d'usage pour évaluer les prix des services fournis par un bien durable au cours d'une période comporte un élément d'investissement. Il convient de noter qu'elle constitue également une *méthode du coût d'opportunité financier* puisque le coût d'opportunité du capital immobilisé dans l'achat (ou dans la conservation) du bien durable est pris en compte. On notera enfin que les coûts d'usage diffèrent des prix des biens non durables ou des services, car le concept du coût d'usage suppose que l'on calcule le prix des biens durables à *deux* moments différents, et non à un seul. De ce fait, l'argent reçu ou payé au premier de ces moments (à supposer que les prix augmentent dans le pays) a davantage de valeur que l'argent reçu ou payé au second, ce qui explique l'intervention des *taux d'intérêt* dans la formule du coût d'usage.

3.47 Pour la même raison, les *prix prévus* peuvent jouer un rôle si l'on calcule le coût d'usage au début de la période considérée, et non pas à la fin. Deux possibilités s'offrent donc au statisticien pour le choix de P_S^1 :

- utiliser le *prix anticipé* du bien durable calculé en début de période pour la fin de la période considérée; ou
- utiliser le *prix de marché réel* d'un bien durable similaire d'occasion à la fin de la période considérée (si tant est qu'il existe).

3.48 L'utilisation d'un prix anticipé aboutit à un *coût d'usage ex ante*, tandis que l'utilisation du prix de marché réel du bien durable usagé à la fin de la période considérée donne un *coût d'usage ex post*. Quel concept appliquer dans la pratique? Dans le contexte qui nous occupe, deux raisons justifient le recours au concept *ex ante* :

- le coût d'usage *ex ante* est probablement plus proche du loyer du bien durable (s'il existe)[23], dont de nombreux statisticiens estiment qu'il correspond mieux aux services rendus par ce dernier au cours de la période considérée, et
- le coût d'usage *ex ante* est plus proche du *coût anticipé* de l'utilisation du bien durable au cours de la période considérée pour l'acquéreur; ce dernier ne peut connaître précisément le prix du bien en fin de période et doit donc formuler des hypothèses à cet égard; il sera donc le coût prévu de l'utilisation des services fournis par le bien durable au cours de la période considérée. Il devrait donc correspondre au prix des services fournis par ce bien qui est susceptible de motiver le comportement du consommateur.

Les moyens de formuler précisément des prévisions du prix de vente d'un bien durable usagé seront examinés plus loin, dans le cadre de l'analyse de la définition des prix des services de logement.

3.49 Compte tenu de ces difficultés, il n'est pas étonnant que de nombreux statisticiens préfèrent ne pas recourir au concept du coût d'usage pour calculer les prix. Néanmoins, il est parfois impossible de l'éviter pour calculer les prix des services fournis par les logements occupés par leur propriétaire dans certaines conditions. On peut exprimer la formule du coût d'usage (3.A1) sous une forme plus habituelle si l'on définit le *taux d'amortissement économique* δ^0 à la fin de la période 0 et le *taux d'inflation des actifs* i^0 sur la période 0. On définit le *taux d'amortissement économique* δ^0 à la fin de la période 1 comme suit :

$$(1 - \delta^0) \equiv P_S^1/P^1 \qquad (3.A4)$$

où P_S^1 est le prix d'un actif usagé à la fin de la période 0 et P^1 le prix d'un actif neuf à la fin de la période 0[24]. Le *taux d'inflation* de l'actif neuf *sur la période 0*, i^0, se définit comme suit :

$$1+i^0 \equiv P^1/P^0 \qquad (3.A5)$$

Si l'on retire P^1 des équations (3.A4) et (3.A5), on obtient la formule suivante qui exprime le *prix de l'actif usagé à la fin de la période 0* :

$$P_S^1 = (1 - \delta^0)(1 + i^0)P^0 \qquad (3.A6)$$

Si l'on introduit l'équation (3.A6) dans l'équation (3.A1), on obtient l'équation suivante qui décrit le *coût d'usage* u^0 *sur la période 0* :

$$u^0 = [(1 + r^0) - (1 - \delta^0)(1 + i^0)]P^0/(1 + r^0) \qquad (3.A7)$$

On notera que $r^0 - i^0$ peut être interprété comme étant le *taux d'intérêt réel* de la période 0, et $\delta^0(1+i^0)$ comme le *taux d'amortissement corrigé de l'inflation.*

3.50 Dans l'équation (3.A7), le coût d'usage u^0 est exprimé en termes de prix actualisés au *début* de la période 0. On peut également exprimer le coût d'usage en termes de prix

22 Diewert (1974; 504) a lui aussi procédé à ce calcul du coût d'usage des biens de consommation durables.

23 Si l'activité d'une entreprise consiste à louer à bail les services d'une automobile pendant une période donnée, elle doit établir des prévisions quant au prix des véhicules usagés à la fin de la période de location pour calculer les tarifs de location ou de crédit-bail de son stock d'automobiles.

24 Si le bien durable acheté (ou détenu) par le ménage au début de la période considérée est un bien usagé, P^1 sera le prix d'un bien durable usagé présentant le même état que le bien initialement détenu sur le marché d'occasion.

dont on calcule la valeur future ou «appréciés» à la *fin* de la période 0[25]. Le *coût d'usage à la fin de la période 0*, p^0, est défini comme suit :

$$p^0 \equiv (1 + r^0)u^0 = [(1 + r^0) - (1 - \delta^0)(1 + i^0)]P^0$$
$$= [r^0 - i^0 + \delta^0(1 + i^0)]P^0 \qquad (3.A8)$$

où la deuxième équation découle de l'équation (3.A7). Si le taux d'intérêt réel r^{0*} est défini comme le taux d'intérêt nominal r^0 moins le taux d'inflation de l'actif i^0 et que l'on néglige le terme généralement petit $\delta^0 i^0$, le coût d'usage à la fin de la période défini par l'équation (3.A8) se réduit à[26] :

$$p^0 = (r^{0*} + \delta^0)P^0 \qquad (3.A9)$$

Si l'on fait abstraction des coûts de transaction, on voit que le coût d'usage à la fin de la période défini par l'équation (3.A9) est un *loyer approximatif*; le loyer de l'utilisation d'un bien durable doit être égal au coût d'opportunité (réel) du capital immobilisé, $r^{0*}P^0$, plus la perte de valeur de l'actif durant la période, $\delta^0 P^0$. Les formules (3.A8) et (3.A9) permettent donc de mieux comprendre les déterminants économiques du montant du loyer ou du crédit-bail des biens de consommation durables.

3.51 Si l'on utilise la formule simplifiée du coût d'usage définie par l'équation (3.A9), il n'est guère plus difficile de calculer un indice des prix fondé sur le coût d'usage d'un bien durable que de calculer un indice des prix fondé sur le prix d'achat de ce bien, P^0. Il suffit au statisticien de :

- formuler une hypothèse raisonnable quant au niveau approprié du taux d'intérêt réel mensuel ou trimestriel $r0^{*}$[27];
- formuler une hypothèse quant au niveau raisonnable du taux d'amortissement mensuel, trimestriel ou annuel δ^0[28];
- relever les prix d'achat P^0 du bien durable et définir le coût d'usage.

25 Ainsi, le coût d'usage u^0 au début de la période actualise tous les coûts et bénéfices monétaires en leur équivalent en unité monétaire au début de la période 0, tandis que p^0 actualise (ou revalorise) tous les coûts et bénéfices monétaires en leur équivalent en unité monétaire à la fin de la période 0. Cela laisse ouverte la question du traitement adéquat des flux correspondants aux transactions effectuées durant la période. Suivant les conventions adoptées dans la comptabilité financière, les transactions effectuées durant l'exercice comptable doivent être considérées comme ayant lieu à la fin de cet exercice. Par conséquent, d'après cette convention, les statisticiens doivent utiliser les coûts d'usage à la fin de la période. Pour d'autres éléments concernant les coûts d'usage en début et en fin de période, voir Diewert (2005; 485).

26 Si l'on prend le ratio de la valeur locative approximative du bien durable, p^0, à sa valeur d'actif, P^0, on obtient le ratio du loyer à la valeur du logement, $p^0/P^0 = r^{0*} + \delta^0$, qui est égal à la somme du taux d'intérêt réel approprié r^{0*}, plus le taux d'amortissement approprié δ^0. Comme les taux d'intérêt réels et les taux d'amortissement sont à peu près constants dans le temps, le ratio du loyer à la valeur du logement le sera également; en conséquence, le ratio historique du loyer à la valeur du logement multiplié par l'indice courant des prix des actifs donnera généralement une approximation adéquate du loyer imputé du bien de consommation durable. Les études portant sur l'immobilier désignent souvent le ratio du loyer à la valeur sous le nom de taux de capitalisation. Voir par exemple Garner et Short (2009; 237), ou Crone, Nakamura et Voith (2009; 70).

27 Cette opération est assez complexe. Il est difficile d'établir précisément le niveau approprié du coût d'opportunité nominal du capital supporté par le ménage et, même si nous parvenons à un accord sur ce point, il sera difficile d'estimer les taux d'inflation anticipés. Au final, cela reviendra éventuellement à choisir un taux d'intérêt réel quelque peu arbitraire, dans une fourchette comprise entre 2 % et 5 % (taux annuels), en fonction des taux récemment enregistrés dans le pays considéré.

28 Le modèle géométrique de l'amortissement ne nécessite qu'un seul taux d'amortissement, mensuel ou trimestriel. D'autres modèles d'amortissement exigent parfois que l'on estime une séquence des taux d'amortissement de chaque génération. Si le taux annuel d'amortissement géométrique estimé est δ_a, on peut obtenir le taux mensuel d'amortissement géométrique correspondant δ en résolvant l'équation $(1 - \delta)^{12} = 1 - \delta_a$. De même, si le taux d'intérêt annuel réel estimé est r_a^*, on peut obtenir le taux d'intérêt mensuel réel correspondant r^* en résolvant l'équation $(1 + r^*)^{12} = 1 + r_a^*$.

3.52 L'évaluation des services rendus par un bien de consommation durable par la méthode du coût d'usage présente d'autres difficultés. L'analyse ci-dessus porte uniquement sur la construction du coût d'usage d'un bien de consommation durable récemment acheté. Il nous faut l'élargir pour évaluer également le prix des services fournis par des unités usagées du bien en question. Pour ce faire, nous devons formuler des hypothèses quant au mode d'amortissement applicable à ce dernier; le flux de services fourni au consommateur demeure-t-il constant durant toute la durée de vie utile du bien ou diminue-t-il à mesure que celui-ci vieillit? S'il demeure constant, la méthode d'amortissement est celle dite du *fiacre centenaire* ou de *l'ampoule électrique*; s'il décline selon un taux linéaire ou géométrique, il s'agit de la *méthode d'amortissement linéaire* ou *géométrique*[29].

3.53 Comment déterminer laquelle de ces deux méthodes (fiacre centenaire ou géométrique) est applicable à un bien de consommation durable donné? On peut établir les deux modèles d'amortissement (et de valorisation) si l'on dispose d'*informations transversales sur la valeur locative* du bien en fonction de l'âge de l'actif loué. Si l'on suppose que le modèle d'amortissement appliqué est celui du fiacre centenaire, le loyer du bien à un moment donné doit être à peu près constant quel que soit l'âge de ce dernier; si, en revanche, il s'agit du modèle d'amortissement géométrique, le loyer doit diminuer selon un taux géométrique en fonction de l'âge du bien durable usagé. Les différents modèles d'amortissement peuvent ainsi être établis s'il existe un marché de la location pour les biens durables. De manière analogue, on peut établir d'autres modèles d'amortissement si l'on dispose d'*informations transversales sur le prix des unités usagées* du bien de consommation durable[30].

Coût d'usage des logements occupés par leur propriétaire

3.54 Un *logement occupé par son propriétaire* diffère d'un bien de consommation durable courant de par son caractère unique. Il sera donc difficile d'utiliser les informations relatives aux prix des actifs usagés pour déterminer le modèle d'amortissement, ce qui est nécessaire pour évaluer le coût d'usage d'un logement de cette nature. Comme signalé dans l'introduction au présent chapitre, un logement spécifique dans un pays donné est unique pour diverses raisons :

- L'*emplacement* de chaque logement est unique et influe sur son prix.

29 Pour des descriptions des méthodes de construction des coûts d'usage en fonction de l'âge de l'actif pour chacun de ces modèles d'amortissement, voir Diewert et Lawrence (2000) ou Diewert (2005; 506–521).

30 Dans le contexte de l'immobilier, où chaque logement peut être considéré comme un actif unique, il convient de formuler des hypothèses supplémentaires pour définir le mode d'amortissement. On supposera par exemple que la même méthode d'amortissement s'applique à tous les logements d'une catégorie donnée de bâtiments. Dans ce cas, les données empiriques conduisent à penser que la méthode du fiacre centenaire ne sera probablement pas appliquée au marché du logement, les locataires étant généralement disposés à payer un loyer majoré pour louer un logement neuf plutôt qu'un logement plus ancien de même type. Pour des données empiriques sur ce point, voir Malpezzi, Ozanne et Thibodeau (1987; 378), et Hoffmann et Kurz (2002; 19).

- Le logement *se déprécie* avec le temps; à moins que la méthode d'amortissement du fiacre centenaire ne soit appliquée, l'utilité d'un logement donné pour le ménage l'occupant diminue généralement au fil du temps sous l'effet du vieillissement du bâtiment.
- Les effets de la dépréciation peuvent en revanche être compensés par des *dépenses de rénovation,* qui augmentent l'utilité du logement.

3.55 Dans certains cas, il convient de décomposer le prix d'un bien immobilier en deux éléments : les terrains et les bâtiments. Pour modéliser cette situation, prenons un logement récemment construit et acheté au début de la période 0. Supposons que son prix d'achat soit V^0. On peut considérer que cette valeur est la somme du coût de production du bâtiment, disons, $P_S{}^0Q_S{}^0$, où $Q_S{}^0$ est le nombre de mètres carrés de surface de plancher du bâtiment et $P_S{}^0$ le prix de la construction par mètre carré au début de la période 0, et du coût du terrain, disons, $P_L{}^0Q_L{}^0$, où $Q_L{}^0$ est le nombre de mètres carrés correspondant à la surface de plancher du bâtiment et au terrain qui l'entoure, et $P_L{}^0$ le prix du terrain par mètre carré au début de la période 0[31]. Par conséquent, au début de la période 0, la *valeur du logement* V^0, se définit comme suit :

$$V^0 = P_S{}^0Q_S{}^0 + P_L{}^0Q_L{}^0 \qquad (3.A10)$$

3.56 Supposons que le prix prévu d'une unité d'un bâtiment neuf au début de la période 1 soit $P_S{}^{1a}$ et que le prix prévu d'une unité de terrain au début de la période 1 soit $P_L{}^{1a}$. On définit les *taux d'inflation anticipés des bâtiments neufs et des terrains au cours de la période 0*, respectivement $i_S{}^0$ et $i_L{}^0$, de la façon suivante :

$$1 + i_S{}^0 \equiv P_S{}^{1a}/P_S{}^0 \qquad (3.A11)$$

$$1 + i_L{}^0 \equiv P_L{}^{1a}/P_L{}^0 \qquad (3.A12)$$

Soit δ^0 le taux d'amortissement du bâtiment au cours de la période 0. La valeur anticipée du bâtiment et du terrain connexe au début de la période 1 est alors égale à :

$$V^{1a} = P_S{}^{1a}(1 - \delta^0)Q_S{}^0 + P_L{}^{1a}Q_L{}^0 \qquad (3.A13)$$

La valeur anticipée de l'unité de logement à la fin de la période 1, V^{1a}, est donc égale au prix prévu (par unité de bâtiment neuf de même qualité) à la fin de la période, $P_S{}^{1a}$, multiplié par un moins le taux d'amortissement de la période 0, $(1 - \delta^0)$, multiplié par la quantité de bâtiments achetés au début de la période 0, $Q_S{}^0$[32], plus le prix prévu des terrains à la fin de la période 0, $P_L{}^{1a}$, multiplié par la quantité de terrains associés au bâtiment, $Q_L{}^0$.

3.57 Calculons maintenant le coût (y compris le coût d'opportunité du capital imputé r^0)[33] de l'achat du logement au début de la période 0 et de sa vente (fictive) à la fin de la période 0. On obtient, à l'aide des équations (3.A11)–(3.A13), le *coût d'usage ou loyer imputé* R^0 du logement *à la fin de la période 0* suivant :

$$\begin{aligned} R^0 &\equiv V^0(1 + r^0) - V^{1a} \qquad (3.A14)\\ &= [P_S{}^0Q_S{}^0 + P_L{}^0Q_L{}^0](1 + r^0) - [P_S{}^{1a}(1 - \delta^0)Q_S{}^0 + P_L{}^{1a}Q_L{}^0]\\ &= [P_S{}^0Q_S{}^0 + P_L{}^0Q_L{}^0](1 + r^0) - [P_S{}^0(1 + i_S{}^0)(1 - \delta^0)Q_S{}^0 + P_L{}^0(1 + i_L{}^0)Q_L{}^0]\\ &= p_S{}^0Q_S{}^0 + p_L{}^0Q_L{}^0 \end{aligned}$$

où les *coûts d'usage* respectifs *du bâtiment et du terrain* au cours de la période 0, $p_S{}^0$ et $p_L{}^0$ se définissent comme suit :

$$\begin{aligned} p_S{}^0 &= [(1 + r^0) - (1 + i_S{}^0)(1 - \delta^0)]P_S{}^0\\ &= [r^0 - i_S{}^0 + \delta^0(1 + i_S{}^0)]P_S{}^0 \qquad (3.A15) \end{aligned}$$

$$p_L{}^0 = [(1 + r^0) - (1 + i_L{}^0)]P_L{}^0 = [r^0 - i_L{}^0]P_L{}^0 \qquad (3.A16)$$

On notera que les formules ci-dessus font apparaître certains des principaux déterminants de la valeur locative de marché des biens immobiliers de location[34]. On peut encore simplifier les formules du coût d'usage définies par les équations (3.A15) et (3.A16) si l'on fait les mêmes approximations qu'à la section précédente (voir l'équation (3.A9) ci-dessus); on suppose donc que le taux d'intérêt réel r^{0*} peut être une approximation des termes $r^0 - i_S{}^0$ et $r^0 - i_L{}^0$, et on néglige le petit terme δ^0 multiplié par $i_S{}^0$ de l'équation (3.A15). Les coûts d'usage définis par les équations (3.A15) et (3.A16) se réduisent alors à :

$$p_S{}^0 = (r^{0*} + \delta^0)P_S{}^0 \qquad (3.A17)$$

$$p_L{}^0 = r^{0*}P_L{}^0 \qquad (3.A18)$$

3.58 L'exposé ci-dessus a passé sous silence deux autres sources de coûts associés à la propriété d'un logement durant la période 0 :

- divers coûts liés à l'entretien et à l'assurance du logement en question et
- les impôts fonciers dont le propriétaire est redevable aux autorités locales ou à l'État.

Supposons que les frais d'entretien et d'assurance de la période 0, $M_S{}^0$, soient essentiellement associés au bâtiment, et non au terrain qui le porte. Supposons ensuite que ces coûts soient payés à la fin de la période 0. On peut les transformer en *frais par unité de bâtiment* $\mu_S{}^0$ selon la formule suivante :

$$\mu_S{}^0 \equiv M_S{}^0/(P_S{}^0Q_S{}^0) \qquad (3.A19)$$

31 Si le logement fait partie d'un bâtiment composé de plusieurs logements, le terrain qui lui est associé sera la part correspondante de la superficie totale du terrain. Cette part pourrait correspondre à 1 divisé par le nombre de logements sur le lotissement, ou à la surface de plancher du logement divisée par la surface de plancher totale du bâtiment. Ces deux répartitions se justifieraient.

32 Ainsi le taux d'amortissement de la période 0, δ^0, est le taux d'amortissement transversal prévu de la fin de la période; autrement dit, δ^0 est défini par l'équation $(1-\delta^0) = V_S{}^{1a}/P_S{}^{1a}Q_S{}^0$, où $V_S{}^{1a}$ est la valeur de marché prévue du bâtiment (amorti) à la fin de la période 0, et $P_S{}^{1a}Q_S{}^0$ la valeur prévue à la fin de la période 0 d'un bâtiment neuf de surface de plancher $Q_S{}^0$.

33 On trouvera des analyses plus approfondies quant au choix du coût d'opportunité du capital approprié, dans le cas où le propriétaire d'un logement aurait financé son achat par un prêt, dans Diewert et Nakamura (2009), Diewert, Nakamura et Nakamura (2009) et Garner et Verbrugge (2009b; 176).

34 Si l'on regarde l'équation (3.A16), on peut observer que le coût d'usage du terrain pourrait être négatif si le taux d'appréciation prévu du prix du terrain, $i_L{}^0$, est supérieur au coût d'opportunité du capital au début de la période, r^0. D'éventuelles solutions à cette difficulté seront analysées plus loin.

Supposons que les taxes foncières qui portent sur le bâtiment, T_S^0, et celles qui portent sur le terrain, T_L^0, sont payées à la fin de la période 0. On peut alors définir le taux d'imposition du bâtiment et du terrain pour la période 0, τ_S^0 et τ_L^0, comme suit :

$$\tau_S^0 \equiv T_S^0/(P_S^0 Q_S^0) \text{ et } \tau_L^0 \equiv T_L^0/(P_L^0 Q_L^0) \quad (3.A20)$$

Ces coûts supplémentaires doivent être ajoutés au loyer imputé pour l'utilisation du logement R^0. L'équation (3.A14) s'exprime donc maintenant comme suit :

$$R^0 \equiv V^0(1 + r^0) - V^{1a} + M_S^0 + T_S^0 + T_L^0 \quad (3.A21)$$
$$= p_S^0 Q_S^0 + p_L^0 Q_L^0$$

où les nouveaux *coûts d'usage* respectifs *du bâtiment et du terrain* pendant la période 0, p_S^0 et p_L^0, sont définis comme suit :

$$p_S^0 = [r^0 - i_S^0 + \delta^0(1 + i_S^0) + \mu_S^0 + \tau_S^0]P_S^0 \quad (3.A22)$$

$$p_L^0 = [r^0 - i_L^0 + \tau_L^0] P_L^0 \quad (3.A23)$$

Le *loyer imputé d'un logement* dérivant de l'application de la méthode du coût d'usage pour évaluer les services de logement se compose donc de six coûts principaux :

- le coût d'opportunité réel du capital financier immobilisé dans le bâtiment, $(r^0 - i_S^0)P_S^0 Q_S^0$;
- le coût d'opportunité réel du capital financier immobilisé dans le terrain, $(r^0 - i_L^0)P_L^0 Q_L^0$;
- le coût d'amortissement du bâtiment, $\delta^0(1 + i_S^0)P_S^0 Q_S^0$;
- les coûts d'entretien et d'assurance associés au bâtiment, $\mu_S^0 P_S^0 Q_S^0$;
- les taxes foncières liées au bâtiment, $\tau_S^0 P_S^0 Q_S^0$;
- les taxes foncières associées au terrain portant et entourant le bâtiment, $\tau_L^0 P_L^0 Q_L^0$.

3.59 La méthode du coût d'usage présentée ci-dessus pour évaluer le prix des services fournis par les logements peut s'appliquer à diverses *strates d'habitations* : logements individuels, maisons en rangée ou duplex et immeubles d'habitation. Pour ces deux dernières catégories, il faut construire la composante «terrain» de chaque logement. Par exemple, si un immeuble d'habitation compte vingt appartements, la part de terrain correspondant à chaque appartement pourrait être établie à 1/20e de la superficie totale occupée par l'immeuble[35]. Les logements peuvent également être regroupés en fonction de leur type de construction, à savoir essentiellement le bois, la brique, le béton ou les matériaux «traditionnels».

3.60 Si un organisme statistique établit des estimations du compte de patrimoine national, on devrait disposer de données relatives à la valeur totale des terrains et des bâtiments résidentiels. Il se peut toutefois que le nombre de terrains résidentiels ne soit pas connu. On peut obtenir des estimations du stock réel de bâtiments résidentiels d'un pays en déflatant la valeur estimée des logements résidentiels dans le compte de patrimoine au moyen du déflateur national des prix des investissements dans le logement résidentiel correspondant.

3.61 La méthode d'évaluation des prix des services de logement selon le concept du coût d'usage peut être utilisée de deux façons au moins :

- on peut comparer les coûts d'usage aux valeurs locatives de marché de logements effectivement loués pendant la période observée;
- les coûts d'usage peuvent servir à évaluer les services fournis par les logements occupés par leur propriétaire.

Comme on le verra plus loin, il se trouve que les coûts d'usage s'approchent des loyers pratiqués sur le marché (tout au moins en ce qui concerne les logements à faible coût aux États-Unis), pourvu que les prévisions en matière d'inflation des prix immobiliers soient établies selon une méthode déterminée.

3.62 Comme indiqué précédemment, deux grandes méthodes ont été proposées pour évaluer les services fournis par les logements occupés par leur propriétaire dans les comptes nationaux : le concept du coût d'usage, que nous venons d'expliquer, et le concept d'équivalent-loyer. La seconde est simple; nous cherchons pour une strate donnée de logements occupés par leur propriétaire des logements similaires loués et imputons la valeur locative de marché aux logements occupés par leurs propriétaires correspondants. Cette méthode fonctionne bien dans de nombreux pays, mais pas dans ceux où les marchés locatifs sont étroits ou les loyers réglementés.

3.63 Si l'on utilise la méthode du coût d'usage pour évaluer les services fournis par les logements occupés par leur propriétaire dans un pays, il convient de retirer le terme correspondant à l'entretien et aux primes d'assurance, μ_S^0, de la formule du coût d'usage des bâtiments (3.A22), les dépenses d'entretien et d'assurance pour ces logements étant généralement comptabilisées ailleurs, dans les dépenses des ménages.

3.64 La méthode de calcul simplifiée du coût d'usage du logement expliquée aux équations (3.A17) et (3.A18) peut l'être encore davantage si l'on suppose que le ratio de la quantité de terrain aux bâtiments est fixe. Le coût d'usage global du logement est alors égal à $[r^{0*} + \delta + \mu + \tau]P_H^0$, où P_H est un indice des prix des logements corrigé de la qualité applicable à l'ensemble du stock immobilier national (bâtiments et terrains les portant) pendant la période considérée et δ, μ et τ sont respectivement un taux d'amortissement, un taux de dépenses d'entretien et d'assurance et un taux d'imposition foncier qui s'applique à l'ensemble composé par les bâtiments et le terrain. Comme nous l'avons vu au paragraphe ci-dessus, le terme μ devrait ici être retiré du coût d'usage simplifié. La *méthode simplifiée* qui en résulte est appliquée en Islande (voir Gudnason (2004) et Gudnason

35 Il n'est pas facile d'allouer le terrain commun aux unités de logement en parts individuelles; par exemple, au lieu de diviser le terrain en parts égales, nous pourrions prendre pour valeur d'imputation la surface de plancher relative de chaque appartement. L'élévation relative des différents appartements pose également des problèmes; le loyer d'un appartement situé à un étage supérieur sera généralement plus élevé que celui d'un appartement situé à un étage inférieur.

et Jónsdóttir (2009)[36]) et dans quelques pays européens; en voir la description détaillée dans Katz (2009)[37]. Le Bureau of Economic Analysis des États-Unis en utilise une variante. Lebow et Rudd (2003; 168) notent qu'aux États-Unis, les comptes nationaux imputent les services fournis par les logements occupés par leur propriétaire en appliquant les *ratios du loyer à la valeur* des logements occupés par des locataires au stock de logements occupés par leur propriétaire qui présentent les mêmes caractéristiques que les logements loués[38]. On peut considérer ce ratio du loyer à la valeur comme une estimation du taux d'intérêt réel applicable majoré du taux d'amortissement, d'un taux de dépenses d'entretien et d'assurance et du taux d'imposition foncier, $r^{0*} + \delta + \mu + \tau$[39].

3.65 Comment faut-il estimer le taux d'intérêt réel, r^{0*}? Une méthode consiste simplement à formuler une hypothèse raisonnable[40] :

«Restait à définir la valeur appropriée du taux de rendement réel. Les données présentées au groupe de travail indiquaient que, dans les pays d'Europe de l'Ouest tout au moins, le taux de rendement réel adapté aux logements occupés par leur propriétaire était inférieur de 2,5 % à 3 % à celui d'autres biens durables. Le groupe a estimé d'un commun accord que, compte tenu de la situation des pays candidats d'Europe de l'Est, il convenait de supposer un taux de rendement réel de 2,5 % sur les logements comme sur les terrains.» Arnold J. Katz (2009; 46).

3.66 Une deuxième méthode consiste à utiliser les taux d'intérêt hypothécaires comme mesure du coût d'opportunité nominal du capital financier immobilisé dans le logement et de recourir aux techniques économétriques de prévision pour estimer l'inflation prévue des prix immobiliers (le taux d'intérêt réel peut alors être fixé à un montant égal au taux d'intérêt nominal moins le taux d'inflation prévu des prix immobiliers). Verbrugge (2008) et Garner et Verbrugge (2009a) (2009b) ont testé plusieurs variantes de cette approche sur des données américaines. Ils ont cependant montré qu'elle n'était pas efficace dans la mesure où les estimations du coût d'usage en résultant étaient extrêmement irrégulières (et souvent négatives) et très éloignées des valeurs locatives de marché correspondantes.

3.67 Une troisième méthode de détermination d'un taux d'intérêt réel approprié pour calculer le coût d'usage des services de logement a été appliquée par Garner et Verbrugge (2009b) à des données américaines. Ces auteurs ont utilisé les taux d'intérêt hypothécaires applicables comme estimation du coût d'opportunité nominal du capital financier et utilisé le taux d'inflation anticipé de l'indice des prix à la consommation pour la période courante comme mesure de l'appréciation prévue des prix immobiliers. À leur grande surprise, ils ont constaté que les coûts d'usage ainsi calculés correspondaient assez bien aux valeurs locatives de marché[41]. On peut conclure que la formulation d'une hypothèse raisonnable quant au taux d'intérêt réel ou le recours au taux d'inflation de l'IPC comme indicateur du taux d'inflation prévu des prix immobiliers aboutissent à des coûts d'usage raisonnables qui devraient être assez similaires aux valeurs locatives de marché, tout au moins pour les logements relativement bon marché.

3.68 À l'évidence, les principaux déterminants des coûts d'usage des bâtiments et des terrains sont l'indice des prix de la construction des logements neufs, P_S^t, et l'indice des prix des terrains résidentiels, P_L^t. La plupart des organismes statistiques calculent un indice des prix à qualité constante des bâtiments résidentiels neufs, car celui-ci est indispensable pour déflater les dépenses d'investissement dans les bâtiments résidentiels dans les comptes nationaux. On peut l'utiliser comme approximation de P_S^t[42].

3.69 Ainsi s'achève le tour d'horizon de la méthode du coût d'usage aux fins d'évaluation des prix des services de logement. La section qui suit présente une autre approche à l'évaluation des prix des services fournis par les logements occupés par leur propriétaire : la *méthode du coût d'opportunité*.

Évaluation des services fournis par les logements occupés par leur propriétaire par la méthode du coût d'opportunité

3.70 Rappelons-nous les deux grandes méthodes d'évaluation des services fournis par les logements occupés par leur propriétaire : la méthode de l'équivalent-loyer et celle du coût d'usage. Dans la première, le propriétaire d'un logement qui choisit d'y vivre (ou tout au moins de ne pas le mettre en location) évalue les services fournis par le logement en fonction du manque à gagner sur le loyer en vigueur sur le marché. Il s'agit là d'un *coût d'opportunité* très *direct* lié à l'utilisation du logement. Le coût d'usage, en revanche, correspond essentiellement au *coût d'opportunité financier* lié à l'utilisation des services fournis par le logement durant la période considérée.

36 Le taux d'intérêt réel retenu est d'environ 4 % par an, et l'on pose en hypothèse un taux d'amortissement global des terrains et des bâtiments de 1,25 % par an. Le taux d'amortissement des seuls bâtiments est estimé à 1,5 % par an. L'IPC islandais comptabilise séparément les impôts fonciers. Les informations sur le prix des logements sont fournies par la Commission nationale d'évaluation, à partir des données sur les ventes de logements neufs et anciens. Cette commission estime également la valeur du stock de logements et des terrains en Islande au moyen d'un modèle de régression hédonique établi à partir des données sur les ventes de biens immobiliers. La valeur du logement de chaque ménage est tirée de l'enquête sur le budget des ménages.

37 Katz (2009) et Garner et Verbrugge (2009b; 176) donnent d'autres références d'études portant sur la méthode simplifiée du coût d'usage.

38 Voir également Crone, Nakamura et Voith (2009) et Garner et Short (2009; 237) pour une description de cette méthode par capitalisation pour déterminer la valeur locative des logements à partir des valeurs correspondantes estimées des actifs. On peut constater que cette méthode permet en fait d'appliquer le concept de l'équivalent-loyer à l'évaluation des services fournis par les logements occupés par leur propriétaire.

39 Si la valeur d'un logement occupé par son propriétaire est V^0 et que le ratio du loyer à la valeur d'un logement loué présentant les mêmes caractéristiques est $\gamma = r^{0*} + \delta + \mu + \tau$, le loyer imputé au logement occupé par son propriétaire est égal à $(\gamma - \mu)V^0 = (r^{0*} + \delta + \tau)V^0$, puisque les dépenses d'assurance et d'entretien qui lui sont associées seront comptabilisées ailleurs dans le Système de comptabilité nationale.

40 Le Bureau australien des statistiques suppose un taux d'intérêt réel constant de 4 % par an pour établir ses estimations des services en capital.

41 Dans le cadre de cette approche, Garner et Verbrugge (2009b; 179) ont également constaté l'absence de coûts d'usage négatifs dans leurs estimations fondées sur le jeu de données américaines.

42 Il est possible que cet indice ne soit qu'une approximation puisqu'il couvre à la fois la construction de biens immobiliers de location et celle de logements occupés par leur propriétaire.

Tableau 3.1. Estimations en pourcentage des ratios loyer/valeur (taux de capitalisation)

	Locataire			
	Alaska	Washington, D.C.	Caraïbes	Hawaï–Pacifique
Prix (en dollars)	(1)	(2)	(3)	(4)
50.000	13,0	8,9	6,3	6,9
100.000	12,0	8,2	5,8	6,4
200.000	10,2	6,9	4,9	5,4
500.000	6,2	4,3	3,0	3,3

Source : Heston et Nakamura (2009a).

Certains ont suggéré que le véritable coût d'opportunité associé à l'utilisation des services fournis par un logement occupé par son propriétaire est le montant *maximal entre le manque à gagner sur le loyer et le coût d'usage :*

«Nous conclurons cette section sur l'observation (contestable) suivante : peut-être le coût d'opportunité «correct» d'un logement occupé par son propriétaire n'est-il pas son coût d'usage interne, mais le montant *maximum* du coût d'usage interne et du prix auquel ce bien pourrait se louer sur le marché. Après tout, le concept du coût d'opportunité est censé représenter le *sacrifice maximum* supporté pour consommer ou utiliser un objet, ce qui semblerait conduire à l'observation ci-dessus.» W. Erwin Diewert (2009b; 113).

Diewert et Nakamura (2009) et Diewert, Nakamura et Nakamura (2009) ont appliqué cette méthode du coût d'opportunité pour évaluer les services fournis par les logements occupés par leur propriétaire plus en détail, mais l'on voit bien qu'elle paraît valable. Elle présente en outre l'avantage de supprimer le problème associé au concept du coût d'usage, à savoir les coûts d'usage *négatifs* que celle-ci peut produire si des taux d'inflation *ex post* ou prévisionnels des prix immobiliers sont utilisés dans la formule.

3.71 Dans la pratique, l'application de la méthode du coût d'opportunité pour évaluer les services fournis par les logements occupés par leur propriétaire peut aboutir à des résultats analogues à ceux de la méthode de l'équivalent-loyer si seulement le taux d'inflation prévisionnel figurant dans la formule du coût d'usage est égal au taux d'inflation de l'IPC; en effet, Garner et Verbrugge (2009b) ont démontré que, pour la plupart des biens immobiliers locatifs de bas de gamme, la méthode de l'équivalent-loyer et celle du coût d'usage obtiennent des résultats très similaires, aux États-Unis en tout cas. Il apparaît toutefois que, pour les biens de haut de gamme, les coûts d'usage peuvent être considérablement supérieurs aux loyers correspondants sur le marché. Le tableau 3.1, tiré de Heston et Nakamura (2009a; 113) (2009b; 277) présente le rapport entre la valeur locative de marché annuelle moyenne et la valeur marchande des biens immobiliers en location dans diverses régions, autrement dit le taux de rendement en fonction de la valeur du bien locatif. Le tableau se fonde sur une étude des fonctionnaires fédéraux menée dans le cadre d'un exercice de protection des données concernant le programme d'indemnité de vie chère administré par l'Office of Personnel Management des États-Unis. Ce programme, lancé en 1948, verse aux fonctionnaires fédéraux de trois régions (Alaska, Caraïbes et Pacifique) une indemnité de vie chère fondée sur l'écart entre les prix immobiliers pratiqués dans ces régions et ceux de la région de la ville de Washington[43].

Deux éléments ressortent du tableau 3.1 :

- les taux de capitalisation diffèrent sensiblement d'une région à l'autre[44], et
- dans toutes les régions, le taux de capitalisation des biens immobiliers haut de gamme est inférieur de moitié environ à celui des biens bas de gamme.

Le second point ressort également de données américaines beaucoup plus étoffées portant sur les loyers annuels des années 2004–06 construits sous forme de fonction des prix correspondants des logements figurant au graphique 1 dans Garner et Verbrugge (2009b; 178). Le loyer annuel moyen correspondant à un logement de 100.000 dollars s'élevait à quelque 10.000 dollars, mais à 30.000 dollars environ pour un logement de 900.000 dollars. Le taux de capitalisation passait donc de 10 % à 3,3 % environ lorsque le prix du logement augmentait de 100.000 dollars à 900.000 dollars.

3.72 Quels sont les facteurs qui peuvent expliquer cette chute considérable du taux de capitalisation entre les biens immobiliers bon marché et les biens plus onéreux? Comme nous l'avons déjà signalé, le ratio loyer/valeur peut être considéré comme une estimation du taux d'intérêt réel applicable plus le taux d'amortissement, plus le taux d'entretien et d'assurance, plus le taux d'imposition foncier, $r^{0^*} + \delta + \mu + \tau$, taux qui ne devraient pas différer sensiblement pour des biens immobiliers de valeurs distinctes. Trois explications au moins sont possibles :

- La part de terrain des biens immobiliers de grande valeur est peut-être nettement supérieure; en conséquence, le taux

43 Ce programme vise à comparer le coût de la vie pour les employés fédéraux vivant dans les régions non continentales des États-Unis à celui de la région de la ville de Washington. La comparaison du prix des logements est l'une des plus importantes et des plus difficiles à effectuer dans le cadre de ce programme. Les régions concernées sont l'Alaska, Guam, Hawaï, Porto Rico, et les îles Vierges des États-Unis, qui présentent une grande diversité de climats et de besoins en matière de logement.

44 Les taux de capitalisation relativement élevés de l'Alaska tiennent peut-être à l'intégration des services de chauffage dans le loyer.

d'amortissement δ, considéré comme la diminution de la valeur du bien sous l'effet du vieillissement du bâtiment, baisse à mesure que le ratio terrain/bâtiment augmente[45].

- Il se peut qu'un pourcentage substantiel des dépenses de suivi, de comptabilité et de facturation d'un bailleur corresponde à des coûts fixes; ces coûts diminuent donc, en pourcentage du loyer, à mesure que la valeur du bien immobilier augmente.
- Les biens immobiliers résidentiels de grande valeur ne sont pas loués sur une base commerciale; ils le sont parfois à titre temporaire, le locataire faisant fonction de «gardien» et payant un loyer dans une certaine mesure subventionné en regard du coût d'opportunité financier du propriétaire.

La détermination imparfaite du taux d'amortissement ne peut probablement pas expliquer la baisse sensible des taux de capitalisation à mesure que la valeur des biens immobiliers augmente; les taux d'amortissement estimés des logements se situent généralement dans une fourchette de 1 % à 2 % par an[46], un niveau trop faible pour expliquer entièrement la diminution des taux de capitalisation. De la même manière, les coûts associés à l'entretien et aux primes d'assurance d'un bien loué inclus dans le terme μ devraient être relativement faibles et ne devraient donc pas expliquer entièrement ce phénomène. La troisième explication pourrait donc être la bonne. Si c'est le cas, l'évaluation des services fournis par les logements occupés par leur propriétaire par la méthode du coût d'opportunité attribuera à ces services une valeur nettement supérieure à celle résultant de la méthode de l'équivalent-loyer[47].

45 Cette explication a été proposée par Diewert (2009a; 486) et par Garner et Verbrugge (2009b; 182).

46 Garner et Verbrugge (2009b; 176) et Garner et Short (2009; 244) supposent des taux d'amortissement (en pourcentage de la valeur du bien immobilier englobant les bâtiments et les terrains) de 1 % par an.

47 De ce fait, le décalage entre l'évaluation des services fournis par les logements occupés par leur propriétaire par la méthode de l'équivalent-loyer et par celle du coût d'opportunité peut ne pas être très prononcé dans le cadre de séries chronologiques, car ces deux mesures évoluent de concert. Mais, dans le cadre de comparaisons internationales, cet argument n'est pas valable, le pourcentage de logements occupés par leur propriétaire variant considérablement selon les pays.

4

Méthodes de stratification ou de poststratification

Indices fondés sur la moyenne simple ou la médiane

4.1 Les mesures les plus simples de la variation des prix du logement reposent dans une certaine mesure sur une tendance centrale de la répartition des prix des logements vendus sur une période donnée, en particulier sur la moyenne ou sur la médiane. Étant donné que la répartition des prix des logements est en général positivement asymétrique (surtout en raison du caractère hétérogène du logement, de l'asymétrie positive dans la répartition du revenu et du plancher zéro pour les prix de transaction), on utilise plus généralement la médiane que la moyenne. Aucune donnée sur les caractéristiques du logement n'étant nécessaire pour calculer la médiane, on peut facilement construire un indice de prix qui suive l'évolution du prix du logement médian vendu d'une période à la suivante. Un autre attrait des indices médians réside dans le fait qu'ils sont faciles à comprendre.

4.2 Un inconvénient majeur des indices basés sur des médianes simples est qu'ils produisent des estimations imprécises de la variation des prix. L'ensemble de logements effectivement échangés sur une période donnée ou un échantillon de cet ensemble est généralement réduit et pas forcément représentatif du stock total de logements. L'évolution de la composition des propriétés vendues affecte donc beaucoup plus le prix médian de l'échantillon que le prix médian du stock de logements. Prenons par exemple une ville comportant deux régions A et B, la région A ayant plus de maisons chères que la région B. Supposons que les maisons médianes vendues en 2006 et en 2008 proviennent de la région A, tandis que celle vendue en 2007 provienne de la région B. Il s'ensuivrait que l'indice de la médiane pourrait enregistrer une forte hausse de 2006 à 2007, puis une forte baisse de 2007 à 2008. Un tel indice serait un très mauvais indicateur de ce qui se produit effectivement sur le marché du logement. Ainsi, un indice de la médiane (ou de la moyenne) serait un guide très inexact de l'évolution des prix lors d'un changement considérable de la composition des maisons vendues d'une période à l'autre. S'il existait une corrélation entre les points de rupture des cycles des prix immobiliers et l'évolution de cette composition, alors une médiane pourrait être particulièrement trompeuse au cours des périodes où la prime à l'exactitude est la plus élevée.

4.3 Un problème sans doute plus important que le bruit à court terme est celui de l'erreur systématique ou du biais. Un indice de la médiane simple est biaisé quand la qualité du stock de logements évolue au fil du temps. L'indice de la médiane est biaisé à la hausse si la qualité moyenne s'améliore au fil des ans. Le biais peut également survenir si certains types de logements sont vendus plus fréquemment que d'autres et, en même temps, si leurs prix évoluent différemment. Par exemple, quand des maisons de meilleure qualité se vendent plus fréquemment et quand leurs prix augmentent plus vite que ceux des maisons de moindre qualité, cela peut créer un biais à la baisse si le nombre de ventes par type de maison ne reflète pas correctement le nombre de maisons dans le stock. C'est ce qu'on appelle parfois le problème de sélection de l'échantillon. Le fait que les maisons vendues constituent souvent une portion congrue qui n'est pas nécessairement représentative du stock total de logements peut également fausser d'autres méthodes de construction de l'indice des prix immobiliers, y compris les méthodes hédoniques et celle des ventes répétées (qui seront examinées aux chapitres 5 et 6).

Stratification

4.4 La poststratification d'un échantillon est une technique générale permettant de réduire le biais relatif à la sélection de l'échantillon. Dans le cas des indices de prix immobiliers, la stratification est l'outil le plus simple de contrôle des variations de la composition ou «du degré de qualité» des propriétés vendues. La stratification est aussi nécessaire si les utilisateurs souhaitent avoir des indices de prix pour différents segments du marché immobilier.

4.5 La stratification n'est autre que la séparation de l'échantillon total de maisons en un certain nombre de sous-échantillons ou strates. Après la construction d'une mesure de l'évolution de la tendance centrale pour chaque strate, telle qu'un indice des prix moyens ou médians, l'IPIR global stratifié est généralement calculé comme une moyenne pondérée des indices pour chaque strate. Avec M strates différentes, l'indice stratifié tel qu'il est calculé en pratique dans divers pays peut se traduire par la formule mathématique suivante :

$$P^{0t} = \sum_{m=1}^{M} w_m^0 P_m^{0t} \qquad (4.1)$$

P_m^{0t} étant l'indice pour la strate m qui compare le prix moyen (médian) de la période courante ou de comparaison t au prix moyen (médian) d'une période antérieure ou de base 0, et w_m^0 dénotant le poids de la strate m. Les poids sont des parts de valeur relatives aux strates. Ils se rapportent à la période de référence, qui est généralement d'un an (tandis que les périodes de comparaison peuvent être des mois ou des trimestres). Pour des raisons pratiques, les poids sont souvent maintenus à un niveau fixe durant plusieurs années, ce qui ne constitue généralement pas une bonne pratique. L'on trouvera ci-dessous d'autres détails sur les problèmes d'agrégation et de pondération dans ce contexte.

4.6 Les types de pondérations en valeurs utilisés dépendent de l'indice cible que l'IPIR est censé estimer. S'il s'agit de suivre l'évolution des prix du stock de logements, il

faudrait évidemment utiliser les pondérations de stocks — les parts des valeurs de stock — des strates. Si par contre la cible est un IPIR des ventes ou des acquisitions, les pondérations des ventes (des dépenses) devraient être appliquées[1].

4.7 L'efficacité de la stratification dépend des variables de stratification utilisées et l'évolution de la composition de l'échantillon n'est contrôlée qu'entre les diverses strates. Par exemple, un indice stratifié selon l'emplacement des maisons ne sera pas sensible aux changements de la composition de l'échantillon des propriétés entre les différents endroits définis. Mais cette mesure sera sensible aux modifications de l'échantillon des propriétés qui ne seront pas liées à l'emplacement. En outre, un indice stratifié ne rend pas compte des variations de la composition des propriétés vendues dans chaque strate et, en l'occurrence, de l'évolution de la composition des propriétés vendues dans les limites de chaque emplacement.

4.8 Une stratification très détaillée selon des caractéristiques des logements telles que la taille du bâtiment, la grandeur du terrain, le type d'habitation, l'emplacement et les commodités rehausse l'homogénéité et réduit ainsi le problème des changements de qualité, même s'il y a de fortes chances que certains changements persistent au niveau de la mesure de la qualité. Toutefois, il y a un compromis à envisager. En augmentant le nombre de strates, on réduit le nombre moyen d'observations par strate, et une stratification très détaillée peut entraîner une augmentation de l'écart-type de l'IPIR global. Il va sans dire que l'on ne peut concevoir un programme de stratification détaillée que si les caractéristiques définissant des strates sont disponibles pour toutes les données de l'échantillon. Un autre problème pratique réside dans la difficulté éventuelle d'obtenir des données exactes sur les pondérations (du stock) pour de petits sous-groupes.

4.9 Quand on utilise uniquement des variables de stratification physiques et géographiques comme celles susmentionnées, la méthode de stratification ne tient pas compte de l'évolution de la qualité de chaque propriété. Par évolution de la qualité, nous entendons l'effet des rénovations et réaménagements effectués en raison de la dépréciation des bâtiments. Ce phénomène peut être également appelé «dépréciation nette». Évidemment, la dépréciation dépend de l'ancienneté du bâtiment, encore que les taux d'amortissement peuvent différer selon le type d'habitation ou selon l'emplacement. C'est pour cette raison que l'ancienneté du bâtiment est considérée au chapitre 3 comme l'un des principaux attributs de la qualité qui déterminent le prix. Par conséquent, la stratification selon la catégorie «ancienneté» peut contribuer à réduire le problème de la variation de la qualité.

1 Les indices de prix des logements établis dans l'UE dans le cadre d'une étude pilote sont des exemples de tels indices des acquisitions (voir Makaronidis et Hayes, 2006, ou Eurostat, 2010).

4.10 Le fait d'adopter la catégorie «ancienneté» comme une autre variable de stratification réduit davantage le nombre moyen d'observations par strate et peut donner lieu à des estimations peu fiables des variations de prix. Dans ces circonstances, les techniques de régression hédonique — examinées au chapitre 4 — sont généralement plus efficaces que la stratification. Comme nous l'avons déjà signalé, une certaine méthode de régression hédoniste sera également nécessaire pour décomposer l'IPIR global en composantes «terrain» et «bâtiments» si cela est nécessaire pour n'importe laquelle des raisons invoquées au chapitre 2. Une telle décomposition ne saurait résulter des méthodes de stratification.

4.11 Les IPIR stratifiés sont établis par de nombreux offices statistiques et d'autres organismes publics, y compris le Ministère de l'environnement du Royaume-Uni (1982) et le Bureau australien des statistiques (ABS, 2006). Si les études théoriques[2] ne traitent guère de la stratification, des travaux de plus en plus nombreux portent sur la segmentation du marché à l'aide de techniques statistiques comme l'analyse typologique et l'analyse factorielle; voir par exemple Dale-Johnson (1982), Goodman et Thibodeau (2003), ainsi que Thibodeau (2003). En principe, ces techniques pourraient servir à la définition de sous-marchés du logement, qui pourraient constituer par la suite des strates pour la construction d'un IPIR stratifié. Le Bureau australien des statistiques a mis à l'essai cette approche (ABS, 2005).

4.12 Prasad et Richards (2006) (2008) ont proposé une nouvelle méthode de stratification et l'ont testée sur une série de données australiennes. Ils ont regroupé des banlieues selon le niveau des prix moyens à long terme des logements qui s'y trouvent, au lieu d'agglomérer simplement des régions géographiques plus petites en régions plus grandes. Leur méthode de stratification a été spécialement conçue pour prendre en compte ce qui pourrait être la plus importante forme de variation de la composition de l'échantillon, à savoir l'évolution de la proportion de maisons vendues dans des régions à prix élevés et à prix faibles sur une période donnée[3]. Il convient de noter que les auteurs ont utilisé des indices des prix de la médiane au niveau des strates. Dans cette foulée, McDonald et Smith (2009) ont construit une mesure similaire en stratifiant des prix médians de logements pour la Nouvelle-Zélande.

2 Néanmoins, plusieurs chercheurs ont utilisé des indices de prix médians stratifiés, surtout à titre de comparaison; voir par exemple Mark et Goldberg (1984), Crone et Voith (1992), Gatzlaff et Ling (1994) ainsi que Wang et Zorn (1997).

3 En règle générale, la stratification selon la variable d'intérêt ne saurait être utilisée dans la mesure où elle peut entraîner des résultats biaisés. La variable expérimentale utilisée par Prasad et Richards (2006) (2008) est l'évolution (à long terme) des prix des logements, et non le niveau des prix des logements, ce qui rend leur méthode de stratification sans doute défendable. Mais on ne sait pas grand-chose des propriétés statistiques de ce type d'indice stratifié et il serait souhaitable d'examiner le problème du biais potentiel avant de produire un tel indice.

Problèmes d'agrégation et de pondération

Première phase d'agrégation

4.13 La stratification est une procédure en deux phases : les indices de prix sont établis au niveau de la strate et sont ensuite agrégés entre les diverses strates. Comme nous l'avons signalé ci-dessus, les indices de la médiane des strates sont couramment utilisés, surtout parce qu'ils sont souvent plus stables que les indices de la moyenne correspondants. Toutefois, nous nous concentrerons davantage sur les moyennes que sur les médianes. La théorie classique des indices traite des problèmes d'agrégation et, en l'occurrence, de l'agrégation des observations des prix des logements à l'intérieur des strates. Contrairement à la médiane, les moyennes sont des fonctions d'agrégation, qui établissent le lien avec la théorie des indices. D'où la question suivante : quel type de moyenne faut-il prendre?

4.14 Le Manuel de l'IPC (2004) formule des recommandations sur la manière de construire des indices de prix à la première phase de l'agrégation si les informations relatives aux quantités ne sont pas disponibles, puis à la seconde phase de l'agrégation quand on dispose d'informations tant sur le prix que sur la valeur (ou la quantité). S'agissant de la première phase, le chapitre 20 du Manuel de l'IPC recommande généralement d'utiliser la moyenne géométrique non pondérée ou l'indice de Jevons pour agréger les différents prix relevés dans un indice. Mais cette recommandation d'ordre général n'est pas applicable dans le contexte actuel.

4.15 Si le but est de construire un indice des prix de vente des propriétés résidentielles, le concept approprié pour exprimer le prix (élémentaire) en une période t pour une strate ou strate homogène du dispositif de stratification est celui de la *valeur unitaire*. Parce que chaque vente de propriété résidentielle est assortie de sa propre quantité, qui est égale à l'unité, la quantité correspondante pour cette strate est la *somme* simple des propriétés échangées durant la période t. Nous pouvons décrire formellement ce phénomène de la manière suivante. Supposons que, dans la période t, $N(t,m)$ ventes de propriétés sont observées dans une strate m, le prix de vente (la valeur) de la propriété n étant égal à V_n^t pour $n = 1,...,N(t,m)$. Ensuite, le prix et la quantité appropriés pour la strate m durant la période t sont :

$$P_m^t \equiv \sum_{n=1}^{N(t,m)} V_n^t / N(t,m) \tag{4.2}$$

$$Q_m^t \circ N(t,m) \tag{4.3}$$

Cette définition étroite du concept de la valeur unitaire est effectivement recommandée dans le Manuel de l'IPC (2004; 356). Si le dispositif de stratification mène à des strates qui ne sont pas assez étroitement définies, il peut évidemment en résulter un biais de la valeur unitaire, ce qui revient à dire qu'un certain biais de changement de la qualité peut persister[4].

Seconde phase d'agrégation

4.16 La prochaine question à résoudre est la suivante : quelle formule de calcul de l'indice faudrait-il utiliser pour agréger les prix et les quantités élémentaires en un IPIR global? Le Manuel de l'IPC analyse dans le menu détail cette question du choix de formule. Plusieurs formules de calcul de l'indice sont recommandées, mais l'indice idéal de Fisher semble être un bon choix dans l'ensemble, car il peut se justifier sous plusieurs angles différents[5]. L'indice de Fisher est la moyenne géométrique des indices de Laspeyres et de Paasche.

4.17 Pour illustrer ce phénomène, supposons que $P^t \equiv [P_1^t,...,P_M^t]$ et $Q^t \equiv [Q_1^t,...,Q_M^t]$ notent les vecteurs des prix et des quantités des strates pour la période t. L'indice des prix de Laspeyres, soit P_L^{st}, qui va de la période (de référence) s à la période (de comparaison) t, peut être défini comme suit :

$$P_L^{st}(P^s,P^t,Q^s) \equiv \frac{\sum_{m=1}^{M} P_m^t Q_m^s}{\sum_{m=1}^{M} P_m^s Q_m^s} \tag{4.4}$$

Il convient de noter que l'équation (4.4) peut être réécrite sous la forme de (4.1) si $s = 0$, les indices de prix des strates étant $P_m^{0t} = P_m^t / P_m^0$ et les parts de valeurs, $w_m^0 = P_m^0 Q_m^0 / \sum_{m=1}^{M} P_m^0 Q_m^0$. L'indice des prix de Paasche, qui va de la période s à la période t, P_P^{st}, est définie comme suit :

$$P_P^{st}(P^s,P^t,Q^t) \equiv \frac{\sum_{m=1}^{M} P_m^t Q_m^t}{\sum_{m=1}^{M} P_m^s Q_m^t} \tag{4.5}$$

L'indice des prix de Fisher pour la période t par rapport à la période s, P_F^{st}, peut être défini comme la moyenne géométrique de (4.4) et (4.5) :

$$P_F^{st}(P^s,P^t,Q^s,Q^t) \tag{4.6}$$
$$\equiv [P_L^{st}(P^s,P^t,Q^s) \times P_P^{st}(P^s,P^t,Q^t)]^{1/2}$$

Rappelons que toutes les quantités exprimées dans ces trois formules sont des nombres de transactions, c'est-à-dire des nombres de prix observés. Pour calculer un indice des prix

4 En pratique, on utilise souvent la stratification brute selon la région et le type de logement. La méthode de stratification selon des fourchettes de prix proposée par Prasad et Richards (2008) pourrait contribuer à réduire le biais de la valeur unitaire. Pour un examen plus général du biais de la valeur unitaire, voir Balk (1998) (2008; 72–74), Silver (2009a) (2009b) (2010), ainsi que Diewert et von der Lippe (2010).

5 Voir dans le Manuel de l'IPC (2004; chapitres 15–18) les différentes justifications du recours à la formule de Fisher.

de Laspeyres, de Paasche, ou de Fisher, l'on a donc besoin des mêmes informations.

4.18 Les indices des prix de Laspeyres, de Paasche ou de Fisher définis par les équations (4.4), (4.5) et (4.6) sont des *indices à base fixe*. Par exemple, s'il existe 3 périodes de données sur les ventes, y compris la période de référence 0, alors la formule de Fisher (4.6) va générer la série d'indices suivante pour ces 3 périodes :

$$1; P_F^{01}(P^0,P^1,Q^0,Q^1); P_F^{02}(P^0,P^2,Q^0,Q^2) \qquad (4.7)$$

Chaînage

4.19 L'utilisation du chaînage est une solution de rechange à la méthode à base fixe. La *méthode du chaînage* consiste à utiliser les données des deux dernières périodes pour calculer un indice-chaîne d'une période à l'autre, qui sert à actualiser le niveau de l'indice à partir de la période précédente. Le chaînage permettrait par exemple de générer la série suivante de l'indice de Fisher pour les trois périodes :

$$1; P_F^{01}(P^0,P^1,Q^0,Q^1); P_F^{12}(P^1,P^2,Q^1,Q^2)P_F^{23}(P^2,P^3,Q^2,Q^3) \qquad (4.8)$$

4.20 La question suivante à examiner est de savoir si les IPIR doivent être construits avec une base fixe ou avec des indices-chaînes. Le Système de comptabilité nationale et le Manuel de l'IPC recommandent tous les deux l'utilisation des indices-chaînes, à condition que les données de prix sous-jacentes affichent des tendances relativement lisses[6]. De plus, si les données sont extrêmement variables, et surtout si elles fluctuent de façon erratique autour d'une tendance, l'utilisation d'indices à base fixe est recommandée. Les variations des prix des propriétés étant généralement assez lisses[7], il est probable que les indices-chaînes fonctionnent bien dans de nombreux cas. Toutefois, un surcroît d'expérimentation avec les données effectives s'impose pour formuler des conseils définitifs sur cette question. Par ailleurs, il peut également y avoir des variations saisonnières des prix des logements, comme le suggère l'exemple de la ville néerlandaise «A» présentée ci-dessous. En l'occurrence aussi, le recours aux indices-chaînes doit être empreint de prudence.

IPIR des stocks

4.21 L'analyse qui précède portait sur la construction d'un indice des prix pour les *ventes* de propriétés résidentielles à l'aide d'une méthode de stratification. Mais comment faudrait-il construire un IPIR pour le *stock* des propriétés résidentielles? En supposant que, pour chaque strate m, les propriétés vendues soient des sélections aléatoires (ou «représentatives») effectuées à partir du stock d'habitations définies par la strate m, on pourrait encore utiliser les prix de valeur unitaire P_m^t définis par (4.2) comme des (estimations des) prix par strate pour un IPIR des stocks. Toutefois, les quantités Q_m^t définies par (4.3) ne seraient plus appropriées; il faudrait les remplacer par le nombre (estimatif) d'unités résidentielles du type défini par la strate m qui font partie du stock de référence au moment t, soit $Q_m^{t^*}$, pour $m = 1,\ldots, M$. Avec ces pondérations quantitatives de la population, les détails résiduels de la construction de l'indice seraient les mêmes que pour l'IPIR relatif aux ventes.

4.22 Pour établir des pondérations de stocks, il faudra procéder à un recensement périodique du stock de logements avec suffisamment de détails sur les propriétés pour que l'on puisse le décomposer en strates appropriées du dispositif de stratification sur une période de référence. Si l'on dispose en temps utile d'informations sur la construction de nouvelles maisons et sur les démolitions, on pourra actualiser les données du recensement et estimer le stock de logements par strate (le $Q_m^{t^*}$), et ce, dans les meilleurs délais. On pourra construire l'IPIR relatif aux stocks à l'aide d'un indice de Fisher (indice-chaîne), à l'instar de l'IPIR relatif aux ventes. En revanche, si les données actuelles relatives aux nouvelles constructions et aux démolitions font défaut, on ne pourra construire qu'un indice de Laspeyres à base fixe en utilisant des données quantitatives provenant du dernier recensement disponible sur le logement (pour la période 0 par exemple), $Q^{0^*} = [Q_1^{0^*},\ldots,Q_M^{0^*}]$, jusqu'à ce que des informations provenant d'un nouveau recensement du logement deviennent disponibles (pour la période T par exemple). L'IPIR de Laspeyres relatif aux stocks se présente donc comme suit :

$$P_L^{0t}(P^0,P^t,Q^{0^*}) \equiv \frac{\sum_{m=1}^{M} P_m^t Q_m^{0^*}}{\sum_{m=1}^{M} P_m^0 Q_m^{0^*}} \qquad (4.9)$$

$$t = 0,\ldots,T$$

4.23 Le chapitre 3 signalait qu'à certaines fins, il est utile d'avoir un IPIR relatif aux stocks pour les logements occupés par leur propriétaire, c'est-à-dire à l'exclusion des logements loués. La construction d'un tel indice se fait de la même manière que pour un IPIR relatif à l'ensemble du stock de logements, sauf que les strates sont désormais limitées aux logements occupés par leur propriétaire. Cela est possible si le recensement périodique de l'habitat recueille des informations indiquant si l'occupant de chaque logement est propriétaire ou locataire.

4.24 Il convient de noter que la construction d'un IPIR (relatif aux stocks ou aux ventes) stratifié devient plus complexe quand des strates sont vides pour certaines périodes. À la fin du présent chapitre, où l'on présente un exemple empirique à partir des données sur les ventes de logements dans la ville néerlandaise «A», on énoncera les grandes lignes d'une démarche axée sur un modèle d'appariement que l'on peut utiliser si certaines strates sont vides.

6 Voir le SCN (2008) et le Manuel de l'IPC (2004; 349).

7 Bien que les fluctuations de prix ne soient pas erratiques dans le secteur immobilier, les quantités sont éminemment variables, surtout si le dispositif de calcul comporte un grand nombre de strates dont chacune contient un nombre limité d'observations. Les variations saisonnières sont également considérables en ce qui concerne les quantités; ainsi, les ventes de propriétés résidentielles chutent de façon spectaculaire durant les mois d'hiver.

Principaux avantages et inconvénients

4.25 Nous résumerons les principaux avantages et inconvénients de l'approche stratifiée utilisant la médiane ou la moyenne. Les avantages sont les suivants :

- Selon le choix des variables de stratification, la méthode tient compte de la variation de la composition des logements.
- La méthode est reproductible, moyennant une liste convenue de variables de stratification.
- Les indices des prix peuvent être construits pour différents types et emplacements de logements.
- La méthode est relativement facile à expliquer aux utilisateurs.

4.26 Les principaux inconvénients de la méthode stratifiée utilisant la médiane ou la moyenne sont les suivants :

- La méthode ne permet pas de bien gérer la dépréciation des logements, à moins que l'ancienneté du bâtiment ne soit une variable de stratification.
- La méthode ne permet pas de bien gérer les logements ayant fait l'objet de réparations ou rénovations profondes (à moins que les rénovations ne soient une variable de stratification).
- La méthode requiert des informations sur les caractéristiques des logements pour que les transactions de vente puissent être affectées aux strates correspondantes.
- Si le système de classification est très rudimentaire, les changements de composition affecteront les indices, de sorte qu'il pourrait y avoir un certain biais de la valeur unitaire au niveau des indices.
- Si le système de classification est très perfectionné, les indices cellulaires pourraient être soumis à une grande variabilité en raison de la petite taille des échantillons, ou des strates pourraient être vides pendant certaines périodes, créant ainsi des difficultés au niveau de l'indice.

4.27 De manière générale, on estime que la méthode de stratification peut être satisfaisante si :

- on choisit un niveau de détail approprié;
- l'ancienneté du bâtiment est l'une des variables de stratification; et
- une décomposition de l'indice en composantes bâtiment et terrain n'est pas nécessaire.

La stratification peut être interprétée comme un cas particulier de régression[8]. Le chapitre 5 examine cette technique plus générale, baptisée régression hédonique quand elle s'applique à la construction d'un indice des prix et à l'ajustement de la qualité.

8 Diewert (2003a) a démontré que les techniques de stratification ou le recours à des variables de substitution peuvent être considérés comme une technique de régression non paramétrique. Dans les études statistiques, ces techniques de partition ou de stratification sont baptisées «analyses des modèles de variance»; voir Scheffé (1959).

Un exemple d'utilisation des données néerlandaises pour la ville de «A»

4.28 Le présent chapitre se terminera par un exemple pratique de la construction d'un indice stratifié à partir des données sur les ventes de maisons individuelles pour une petite ville (d'une population d'environ 60.000 habitants) des Pays-Bas : la ville «A», pour 14 trimestres, du premier trimestre 2005 au deuxième trimestre 2008. Le même ensemble de données sera exploité aux chapitres 5, 6, 7 et 8 pour illustrer les autres méthodes de construction des indices de prix des logements ainsi que les différences numériques pouvant survenir en pratique[9].

4.29 Une unité d'habitation présente un certain nombre de *caractéristiques importantes qui en déterminent le prix* **:**

- la superficie du terrain de la propriété;
- la surface de plancher du bâtiment, c'est-à-dire la taille de la construction implantée sur le terrain qui sous-tend et entoure ce bâtiment;
- l'ancienneté du bâtiment, qui détermine (en moyenne) le degré de détérioration physique ou de dépréciation de la construction;
- l'ampleur des rénovations effectuées sur le bâtiment;
- l'emplacement du bâtiment, c'est-à-dire sa distance par rapport à des commodités telles que les centres commerciaux, les écoles, les restaurants et les lieux de travail;
- le type de bâtiment, c'est-à-dire maison individuelle, maison en rangée, immeuble d'appartements de petite ou de grande hauteur ou immeuble en copropriété;
- le type de construction utilisé pour le bâtiment;
- d'autres caractéristiques spéciales qui déterminent le prix et qui marquent la différence par rapport aux unités résidentielles «moyennes» dans le même emplacement général : piscine, climatisation, aménagement paysager, hauteur du bâtiment, vue sur la mer ou sur une rivière.

Les variables utilisées dans cette étude peuvent être décrites de la manière suivante :

- V_n^t est le prix de vente de la propriété n au trimestre t en euros;
- L_n^t est la superficie du terrain pour la vente de la propriété n au trimestre t en mètres carrés;
- S_n^t est la superficie de l'espace habitable du bâtiment pour la vente de la propriété n au trimestre t en mètres carrés;
- A_n^t est l'ancienneté approximative (en décennies) du bâtiment sur la propriété n au trimestre t.

9 Informations tirées de Diewert (2010).

4.30 À l'évidence, toutes les caractéristiques susmentionnées déterminant le prix n'ont pas été utilisées dans la présente étude. En particulier, les cinq dernières séries de caractéristiques de la propriété ont été négligées. On présume implicitement que les variations trimestrielles de la quantité de rénovations qui ont été effectuées sur les bâtiments, l'emplacement du logement, le type de bâtiment, le type de construction et toute autre caractéristique influant sur le prix des propriétés vendues durant le trimestre n'ont pas changé suffisamment pour constituer un déterminant majeur du prix moyen des propriétés vendues dès lors qu'ont été prises en compte les variations concernant la taille du terrain, la taille du bâtiment et l'ancienneté de la construction[10].

4.31 Il convient d'expliquer la détermination des valeurs de la variable ancienneté A_n^t. Les données initiales ont été codifiées de la manière suivante : si le bâtiment a été construit en 1960–70, l'observation a été assortie de la variable indiquant la décennie, soit BP = 5; 1971–80, BP = 6; 1981–90, BP = 7; 1991–2000, BP = 8; 2001–08, BP = 9. Dans cette étude, la variable ancienneté a été fixée à un niveau égal à 9 – BP. Pour un bâtiment de construction récente n au trimestre t, $A_n^t = 0$. Ainsi, la variable ancienneté donne l'ancienneté (approximative) du bâtiment en décennies.

4.32 Les maisons qui avaient plus de 50 ans au moment de la vente ont été éliminées de la série de données. Deux observations qui avaient des prix de vente anormalement faibles (36.000 et 40.000 €) ont été supprimées, de même que 28 observations dont la superficie dépassait 1.200 m². Aucun autre point aberrant n'a été supprimé de l'échantillon. Après l'apurement des données, nous nous sommes retrouvés avec 2.289 observations sur les 14 trimestres de l'échantillon, soit en moyenne 163,5 ventes d'unités d'habitation individuelles par trimestre. Le prix de vente moyen sur l'ensemble de l'échantillon était de 190.130 €, tandis que le prix médian était de 167.500 €. La taille moyenne du terrain était de 257,6 m² et la taille moyenne du bâtiment (superficie de l'espace habitable) était de 127,2 m². L'ancienneté moyenne des propriétés vendues était d'environ 18,5 ans.

4.33 La méthode de stratification pour la construction d'un indice des prix des logements est en principe très simple : pour chacune des caractéristiques importantes expliquant le prix, on divise les ventes en groupes relativement homogènes. Ainsi, dans cet exemple, les ventes ont été classées en 45 groupes ou strates, soit 3 catégories pour la superficie du terrain L, 3 catégories pour la superficie du bâtiment S, et 5 catégories pour l'ancienneté A (en décennies) du bâtiment ($3 \times 3 \times 5 = 45$ strates distinctes). Une fois les ventes trimestrielles classées dans ces 45 strates, les ventes dans chaque strate pour chaque trimestre ont été additionnées, puis divisées par le nombre d'unités vendues dans cette strate afin d'estimer le prix de la strate par la valeur unitaire P_m^t. Ensuite, ces valeurs unitaires ont été combinées avec le nombre d'unités vendues dans chaque strate, le Q_m^t, pour former les p et les q qui peuvent être insérés dans une formule symétrique de calcul de l'indice, comme les formules idéales de Laspeyres, de Paasche et de Fisher définies par (4.4)–(4.6) ci-dessus[11], produisant un indice stratifié des prix des logements de chacun de ces types. Cependant, puisqu'il n'y a que 163 observations environ pour chaque trimestre et 45 strates à remplir, chaque strate n'avait que 3 observations environ pour chaque trimestre, certaines strates étant vides pour certains trimestres. Nous y reviendrons ultérieurement.

4.34 Comment choisir les limites de taille pour les catégories L et S? Une démarche consisterait à diviser la gamme de L et de S par trois et de créer trois strates d'égale grandeur. Toutefois, cette démarche requerrait un grand nombre d'observations dans les strates du milieu. Dans la présente étude, les limites de taille ont donc été choisies de façon à ce que 50 % des observations environ appartiennent aux catégories moyennes et 25 % environ aux petites et grandes catégories. Pour la variable taille du terrain L, les limites choisies étaient de 160 m² et 300 m², et de 110 m² et 140 m² pour la variable taille du bâtiment S. Ainsi, si $L < 160$ m², l'observation s'inscrivait dans la strate consacrée aux terrains de petite taille; si $160 \text{ m}^2 \leq L < 300 \text{ m}^2$, l'observation s'inscrivait dans la strate consacrée aux terrains de taille moyenne et si $300 \text{ m}^2 \leq L$, l'observation s'inscrivait dans la strate consacrée aux terrains de grande taille. Les probabilités que les échantillons s'inscrivent dans ces trois strates L au cours des 14 trimestres étaient respectivement de 0,24; 0,51 et 0,25. De même, si $S < 110$ m², l'observation s'inscrivait dans la strate consacrée aux petits bâtiments; si $110 \text{ m}^2 \leq S < 140 \text{ m}^2$, alors l'observation s'inscrivait dans la strate consacrée aux bâtiments de taille moyenne, et si $140 \text{ m}^2 \leq S$, alors l'observation s'inscrivait dans la strate consacrée aux bâtiments de grande taille. En conséquence, les probabilités d'appartenir à ces trois strates S sur les 14 trimestres étaient respectivement de 0,21; 0,52 et 0,27.

4.35 Comme on l'a déjà indiqué, les données utilisées n'indiquaient pas l'ancienneté exacte du bâtiment, car on n'a enregistré que la décennie au cours de laquelle le bâtiment a été construit. Il n'y avait donc aucune possibilité de choisir des dates-butoirs exactes pour l'ancienneté du bâtiment. Ainsi, $A = 0$ correspond aux logements qui ont été construits dans les années 2001–08; $A = 1$ aux logements construits en

10 À l'appui de cette hypothèse, il convient de noter que les modèles de régression hédonique examinés dans les chapitres suivants ont justifié 80 % à 90 % de la variation des données relatives aux prix qu'avec trois grandes variables explicatives : L, S et A. Le R^2 entre les prix de vente effectifs et prévus variait de 0,83 à 0,89. Le fait qu'il n'ait pas été nécessaire d'introduire d'autres caractéristiques déterminant le prix pour cette série de données peut sans doute s'expliquer par l'emplacement de la ville «A» dans une plaine monotone et par sa taille relativement petite; l'emplacement n'était pas un facteur important de détermination du prix, tous les emplacements ayant plus ou moins le même accès aux commodités.

11 Les manuels internationaux sur la mesure des prix recommandent cette approche de la valeur unitaire pour la construction des indices de prix au premier stade de l'agrégation; voir le Manuel de l'IPC (2004), le Manuel de l'IPP (2004) et le Manuel XMPI (*Export and Import Price Index Manual,* 2009). Cependant, l'agrégation de la valeur unitaire doit se faire à partir d'éléments homogènes et cette hypothèse pourrait ne pas s'appliquer au contexte actuel, L, S et A étant assez variables dans chaque strate. Mais comme il n'existe déjà qu'un petit nombre d'observations dans chaque strate pour la série de données envisagée, il serait difficile d'introduire d'autres strates pour améliorer l'homogénéité, car cela entraînerait une augmentation du nombre de strates vides et un manque d'harmonie entre les strates.

1991–2000; $A = 2$ aux logements construits en 1981–90, $A = 3$ aux logements construits en 1971–80; et $A = 4$ aux logements construits en 1961–70. En conséquence, les probabilités pour l'échantillon de s'inscrire dans ces cinq strates sur les 14 trimestres étaient respectivement de 0,15; 0,32; 0,21; 0,20 et 0,13. Voir au tableau 4.1 les probabilités conjointes, au niveau de l'échantillon, que la vente d'un logement s'inscrive dans chacune des 45 strates.

4.36 Plusieurs informations présentées au tableau 4.1 méritent d'être signalées :

- Les maisons construites dans les années 60 ($A = 4$), grands bâtiments (S = large) implantées sur de petits terrains (L = petit), ne font l'objet d'aucune observation; cette strate est donc tout à fait vide.
- De nombreuses strates sont pratiquement vides; en particulier, la probabilité de vendre un grand terrain avec une petite maison est très faible, de même que la probabilité de vendre un petit terrain avec une grande maison[12].
- Le «modèle le plus représentatif» vendu sur la période d'échantillonnage correspond à un terrain de taille moyenne, un bâtiment de taille moyenne et une maison qui a été construite dans les années 90 (A =1). Dans le cadre de l'échantillon, la probabilité que la vente d'une maison s'inscrive dans cette strate la plus forte est de 0,09262.

4.37 Le prix de vente moyen de la maison représentative, qui s'inscrit dans la moyenne de L, la moyenne de S et la catégorie A =1, est représenté au graphique 4.1, de même que l'échantillon global du prix moyen et médian de chaque trimestre. Ces prix moyens ont été convertis en indices qui commencent à 1 pour le trimestre 1, le premier trimestre de 2005. Il convient de signaler que ces trois indices des prix du logement sont assez variables.

12 Ainsi, la taille du terrain et la taille du bâtiment ont une corrélation positive avec un coefficient de corrélation de 0,6459. *L* et *S* ont une assez forte corrélation avec la variable *P* du prix de vente : la corrélation entre *P* et *L* est de 0,8234 et entre *P* et *S*, de 0,8100. Ces fortes corrélations posent des problèmes de multicolinéarité dans les modèles de régression hédonique qui seront examinés ultérieurement.

4.38 Certains autres indices sont représentés au graphique 4.1, y compris un indice de Fisher à base fixe et à produits appariés et un indice-chaîne de Fisher à produits appariés. Il est nécessaire d'expliquer la signification de «produits appariés» dans ce contexte. Si au moins une maison était vendue chaque trimestre pour chacune des 45 strates, les indices ordinaires de Laspeyres, de Paasche et de Fisher comparant les prix du trimestre t à ceux du trimestre s seraient définis par les équations (4.4)–(4.6) respectivement, où M = 45. Ce principe d'algèbre est applicable à la situation où il existe des transactions dans toutes les strates pour les deux trimestres qui sont comparés. Mais pour la présente série de données, il n'existe en moyenne que 30 strates environ sur 45 qui puissent être appariées entre deux trimestres, les formules (4.4)–(4.6) devant être modifiées afin de résoudre ce *problème de discordance*. Ainsi, en envisageant les modalités de comparaison d'un indice entre les trimestres t et s, il faudrait définir la série de strates m qui ont au moins une transaction pour chacun des trimestres t et s comme la série $S(s,t)$. Ensuite, les *contreparties de la méthode des produits appariés*, P_{ML}^{st}, P_{MP}^{st} et P_{MF}^{st}, à la méthode des indices courants de Laspeyres, de Paasche et de Fisher entre les trimestres s et t donnés par (4.4), (4.5) et (4.6) sont définies de la manière suivante[13] :

$$P_{ML}^{st} \equiv \frac{\sum_{m \in S(s,t)} P_m^t Q_m^s}{\sum_{m \in S(s,t)} P_m^s Q_m^s} \tag{4.10}$$

$$P_{MP}^{st} \equiv \frac{\sum_{m \in S(s,t)} P_m^t Q_m^t}{\sum_{m \in S(s,t)} P_m^s Q_m^t} \tag{4.11}$$

$$P_{MF}^{st} \equiv [P_{ML}^{st} P_{MP}^{st}]^{1/2} \tag{4.12}$$

13 On trouvera une justification de cette démarche de résolution d'une discordance dans le contexte de la théorie des indices bilatéraux, dans l'analyse de Diewert (1980; 498–501) sur le problème connexe du traitement des biens nouveaux et en voie de disparition. D'autres démarches sont également possibles. Pour celles fondées sur une concordance maximum sur toutes les paires de période, voir Ivancic, Diewert et Fox (2011) ainsi que de Haan et van der Grient (2011) en ce qui concerne les démarches fondées sur les méthodes d'imputation; voir Alterman, Diewert and Feenstra (1999). Une méthode d'imputation utile pourrait consister à imputer des prix pour les strates vides en utilisant des régressions hédoniques. Nous y reviendrons après l'examen de diverses méthodes de régression hédonique.

Tableau 4.1. Probabilité de vente dans chaque strate pour l'échantillon

L	*S*	*A* = 0	*A* = 1	*A* = 2	*A* = 3	*A* = 4
Petit	Petit	0,00437	0,02665	0,01660	0,02053	0,02097
Moyen	Petit	0,00349	0,02840	0,01966	0,01092	0,03888
Large	Petit	0,00087	0,00175	0,00044	0,00218	0,00612
Petit	Moyen	0,01223	0,05242	0,04281	0,02053	0,00699
Moyen	Moyen	0,03277	0,09262	0,08869	0,07907	0,02141
Large	Moyen	0,00786	0,02315	0,01005	0,01442	0,01398
Petit	Large	0,00306	0,00218	0,00175	0,00568	0,00000
Moyen	Large	0,03145	0,03495	0,00786	0,02097	0,00306
Large	Large	0,04893	0,05461	0,02315	0,02490	0,01660

Source : calculs des auteurs basés sur les données du cadastre néerlandais.

Au graphique 4.1, l'indice de Fisher à base fixe est l'indice des prix de Fisher à produits appariés défini par (4.12), où la période de référence *s* est maintenue fixe au trimestre 1; autrement dit, les indices $P_{MF}^{1,1}$, $P_{MF}^{1,2}$,…, $P_{MF}^{1,14}$ sont calculés et libellés indice de Fisher à base fixe, P_{FFB}. L'indice qui est libellé indice-chaîne de Fisher à produits appariés, P_{FCH}, est l'indice des prix $P_{MF}^{1,1}$, $P_{MF}^{1,1}$ $P_{MF}^{1,2}$, $P_{MF}^{1,1}$ $P_{MF}^{1,2}$ $P_{MF}^{2,3}$, …, $P_{MF}^{1,1}$ $P_{MF}^{1,2}$ … $P_{MF}^{13,14}$. Il convient de noter que les indices-chaîne et à base fixe de Fisher (produits appariés) sont assez proches l'un de l'autre et sont beaucoup plus lisses que les indices fondés sur la moyenne, la médiane et un produit représentatif[14]. La liste des données pour ces 5 séries est représentée au graphique 4.1 et au tableau 4.2.

4.39 Les indices de Fisher à produits appariés doivent être considérés comme étant plus exacts que les autres indices qui n'utilisent qu'un volume limité d'informations disponibles sur les prix et les quantités. La tendance des indices de Fisher étant assez lisse, l'indice-chaîne de Fisher doit être privilégié par rapport à l'indice de Fisher à base fixe, d'après les recommandations de Hill (1988) (1993) et du Manuel de l'IPC (2004). Il convient de rappeler qu'en l'occurrence, il n'est pas nécessaire d'utiliser les indices de Laspeyres ou de Paasche, car les données sur les ventes de logements contiennent à la fois des informations sur les valeurs et les quantités. Dans ces conditions, les indices de Fisher sont préférables à ceux de Laspeyres et de Paasche (qui n'utilisent pas toutes les informations disponibles sur les prix et les quantités pour les deux périodes comparées).

4.40 Compte tenu de l'hétérogénéité considérable dans chaque strate, il est fort possible qu'il existe un certain biais de la valeur unitaire dans les indices de Fisher à produits appariés. Mais si l'on utilisait une stratification plus fine, le degré d'appariement baisserait de façon spectaculaire. Déjà, avec la stratification actuelle, près des deux tiers seulement des strates pourraient être appariées sur deux trimestres. Il y a un arbitrage à faire entre avoir trop peu de strates avec la possibilité d'un biais de la valeur unitaire et avoir un dispositif de stratification plus détaillé, mais avec beaucoup moins d'appariement des données dans les strates sur les deux périodes comparées.

4.41 En regardant le tableau 4.2 et le graphique 4.1, on constate que l'indice-chaîne de Fisher fait état d'une baisse des prix des logements durant les quatre trimestres de 2005, 2006 et 2007. Il est possible que les prix des logements baissent pour des raisons saisonnières au quatrième trimestre d'une année. Pour parer à cette éventualité, on construira dans la prochaine section un indice de Fisher à produits appariés à glissement annuel.

14 Les moyennes (et les écarts-types) des 5 séries mentionnées jusqu'ici sont les suivantes : P_{FCH} = 1,0737 (0,0375), P_{FFB} = 1,0737 (0,0370), P_{Moy} = 1,0785 (0,0454), $P_{Médiane}$ = 1,0785 (0,0510), et P_{Rep} = 1,0586 (0,0366). Ainsi, l'indice des prix à modèles représentatifs a une moindre variance que les deux indices de Fisher à produits appariés, mais il a un biais substantiel par rapport aux deux indices de Fisher à produits appariés : l'indice des prix à modèles représentatifs est nettement inférieur aux indices de Fisher pour la plus grande partie de la période d'échantillonnage.

Traitement de la saisonnalité pour l'exemple néerlandais

4.42 La méthode la plus simple et la plus satisfaisante sur le plan théorique en ce qui concerne le traitement des biens saisonniers consiste à supposer que chaque produit de base à chaque saison de l'année est un produit «annuel» distinct quand l'objectif est de construire des indices annuels des prix et des quantités. Cette idée remonte à Mudgett pour les prix à la consommation et à Stone pour les prix à la production :

«L'indice de base est un indice annuel et, en tant qu'indice des prix ou des quantités, il appartient à la même catégorie que ceux sur lesquels des livres et des brochures ont été publiés en quantité au fil des ans.» Bruce D. Mudgett (1955; 97).

«L'existence d'une saisonnalité régulière des prix qui se répète plus ou moins d'une année à l'autre suggère très fortement que les variétés d'un produit disponible à des saisons différentes ne peuvent pas être transformées les unes en les autres sans coût et, par conséquent, dans tous les cas où les variations saisonnières des prix sont significatives, les variétés disponibles à différents moments de l'année devraient être traitées en principe comme des produits distincts.» Richard Stone (1956; 74–75).

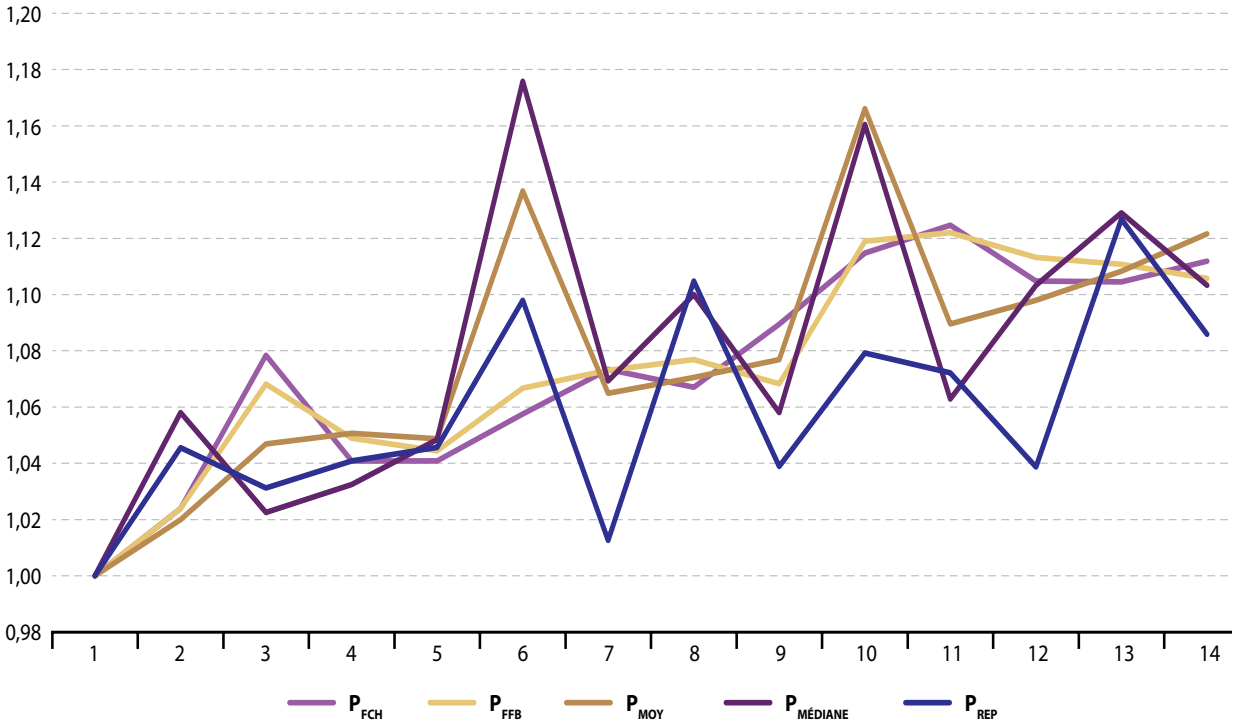
Diewert (1983) a généralisé le cadre annuel de Mudgett–Stone pour permettre des *comparaisons annuelles glissantes* de données pour 12 mois consécutifs avec une année de base de 12 mois de données ou pour des comparaisons de quatre trimestres consécutifs de données, avec une année de base de quatre trimestres consécutifs. Autrement dit, l'idée fondatrice est de comparer les prix et les quantités de l'année glissante en cours avec les prix et les quantités de la même saison d'une année de base[15]. Dans le contexte actuel, nous avons en principe[16] des données sur les prix et les quantités, chaque trimestre, pour 45 catégories de logement. Si la vente d'un logement pour chaque saison est traitée comme un produit distinct, alors il en existera 180 pour l'année.

4.43 Pour la première valeur de l'indice, les quatre trimestres de données sur les prix et les quantités pour les ventes de logements individuels dans la ville «A» (180 séries) sont comparés aux mêmes données à l'aide de la formule idéale de Fisher. Naturellement, l'indice qui en résulte est égal à 1. Pour la valeur suivante de l'indice, les données pour

15 On trouvera des considérations théoriques et des exemples supplémentaires sur cette méthode de l'année glissante, dans les chapitres du Manuel de l'IPC (2004) et du Manuel de l'IPP (2004) consacrés à la saisonnalité et dans Diewert (1998) et Balk (2008; 151–169). Pour justifier les indices en glissement annuel dans la perspective économique de la théorie des indices, certaines restrictions sur les préférences sont nécessaires; voir les détails dans Diewert (1999; 56–61). Il convient de noter que les intempéries et la non-fixité du jour de Pâques peuvent rendre les «saisons» variables et la méthode inefficace; voir Diewert, Finkel et Artsev (2009). Cependant, avec des données trimestrielles, ces limites de l'indice en glissement annuel sont moins importantes.

16 En pratique, comme nous l'avons vu dans la section précédente, beaucoup de strates sont vides dans chaque période.

Graphique 4.1. Indices-chaînes de Fisher à produits appariés et à base fixe, et indices des prix moyens, médians et à modèles représentatifs



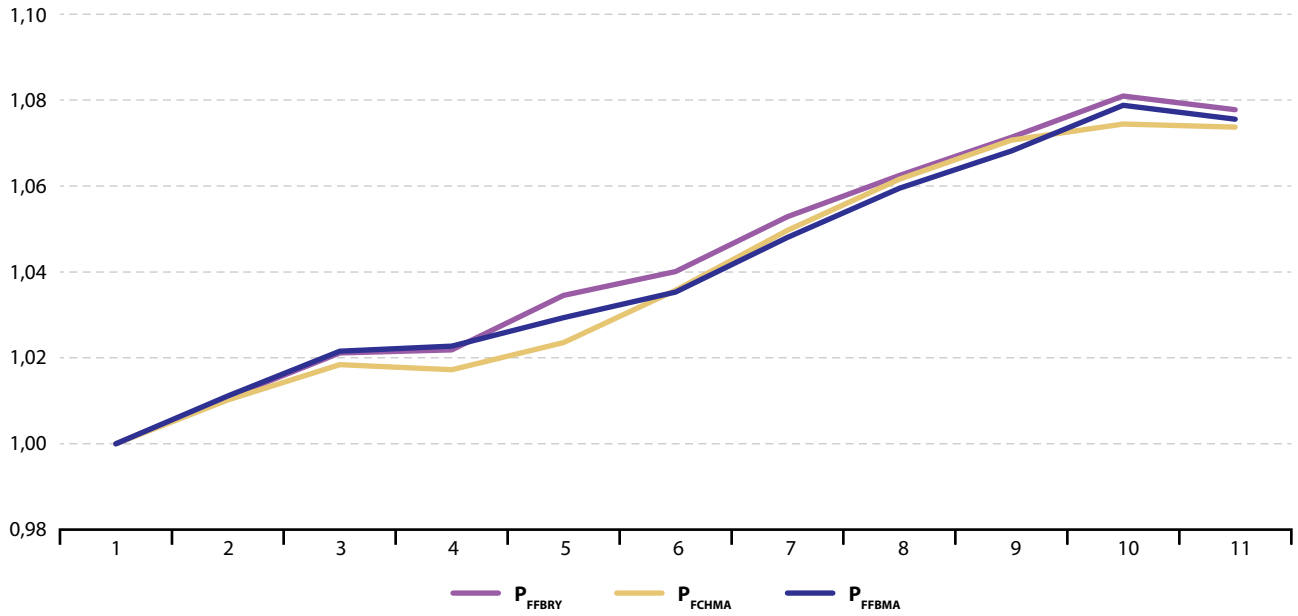
Source : calculs des auteurs basés sur les données du cadastre néerlandais.

Tableau 4.2. Indices-chaînes de Fisher à produits appariés et à base fixe, et indices des prix moyens, médians et à modèles représentatifs

Trimestre	P_{FCH}	P_{FFB}	P_{Moy}	$P_{Médiane}$	P_{Rep}
1	1,00000	1,00000	1,00000	1,00000	1,00000
2	1,02396	1,02396	1,02003	1,05806	1,04556
3	1,07840	1,06815	1,04693	1,02258	1,03119
4	1,04081	1,04899	1,05067	1,03242	1,04083
5	1,04083	1,04444	1,04878	1,04839	1,04564
6	1,05754	1,06676	1,13679	1,17581	1,09792
7	1,07340	1,07310	1,06490	1,06935	1,01259
8	1,06706	1,07684	1,07056	1,10000	1,10481
9	1,08950	1,06828	1,07685	1,05806	1,03887
10	1,11476	1,11891	1,16612	1,16048	1,07922
11	1,12471	1,12196	1,08952	1,06290	1,07217
12	1,10483	1,11321	1,09792	1,10323	1,03870
13	1,10450	1,11074	1,10824	1,12903	1,12684
14	1,11189	1,10577	1,12160	1,10323	1,08587

Source : calculs des auteurs basés sur les données du cadastre néerlandais.

Graphique 4.2. Indice de Fisher à base fixe pour l'année glissante, indice-chaîne de Fisher à moyenne mobile, et indice de Fisher des prix à base fixe et à moyenne mobile



Source : calculs des auteurs basés sur les données du cadastre néerlandais.

le premier trimestre de 2005 sont écartées et celles du premier trimestre de 2006 sont jointes aux données des trimestres 2–4 de 2005. L'indice de Fisher qui en résulte est la deuxième courbe illustrée au graphique 4.2, celle de la série à produits appariés pour l'année glissante (RY). Mais, comme on l'a vu pour les indices-chaînes de Fisher et à base fixe présentés au graphique 4.1, toutes les strates n'ont pas pu être appariées avec la méthode de l'année glissante; en d'autres termes, certaines strates étaient vides au premier trimestre de 2006 et correspondaient à des strates du premier trimestre de 2005 qui n'étaient pas vides, et vice versa. Ainsi, en construisant l'indice de l'année glissante P_{RY} présenté au graphique 4.2, la comparaison entre l'année glissante et les données relatives à 2005 a été limitée au groupe de strates qui n'étaient pas vides pour les deux années; autrement dit, les indices de Fisher pour l'année glissante présentés au graphique 4.2 sont des indices à produits appariés. Les modèles non appariés sont omis de la comparaison des indices[17].

4.44 Les résultats sont présentés au graphique 4.2. Signalons qu'il y a un net fléchissement à la fin de la période d'échantillonnage, mais que les baisses figurant au graphique 4.1 pour les trimestres 4 et 8 peuvent être considérées comme saisonnières; cela signifie que les indices de l'année glissante au graphique 4.2 n'auront (définitivement) baissé qu'à la fin de la période d'échantillonnage. Signalons aussi que la valeur de l'indice pour l'observation 5 compare les données de l'année civile 2006 aux données correspondantes de l'année civile 2005 et la valeur de l'indice pour l'observation 9 compare les données de l'année civile 2007 aux données correspondantes de l'année civile 2005; en d'autres termes, ces valeurs indiciaires correspondent aux indices annuels de Mudgett–Stone.

4.45 La construction d'indices de Fisher à produits appariés pour l'année glissante est plutôt un travail à forte intensité de main-d'œuvre, car les strates qui sont appariées sur deux périodes quelconques varient selon les périodes. Une méthode raccourcie (et moins précise) de dessaisonnalisation d'une série, telle que l'indice-chaîne de Fisher à produits appariés P_{FCH} et l'indice de Fisher à base fixe P_{FFB} présentés au tableau 4.2, consiste simplement à prendre une *moyenne mobile* de ces séries sur 4 trimestres. La série qui en résulte pour l'année glissante, P_{FCHMA} et P_{FFBMA}, peut être comparée à la série de Mudgett–Stone–Diewert P_{RY} pour l'année glissante; voir le graphique 4.2. Les données correspondant au graphique 4.2 sont présentées au tableau 4.3.

4.46 On peut voir qu'une moyenne mobile des indices-chaînes de Fisher et à base fixe d'un trimestre à l'autre, P_{FCH} et P_{FFB}, présentés au tableau 4.2, rend assez bien compte de l'indice de Fisher P_{FFBRY} à base fixe pour l'année glissante.

17 Onze comparaisons d'années glissantes sont possibles à partir des données des 14 trimestres qui sont disponibles. Les nombres de strates non appariées ou vides pour les années glissantes 2, 3, …, 11 sont les suivants : 50, 52, 55, 59, 60, 61, 65, 65, 66, 67. Le nombre relativement faible de strates non appariées ou vides pour les années glissantes 2, 3 et 4 s'explique par le fait que ¾ des données sont appariées pour l'année glissante 2, la moitié des données pour l'année glissante 3, et ¼ des données pour l'année glissante 4.

Tableau 4.3. Indice de Fisher à base fixe pour l'année glissante, indice-chaîne de Fisher à moyenne mobile, et indice de Fisher des prix à base fixe et à moyenne mobile

Année glissante	P_{FFBRY}	P_{FCHMA}	P_{FFBMA}
1	1,00000	1,00000	1,00000
2	1,01078	1,01021	1,01111
3	1,02111	1,01841	1,02156
4	1,02185	1,01725	1,02272
5	1,03453	1,02355	1,02936
6	1,04008	1,03572	1,03532
7	1,05287	1,04969	1,04805
8	1,06245	1,06159	1,05948
9	1,07135	1,07066	1,06815
10	1,08092	1,07441	1,07877
11	1,07774	1,07371	1,07556

Source : calculs des auteurs basés sur les données du cadastre néerlandais.

Mais il existe des différences pouvant atteindre 1 % entre l'indice préféré pour l'année glissante et l'indice à moyenne mobile. Il convient de rappeler que l'indice de Fisher à base fixe comparait les données des trimestres 1 à 14 aux données correspondantes du trimestre 1. Ainsi, les observations relatives aux trimestres 2 et 1, 3 et 1, et 4 et 1 n'ont pas autant de chances d'être comparables que les indices sur année glissante où les données de n'importe quel trimestre sont toujours comparées à celles du trimestre correspondant de l'année de référence. On peut en dire autant de l'indice de la moyenne mobile P_{FCHMA}; les comparaisons qui entrent dans la confection des liens dans cet indice proviennent des données trimestrielles et ne sont probablement pas aussi exactes que les comparaisons annuelles pour le même trimestre[18].

18 Plus la saisonnalité est forte, plus fort sera cet argument en faveur de l'exactitude de l'indice en glissement annuel. On peut voir la force de cet argument si tous les prix de vente de logements pour chaque strate s'avèrent éminemment saisonniers; en d'autres termes, les ventes pour chaque strate donnée ne surviennent qu'au cours d'un trimestre chaque année. Les comparaisons trimestrielles sont évidemment impossibles dans cette situation, mais les indices en glissement annuel seront parfaitement bien définis.

5 Méthodes de régression hédonique

Modélisation et estimation hédoniques

5.1 La méthode de régression hédonique reconnaît que des biens hétérogènes peuvent être décrits par leurs attributs ou leurs caractéristiques, c'est-à-dire qu'un bien est essentiellement un ensemble de caractéristiques (de performance)[1]. Dans le contexte du logement, cet ensemble peut inclure des attributs concernant aussi bien la structure des bâtiments que l'emplacement des biens immobiliers. Il n'existe pas de marché pour les caractéristiques puisque celles-ci ne peuvent être vendues séparément et, de ce fait, leurs prix ne sont pas observés indépendamment. La demande et l'offre des biens déterminent implicitement les contributions marginales de leurs caractéristiques à leur prix. Ces contributions marginales ou prix virtuels peuvent être estimés à l'aide de techniques de régression. La méthode hédonique peut notamment permettre d'estimer la disposition à payer pour les différentes caractéristiques ou le coût marginal de la production de celles-ci. Nous nous intéressons plus particulièrement ici au deuxième grand objectif du manuel, à savoir la construction d'indices de prix ajustés pour tenir compte de la qualité.

Modélisation hédonique

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

$$p_n^t = f(z_{n1}^t,...,z_{nK}^t,\varepsilon_n^t) \qquad (5.1)$$
$$t = 0,...,T$$

où ε_n^t est un terme d'erreur aléatoire (bruit blanc). Pour pouvoir estimer les contributions marginales des caractéristiques à l'aide des techniques de régression courantes, l'équation (5.1) doit être présentée sous forme d'une équation paramétrique. Les deux spécifications hédoniques les plus connues sont le modèle linéaire

$$p_n^t = \beta_0^t + \sum_{k=1}^{K} \beta_k^t z_{nk}^t + \varepsilon_n^t \qquad (5.2)$$

et le modèle log-linéaire

$$\ln p_n^t = \beta_0^t + \sum_{k=1}^{K} \beta_k^t z_{nk}^t + \varepsilon_n^t \qquad (5.3)$$

où β_0^t et β_k^t sont la constante et les paramètres des caractéristiques à estimer. Dans les deux spécifications, les caractéristiques peuvent être des transformations, logarithmiques par exemple, de variables continues. Dans la pratique, les variables explicatives seront souvent plutôt catégorielles que continues et représentées par une série de variables indicatrices qui sont égales à 1, lorsque le bien appartient à la catégorie considérée, et égales à 0 dans le cas contraire.

5.3 Pour les produits comme les biens de haute technologie, le modèle log-linéaire (5.3) est utilisé de préférence, car c'est celui qui a le plus de chances de réduire le problème d'hétéroscédasticité (variabilité de la variance des erreurs), les prix ayant tendance à obéir à une distribution log-normale (Diewert, 2003b). Dans le contexte du logement, par contre, le modèle linéaire présente de nombreux avantages. On a vu au chapitre 3 que la taille du bâtiment et la superficie du terrain sur lequel il repose constituent deux variables importantes pour la détermination du prix. La valeur d'un bien immobilier correspondant généralement à la *somme* du prix du bâtiment et du prix du terrain, on peut estimer souhaitable d'inclure de façon linéaire les terrains et les bâtiments dans le modèle si l'on dispose des données nécessaires. Cette question sera examinée de façon plus approfondie au chapitre 8, notamment en décomposant l'indice hédonique en ses composantes «terrain» et «bâtiments». Toutes les sources de données ne contiennent malheureusement pas d'informations sur la taille des terrains et des bâtiments. Celles concernant la taille des terrains en particulier risquent de faire défaut. Quand la taille du terrain (ou du bâtiment) n'est pas incluse comme variable explicative, le recours à des modèles log-linéaires a été jugé relativement satisfaisant dans de nombreuses études empiriques.

5.4 Les paramètres des caractéristiques β_k^t dans les équations (5.2) et (5.3) peuvent varier dans le temps. Cela cadre avec l'idée selon laquelle les conditions du marché du logement déterminent les contributions marginales des caractéristiques : il n'y a en effet aucune raison de penser *a priori* que ces contributions sont constantes quand les conditions de l'offre et de la demande changent (Pakes, 2003). Les conditions du marché ont cependant de très fortes chances de changer de façon progressive. On peut donc formuler sans hésiter, au moins peut-être pour le court terme, l'hypothèse simplificatrice de la stabilité dans le temps des paramètres des caractéristiques (mais non de la constante). Dans le cas du modèle log-linéaire, cela conduirait à la version avec contrainte suivante de l'équation (5.3) :

$$\ln p_n^t = \beta_0^t + \sum_{k=1}^{K} \beta_k z_{nk}^t + \varepsilon_n^t \qquad (5.4)$$

Comme on le verra plus loin, les constantes dépendantes du temps (β_0^t) peuvent être converties en un indice des prix à qualité constante.

5.5 Supposons que nous disposons de données sur les prix de vente et les caractéristiques pour les échantillons $S(0),S(1),...,S(T)$ de biens immobiliers de taille $N(0),N(1),...,N(T)$ vendus pendant les périodes $t = 0,...,T$. Dans les hypothèses d'erreur classiques, en particulier celles de moyenne nulle et de variance constante, on peut estimer les paramètres des modèles hédoniques (5.2) et (5.3) à l'aide

1 La méthode de régression hédonique remonte au moins à Court (1939) et Griliches (1961). Lancaster (1966) et Rosen (1974) en ont établi les fondements conceptuels. Colwell et Dilmore (1999) estiment que la première étude hédonique jamais publiée est une thèse de maîtrise de l'université du Minnesota sur la valeur des terres agricoles, parue en 1922.

d'une régression effectuée selon la méthode des moindres carrés ordinaires (MCO) sur les données des échantillons de chaque période, séparément. La version avec contrainte (5.4) peut être estimée à partir des données regroupées des différentes périodes à condition d'inclure des variables indicatrices des périodes temporelles (moins une pour éviter une colinéarité parfaite). L'équation d'estimation pour le modèle log-linéaire avec contrainte (5.4), généralement appelé *modèle hédonique à variables indicatrices*, devient ainsi

$$\ln p_n^t = \beta_0 + \sum_{\tau=1}^{T} \delta^\tau D_n^\tau + \sum_{k=1}^{K} \beta_k z_{nk}^t + \varepsilon_n^t \qquad (5.5)$$

où la variable indicatrice temporelle D_n^τ prend la valeur 1 si l'observation correspond à la période τ et la valeur 0 dans le cas contraire; une variable indicatrice temporelle pour la période de référence 0 est omise. Il est aussi possible, même si cela n'est que rarement pratiqué, de définir un modèle à variables indicatrices temporelles en utilisant le prix non transformé comme variable dépendante. Cette spécification de modèle sera examinée dans l'exemple empirique donné à la fin du présent chapitre.

Questions d'ordre pratique

5.6 Une question importante est celle du choix de la série de variables explicatives à inclure dans l'équation hédonique. Si des variables pertinentes — les caractéristiques qui semblent devoir affecter le prix d'un bien immobilier (énumérées au chapitre 3) — sont omises, les paramètres estimés des caractéristiques incluses seront, de ce fait, affectés d'un biais. Celui-ci se répercute sur les prix prédits, calculés à partir des coefficients de régression, ainsi que sur les indices hédoniques. Chaque bien immobilier peut être considéré comme un bien unique du fait principalement de son emplacement. Il peut toutefois être difficile d'obtenir des informations précises sur l'emplacement des biens et leur voisinage (Case, Pollakowski et Wachter, 1991). Des informations sur d'autres caractéristiques des biens peuvent également faire défaut et celles-ci pourraient parfois être difficiles à mesurer directement. On doit donc reconnaître qu'en pratique, toute estimation d'un modèle hédonique pour le logement sera affectée d'un biais dû à l'omission de variables[2]. Le signe et l'ampleur du biais ainsi que son effet sur l'indice des prix sont difficiles à prédire. L'ampleur du biais dépend notamment de la corrélation entre les variables omises et incluses.

5.7 L'importance de l'emplacement a conduit les chercheurs à utiliser les coordonnées géographiques, en longitudes et latitudes, des biens immobiliers dans les régressions hédoniques. Pour cela, une matrice des distances entre tous les biens couverts par l'ensemble de données est généralement construite, puis il est fait appel à des méthodes économétriques appropriées (bien qu'assez spécialisées) pour tenir compte de la dépendance spatiale dans l'équation estimée. La prise en compte explicite de la dépendance spatiale peut contribuer à résoudre le problème que pose l'omission des variables territoriales. La dépendance spatiale peut être prise en compte dans les variables explicatives ou le résidu. La première méthode, c'est-à-dire celle consistant à inclure l'emplacement comme une variable explicative en utilisant les données géospatiales, est la plus simple. Cela peut être effectué de façon paramétrique ou non paramétrique en recourant par exemple à des splines, comme l'ont montré Hill, Melser et Reid (2010). Nous renvoyons le lecteur à Hill (2011) pour une analyse poussée et un inventaire des travaux publiés sur la dépendance spatiale, l'utilisation des données géospatiales ainsi que sur l'estimation non paramétrique[3].

5.8 Le problème de la multicolinéarité est un problème bien connu des régressions hédoniques. Une forte corrélation entre certaines des variables incluses accroît l'écart-type des coefficients des régressions; les coefficients deviennent instables. Là encore, il est difficile de prédire *a priori* l'effet que cela aura sur les indices hédoniques. La multicolinéarité peut ne pas poser trop de problèmes dans certains cas. Par exemple, si l'on est moins intéressé par les valeurs des paramètres que par les prévisions de prix à utiliser dans l'estimation de l'indice global des prix des logements ajusté pour tenir compte de la qualité, le problème de la multicolinéarité ne devrait pas être trop aigu. Il est préférable, dans ce cas, d'inclure une variable pertinente, au risque de créer un problème de multicolinéarité, son exclusion conduisant à un biais d'omission de variables. Mais quand les valeurs des paramètres sont intéressantes en tant que telles, si l'on essaie, par exemple, de décomposer les prix des biens immobiliers en leurs composantes «terrain» et «bâtiments», la multicolinéarité pose alors problème. On en aura confirmation au chapitre 8.

5.9 Un apurement des données peut s'avérer nécessaire, comme lorsque d'autres méthodes sont utilisées. Les erreurs de saisie évidentes doivent être supprimées, mais une certaine prudence s'impose. L'élimination des données aberrantes d'une régression en vue d'obtenir des coefficients plus stables (et donc des indices de prix plus stables) est souvent arbitraire et susceptible de conduire à des estimations biaisées. Pour utiliser les techniques hédoniques, il faut disposer de données sur toutes les caractéristiques incluses dans le modèle. Or beaucoup de séries de données contiennent malheureusement des non-réponses partielles, c'est-à-dire que les informations concernant une ou plusieurs caractéristiques peuvent faire défaut pour une partie de l'échantillon. Des procédures d'imputation des données manquantes ont été mises au point, mais là encore il est important d'éviter de faire des choix arbitraires qui peuvent affecter les résultats.

2 Il convient de mentionner aussi à cet égard que les informations relatives aux caractéristiques de chaque logement inclus dans l'échantillon devraient être disponibles en temps réel. Les caractéristiques des logements peuvent changer dans le temps (c'est en fait pourquoi elles sont affectées d'un exposant pour le temps t dans les modèles hédoniques qui précèdent). Le fait de maintenir fixes les caractéristiques implique que l'indice des prix hédonique n'est pas ajusté pour tenir compte de ces changements qualitatifs.

3 Colwell (1998) a proposé une méthode d'interpolation spatiale non paramétrique qui semble bien adaptée pour modéliser les prix des terrains comme une fonction des coordonnées géographiques bidimensionnelles des biens immobiliers.

5.10 Nous allons maintenant examiner, dans les deux prochaines sections, les deux principales méthodes hédoniques utilisables pour construire des indices des prix des logements, ajustés pour tenir compte de la qualité, à savoir la méthode des variables indicatrices temporelles et celle des imputations. Sans nier les problèmes économétriques potentiels, nous nous intéresserons plus particulièrement à l'utilisation des régressions par la méthode des moindres carrés pour estimer les modèles.

Méthode des variables indicatrices temporelles

5.11 La méthode des variables indicatrices temporelles a été souvent utilisée dans les travaux universitaires, mais moins par les organismes statistiques, pour construire un indice hédonique des prix des logements[4]. L'un de ses avantages est sa simplicité; l'indice des prix découle directement de l'équation estimée de la régression sur les variables indicatrices temporelles regroupées (5.5). En effectuant une régression globale sur les données regroupées des échantillons $S(0), S(1),..., S(T)$ concernant les périodes $t = 0,...,T$ (avec les superficies $N(0), N(1),..., N(T)$), on obtient les coefficients $\hat{\beta}^0$, $\hat{\delta}^t$ $(t = 1,...,T)$ et $\hat{\beta}_k$ $(k = 1,...,K)$. Le paramètre de la variable indicatrice temporelle fait monter ou descendre la superficie hédonique et mesure l'effet du «temps» sur le logarithme du prix. L'exponentiation des coefficients des variables indicatrices temporelles permet donc de neutraliser les effets des variations quantitatives des caractéristiques et donne une indication de l'évolution des prix des logements ajustés pour tenir compte des changements de qualité entre la période de référence 0 et chaque période de comparaison t. Autrement dit, l'indice à variables indicatrices temporelles entre la période 0 et la période t est donné par[5]

$$P_{TD}^{0t} = \exp(\hat{\delta}^t) \tag{5.6}$$

5.12 Le regroupement des données transversales préserve une certaine liberté. Les coefficients de régression $\hat{\beta}_k$ seront généralement affectés de moindres écarts-types que les coefficients $\hat{\beta}_k^t$ qui seraient obtenus en estimant le modèle (5.19) séparément à l'aide des données des échantillons $S(0), S(1),..., S(T)$. Si cette plus grande efficience peut être considérée comme un avantage, elle a aussi un prix : l'hypothèse de la fixité des paramètres des caractéristiques est un inconvénient de la méthode hédonique des variables indicatrices temporelles.

5.13 Quand on utilise les moindres carrés ordinaires (MCO), l'indice hédonique à variables indicatrices temporelles peut être exprimé sous la forme suivante (voir, par exemple, Diewert, Heravi et Silver, 2009; de Haan, 2010a)

$$P_{TD}^{0t} = \frac{\prod_{n\in S(t)} (p_n^t)^{1/N(t)}}{\prod_{n\in S(0)} (p_n^0)^{1/N(0)}} \exp\left[\sum_{k=1}^{K} \hat{\beta}_k (\bar{z}_k^0 - \bar{z}_k^t)\right] \tag{5.7}$$

où $\bar{z}_k^s = \sum_{n\in S(s)} z_{nk}^s / N(s)$ est la moyenne empirique de la caractéristique k pendant la période s $(s = 0, t)$. L'équation (5.7) nous dit que l'indice à variables indicatrices temporelles est essentiellement le produit de deux facteurs. Le premier facteur est le ratio des prix exprimé sous forme de moyenne géométrique pendant les périodes t et 0. Le deuxième facteur, $\exp[\sum_{k=1}^{K} \hat{\beta}_k (\bar{z}_k^0 - \bar{z}_k^t)]$, ajuste ce ratio des moyennes empiriques brutes pour tenir compte des différences entre les caractéristiques moyennes $\bar{z}_k^0$ et $\bar{z}_k^t$; il fait fonction de facteur d'ajustement pour la qualité qui prend en compte à la fois l'évolution des composantes de la qualité et les changements observés dans la qualité des biens immobiliers (à condition que tous les attributs pertinents déterminant la qualité soient inclus dans le modèle hédonique). Il convient de noter que l'indice des prix à variables indicatrices temporelles se réduit au ratio des prix exprimé sous forme de moyenne géométrique si $\bar{z}_k^t = \bar{z}_k^0$, c'est-à-dire s'il se trouve que les caractéristiques moyennes sont égales pendant la période t et la période 0.

5.14 Supposons pour simplifier que le stock de logements est constant, en ce sens qu'il comporte toujours le même nombre de logements, et que la qualité des différents biens immobiliers ne change pas. Supposons, en outre, que $S(0)$ et $S(t)$ correspondent à des échantillons aléatoires ou «représentatifs» du stock de logements. Dans ce cas, la méthode des variables indicatrices temporelles vise implicitement à obtenir un rapport de prix moyens calculés par moyenne géométrique pour l'ensemble du stock, qui est égal à la moyenne géométrique des rapports de prix[6]. Le concept que l'on souhaite mesurer peut certes varier en fonction des fins visées, mais il est difficile de voir à quelles fins servirait un indice des prix du stock immobilier résidentiel reposant sur la moyenne géométrique. Des indices de prix de l'immobilier résidentiel reposant sur une moyenne arithmétique comme un indice suivant l'évolution de la valeur d'un stock fixe de logements dans le temps semblent plus appropriés (voir aussi les chapitres 4 et 8).

5.15 Les échantillons de logements achetés ou vendus, $S(0)$ et $S(t)$, peuvent ne pas être représentatifs du stock total de logements (ou du nombre total de logements vendus). Une solution pourrait être de pondérer les échantillons de manière à les rendre représentatifs. Effectuer une régression des MCO sur l'ensemble de données pondérées (regroupées) équivaut à

4 Cette méthode a d'abord été mise au point par Court (1939; 109–111) et elle en constitue la deuxième suggestion hédonique. La terminologie que nous utilisons ici n'est pas employée par tous les auteurs d'études sur les biens immobiliers. Par exemple, Crone et Voith (1992) appellent la méthode des variables indicatrices temporelles la «méthode hédonique avec contrainte». Gatzlaff et Ling (1994) l'appellent, quant à eux, la «méthode explicite des variables temporelles» tandis que Knight, Dombrow et Sirmans (1995) parlent de «méthode des paramètres variables». D'autres formulations encore figurant dans les publications, les avis émis sur les mérites relatifs des diverses méthodes hédoniques doivent être interprétés avec précaution.

5 L'espérance de l'exponentielle du coefficient de la variable indicatrice temporelle ne correspond pas exactement au paramètre de cette variable. Le biais lié à cet écart est souvent appelé le biais du petit échantillon : il diminue en effet quand la taille de l'échantillon augmente. Sauf si l'échantillon est extraordinairement petit, ce biais sera faible par rapport à l'écart-type et il peut généralement être négligé.

6 Dans le cadre de la théorie des indices, ce type d'indice est connu sous le nom d'indice de Jevons.

réaliser une régression selon la méthode des moindres carrés pondérés (MCP) sur la série initiale de données. Les manuels d'économétrie ne suggèrent pas le recours aux MCP dans l'hypothèse d'une variance constante des termes d'erreur du fait que cela serait source d'hétéroscédasticité. Il convient de noter que l'utilisation de la méthode des variables indicatrices temporelles avec recours aux MCP permettra aussi d'obtenir un indice reposant sur la moyenne géométrique qui, dans ce cas, sera pondéré.

5.16 Plutôt que de recourir aux régressions des MCP, il pourrait être préférable de stratifier les échantillons, d'effectuer des régressions des MCO sur les données des différentes strates, puis de pondérer explicitement les indices hédoniques de chaque strate à l'aide des coefficients de pondération du stock (ou des ventes) de logements pour construire un indice global des prix de l'immobilier résidentiel avec une moyenne arithmétique au niveau d'agrégation le plus poussé. Cette approche hédonique stratifiée présente plusieurs autres avantages, comme on le verra plus loin.

5.17 Un problème que pose la méthode des variables indicatrices temporelles est la révision qu'elle implique. Si la série temporelle est étendue de T à $T+1$ avec de nouvelles données d'observations ajoutées, les coefficients des caractéristiques s'en trouveront modifiés. Les nouveaux chiffres de l'indice des prix calculés pour les périodes $t = 1,..., T$ différeront en conséquence de ceux antérieurement obtenus[7]. Lorsque de nouvelles données sont disponibles, l'efficience liée au regroupement des données augmente et de meilleures estimations peuvent être effectuées, ce qui peut être en fait considéré comme un avantage plutôt que comme un inconvénient de la méthode. D'un autre côté, les organismes statistiques et leurs utilisateurs seront très probablement peu disposés à accepter que les chiffres déjà publiés soient constamment révisés.

5.18 La méthode des variables indicatrices temporelles couvrant plusieurs périodes semble donc présenter un intérêt limité pour la production d'indices officiels des prix des logements, même si le problème des révisions peut être surmonté de plusieurs façons. L'une d'elles serait d'estimer les indices à variables indicatrices temporelles pour les périodes adjacentes t-1 et t, puis de les multiplier pour obtenir une série temporelle ne nécessitant pas de révisions. Ce chaînage des données de fréquence élevée présente en outre l'avantage d'assouplir l'hypothèse des paramètres fixes. Il n'est cependant pas totalement exempt de problèmes lui-même. Une dérive peut se produire dans l'indice quand les données accusent des fluctuations systématiques comme des fluctuations saisonnières[8].

7 Comme l'a dit Hill (2004), la méthode des variables indicatrices temporelles viole la conservation du passé.

8 Une autre solution consisterait à utiliser une fenêtre glissante. Supposons, par exemple, que nous avons estimé un indice à variables indicatrices temporelles à partir de données couvrant une période de douze mois. Nous pouvons alors supprimer les données du premier mois et ajouter celles du treizième mois et estimer l'indice à indicatrices temporelles à partir de cette nouvelle série de données et ainsi de suite. En multipliant (chaînant) les derniers changements d'un mois à l'autre, on obtient une série temporelle non révisée. Voir Shimizu, Nishimura et Watanabe (2010) pour une application. Dans l'exemple donné pour la ville de «A» à la fin du présent chapitre, aucun problème sérieux de dérive ne semble se poser; la méthode de la fenêtre glissante donne à peu près les mêmes résultats que la régression avec variables indicatrices temporelles couvrant plusieurs périodes.

Prix des caractéristiques et méthodes d'imputation

5.19 Dans le cadre de la deuxième grande méthode d'élaboration d'un indice de prix hédonique, des régressions sont effectuées séparément pour chaque période et l'indice est construit en utilisant les prix prédits à partir des coefficients de régression. Cette méthode est plus souple que celle des variables indicatrices temporelles du fait que les prix implicites des caractéristiques peuvent varier dans le temps. Il en existe deux variantes : la *méthode des prix des caractéristiques* et la *méthode des imputations*. On verra que ces deux méthodes sont équivalentes dans certaines circonstances. Nous allons commencer par examiner la première[9].

Méthode des prix des caractéristiques

5.20 Pour illustrer cette méthode, supposons comme plus tôt que des données d'observation sont disponibles sur les prix et sur les caractéristiques pertinentes des logements vendus pendant la période de référence 0 et chaque période de comparaison t. Nous supposerons aussi pour commencer que le modèle hédonique linéaire (5.2) est applicable et qu'il est estimé sur la base des données de la période 0 et de la période t séparément. On obtient ainsi les coefficients de régression $\hat{\beta}_0^s$ et $\hat{\beta}_k^s$ $(k = 1,..., K)$ pour $s = 0, t$. Les prix prédits pour chaque bien immobilier sont $\hat{p}_n^0 = \hat{\beta}_0^0 + \sum_{k=1}^{K} \hat{\beta}_k^0 z_{nk}^0$ et $\hat{p}_n^t = \hat{\beta}_0^t + \sum_{k=1}^{K} \hat{\beta}_k^t z_{nk}^t$. Il est aussi possible de calculer les prix prédits pour la période 0 et la période t pour un bien «normalisé» présentant un certain nombre (fixe) de caractéristiques z_k^*. Le ratio des prix estimés ainsi obtenu s'établit comme suit :

$$\frac{\hat{p}^t}{\hat{p}^0} = \frac{\hat{\beta}_0^t + \sum_{k=1}^{K} \hat{\beta}_k^t z_k^*}{\hat{\beta}_0^0 + \sum_{k=1}^{K} \hat{\beta}_k^0 z_k^*} \qquad (5.8)$$

L'expression (5.8) est un indice des prix ajusté pour tenir compte de la qualité puisque les caractéristiques restent fixes. Mais diverses valeurs de z_k^* donneront des nombres-indices différents. Quelle serait donc l'option préférée?

5.21 Supposons que nous cherchons à obtenir un indice des prix de l'immobilier résidentiel reposant sur les ventes. Deux choix naturels s'offrent pour z_k^* dans l'expression (5.8) : les caractéristiques moyennes de l'échantillon pour la période de référence $\bar{z}_k^0$ et les moyennes de l'échantillon pour la période de comparaison t $(t = 1,..., T)$, $\bar{z}_k^t$. La solution généralement adoptée dans le cadre de la théorie des

9 Là encore, la terminologie varie d'un auteur à l'autre. Par exemple, Crone et Voith (1992) et Knight, Dombrow et Sirmans (1995) qualifient cette méthode de «méthode hédonique» (par opposition à la «méthode hédonique avec contrainte» ou à la «méthode des paramètres variables» que nous avons appelée la méthode des variables indicatrices temporelles) tandis que Gatzlaff et Ling (1994) la désignent sous le nom de «méthode strictement transversale».

nombres-indices consiste à traiter les indices de prix obtenus — qui sont tout aussi valables — de façon symétrique en utilisant la moyenne géométrique. En posant $z_k^* = \bar{z}_k^0$ dans l'expression (5.8), on obtient un indice des prix des caractéristiques, de type Laspeyres :

$$P_{CPL}^{0t} = \frac{\hat{\beta}_0^t + \sum_{k=1}^{K} \hat{\beta}_k^t \bar{z}_k^0}{\hat{\beta}_0^0 + \sum_{k=1}^{K} \hat{\beta}_k^0 \bar{z}_k^0} \tag{5.9}$$

En posant $z_k^* = \bar{z}_k^t$ dans l'expression (5.8), on obtient un indice de type Paasche :

$$P_{CPP}^{0t} = \frac{\hat{\beta}_0^t + \sum_{k=1}^{K} \hat{\beta}_k^t \bar{z}_k^t}{\hat{\beta}_0^0 + \sum_{k=1}^{K} \hat{\beta}_k^0 \bar{z}_k^t} \tag{5.10}$$

En calculant la moyenne géométrique des équations (5.9) et (5.10), on obtient l'indice des prix des caractéristiques de type Fisher :

$$P_{CPF}^{0t} = \left[P_{CPL}^{0t} P_{CPP}^{0t}\right]^{1/2} \tag{5.11}$$

5.22 La méthode des prix des caractéristiques peut aussi être appliquée conjointement avec le modèle log-linéaire correspondant à l'équation (5.3). En effectuant des régressions séparées de ce modèle sur les données d'observation pour les périodes 0 et *t*, on obtient les prix prédits (après exponentiation) $\hat{p}_n^0 = \exp(\hat{\beta}_0^0)\exp[\sum_{k=1}^{K} \hat{\beta}_k^0 z_{nk}^0]$ et $\hat{p}_n^t = \exp(\hat{\beta}_0^t)\exp[\sum_{k=1}^{K} \hat{\beta}_k^t z_{nk}^t]$. Comme on l'a fait dans l'expression (5.8) pour le modèle linéaire, on peut prédire le prix d'un logement normalisé. En utilisant les moyennes des caractéristiques de l'échantillon pendant la période de référence pour définir ce logement, on obtient la contrepartie géométrique de l'indice des prix des caractéristiques, de type Laspeyres (5.9) :

$$P_{CPGL}^{0t} = \frac{\exp(\hat{\beta}_0^t)\exp\left[\sum_{k=1}^{K} \hat{\beta}_k^t \bar{z}_k^0\right]}{\exp(\hat{\beta}_0^0)\exp\left[\sum_{k=1}^{K} \hat{\beta}_k^0 \bar{z}_k^0\right]} = \exp(\hat{\beta}_0^t - \hat{\beta}_0^0)\exp\left[\sum_{k=1}^{K} (\hat{\beta}_k^t - \hat{\beta}_k^0)\bar{z}_k^0\right] \tag{5.12}$$

On obtient la contrepartie géométrique de l'indice hédonique de type Paasche (5.10) en utilisant les moyennes des caractéristiques de l'échantillon pendant la période de comparaison :

$$P_{CPGP}^{0t} = \frac{\exp(\hat{\beta}_0^t)\exp\left[\sum_{k=1}^{K} \hat{\beta}_k^t \bar{z}_k^t\right]}{\exp(\hat{\beta}_0^0)\exp\left[\sum_{k=1}^{K} \hat{\beta}_k^0 \bar{z}_k^t\right]} = \exp(\hat{\beta}_0^t - \hat{\beta}_0^0)\exp\left[\sum_{k=1}^{K} (\hat{\beta}_k^t - \hat{\beta}_k^0)\bar{z}_k^t\right] \tag{5.13}$$

En calculant la moyenne géométrique des équations (5.12) et (5.13), on obtient

$$P_{CPGF}^{0t} = \left[P_{CPGL}^{0t} P_{CPCG}^{0t}\right]^{1/2} = \exp(\hat{\beta}_0^t - \hat{\beta}_0^0)\exp\left[\sum_{k=1}^{K} (\hat{\beta}_k^t - \hat{\beta}_k^0)\bar{z}_k^{0t}\right] \tag{5.14}$$

où $\bar{z}_k^{0t} = (\bar{z}_k^0 + \bar{z}_k^t)/2$ dans l'équation (5.14) correspond à la moyenne des caractéristiques moyennes pendant la période de référence et la période de comparaison.

5.23 Si l'on vise à établir un indice des prix de l'immobilier résidentiel reposant sur le stock plutôt que sur les ventes de logements, les deux choix naturels pour les caractéristiques z_k^* dans l'équation (5.8) sont les caractéristiques moyennes du *stock* de logements pendant la période de référence et pendant la période de comparaison. Dans le premier cas, on obtiendra un indice des prix du stock de logements de type Laspeyres et, dans le second, un indice des prix du stock de logements de type Paasche. Ces deux indices mesurent l'évolution de la valeur du stock de logements ajustée en fonction de la qualité, mais ils ne donnent généralement pas les mêmes résultats. Non seulement la qualité moyenne du stock de logements évolue dans le temps, mais l'indice de type Laspeyres ignore les nouveaux biens immobiliers entrés sur le marché du logement tandis que l'indice de type Paasche ne tient pas compte des biens qui en sont sortis.

5.24 Il est bien entendu irréaliste de supposer que les moyennes sont disponibles pour les caractéristiques de l'ensemble du stock de biens immobiliers inclus dans le modèle hédonique. Il faut, dans la plupart des cas, recourir à des estimations, c'est-à-dire aux moyennes de l'échantillon $\bar{z}_k^0$ et $\bar{z}_k^t$ qui reposent sur les mêmes données concernant les caractéristiques que celles utilisées pour estimer les équations hédoniques. Cela conduit aux formules (5.9) et (5.10) ou à la moyenne géométrique (5.11) qui décrivent les indices de prix de l'immobilier résidentiel reposant sur les ventes. Cela nous rappelle une fois de plus que ces indices peuvent être considérés comme des estimateurs des indices des prix du stock de logements, à condition que les échantillons soient représentatifs de l'ensemble du stock. Cette condition ayant toutefois peu de chances d'être remplie, il est généralement procédé à la stratification des échantillons et à la pondération des indices estimés des strates à l'aide des coefficients de pondération des stocks.

La méthode à imputation hédonique

5.25 La question se pose de savoir comment articuler la méthode des prix des caractéristiques décrite plus haut avec la méthode courante (d'appariement de modèles) pour construire les indices de prix. Du point de vue des nombres-indices, nous pouvons examiner la question de la façon suivante. Les prix des biens immobiliers vendus pendant la période 0 ne peuvent être observés pendant la période *t* et sont donc «manquants» du fait que ces biens, ou du moins, la majeure partie d'entre eux, ne seront pas revendus pendant la période *t*. Les prix pendant la période 0 des biens vendus pendant la période *t*

ne sont pas non plus observables. Ces «prix manquants» doivent être imputés pour appliquer les formules types des nombres-indices[10]. Les indices reposant sur l'imputation hédonique utilisent pour cela les prix prédits, évalués pour des caractéristiques fixes, sur la base des régressions hédoniques effectuées pour toutes les périodes temporelles.

Indices avec imputation reposant sur la moyenne arithmétique

5.26 L'indice de Laspeyres avec imputation impute les prix de la période t pour les biens immobiliers appartenant à l'échantillon de la période de référence $S(0)$, évalués en fonction des caractéristiques de la période de référence pour neutraliser les effets des changements de qualité. En utilisant le modèle linéaire (5.1), les prix imputés sont $\hat{p}_n^t(0) = \hat{\beta}_0^t + \sum_{k=1}^{K} \hat{\beta}_k^t z_{nk}^0$ et l'indice de Laspeyres avec imputation hédonique devient

$$P_{HIL}^{0t} = \frac{\sum_{n\in S(0)} 1\hat{p}_n^t(0)}{\sum_{n\in S(0)} 1p_n^0} = \frac{\sum_{n\in S(0)} \left[\hat{\beta}_0^t + \sum_{k=1}^{K} \hat{\beta}_k^t z_{nk}^0\right]}{\sum_{n\in S(0)} p_n^0}$$

$$= \frac{\hat{\beta}_0^t + \sum_{k=1}^{K} \hat{\beta}_k^t \bar{z}_k^0}{\sum_{n\in S(0)} p_n^0 / N(0)} \qquad (5.15)$$

Il convient de noter que la quantité associée à chaque prix est égale à 1, ce qui veut essentiellement dire que chaque logement est unique et ne peut être apparié qu'en recourant à un modèle.

5.27 L'indice de Laspeyres avec imputation hédonique (5.15) est un exemple d'indice à *imputation unique* dans lequel les prix observés restent inchangés. On peut soutenir qu'il serait préférable d'utiliser la méthode de la *double imputation* en vertu de laquelle les prix observés sont remplacés par les valeurs prédites du fait que les biais dus aux variables omises qui affectent les estimations de la période 0 et de la période t ont des chances de se compenser, au moins jusqu'à un certain point; voir Hill, 2011, par exemple. En utilisant $\hat{p}_n^0 = \hat{\beta}_0^0 + \sum_{k=1}^{K} \hat{\beta}_k^0 z_{nk}^0$, l'indice de prix hédonique de Laspeyres à double imputation (DI) s'établit comme suit :

$$P_{HDIL}^{0t} = \frac{\sum_{n\in S(0)} 1\hat{p}_n^t(0)}{\sum_{n\in S(0)} 1\hat{p}_n^0} = \frac{\sum_{n\in S(0)} \left[\hat{\beta}_0^t + \sum_{k=1}^{K} \hat{\beta}_k^t z_{nk}^0\right]}{\sum_{n\in S(0)} \left[\hat{\beta}_0^0 + \sum_{k=1}^{K} \hat{\beta}_k^0 z_{nk}^0\right]}$$

$$= \frac{\hat{\beta}_0^t + \sum_{k=1}^{K} \hat{\beta}_k^t \bar{z}_k^0}{\hat{\beta}_0^0 + \sum_{k=1}^{K} \hat{\beta}_k^0 \bar{z}_k^0} = P_{CPL}^{0t} \qquad (5.16)$$

10 Comme on l'a déjà dit, la théorie hédonique remonte au moins à Court (1939; 108). L'imputation était sa suggestion hédonique numéro 1. Celle-ci a été reprise par Griliches (1971a; 59–60) (1971b; 6) et Triplett et McDonald (1977; 144). Parmi les études plus récentes sur les imputations hédoniques, ont peut citer entre autres celles de Diewert (2003b), de Haan (2004) (2009) (2010a), Triplett (2004) et Diewert, Heravi et Silver (2009). Dans le contexte du logement, la méthode à imputation hédonique est examinée de façon approfondie par Hill et Melser (2008) et Hill (2011).

Il ressort d'une comparaison entre cette équation et l'équation (5.12) qu'en utilisant le modèle linéaire, l'indice à double imputation équivaut à l'indice des prix des caractéristiques de type Laspeyres. Ce résultat ne dépend pas de la méthode d'estimation utilisée. Si nous utilisions une régression de la méthode des MCO pour estimer le modèle linéaire, l'indice à imputation unique équivaudrait à l'indice à double imputation et il coïnciderait aussi avec l'indice des prix des caractéristiques du fait que dans ce cas $\sum_{n\in S(0)} p_n^0 = \sum_{n\in S(0)} \hat{p}_n^0$, le modèle hédonique incluant une constante, la somme des résidus de la régression des MCO est nulle.

5.28 L'indice hédonique de Paasche à imputation unique impute les prix, pour la période de référence, des biens immobiliers appartenant à l'échantillon $S(t)$ de la période t, évalués sur la base des caractéristiques correspondant à cette dernière période. En utilisant à nouveau le modèle linéaire (5.1), ces prix imputés nous sont donnés par $\hat{p}_n^0(t) = \hat{\beta}_0^0 + \sum_{k=1}^{K} \hat{\beta}_k^0 z_{nk}^t$. Pour gagner de la place, nous n'exposerons ici que la variante de la double imputation. Les prix observés (période t) sont remplacés par ceux prédits sur la base du modèle $\hat{p}_n^t = \hat{\beta}_0^t + \sum_{k=1}^{K} \hat{\beta}_k^t z_{nk}^t$. L'indice hédonique de Paasche à double imputation s'établit alors comme suit :

$$P_{HDIP}^{0t} = \frac{\sum_{n\in S(t)} 1\hat{p}_n^t}{\sum_{n\in S(t)} 1\hat{p}_n^0(t)} = \frac{\sum_{n\in S(t)} \left[\hat{\beta}_0^t + \sum_{k=1}^{K} \hat{\beta}_k^t z_{nk}^t\right]}{\sum_{n\in S(t)} \left[\hat{\beta}_0^0 + \sum_{k=1}^{K} \hat{\beta}_k^0 z_{nk}^t\right]}$$

$$= \frac{\hat{\beta}_0^t + \sum_{k=1}^{K} \hat{\beta}_k^t \bar{z}_k^t}{\hat{\beta}_0^0 + \sum_{k=1}^{K} \hat{\beta}_k^0 \bar{z}_k^t} = P_{CPP}^{0t} \qquad (5.17)$$

Il coïncide avec l'indice des prix des caractéristiques de type Paasche. En cas de recours à la régression des MCO, il équivaut à l'indice de Paasche à imputation unique du fait que dans ce cas précis, le numérateur égale $\sum_{n\in S(t)} p_n^t$. Il ne sera donc pas nécessaire d'estimer les équations hédoniques pour les périodes de comparaison $t = 1,\ldots,T$; il suffira d'estimer l'équation hédonique de la période de référence pour obtenir les valeurs imputées pour cette période.

5.29 On obtient l'indice hédonique de Fisher à double imputation en calculant la moyenne géométrique des équations (5.16) et (5.17) :

$$P_{HDIF}^{0t} = \left[P_{HDIL}^{0t} P_{HDIP}^{0t}\right]^{1/2} \qquad (5.18)$$

Les indices avec imputation qui précèdent peuvent être interprétés de deux façons. Ils peuvent être considérés comme des estimateurs de l'évolution de la valeur ajustée en fonction de la qualité de l'ensemble du stock de logements, c'est-à-dire comme des indices des prix de l'immobilier résidentiel reposant sur les stocks, ou comme des estimateurs des indices des prix de l'immobilier résidentiel reposant sur les ventes et ajustés pour tenir compte de la qualité. Dans le premier cas, pour obtenir des résultats pratiquement non biaisés, chaque échantillon

devrait être un échantillon aléatoire ou représentatif du stock de logements. Les problèmes de biais d'échantillonnage pourraient être moins sérieux dans le second cas, mais tout dépend de la façon dont l'échantillonnage est conçu[11].

Indices avec imputation reposant sur la moyenne géométrique

5.30 La méthode de l'imputation peut aussi être appliquée aux formules d'indices des prix reposant sur la moyenne géométrique. Commençons par ce que nous pourrions appeler la contrepartie géométrique de l'indice des prix de Laspeyres avec imputation (5.15). Aux fins de «cohérence», les imputations seront calculées à l'aide du modèle hédonique log-linéaire (5.3), et non plus à l'aide du modèle linéaire. Les prix imputés de la période t pour les biens immobiliers figurant dans l'échantillon de la période de référence $S(0)$, évalués en fonction des caractéristiques de la période de référence, sont $\hat{p}_n^t(0) = \exp(\hat{\beta}_0^t)\exp[\sum_{k=1}^K \hat{\beta}_k^t z_{nk}^0]$. L'indice géométrique non pondéré avec double imputation, dans lequel les prix de la période de référence sont remplacés par les valeurs prédites $\hat{p}_n^0 = \exp(\hat{\beta}_0^0)\exp[\sum_{k=1}^K \hat{\beta}_k^0 z_{nk}^0]$ s'établit donc comme suit :

$$P_{HDIGL}^{0t} = \frac{\prod_{n\in S(0)}(\hat{p}_n^t(0))^{1/N(0)}}{\prod_{n\in S(0)}(\hat{p}_n^0)^{1/N(0)}} = \exp(\hat{\beta}_0^t - \hat{\beta}_0^0)\exp\left[\sum_{k=1}^K(\hat{\beta}_k^t - \hat{\beta}_k^0)\bar{z}_k^0\right] = P_{CPGL}^{0t} \quad (5.19)$$

On obtient, de même, la contrepartie géométrique de l'indice des prix de Paasche avec imputation (5.16) en imputant les prix de la période 0 pour les biens immobiliers figurant dans l'échantillon $S(t)$ de la période t, qui sont donnés par $\hat{p}_n^0(t) = \exp(\hat{\beta}_0^0)\exp[\sum_{k=1}^K \hat{\beta}_k^0 z_{nk}^t]$, et en remplaçant les prix observés de la période t par les prix prédits $\hat{p}_n^t = \exp(\hat{\beta}_0^t)\exp[\sum_{k=1}^K \hat{\beta}_k^t z_{nk}^t]$. On a donc :

$$P_{HDIGP}^{0t} = \frac{\prod_{n\in S(t)}(\hat{p}_n^t)^{1/N(t)}}{\prod_{n\in S(t)}(\hat{p}_n^0(t))^{1/N(t)}} = \exp(\hat{\beta}_0^t - \hat{\beta}_0^0)\exp\left[\sum_{k=1}^K(\hat{\beta}_k^t - \hat{\beta}_k^0)\bar{z}_k^t\right] = P_{CPGP}^{0t} \quad (5.20)$$

5.31 Si l'on recourt à la méthode des MCO pour estimer les équations de la régression log-linéaire, le dénominateur de l'équation (5.19) et le numérateur de l'équation (5.20) équivalent à la moyenne géométrique des prix de l'échantillon pour la période 0 et la période t, respectivement, et les indices à double imputation coïncident avec ceux à imputation unique. En calculant la moyenne géométrique des équations (5.19) et (5.20), on obtient :

$$P_{HDIGF}^{0t} = \left[P_{HDIGL}^{0t}\, P_{HDIGP}^{0t}\right]^{1/2} = \exp(\hat{\beta}_0^t - \hat{\beta}_0^0)\exp\left[\sum_{k=1}^K(\hat{\beta}_k^t - \hat{\beta}_k^0)\bar{z}_k^{0t}\right] = P_{CPGF}^{0t} \quad (5.21)$$

où $\bar{z}_k^{0t} = (\bar{z}_k^0 + \bar{z}_k^t)/2$ représente la moyenne des caractéristiques moyennes pour les périodes 0 et t, comme auparavant.

5.32 L'équation de l'indice avec imputation symétrique (5.21) peut être réécrite sous une forme étonnamment proche de l'équation (5.7) correspondant à l'indice à variables indicatrices temporelles quand on recourt à la méthode des MCO pour estimer les équations hédoniques (voir Diewert, Heravi et Silver, 2009, et de Haan, 2010a) :

$$P_{HDIGF}^{0t} = \frac{\prod_{n\in S(t)}(p_n^t)^{1/N(t)}}{\prod_{n\in S(0)}(p_n^0)^{1/N(0)}}\exp\left[\sum_{k=1}^K \hat{\beta}_k^{0t}(\bar{z}_k^0 - \bar{z}_k^t)\right] \quad (5.22)$$

où $\hat{\beta}_k^{0t} = (\hat{\beta}_k^0 + \hat{\beta}_k^t)/2$ représente la valeur moyenne du k^e coefficient pour les périodes 0 et t. L'équation (5.22) ajuste le ratio des prix observés exprimés sous forme de moyenne géométrique pour tenir compte de différences éventuelles dans les caractéristiques moyennes de l'échantillon. Triplett (2006) parle à ce propos d'«ajustement hédonique de la qualité». En effectuant un rapprochement avec l'équation (5.7), on constate qu'en cas de non-variation des moyennes de l'ensemble des caractéristiques de l'échantillon ($\bar{z}_k^0 = \bar{z}_k^t$), l'indice avec imputation hédonique symétrique et l'indice à variables indicatrices temporelles coïncident et équivalent au ratio des prix observés, exprimés sous forme de moyenne géométrique, mais il va sans dire que cela se produit rarement. Ces deux types d'indices hédoniques coïncident aussi si, pour chaque caractéristique, le coefficient moyen $\hat{\beta}_k^{0t}$ des deux régressions distinctes est égal au coefficient $\hat{\beta}_k$ de la régression avec variables indicatrices temporelles. C'est également rare, mais cela permet de penser que les deux méthodes produisent des résultats similaires si les paramètres des caractéristiques restent à peu près constants dans le temps.

5.33 Si l'on peut supposer constants dans le temps les paramètres des caractéristiques, on peut remplacer les coefficients moyens $\hat{\beta}_k^{0t}$ dans l'équation (5.22) par les coefficients $\hat{\beta}_k^0$ de la période de référence. Il ne sera pas nécessaire, dans ce cas, d'effectuer une régression pour chaque période temporelle et on utilisera en fait l'indice de prix avec imputation non symétrique correspondant à l'équation (5.13)[12]. La régression pour la période de référence pourra être effectuée sur un plus grand ensemble de données pour augmenter la stabilité des coefficients. Il est recommandé de vérifier régulièrement si les coefficients ont sensiblement changé et de les actualiser s'il y a lieu.

11 Si toutes les transactions immobilières étaient observées, il ne serait procédé à aucun échantillonnage du point de vue des ventes et le problème de biais d'échantillonnage ne se poserait pas. Dans beaucoup de pays, l'administration du cadastre enregistre toutes les transactions, du moins celles concernant la revente de maisons, mais l'ensemble de données ainsi réuni ne contient généralement que peu d'informations sur les caractéristiques des logements; voir, par exemple, Lim et Pavlou (2007) ou Academetrics (2009).

12 On parle en Europe de «réévaluation hédonique des prix» pour ce type d'ajustement hédonique de la qualité, surtout quand la taille de l'échantillon est fixe (Destatis, 2009).

5.34 Comme on l'a déjà dit, les indices de prix reposant sur la moyenne géométrique se prêtent moins bien au rôle d'estimateurs des indices des prix de l'immobilier résidentiel, ajustés pour tenir compte de la qualité. Nous n'allons toutefois pas jusqu'à dire qu'ils ne devraient jamais être utilisés. Dans le contexte de la stratification, le recours à l'équation (5.21) pourrait donner des résultats satisfaisants puisqu'il permettrait d'associer un ajustement en fonction de la qualité (en utilisant un modèle de régression hédonique log-linéaire) et une formule d'indice symétrique au sein des différentes strates à un ajustement des composantes dans l'ensemble des strates. Cette approche hédonique stratifiée est examinée dans la prochaine section.

Indices hédoniques stratifiés

5.35 La stratification, ou la fixation de la structure, a été considérée au chapitre 4. La stratification est un moyen simple et efficace d'éliminer l'incidence des modifications de la composition en éléments de qualité des biens immobiliers vendus. Néanmoins, certains des changements survenus au sein même des strates ne seront probablement pas tous éliminés du fait essentiellement que chaque bien immobilier est unique. Un biais pourrait donc exister au niveau des valeurs unitaires. Le recours à un système de stratification plus poussé pourrait être difficilement envisageable surtout si le nombre d'observations disponibles était relativement limité. Il pourrait valoir la peine, si l'on disposait des données nécessaires sur les caractéristiques, d'utiliser une stratification moins fine et de recourir à une régression hédonique au niveau des strates pour éliminer l'incidence des changements survenus dans les composantes qualitatives. Cette approche en deux temps associerait l'utilisation de la méthode hédonique au niveau des strates à une pondération explicite au niveau global pour obtenir un indice général des prix de l'immobilier résidentiel.

5.36 Deux avantages de la stratification ont déjà été mentionnés. Premièrement, celle-ci permet aux organismes statistiques de publier des indices de prix de l'immobilier résidentiel pour différents segments du marché, ce qui sera positif pour les utilisateurs puisqu'il est bien connu que l'évolution des prix peut varier assez fortement entre les types de logement, les régions, etc. Deuxièmement, elle peut permettre de réduire des biais d'échantillonnage, notamment celui lié à la non-réponse dans le cas notamment de l'indice des prix de l'immobilier résidentiel reposant sur le stock de logements.

5.37 Il est fortement recommandé de recourir à la stratification en cas d'utilisation des techniques de régression hédonique pour éliminer l'incidence des changements affectant les composantes qualitatives. Il est très peu probable qu'un seul modèle hédonique soit valable pour tous les segments du marché et il est donc souhaitable d'effectuer des régressions distinctes pour les divers types de biens, les divers emplacements, etc. Deux problèmes sont en fait en jeu ici. Le plus important peut-être tient au fait qu'à chaque segment du marché correspond un ensemble particulier de caractéristiques. Par exemple, celles concernant les maisons individuelles ne sont pas les mêmes que celles concernant les appartements situés dans des tours, ne serait-ce qu'en raison du fait que l'étage auquel se trouve un appartement est considéré comme une variable importante de la détermination de son prix. Le deuxième problème en jeu, certes moins important sans doute, est que la valeur des paramètres peut varier d'un segment à l'autre du marché pour les mêmes caractéristiques. N'importe quel manuel d'économétrie contient des tests statistiques pour les différences observées, entre les sous-échantillons, dans les valeurs des paramètres.

5.38 C'est par rapport à la méthode d'imputation que l'approche hédonique stratifiée peut être le plus facilement illustrée, surtout quand on l'associe à la formule de l'indice de Laspeyres. Reprenons la troisième expression de la partie droite de l'équation correspondant à l'indice des prix de Laspeyres avec imputation hédonique unique (5.15), où les prix de la période t «manquent» pour les logements figurant dans l'échantillon de la période de référence $S(0)$, et sont imputés par $\hat{p}_n^t(0)$ (à l'aide du modèle de régression hédonique estimé pour la période t). Supposons, comme au chapitre 4, que l'ensemble de l'échantillon est stratifié (*a posteriori*) en M sous-échantillons $S_m(0)$. L'équation (5.15) peut alors être réécrite sous la forme suivante :

$$P_{HIL}^{0t} = \frac{\sum_{n\in S(0)} \hat{p}_n^t(0)}{\sum_{n\in S(0)} p_n^0} = \frac{\sum_{m=1}^{M}\sum_{n\in S_m(0)} \hat{p}_n^t(0)}{\sum_{m=1}^{M}\sum_{n\in S_m(0)} p_n^0}$$

$$= \frac{\sum_{m=1}^{M}\sum_{n\in S_m(0)} p_n^0\left[\sum_{n\in S_m(0)} \hat{p}_n^t(0) \Big/ \sum_{n\in S_m(0)} p_n^0\right]}{\sum_{m=1}^{M}\sum_{n\in S_m(0)} p_n^0} = \sum_{m=1}^{M} s_m^0 P_{HIL,m}^{0t} \qquad (5.23)$$

où $P_{HIL,m}^{0t} = \sum_{n\in S_m(0)} \hat{p}_n^t(0) / \sum_{n\in S_m(0)} p_n^0$ représente l'indice des prix de Laspeyres avec imputation hédonique (unique) entre la période de référence et la période t pour la cellule m; $s_m^0 = \sum_{n\in S_m(0)} p_n^0 / \sum_{n\in S(0)} p_n^0$ est la part correspondante de la valeur des ventes qui sert de coefficient de pondération pour $P_{HIL,m}^{0t}$. Il convient de noter que la dernière expression de l'équation (5.23) a une structure comparable à celle de l'indice découpé en cellules homogènes correspondant à l'équation (4.1), mais dans le cas présent, les indices des cellules reposent sur des imputations hédoniques plutôt que sur des valeurs unitaires.

5.39 L'équation (5.23) montre que si les prix imputés $\hat{p}_n^t(0)$ pour tous les logements de l'échantillon $S(0)$ reposent sur une seule régression hédonique globale, l'indice agrégé de Laspeyres avec imputation hédonique peut être écrit sous la forme d'un indice stratifié. Mais c'est simplement une autre façon d'écrire les choses, et non ce que l'on entend par «approche hédonique stratifiée». En outre, comme on l'a indiqué plus haut, il n'est pas du tout réaliste de penser pouvoir utiliser un seul modèle. Il ne faut donc pas effectuer une seule grande

régression hédonique, mais des régressions séparées sur les données des sous-échantillons de chaque période pour obtenir les prix imputés (de la période *t*) et les indices des cellules avec imputation. C'est ainsi qu'on obtiendra un indice stratifié avec imputation hédonique de type Laspeyres.

5.40 Il serait préférable d'estimer un indice hédonique stratifié de Fisher plutôt qu'un indice de Laspeyres. C'est tout à fait possible pour un indice des prix de vente de l'immobilier résidentiel, mais peut-être pas pour un indice des prix du stock immobilier résidentiel, comme on l'a déjà dit au chapitre 3 du fait que l'on manque souvent de données actualisées sur le nombre de biens immobiliers recensés.

Principaux avantages et inconvénients

5.41 Cette section du manuel fait le point des avantages et des inconvénients que présente le recours aux régressions hédoniques pour construire un indice des prix de l'immobilier résidentiel. Ses principaux avantages sont les suivants :

- Si la liste disponible des caractéristiques des biens immobiliers est assez détaillée, les méthodes hédoniques peuvent, en principe, permettre d'éliminer l'incidence des changements survenus dans la composition de l'échantillon ainsi que dans la qualité des divers biens immobiliers.
- Des indices de prix peuvent être construits pour différents types de logements et différents emplacements si l'échantillon est stratifié comme il convient. La stratification présente elle-même plusieurs autres avantages.
- La méthode hédonique est probablement celle qui permet d'exploiter au mieux les données disponibles.
- La variante de la méthode des régressions hédoniques qui fait appel à l'imputation est analogue à la méthode de la fixation d'un échantillon qui est largement utilisée pour construire des indices de prix.

5.42 Les principaux inconvénients des régressions hédoniques sont les suivants :

- Il peut s'avérer difficile de neutraliser suffisamment l'effet de l'emplacement si les prix des biens immobiliers et leur évolution varient au sein des régions. L'adoption d'une approche stratifiée des régressions hédoniques permettra toutefois de surmonter en partie ce problème.
- L'approche hédonique nécessite beaucoup de données puisque, pour l'appliquer, il faut disposer de données sur toutes les caractéristiques pertinentes des biens immobiliers, ce qui rend son utilisation assez coûteuse[13].
- La méthode hédonique est, pour l'essentiel, reproductible, mais divers choix peuvent être opérés en ce qui concerne notamment l'ensemble des caractéristiques incluses dans le modèle, la forme fonctionnelle, les transformations possibles de la variable dépendante[14] et la spécification stochastique, etc., qui pourraient se traduire par des estimations divergentes de l'évolution générale des prix. Cela implique donc sans doute la nécessité de disposer d'un grand nombre de métadonnées.
- L'idée générale de la méthode hédonique est facilement compréhensible, mais certains de ses aspects techniques risquent d'être difficiles à expliquer aux utilisateurs.

5.43 Globalement donc, les régressions hédoniques constituent probablement la meilleure approche à suivre pour construire des indices de prix de l'immobilier résidentiel à qualité constante pour divers types de biens immobiliers[15]. Nous sommes favorables à l'utilisation de la variante avec (double) imputation parce qu'elle constitue l'approche hédonique la plus flexible et qu'elle est analogue à la méthode des produits appariés standard adoptée pour construire des indices de prix.

5.44 Dans les trois prochaines sections, nous allons illustrer les diverses méthodes de régression hédonique à l'aide des données relatives à la ville de «A» qui a été décrite à la fin du chapitre 4. Les deux sections suivantes présentent les résultats des régressions hédoniques à variables indicatrices temporelles utilisant, respectivement, le logarithme du prix de vente comme variable dépendante et le prix de vente non transformé. La dernière section illustre la méthode de l'imputation hédonique. Tous les indices de prix ainsi obtenus concernent les *ventes* de maisons individuelles; certains résultats obtenus à partir des données concernant la ville de «A» pour les indices relatifs au *stock* de logements ne seront présentés qu'au chapitre 8.

Modèles à variables indicatrices temporelles utilisant le logarithme du prix comme variable dépendante

Le modèle log-linéaire à variables indicatrices temporelles

5.45 Reprenons la description des données concernant les ventes de maisons individuelles dans la ville néerlandaise de «A». Pendant le trimestre *t*, $N(t)$ ventes de maisons individuelles ont été effectuées dans cette ville avec p_n^t correspondant au

13 Toutefois, comme on le verra avec l'exemple néerlandais donné plus loin, des informations sur l'emplacement du bien immobilier, son type, son ancienneté, sa surface de plancher et la superficie du terrain peuvent permettre d'expliquer pour l'essentiel la variation du prix de vente.

14 Par exemple, la variable dépendante pourrait, entre autres, être le prix de vente du bien ou son logarithme, ou le prix de vente divisé par la superficie du bâtiment.

15 Cette évaluation concorde avec celle de Hoffmann et Lorenz (2006; 15) : «Pour ce qui est de l'ajustement pour tenir compte de la qualité, les méthodes hédoniques sont certainement la solution d'avenir.» Gouriéroux et Laferrère (2009) ont montré qu'il est possible de construire un modèle officiel national de régression hédonique crédible pour les biens immobiliers.

prix de vente de la maison *n* vendue pendant le trimestre *t*. Des informations sont disponibles sur trois caractéristiques de cette maison *n* vendue pendant la période *t* : L_n^t est la superficie du terrain en mètres carrés (m²); S_n^t est la surface de plancher du bâtiment en m² et A_n^t est l'ancienneté de la maison *n* en dizaines d'années pendant la période *t*. En utilisant ces variables, le modèle courant *log-linéaire de régression hédonique à variables indicatrices temporelles* est défini par le système suivant d'équations de régression[16] :

$$\ln p_n^t = \alpha + \beta L_n^t + \gamma S_n^t + \delta A_n^t + \tau^t + \varepsilon_n^t \qquad (5.24)$$

$$t = 1,...,14;\ n = 1,...,N(t);\ \tau^1 \equiv 0$$

où τ^t est un paramètre qui fait monter ou descendre le prix hédonique au trimestre *t* par rapport au prix du trimestre 1[17].

5.46 Il est facile de construire un indice des prix à l'aide du modèle log-linéaire de régression hédonique à variables indicatrices temporelles (5.24). En prenant l'exponentielle des deux membres de l'équation (5.24) et en négligeant le terme d'erreur, on obtient $p_n^t = \exp(\alpha)[\exp(L_n^t)]^\beta[\exp(S_n^t)]^\gamma[\exp(A_n^t)]^\delta \exp(\tau^t)$. Si l'on pouvait observer un bien immobilier présentant les *mêmes caractéristiques* pendant la période de référence 1 et une période de comparaison $t(>1)$, le ratio de prix correspondant (en négligeant à nouveau les termes d'erreur) serait simplement égal à $\exp(\tau^t)$. Pendant deux périodes consécutives *t* et *t*+1, le ratio de prix (toujours en négligeant les termes d'erreur) serait égal à $\exp(\tau^{t+1})/\exp(\tau^t)$ et pourrait servir de maillon dans un indice-chaîne de prix. Le graphique 5.1 montre l'indice ainsi obtenu, appelé P_{H1} (indice hédonique n° 1) tandis que le tableau 5.1 expose les nombres-indices. La valeur du R^2 obtenue pour ce modèle était de 0,8420, ce qui est assez satisfaisant pour un modèle de régression hédonique ne comportant que trois variables explicatives[18]. Il convient de noter, à des fins de comparaison ultérieure, que la log-vraisemblance était de 1.407,6.

5.47 Un aspect négatif de ce modèle est que le modèle sous-jacent de la formation des prix ne paraît pas plausible : *S* et *L* interagissent de façon multiplicative pour la détermination du prix global de la maison alors qu'il semble très probable que la taille du terrain *L* et la taille de la maison *S* interagissent de façon plutôt additive dans la détermination du prix global.

5.48 Un autre aspect négatif du modèle de régression (5.24) est que l'ancienneté du logement est prise en compte de manière additive. Or, si l'on peut s'attendre à ce que l'ancienneté interagisse directement avec la variable des bâtiments *S* en tant que variable de la dépréciation (nette), ce n'est pas le cas pour la variable représentative du terrain *L*, du fait qu'un terrain ne se déprécie pas. Cette interaction directe de l'ancienneté avec les bâtiments sera prise en compte dans le modèle suivant.

Le modèle log-linéaire à variables indicatrices temporelles avec ajustement de la qualité des bâtiments

5.49 Si l'ancienneté *A* interagit avec la taille des bâtiments *S* de manière multiplicative, une variable explicative appropriée pour le prix de vente d'une maison serait $\gamma(1-\delta)^A S$ (c'est-à-dire la dépréciation géométrique où δ est le taux d'amortissement géométrique sur dix ans) ou $\gamma(1-\delta A)S$ (dépréciation linéaire où δ est le taux d'amortissement linéaire sur dix ans) plutôt que la spécification selon la méthode additive $\gamma S + \delta A$. La variante linéaire de cette catégorie de modèles va être estimée par la suite[19]. Le *modèle log-linéaire de régression hédonique à variables indicatrices temporelles avec ajustement de la qualité des bâtiments* devient

$$\ln p_n^t = \alpha + \beta L_n^t + \gamma(1-\delta A_n^t)S_n^t + \tau^t + \varepsilon_n^t \qquad (5.25)$$

$$t = 1,...,14;\ n = 1,...,N(t);\ \tau^1 \equiv 0$$

5.50 Le modèle de régression (5.25) a été appliqué en utilisant les 14 trimestres de données sur les ventes pour la ville de «A». Il convient de noter qu'un seul taux d'amortissement linéaire est estimé. Le taux d'amortissement (net) sur dix ans[20] a été estimé à $\hat{\delta} = 11{,}94$ % (soit environ 1,2 % par an), ce qui est très raisonnable. De même que pour le modèle (5.24), si une maison présente les *mêmes caractéristiques* pendant deux périodes consécutives *t* et *t*+1, le ratio de prix correspondant (en négligeant les termes d'erreur) $\exp(\tau^{t+1})/\exp(\tau^t)$ peut servir de maillon dans un indice-chaîne de prix; voir le graphique 5.1 et le tableau 5.1 pour l'indice ainsi obtenu, appelé P_{H2}. La valeur du R^2 obtenue pour ce modèle était de 0,8345, ce qui est un peu moins que pour le modèle précédent, et la log-vraisemblance était de 1.354,9, soit une baisse non négligeable par rapport au chiffre de 1.407,6 précédemment obtenu[21].

5.51 Il semble que la prise en compte d'autres considérations théoriques — concernant le traitement de l'ancienneté de la maison — a conduit à un moins bon ajustement du modèle aux données empiriques. Ce modèle et le précédent

16 L'équation d'estimation pour l'ensemble de données regroupées inclura des variables indicatrices temporelles pour désigner les trimestres. Pour tous les modèles estimés pour la ville de «A», on suppose que les termes d'erreur ε_n^t sont des variables normales indépendamment distribuées de moyenne nulle et de variance constante. Les paramètres inconnus de chaque modèle de régression sont estimés par la méthode du maximum de vraisemblance. L'estimation proprement dite a été effectuée en utilisant l'option non linéaire du logiciel économétrique Shazam.

17 Les 15 paramètres α, $\tau^1,...,\tau^{14}$ correspondent à des variables qui sont exactement colinéaires dans la régression (5.24) et la restriction $\tau^1 = 0$ est donc imposée pour déterminer les paramètres restants.

18 Plus loin dans ce chapitre et au chapitre 8, des régressions hédoniques seront effectuées en utilisant les prix p_n^t comme variables dépendantes plutôt que les logarithmes des prix. Pour faciliter les comparaisons de la qualité de l'ajustement des modèles, nous transformerons les valeurs prédites pour les modèles logarithmiques des prix en niveaux de prix prédits en prenant l'exponentielle des prix prédits puis en calculant le coefficient de corrélation entre ces niveaux de prix prédits et les prix effectifs. En élevant au carré ce coefficient de corrélation, on obtient une *mesure adaptée du niveau de la qualité de l'ajustement* des modèles logarithmiques des prix qui est représentée par R^{*2}. Pour ce modèle particulier, $R^{*2} = 0{,}8061$.

19 Cette régression est essentiellement linéaire dans les paramètres inconnus et elle est donc très facile à estimer.

20 C'est le taux d'amortissement nette qui a été utilisé, car aucune information n'était disponible sur les dépenses de rénovation et donc δ correspond à la dépréciation brute pour usure de la maison, déduction faite des dépenses moyennes consacrées aux travaux de rénovation et de réparation.

21 La valeur du R^2 adapté pour ce modèle était $R^{*2} = 0{,}7647$, ce qui là aussi représente une baisse non négligeable par rapport au niveau correspondant du R^2 observé pour le précédent modèle logarithmique des prix.

ne sont cependant probablement pas bien spécifiés[22] : ils multiplient tous les deux la superficie du terrain par celle du bâtiment pour déterminer le prix de la maison alors qu'une interaction additive entre L et S serait sans doute plus appropriée qu'une interaction multiplicative.

5.52 Il convient de noter qu'étant donné le taux d'amortissement δ, les *bâtiments après ajustement de leur qualité* (pour tenir compte de leur ancienneté) des différentes maisons n pendant chaque trimestre t peuvent être définis comme suit :

$$S_n^{t^*} \equiv (1-\delta A_n^t)S_n^t \qquad (5.26)$$

$$t = 1,...,14;\ n = 1,...,N(t)$$

Le modèle log-log à variables indicatrices temporelles avec ajustement de la qualité des bâtiments en fonction de leur ancienneté

5.53 Dans le reste de cette section, nous utiliserons comme variable explicative les bâtiments après ajustement de leur qualité (en fonction de leur ancienneté), $(1-\delta A)S$, plutôt que leur surface non ajustée, S. Le modèle log-log est similaire au précédent modèle log-linéaire si ce n'est que maintenant, au lieu d'utiliser L et $(1-\delta A)S$ comme variables explicatives dans le modèle de régression, ce sont les logarithmes de la superficie du terrain et de la surface des bâtiments après ajustement de leur qualité qui sont utilisés comme variables indépendantes. Le *modèle log-log de régression hédonique à variables indicatrices temporelles avec ajustement de la qualité des bâtiments* se présente comme suit[23] :

$$\ln p_n^t = \alpha + \beta \ln L_n^t + \gamma \ln[(1-\delta A_n^t)S_n^t] + \tau^t + \varepsilon_n^t \qquad (5.27)$$

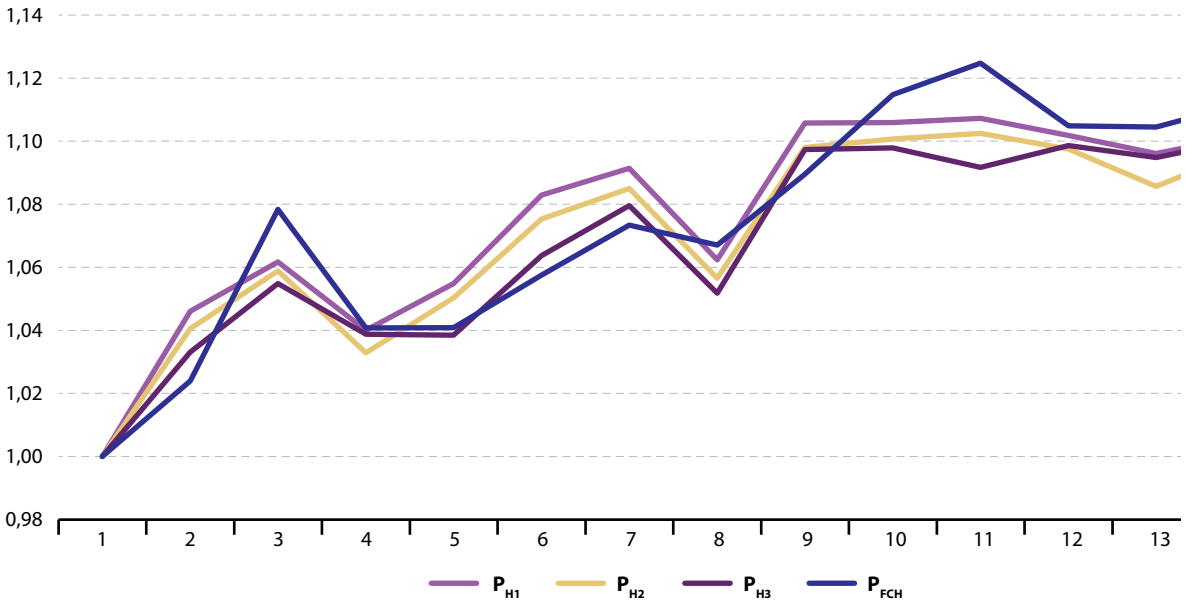
$$t = 1,...,14;\ n = 1,...,N(t);\ \tau^1 \equiv 0$$

5.54 Sur la base des données concernant la ville néerlandaise de «A», le taux d'amortissement (net) estimé sur dix ans était $\hat{\delta} = 0{,}1050$ (erreur-type de 0,00374). Si les deux membres de l'équation (5.27) étaient passés à l'exponentielle et les termes d'erreur négligés, le prix p_n^t de la maison égalerait $\exp(\alpha)[L_n^t]^\beta[S_n^{t^*}]^\gamma \exp(\tau^t)$, $S_n^{t^*}$ correspondant aux bâtiments après ajustement de leur qualité, tels qu'ils ont été définis par l'équation (5.26). Aussi, si nous pouvions observer une maison présentant les *mêmes caractéristiques* pendant deux périodes consécutives t et $t+1$, le ratio de prix correspondant (en négligeant les termes d'erreur) serait égal à $\exp(\tau^{t+1})/\exp(\tau^t)$ et il pourrait là aussi servir de maillon pour un l'indice-chaîne de prix; voir le graphique 5.1 et le tableau 5.1 pour l'indice ainsi obtenu, appelé P_{H3}. La valeur du R^2 obtenue pour ce modèle était de 0,8599 (et la mesure

22 Si la variation des variables indépendantes est relativement faible, la différence entre les indices liée aux divers modèles de régression hédonique considérés dans la présente section et les deux suivantes sera probablement faible elle aussi du fait que pratiquement tous ces modèles peuvent permettre d'obtenir une approximation linéaire de la «vérité». Mais quand la variation des variables indépendantes est importante, comme c'est le cas dans le contexte actuel du logement, le choix de la forme fonctionnelle peut avoir un effet important. Le choix des variables indépendantes de la régression comme celui de la forme fonctionnelle doit donc faire l'objet d'un raisonnement *a priori*. Pour une analyse plus approfondie des questions concernant la forme fonctionnelle, voir Diewert (2003a).

23 Ce modèle de régression hédonique s'avère être une variante de l'*approche orientée vers le consommateur des modèles hédoniques pour le logement* de McMillen (2003). Son cadre théorique, qui s'appuie sur les travaux de Muth (1971), est décrit dans ses grandes lignes dans Diewert, de Haan et Hendriks (2010). Voir aussi McDonald (1981).

Graphique 5.1. Indices de prix log-linéaires à variables indicatrices temporelles et indice-chaîne de Fisher des prix moyens à échantillon stratifié



Source : calculs des auteurs basés sur les données du cadastre néerlandais.

Tableau 5.1. Indices de prix log-linéaires à variables indicatrices temporelles et indice-chaîne de Fisher des prix moyens à échantillon stratifié

Trimestre	P_{H1}	P_{H2}	P_{H3}	P_{FCH}
1	1,00000	1,00000	1,00000	1,00000
2	1,04609	1,04059	1,03314	1,02396
3	1,06168	1,05888	1,05482	1,07840
4	1,04007	1,03287	1,03876	1,04081
5	1,05484	1,05032	1,03848	1,04083
6	1,08290	1,07532	1,06369	1,05754
7	1,09142	1,08502	1,07957	1,07340
8	1,06237	1,05655	1,05181	1,06706
9	1,10572	1,09799	1,09736	1,08950
10	1,10590	1,10071	1,09786	1,11476
11	1,10722	1,10244	1,09167	1,12471
12	1,10177	1,09747	1,09859	1,10483
13	1,09605	1,08568	1,09482	1,10450
14	1,10166	1,09694	1,10057	1,11189

Source : calculs des auteurs basés sur les données du cadastre néerlandais.

adaptée de la qualité de l'ajustement était $R^{*2} = 0{,}8880$), soit une valeur plus élevée que celle obtenue pour les modèles (5.25) et (5.26); la log-vraisemblance était de 1.545,4, ce qui était bien supérieur à celle observée pour les deux autres modèles (1.407,6 et 1.354,9).

5.55 Les séries de prix des maisons obtenues par les trois régressions log-linéaires à variables indicatrices temporelles considérées dans la présente section, P_{H1}, P_{H2} et P_{H3}, sont indiquées dans le graphique 5.1 avec l'indice-chaîne de Fisher des prix moyens reposant sur un échantillon stratifié, P_{FCH}. Ces quatre séries de prix de maisons sont exposées dans le tableau 5.1. Les quatre indices reflètent la même tendance, mais des écarts de plus de 2 % sont observables pour certains trimestres. Il convient de noter que tous ces indices évoluent dans le même sens d'un trimestre à l'autre et accusent des baisses aux trimestres 4, 8, 12 et 13, mais l'indice P_{H3} — celui qui correspond au modèle log-log à variables indicatrices temporelles — enregistre une augmentation au trimestre 12.

5.56 Si le modèle (5.27) est le plus satisfaisant parmi les modèles hédoniques simples considérés jusqu'à présent, il se distingue malheureusement par le fait que les quantités de terrains et de bâtiments après ajustement de leur qualité déterminent le prix d'un bien immobilier de *manière multiplicative*. Or il est plus probable que le prix des maisons est déterminé par la *somme* pondérée des montants de leur terrain et de leur bâtiment après ajustement pour tenir compte de la qualité, c'est pourquoi nous allons procéder dans la prochaine section à l'estimation d'un modèle additif à variables indicatrices temporelles. Ce modèle devrait être mieux adapté aux données.

Modèles de régression hédonique à variables indicatrices temporelles utilisant le prix comme variable dépendante

Le modèle de régression hédonique linéaire à variables indicatrices temporelles

5.57 Il est permis de penser que le prix de vente d'un bien immobilier est lié de manière linéaire à la fois à la superficie du terrain et à celle du bâtiment du fait du caractère concurrentiel du secteur de la construction résidentielle[24]. Si l'ancienneté du bâtiment est traitée comme une autre caractéristique qui joue un rôle important dans la détermination du prix du bien immobilier, le *modèle de régression hédonique linéaire à variables indicatrices temporelles* qui suit pourrait convenir :

$$p_n^t = \alpha + \beta L_n^t + \gamma S_n^t + \delta A_n^t + \tau^t + \varepsilon_n^t \qquad (5.28)$$

$$t = 1,...,14;\ n = 1,...,N(t);\ \tau^1 \equiv 0$$

5.58 Le modèle de régression linéaire qui précède a été appliqué en utilisant les données afférentes à la ville de «A». La valeur du R^2 obtenue pour lui a été de 0,8687, soit une valeur beaucoup plus élevée que celles obtenues dans les

24 Voir Clapp (1980), Francke et Vos (2004), Gyourko et Saiz (2004), Bostic, Longhofer et Redfearn (2007), Davis et Heathcote (2007), Francke (2008), Diewert (2009b), Koev et Santos Silva (2008), Institut national de la statistique du Portugal (2009), Diewert, de Haan et Hendriks (2010), Diewert (2010) et le chapitre 8 plus loin.

précédentes régressions[25]; la log-vraisemblance a été, quant à elle, de −1.0790,4 (chiffre qui n'est pas facilement comparable à ceux précédemment observés du fait que la variable dépendante utilisée dans ce cas a été uniquement le prix, et non pas le logarithme du prix[26]).

5.59 Il est un peu plus difficile d'utiliser le modèle linéaire défini par les équations (5.28) pour construire un indice global du prix des maisons que d'utiliser le précédent modèle de régression log-linéaire ou log-log à variables indicatrices temporelles. Dans la précédente section, si l'on maintient constantes les caractéristiques et si l'on néglige les termes d'erreur, le *prix relatif* d'une maison sur deux périodes quelconques s'avère constant, ce qui conduit à un indice global non ambigu. Dans le cas présent, lorsqu'on maintient les caractéristiques constantes et que l'on néglige les termes d'erreur, la *différence de prix* pour une même maison s'avère constante, mais les *prix relatifs* de différentes maisons ne le sont généralement pas. Un indice global sera donc construit en utilisant les prix générés par les paramètres estimés pour le modèle (5.28) et évalués à l'aide des superficies moyennes de l'échantillon pour les terrains L et les bâtiments S et de l'ancienneté moyenne A des maisons de l'échantillon[27]. Les prix trimestriels ainsi obtenus pour cette maison «moyenne» ont été convertis en un indice, l'indice P_{H4}, qui est exposé dans le tableau 5.2 et représenté graphiquement dans le graphique 5.2.

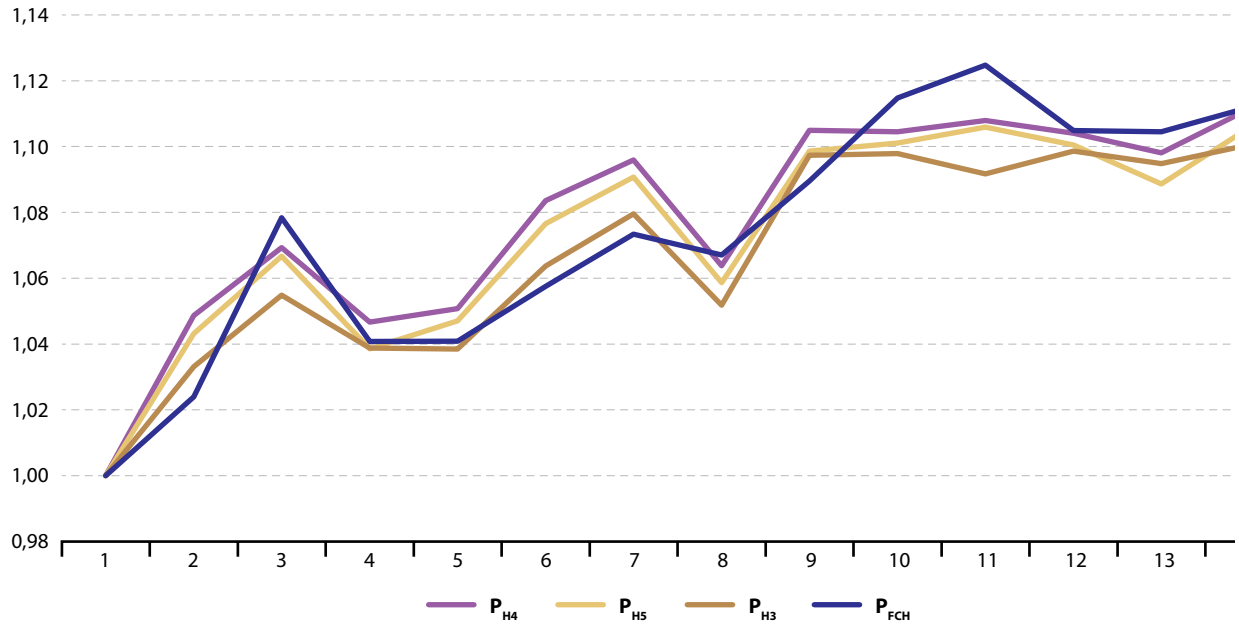
5.60 Le modèle de régression hédonique défini par les équations (5.28) est peut-être le plus simple possible, mais il est en réalité un peu trop simple, car il néglige le fait que les interactions entre l'ancienneté d'un bien immobilier et son prix de vente s'exercent par la voie d'une interaction multiplicative avec la variable des bâtiments et non par celle d'un

25 Il ne faut toutefois pas oublier que la mesure ajustée du degré d'adéquation du modèle log-log correspondant à l'équation (5.27) était de 0,8880 et qu'elle dépassait donc la valeur de 0,8687 obtenue ici.

26 Marc Francke a fait observer qu'il est possible de comparer les log-vraisemblances de deux modèles lorsque la variable dépendante a été transformée par une fonction connue dans le second modèle; voir Davidson et McKinnon (1993; 491) qui recourent à un ajustement jacobien pour pouvoir comparer les log-vraisemblances de deux modèles.

27 Les superficies moyennes de l'échantillon pour L et S étaient respectivement de 257,6 m² et 127,2 m², et l'ancienneté moyenne des maisons individuelles vendues pendant la période couverte par l'échantillon était de 18,5 ans.

Graphique 5.2. Indices de prix linéaires à variables indicatrices temporelles, indice de prix log-log à variables indicatrices temporelles et indice-chaîne de Fisher des prix moyens à échantillon stratifié



Source : calculs des auteurs basés sur les données du cadastre néerlandais.

Tableau 5.2. Indices de prix linéaires à variables indicatrices temporelles, indice de prix log-log à variables indicatrices temporelles et indice-chaîne de Fisher des prix moyens à échantillon stratifié

Trimestre	P_{H4}	P_{H5}	P_{H3}	P_{FCH}
1	1,00000	1,00000	1,00000	1,00000
2	1,04864	1,04313	1,03314	1,02396
3	1,06929	1,06667	1,05482	1,07840
4	1,04664	1,03855	1,03876	1,04081
5	1,05077	1,04706	1,03848	1,04083
6	1,08360	1,07661	1,06369	1,05754
7	1,09593	1,09068	1,07957	1,07340
8	1,06379	1,05864	1,05181	1,06706
9	1,10496	1,09861	1,09736	1,08950
10	1,10450	1,10107	1,09786	1,11476
11	1,10788	1,10588	1,09167	1,12471
12	1,10403	1,10044	1,09859	1,10483
13	1,09805	1,08864	1,09482	1,10450
14	1,11150	1,10572	1,10057	1,11189

Source : calculs des auteurs basés sur les données du cadastre néerlandais.

facteur additif général. Le modèle (5.28) fait l'objet ci-après d'une nouvelle estimation pour laquelle on utilise comme variable explicative les bâtiments ajustés pour tenir compte de la qualité au lieu de se contenter d'intégrer séparément l'ancienneté A comme une caractéristique indépendante.

Le modèle linéaire à variables indicatrices temporelles avec ajustement de la qualité des bâtiments

5.61 Le *modèle hédonique linéaire avec ajustement de la qualité des bâtiments* peut être exprimé de la manière suivante

$$p_n^t = \alpha + \beta L_n^t + \gamma(1-\delta A_n^t)S_n^t + \tau^t + \varepsilon_n^t \qquad (5.29)$$

$$t = 1,...,14;\ n = 1,...,N(t);\ \tau^1 \equiv 0$$

C'est le modèle hédonique le plus plausible considéré jusqu'à présent. Il fonctionne avec des bâtiments S^* dont la qualité est ajustée (en fonction de l'ancienneté) égaux à $(1-\delta A)S$, au lieu que A et S soient intégrés dans la régression de façon linéaire comme des variables totalement indépendantes.

5.62 Les résultats obtenus avec ce modèle ont été nettement plus satisfaisants que ceux obtenus avec le modèle (5.28). La log-vraisemblance a augmenté de 92 unités pour atteindre –10.697,8 et la valeur du R^2 est passée de 0,8687 à 0,8789. Le taux d'amortissement estimé sur dix ans était de $\hat{\delta} = 0,1119$ (0,00418), ce qui reste raisonnable. Ce modèle de régression linéaire a la même particularité que le modèle (5.28) : les *différences* de prix entre les maisons sont constantes dans le temps pour tous les modèles à caractéristiques constantes, mais les *ratios* des prix des maisons ne sont pas constants. Un indice global a donc à nouveau été construit en utilisant les prix générés par les paramètres estimés du modèle (5.29) et évalués à l'aide des superficies moyennes de l'échantillon pour les terrains L et les bâtiments S et de l'ancienneté moyenne A des maisons de l'échantillon. Les prix trimestriels ainsi obtenus pour ce modèle «moyen» ont été convertis en un indice, l'indice P_{H5}, qui est exposé dans le tableau 5.2 et représenté dans le graphique 5.2. À des fins de comparaison, l'indice P_{H3} (l'indice du modèle log-log à variables indicatrices temporelles) et l'indice P_{FCH} (l'indice-chaîne de Fisher des prix moyens reposant sur un échantillon stratifié) sont représentés dans ce graphique avec les indices P_{H4} et P_{H5}. Les indices préférés jusqu'à présent sont les indices P_{FCH} et P_{H5}.

5.63 On peut à nouveau constater que les quatre indices reflètent la même tendance, mais qu'ils accusent pour certains trimestres des différences supérieures à 2 %. Il convient de noter que tous ces indices évoluent dans le même sens d'un trimestre à l'autre et qu'ils accusent des baisses aux trimestres 4, 8, 12 et 13, mais que l'indice P_{H3} enregistre une augmentation au trimestre 12.

5.64 Un problème qui se pose avec les modèles de régression hédonique à variables indicatrices temporelles considérés jusqu'à présent est que les prix des terrains et des bâtiments après ajustement de leur qualité ne peuvent évoluer librement d'une période à l'autre. Ce problème n'affecte pas le type de modèles hédoniques que nous allons considérer dans la prochaine section.

Modèles de régression avec imputation hédonique

5.65 La théorie des *indices avec imputation hédonique* expliquée plus haut est appliquée comme suit dans le cas présent. Une régression linéaire est effectuée sous la forme suivante pour chaque période :

$$p_n^t = \alpha^t + \beta^t L_n^t + \gamma^t (1 - \delta^t A_n^t) S_n^t + \varepsilon_n^t \qquad (5.30)$$

$$t = 1,...,14;\ n = 1,...,N(t)$$

En utilisant les données pour la ville de «A», quatre paramètres seulement doivent être estimés pour chaque trimestre : α^t, β^t, γ^t et δ^t pour $t = 1,...,14$. Il convient de noter que le modèle correspondant aux équations (5.30) est de forme analogue à celui défini par les équations (5.29) en dehors des différences importantes suivantes :

- Un seul paramètre de dépréciation est estimé dans le modèle défini par les équations (5.29) alors que, dans le présent modèle, 14 paramètres de dépréciation sont estimés (un par trimestre).
- De même, le modèle (5.29) ne comportait qu'un paramètre α, β et γ alors que dans le modèle (5.30), il y a 14 paramètres α^t,14 β^t et 14 γ^t à estimer. En revanche, le modèle (5.29) comportait 13 paramètres de décalage temporel (le τ^t) de plus à estimer.

Le modèle avec imputation hédonique implique donc l'estimation de 56 paramètres contre 17 seulement pour le modèle à variables indicatrices temporelles, ce qui permet de penser qu'il sera bien mieux ajusté aux données.

5.66 Dans le contexte du logement, le suivi d'un ensemble de logements exactement identiques sur les périodes considérées est impossible; la dépréciation et les activités de rénovation font qu'un même logement ne se prête pas à des comparaisons d'une période à l'autre. On peut remédier à ce manque de correspondance entre, disons, les trimestres t et t+1 en utilisant les paramètres estimés à l'aide de la régression hédonique pour le trimestre t+1 et en estimant les prix de tous les logements vendus pendant le trimestre t. On obtient ainsi les *prix prédits pour le trimestre t+1 des logements du trimestre t*, $\hat{p}_n^{t+1}(t)$, c'est-à-dire :

$$\hat{p}_n^{t+1}(t) \equiv \hat{\alpha}^{t+1} + \hat{\beta}^{t+1} L_n^t + \hat{\gamma}^{t+1} (1 - \hat{\delta}^{t+1} A_n^t) S_n^t \qquad (5.31)$$

$$t = 1,...,13;\ n = 1,...,N(t)$$

où $\hat{\alpha}^t$, $\hat{\beta}^t$, $\hat{\gamma}^t$ et $\hat{\delta}^t$ sont les estimations des paramètres du modèle (5.30) pour $t = 1,...,14$. On a maintenant un ensemble de prix «pseudo appariés» entre les trimestres t et t+1 pour les logements vendus au trimestre t et on peut construire l'*indice suivant avec imputation hédonique de type Laspeyres (ou modèle «pseudo apparié»)* allant du trimestre t au trimestre t+1[28] :

$$P_{HIL}^{t,t+1} \equiv \frac{\sum_{n=1}^{N(t)} 1 \hat{p}_n^{t+1}(t)}{\sum_{n=1}^{N(t)} 1 p_n^t} \qquad (5.32)$$

$$t = 1,...,13$$

Comme on l'a dit plus haut, chaque prix n'est associé qu'à une unité du fait que chaque logement est essentiellement unique et ne peut être apparié qu'à l'aide d'un modèle.

5.67 La même méthode peut être appliquée rétroactivement en partant des ventes de logements qui ont eu lieu au trimestre t+1; on prend alors les paramètres de la régression hédonique pour le trimestre t et on estime les prix en t+1, $\hat{p}_n^t(t+1)$ de tous les logements vendus pendant le trimestre t+1.

$$\hat{p}_n^t(t+1) \equiv \hat{\alpha}^t + \hat{\beta}^t L_n^{t+1} + \hat{\gamma}^t (1 - \hat{\delta}^t A_n^{t+1}) S_n^{t+1} \qquad (5.33)$$

$$t = 1,...,13;\ n = 1,...,N(t+1)$$

On dispose alors d'un ensemble de prix «appariés» pour le trimestre t des logements vendus pendant la période t+1 et on peut construire l'*indice suivant avec imputation hédonique de type Paasche (ou modèle «pseudo apparié»)* allant du trimestre t au trimestre t+1 :

$$P_{HIP}^{t,t+1} \equiv \frac{\sum_{n=1}^{N(t+1)} 1 p_n^{t+1}}{\sum_{n=1}^{N(t+1)} 1 \hat{p}_n^t(t+1)} \qquad (5.34)$$

$$t = 1,...,13$$

5.68 Une fois calculés les indices de prix de Laspeyres et de Paasche avec imputation hédonique, on peut construire l'*indice avec imputation hédonique de type Fisher* correspondant, allant de la période t à la période t+1 en calculant la moyenne géométrique des deux indices définis par les équations (5.32) et (5.34) :

$$P_{HIF}^{t,t+1} \equiv \left[P_{HIL}^{t,t+1} P_{HIP}^{t,t+1} \right]^{1/2} \qquad (5.35)$$

$$t = 1,...,13$$

5.69 Les indices-chaînes de prix à imputation de Laspeyres, Paasche et Fisher, P_{HIL}, P_{HIP} et P_{HIF}, ainsi obtenus à partir des données afférentes à la ville de «A», sont représentés dans le graphique 5.3 ci-dessous et énumérés dans le tableau 5.3. Ces trois indices avec imputation sont étonnamment proches les uns des autres. Celui de Fisher est l'indice

28 Du fait que les régressions définies par les équations (5.30) comportent une constante et sont essentiellement linéaires au niveau des variables explicatives, la somme des résidus de l'échantillon sera nulle dans chacune des régressions. La somme des prix prédits égalera donc la somme des prix observés pour chaque période. En conséquence, la somme des prix observés figurant au dénominateur de l'équation (5.32) sera égale à la somme des prix prédits correspondants et, de même, la somme des prix observés figurant au numérateur de l'équation (5.34) sera égale à la somme correspondante des prix prédits.

Graphique 5.3. Indices-chaînes des prix à imputation hédonique de Laspeyres, Paasche et Fisher

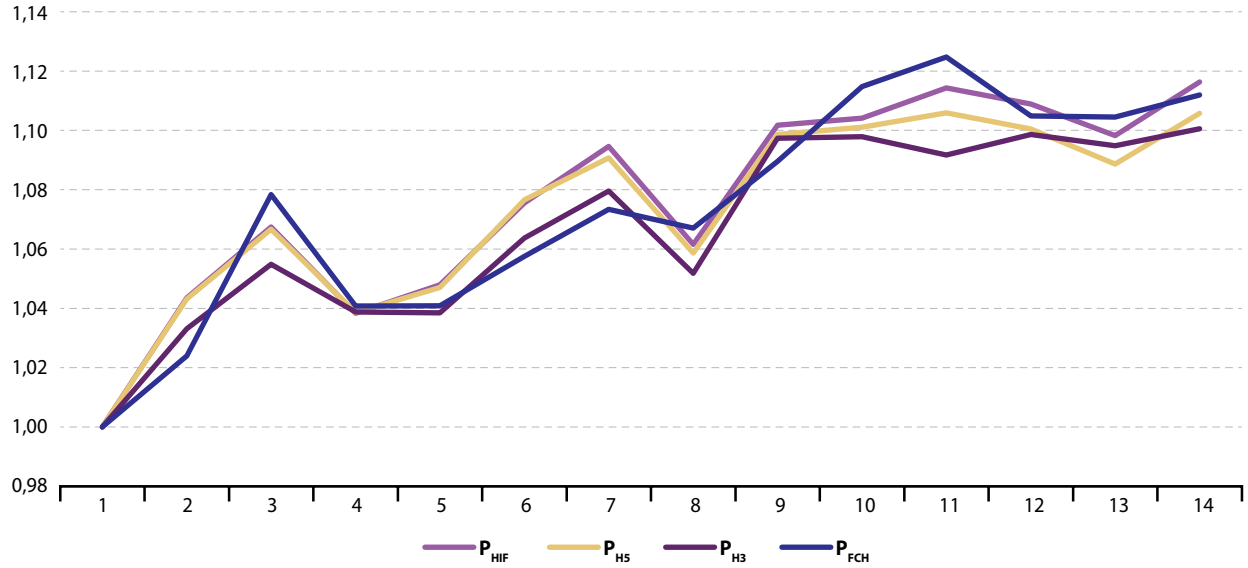
Source : calculs des auteurs basés sur les données du cadastre néerlandais.

Tableau 5.3. Indices-chaînes des prix à imputation hédonique de Laspeyres, Paasche et Fisher

Trimestre	P_{HIL}	P_{HIP}	P_{HIF}
1	1,00000	1,00000	1,00000
2	1,04234	1,04479	1,04356
3	1,06639	1,06853	1,06746
4	1,03912	1,03755	1,03834
5	1,04942	1,04647	1,04794
6	1,07267	1,07840	1,07553
7	1,08923	1,10001	1,09460
8	1,05689	1,06628	1,06158
9	1,09635	1,10716	1,10174
10	1,09945	1,10879	1,10411
11	1,11062	1,11801	1,11430
12	1,10665	1,11112	1,10888
13	1,09830	1,09819	1,09824
14	1,11981	1,11280	1,11630

Source : calculs des auteurs basés sur les données du cadastre néerlandais.

Graphique 5.4. Indice de Fisher à imputation des prix, indice-chaîne de Fisher des prix moyens à échantillon stratifié, indice de prix linéaires à variables indicatrices temporelles et indice de prix log-log à variables indicatrices temporelles



Source : calculs des auteurs basés sur les données du cadastre néerlandais.

de prix obtenu par imputation hédonique que nous préférons jusqu'à présent; il est préférable aux indices à variables indicatrices temporelles du fait que l'imputation permet aux prix des terrains et des bâtiments après ajustement de leur qualité d'évoluer indépendamment dans le temps, alors que les indices à variables indicatrices temporelles font évoluer parallèlement la surface hédonique. Les résultats empiriques indiquent qu'au moins pour la présente série de données pour la ville de «A», l'indice de Laspeyres avec imputation donne un résultat très proche de celui offert par l'indice de Fisher avec imputation que nous préférons.

5.70 Pour conclure, nos deux «meilleurs» indices sont l'indice de Fisher à imputation P_{HIF} et l'indice-chaîne de Fisher à échantillon stratifié. P_{FCH}. Dans l'ensemble, l'indice à imputation P_{HIF} doit probablement être préféré à l'indice P_{FCH} du fait que les indices reposant sur un échantillon stratifié présenteront un certain biais lié aux valeurs unitaires qui sera très probablement plus important que n'importe quel biais lié à la forme fonctionnelle dans l'indice P_{HIF}. Ces deux «meilleurs» indices sont représentés dans le graphique 5.4 avec l'indice log-log à variables indicatrices temporelles P_{H3} et l'indice linéaire à variables indicatrices temporelles avec ajustement de la qualité des bâtiments P_{H5}. Tous les indices de prix, à l'exception de l'indice P_{H3}, accusent une baisse aux trimestres 4, 8, 12 et 13 et une hausse aux autres trimestres; l'indice P_{H3} évolue en hausse au trimestre 12 au lieu d'évoluer en baisse comme les autres indices.

6

Méthodes des ventes répétées

Le modèle de base des ventes répétées

6.1 La méthode des ventes répétées a été proposée à l'origine par Bailey, Muth et Nourse (1963). À leurs yeux, cette procédure revenait à généraliser la *méthode du panier fixe combinée au chaînage,* appliquée par des précurseurs dans la construction d'indices de prix immobiliers comme Wyngarden (1927) et Wenzlick (1952). Les indices reposant sur la méthode des ventes répétées les plus connus sont les indices américains des prix des logements de Standard and Poor's et Case-Shiller, qui sont calculés pour 20 villes (Standard and Poor's, 2009). L'Agence fédérale de financement du logement (FHFA) calcule aussi pour les États-Unis un indice reposant sur la méthode des ventes répétées[1], mais suivant une approche légèrement différente. Residex et le cadastre britannique calculent des indices des ventes répétées pour les villes australiennes et le Royaume-Uni, respectivement[2].

6.2 Comme son nom l'indique, la méthode des ventes répétées utilise les informations concernant les biens immobiliers qui ont été vendus plus d'une fois. Du fait que c'est une méthode de calcul avec les mêmes biens immobiliers d'une période à l'autre, il n'est pas nécessaire de neutraliser les effets des différences observées dans l'échantillon de biens. Cependant, étant donné la faible incidence des unités revendues parfois, il ne serait pas très utile de calculer un indice des prix de l'immobilier résidentiel reposant sur les ventes répétées en utilisant la méthode courante du panier fixe et les formules classiques d'indices. Un modèle stochastique est en conséquence postulé qui «explique» l'évolution du prix des maisons qui ont été vendues à plusieurs reprises. Ce modèle de régression (à variables indicatrices) est ensuite estimé sur la base des données regroupées (c'est-à-dire de l'évolution globale des prix) pour l'ensemble de la période couverte par l'échantillon.

6.3 Les seules informations nécessaires pour estimer une équation de régression courante dans le cadre de la méthode des ventes répétées sont le prix, la date de la vente et l'adresse des biens concernés, ce qui fait que cette méthode requiert beaucoup moins de données que les méthodes hédoniques. De plus, la méthode des ventes répétées neutralise par défaut l'effet de l'emplacement au niveau de détail le plus poussé (l'adresse), ce que les méthodes de régression hédonique ne sont souvent pas en mesure de faire avec une grande précision[3]. Un inconvénient de la méthode des ventes répétées, toutefois, est qu'un logement qui est vendu à deux moments différents n'est pas nécessairement, exactement le même du fait de facteurs comme la dépréciation et les rénovations effectuées. Par conséquent, plus l'écart temporel entre les ventes est important, plus est contestable l'hypothèse de constance de la qualité sur laquelle repose l'approche de la méthode des ventes répétées.

6.4 On trouve, dans des publications, le modèle stochastique suivant qui explique le logarithme de la valeur (prix) p_n^t du bien immobilier n pendant la période t :

$$\ln p_n^t = P^t + H_n^t + \varepsilon_n^t \tag{6.1}$$

où P^t est un terme commun pour tous les biens immobiliers (le logarithme du «niveau de prix» dans une région ou une ville), H_n^t est un processus aléatoire gaussien qui représente le glissement de la valeur des logements individuels au cours du temps et ε_n^t est un terme d'erreur aléatoire ou bruit blanc. Le modèle (6.1) sert souvent de point de départ à l'établissement de l'équation d'estimation des ventes répétées.

6.5 Un autre point de départ pourrait être le modèle hédonique log-linéaire avec contrainte (5.4) dans lequel les paramètres β_k qui déterminent les prix des caractéristiques devraient être fixes dans le temps. La comparaison de biens «identiques» implique une deuxième restriction : les (quantités des) caractéristiques d'un bien immobilier sont aussi supposées être fixes dans le temps. Si l'on représente la k^e caractéristique du bien n par z_{nk}, le modèle log-linéaire avec contrainte devient alors

$$\ln p_n^t = \beta_0^t + \sum_{k=1}^{K} \beta_k z_{nk} + \varepsilon_n^t \tag{6.2}$$

6.6 Un modèle est obtenu pour le logarithme du changement de valeur du bien n entre deux périodes, disons s et t $(0 \le s < t \le T)$, en soustrayant (6.2) pour ces périodes. Il s'ensuit que

$$\begin{aligned} \ln p_n^t - \ln p_n^s &= \ln(p_n^t / p_n^s) = (\beta_0^t - \beta_0^s) + (\varepsilon_n^t - \varepsilon_n^s) \\ &= \ln P^{st} + (\varepsilon_n^t - \varepsilon_n^s) \end{aligned} \tag{6.3}$$

Le modèle (6.3) indique essentiellement que si l'on néglige le terme d'erreur $\varepsilon_n^t - \varepsilon_n^s$, le logarithme du changement de prix est le même pour tous les biens représentés par P^{st}.

6.7 Supposons maintenant que nous avons un échantillon de maisons qui ont été vendues plus d'une fois sur la période couverte par l'échantillon $t = 0,...,T$ et pour lesquelles nous disposons de données sur les prix de transaction, et donc sur l'évolution de leur prix. La période (de possession) entre les ventes ultérieures variera entre ces maisons. Étant donné, toutefois, que dans le modèle (6.3), les prix de tous les biens immobiliers sont supposés changer au même rythme (en dehors de perturbations aléatoires), les données sur les ventes répétées peuvent être regroupées et le modèle être estimé à l'aide de l'*équation-type des ventes répétées*

$$\ln(p_n^t / p_n^s) = \sum_{t=0}^{T} \gamma^t D_n^t + \mu_n^t \tag{6.4}$$

où D_n^t est une variable indicatrice égale à 1 pendant la période au cours de laquelle la revente a eu lieu, à -1 pendant la période au cours de laquelle la vente précédente a eu lieu,

1 La FHFA, créée en 2008, résulte de la fusion du Bureau national de surveillance des sociétés immobilières (OFHEO), qui publiait jusqu'alors l'indice des ventes répétées, et du Conseil fédéral de financement du logement (FHFB).

2 Le cadastre néerlandais calculait un indice des ventes répétées pour les Pays-Bas jusqu'en 2007, quand il a commencé à publier un indice reposant sur la méthode RPVE d'évaluation des prix de vente conjointement avec Statistics Netherlands. Voir le chapitre 7 pour la méthode RPVE.

3 Toutefois, l'utilisation de données géospatiales pour tenir compte de la dépendance spatiale dans l'équation hédonique pourrait permettre de remédier au problème de l'omission des variables géographiques; voir le chapitre 5 et Hill (2011) pour plus de détails.

et à 0 autrement, μ_i^t étant à nouveau un terme d'erreur[4]. Sous les hypothèses dites classiques, notamment celle selon laquelle les erreurs sont de moyenne nulle et de variance constante, l'équation (6.4) peut être estimée par une régression effectuée selon la méthode des moindres carrés ordinaires (MCO). Un problème de multicolinéarité peut affecter les données, mais les solutions pour y remédier sont limitées, si c'est le cas.

6.8 L'indice des ventes répétées pour la période comprise entre la période 0 et la période t est obtenu en prenant l'exponentielle des coefficients correspondants de la régression $\hat{\gamma}^t$:

$$P_{RS}^{0t} = \exp(\hat{\gamma}^t) \qquad (6.5)$$

La simplicité et l'attrait du modèle-type des ventes répétées résident dans le fait qu'il ne nécessite que des variables indicatrices; aucune donnée sur les caractéristiques n'est en effet nécessaire en dehors de l'emplacement (adresse)[5]. Cela, réuni à la facilité du calcul de l'indice des ventes répétées, explique sans doute en partie pourquoi cette méthode est très répandue dans les publications concernant l'immobilier et le logement.

6.9 Wang et Zorn (1997) ont élaboré une expression analytique pour l'indice des ventes répétées. Il semble avoir une structure géométrique assez complexe. Il peut par conséquent s'avérer difficile d'expliquer la méthode en détail bien que l'idée de l'appariement soit facilement comprise. De plus, comme on l'a indiqué plus haut, il n'est peut-être pas souhaitable de s'orienter vers un indice des prix de l'immobilier reposant sur une moyenne géométrique, surtout pour un indice des prix concernant le stock de logements. Une solution pourrait être d'utiliser une version de la méthode des ventes répétées reposant sur la moyenne arithmétique, comme l'a suggéré Shiller (1991). Les indices des prix du logement de Standard and Poor's (Case-Shiller) sont calculés à l'aide de la méthode des ventes répétées reposant sur la moyenne arithmétique (voir Standard and Poor's, 2009).

Problèmes qui se posent et améliorations qui peuvent être apportées au modèle de base

6.10 Nous allons examiner, dans cette section, un certain nombre de problèmes liés à la méthode des ventes répétées et donner un bref aperçu général des améliorations qui ont été suggérées dans la littérature.

4 Les reventes multiples sont traitées comme des observations indépendantes. Cela ne devrait pas poser trop de problèmes, comme Shiller (1991) l'a fait observer, du fait qu'il n'y a pas de chevauchement entre les périodes de détention dans le cas de reventes multiples.

5 Dans certains pays, comme le Royaume-Uni et les Pays-Bas, le cadastre collecte toutes les données sur les prix de transaction, mais un nombre très limité d'informations sur les caractéristiques, comme le type de logement et l'adresse bien sûr. Il n'est donc pas surprenant de constater que, dans ces pays, des indices des ventes répétées ont été calculés sur la base des données du cadastre. Il convient de noter qu'aux États-Unis, l'indice des ventes répétées du FHFE repose sur les données obtenues auprès de Fannie Mae et Freddie Mac pour les prêts hypothécaires.

Apurement des données

6.11 Dans la pratique, les biens immobiliers revendus très rapidement et ceux non revendus pendant longtemps ont parfois été exclus de régressions effectuées dans le cadre de la méthode des ventes répétées du fait que ces situations pouvaient être «atypiques» et donc biaiser l'indice de prix calculé. Clapp et Giacotto (1998) ainsi que Steele et Goy (1997) ont suggéré d'éliminer les biens immobiliers détenus pendant très peu de temps des séries de données, parce qu'ils pouvaient être liés à des ventes réalisées en catastrophe à la suite d'un divorce ou d'une perte d'emploi, par exemple, ou à des opérations spéculatives. Jansen *et al.* (2008) ont constaté, sur la base des données du cadastre néerlandais, des hausses de prix assez marquées pour les maisons revendues en moins d'un an.

6.12 L'un des avantages de la méthode des ventes répétées est qu'elle est reproductible. Sa reproductibilité risque toutefois d'être compromise si la procédure d'exclusion des observations «atypiques» varie.

Hétéroscédasticité

6.13 Case et Shiller (1987, 1989) ont soutenu que l'évolution des prix des maisons inclut des aspects dont la variance augmente avec l'intervalle séparant les ventes, ce qui contredit l'hypothèse de la variance constante des erreurs. Ils ont suggéré de suivre une approche reposant sur les moindres carrés pondérés (MCP) pour remédier à ce type d'hétéroscédasticité. Les coefficients de pondération sont calculés en régressant les résidus au carré de la régression courante des ventes répétées effectuée selon la méthode des MCO sur une constante et l'intervalle séparant les ventes. Aux États-Unis, la FHFA utilise une version modifiée de son approche pondérée des ventes répétées pour construire des indices de prix trimestriels pour les maisons unifamiliales. Il peut être argué que la variance de l'erreur sera non linéaire dans les intervalles de temps (Calhoun, 1996) et, de ce fait, on régresse les résidus des MCO au carré, sur une constante, l'intervalle de temps et le carré de l'intervalle de temps.

6.14 Des études ont abouti à des résultats ambigus pour l'ajustement de l'hétéroscédasticité. Leishman, Watkins et Fraser (2002), avec des données écossaises, et Jansen *et al.* (2008), avec des données néerlandaises, ont appliqué la méthode courante avec les MCO des ventes répétées et diverses méthodes pondérées. Les deux études ont conclu que la méthode courante n'était pas moins satisfaisante.

Biais de sélection

6.15 Un problème important qui se pose avec les indices des ventes répétées est le risque de biais au niveau de la sélection de l'échantillon. Certains types de maisons peuvent en effet s'échanger plus fréquemment que d'autres sur le marché et être, de ce fait, «surreprésentés» dans l'échantillon des

ventes répétées (tant dans le cas du stock de maisons que dans celui des ventes effectuées pendant une période donnée). Quand ces types de maisons connaissent des évolutions de prix différentes, l'indice des ventes répétées est en général biaisé. Par exemple, si les maisons de médiocre qualité sont revendues plus fréquemment que celles de haute qualité et si le prix de ces dernières augmente plus lentement, l'indice des ventes répétées accusera généralement un biais positif.

6.16 Plusieurs raisons peuvent expliquer la distribution inégale de la durée de détention des biens immobiliers. L'application des théories du cycle de vie à la durée de détention des biens immobiliers conduit à penser que les maisons moins coûteuses changent de propriétaire plus souvent, les gens ayant tendance à moins changer de maison quand ils gravissent les échelons de la propriété. Les plus faibles coûts de transaction qu'impliquent les biens immobiliers moins onéreux du fait notamment des moindres droits de timbre qui leur sont appliqués peuvent aussi se traduire par un taux de rotation plus élevé pour ces biens. En outre, le marché de l'investissement locatif est plus actif, dans certains pays, dans les segments bas de gamme.

6.17 Un certain nombre d'études se sont intéressées à la question de la durée de la détention et du biais de sélection dans les indices de prix reposant sur la méthode des ventes répétées; voir, par exemple, Case, Pollakowski et Wachter (1991) (1997), Cho (1996), Clapp, Giacotto et Tirtiroglu (1991), Gatzlaff et Haurin (1997), Hwang et Quigley (2004) et enfin Steele et Goy (1997). Toutes ces études n'ont pas observé des signes flagrants de l'existence d'un biais de sélection. Clapp, Giacotto et Tirtiroglu (1991) n'ont pas constaté de différences systématiques entre l'échantillon des ventes répétées et l'échantillon complet des transactions sur le long terme. Ils ont estimé que l'arbitrage forçait généralement les prix de l'échantillon des ventes répétées à progresser au même rythme que ceux de l'échantillon complet. Wallace et Meese (1997) ont conclu que leur échantillon de ventes répétées était assez représentatif de l'ensemble des ventes effectuées pendant la période qu'il couvrait. L' «échantillon» de l'ensemble des ventes de logements peut toutefois ne pas être représentatif de l'ensemble du stock de logements.

6.18 Les problèmes potentiels liés au choix de l'échantillon sont inhérents à la méthode des ventes répétées. Il est possible d'y remédier en partie en stratifiant l'échantillon. Un problème qui se pose alors est que les sous-échantillons peuvent être trop petits et produire des indices volatils. Il peut donc être souhaitable de lisser les indices. On peut en outre avancer que les prix de ventes ne représentent pas toujours exactement les valeurs du marché des biens immobiliers, qui peuvent être considérées comme une variable latente. Un *bruit de transaction* peut provoquer la volatilité des indices de prix mesurés. Francke (2010) a suggéré une procédure de lissage qui tiendrait compte du fait que les prix de vente de biens immobiliers revendus à maintes reprises dépendent de l'intervalle entre les ventes ultérieures.

Manque d'efficacité et révision

6.19 Il est souvent reproché à la méthode des ventes répétées de manquer d'efficience du fait que, de par sa nature, elle sous-exploite les données. C'est effectivement le cas par rapport à la méthode hédonique des variables indicatrices temporelles couvrant plusieurs périodes, puisqu'elle ne considère que les logements vendus plus d'une fois et utilise, de ce fait, un ensemble de données généralement beaucoup moins important que l'échantillon des transactions effectuées sur une période donnée. En revanche, plus la période couverte par l'échantillon est longue, plus elle utilise de données (le nombre de maisons revendues augmentant en conséquence). Il s'ensuit donc que l'efficience de la méthode des ventes répétées augmente plus vite que celle de l'approche hédonique quand la période couverte par l'échantillon est plus longue et que davantage de données sont considérées. La méthode des ventes répétées est en outre efficiente en ce sens qu'elle n'utilise comme caractéristique d'un logement que son adresse.

6.20 L'ensemble de données sur lequel s'appuie la méthode des ventes répétées peut être élargi en utilisant les données sur les évaluations comme approximations de la valeur passée ou présente des maisons qui n'ont pas été revendues pendant la période couverte par l'échantillon. L'indice des ventes répétées repose dans ce cas sur des pseudo-données plutôt que sur des données authentiques. La plupart des études empiriques sur cette question s'appuient sur les évaluations de logements sur le point d'être refinancés. Il a été suggéré que les évaluations surestiment généralement le prix de vente effectif des biens immobiliers. L'importance de ce biais pourrait toutefois varier suivant la raison pour laquelle l'information sur les évaluations est recueillie. De Vries *et al.* (2009) ont examiné la fiabilité des données néerlandaises relatives aux évaluations qui sont recueillies pour le compte de l'État pour les besoins de l'impôt sur le revenu et de la fiscalité locale et ils ont conclu que la qualité de ces données était tout à faite satisfaisante et même qu'elle s'améliorait avec le temps. Pour de plus amples informations sur l'utilisation des informations concernant les évaluations dans un indice des ventes répétées et la suppression du biais d'évaluation, voir, par exemple, Geltner (1996), Edelstein et Quan (2006), et Leventis (2006).

6.21 Comme la méthode des variables indicatrices temporelles qui porte sur plusieurs périodes, la méthode des ventes répétées pâtit de la révision des indices précédemment calculés : quand de nouvelles informations sur les ventes répétées deviennent disponibles, la révision des estimations se traduit par une modification des coefficients estimés et donc des indices de prix déduits. Peu d'études empiriques ont été consacrées à cette question jusqu'à présent en dehors de celles, par exemple, de Clapp et Giaccotto (1999), Butler, Chang et Crews Cutts (2005) et Clapham *et al.* (2006). Ces derniers auteurs ont fait des observations qui leur ont donné à penser que les indices reposant sur la méthode des ventes répétées

sont relativement moins stables que les indices hédoniques à variables indicatrices temporelles. Il convient de noter que les révisions peuvent être liées au biais de sélection; il se peut que celui-ci diminue quand la période couverte par l'échantillon est plus longue et l'estimation des coefficients est revue du fait que le nombre de ventes répétées observées est plus important.

Modification de la qualité

6.22 Les indices reposant sur la méthode des ventes répétées sont estimés en postulant que la qualité des biens immobiliers (mesurée par leurs caractéristiques) ne varie pas au cours du temps. Il est parfois soutenu que, globalement, la valeur des rénovations équivaut à peu près à celle de la dépréciation. Cela ne peut toutefois être vrai pour les unités d'habitation considérées isolément, car au fil du temps, nombre d'entre elles sont démolies. Une façon d'éviter ce problème est de limiter l'échantillon des observations de ventes répétées aux unités dont la qualité est jugée relativement constante d'une période de vente à l'autre. Case et Shiller (1989), par exemple, «ont extrait […] les données sur les maisons vendues à deux reprises dont la qualité n'avait apparemment pas changé». Le problème est que les changements de prix inférés peuvent ne pas être indicatifs des changements de prix observés pour l'ensemble de l'échantillon de ventes répétées et aggraver le biais de sélection[6].

6.23 Si des informations étaient disponibles au niveau microéconomique sur les frais d'entretien et de rénovation, elles pourraient être utilisées pour estimer un modèle de régression pour le logement reposant sur la méthode des ventes répétées (ou l'approche hédonique). Dans la pratique, ce genre d'information fait cependant souvent défaut. Abraham et Schauman (1991) ont suggéré d'ajuster l'indice des ventes répétées sur la base des données globales sur les frais de rénovation, corrigées pour tenir compte de la dépréciation des bâtiments; voir aussi Palmquist (1980) (1982). Cette façon de mesurer la dépréciation nette semble toutefois trop approximative et arbitraire pour être utilisée pour l'élaboration de statistiques officielles.

6.24 Shimizu, Nishimura et Watanabe (2010) ont mis au point récemment une version de la méthode des ventes répétées qui tient compte de la dépréciation nette. Leur méthode s'appuie sur un paramètre inconnu des préférences qui doit faire l'objet d'une conjecture. S'il semble préférable de procéder à un ajustement pour tenir compte du problème de la dépréciation (nette) que de l'ignorer totalement, les organismes statistiques risquent de ne pas être tentés de se livrer à ce type de conjecture.

6.25 Shiller (1993a) a élaboré une méthode des ventes répétées qui tient compte des modifications que les caractéristiques d'un logement peuvent avoir subies entre la première et la deuxième vente. Cette méthode implique d'inclure les caractéristiques dans un modèle traditionnel des ventes répétées. Clapp et Giaccotto (1998) ont préconisé d'utiliser les valeurs estimées au moment de la première et de la deuxième vente comme une façon parcimonieuse de tenir compte des changements survenus dans la qualité de biens immobiliers. Goetzmann et Spiegel (1997) ont suggéré d'inclure un terme constant dans la régression des ventes répétées pour refléter l'évolution moyenne de la qualité de l'ensemble des caractéristiques sur la durée moyenne de la détention des biens immobiliers.

6.26 Case et Quigley (1991) ont été les premiers à conseiller l'emploi de *modèles hybrides*. Ceux-ci exploitent toutes les données sur les ventes en combinant les régressions de la méthode des ventes répétées et les régressions hédoniques et ils apportent une réponse non seulement au problème de la modification de la qualité, mais aussi à ceux du biais de sélection et du manque d'efficience. Case et Quigley (1991) et Quigley (1995) ont utilisé des échantillons de propriétés vendues une seule fois et de propriétés vendues à plusieurs reprises pour estimer conjointement les indices de prix à l'aide d'une régression effectuée selon la méthode des moindres carrés généralisés. Hill, Knight et Sirmans (1997) se sont livrés à un exercice comparable, mais de nature plus générale. Leur modèle empile deux équations, un modèle hédonique à variables indicatrices temporelles (incluant l'âge du logement) et un modèle des ventes répétées, qui sont estimées ensemble par la méthode du maximum de vraisemblance. Ces auteurs ont utilisé les prix des caractéristiques pour calculer les indices de prix; voir l'équation (5.9) du chapitre 5[7].

6.27 L'objectif des méthodes hybrides est d'essayer de combiner les meilleurs aspects de l'approche des ventes répétées et de l'approche hédonique. En combinant les deux approches, aucune donnée n'est écartée tandis que la méthode des ventes répétées continue de jouer un rôle majeur dans la construction de l'indice. Nous partageons toutefois l'avis de Hill (2011), qui a du mal à accepter qu'un ratio de prix reposant sur la méthode des ventes répétées soit préféré à un ratio de prix hédonique obtenu par (double) imputation. Il fait observer que : «si les ratios de prix reposant sur la méthode des ventes répétées ne sont pas jugés plus fiables que ceux résultant de la double imputation, il n'y a aucune raison de préférer les méthodes hybrides aux méthodes hédoniques». Au bout du compte, c'est surtout leur complexité qui risque de rendre les modèles hybrides inapplicables par les organismes statistiques.

6 Meese et Wallace (1997) signalent que les unités d'habitation ayant fait l'objet de ventes répétées et dont les caractéristiques ont changé sont généralement plus grandes et en moins bon état que la moyenne des unités n'ayant fait l'objet que d'une transaction.

7 Parmi les autres études portant sur le recours aux modèles hybrides, on peut citer celles de Clapp et Giaccotto (1992), Knight, Dombrow et Sirmans (1995), Englund, Quigley et Redfearn (1998), et Hwang et Quigley (2004).

Principaux avantages et inconvénients

6.28 Les principaux avantages et inconvénients de la méthode des ventes répétées sont exposés ci-après. Ses principaux avantages sont les suivants :

- Dans sa version de base, l'adresse est la seule information sur les caractéristiques des biens immobiliers ayant fait l'objet de plus d'une transaction pendant la période couverte par l'échantillon que requiert la méthode des ventes répétées. Cette information peut être obtenue auprès d'administrations comme celle du cadastre.
- Les régressions sont faciles à effectuer et les indices faciles à construire sur la base de cette méthode.
- C'est un type de méthode d'appariement de modèles sans imputation qui, de façon intrinsèque, neutralise automatiquement l'effet de l'emplacement.
- Ses résultats sont essentiellement reproductibles à condition que le traitement des valeurs aberrantes et les corrections éventuelles de l'hétéroscédasticité (ainsi que le choix entre la moyenne géométrique ou la moyenne arithmétique) soient clairement décrits.

6.29 Les principaux inconvénients de la méthode des ventes répétées sont les suivants :

- Elle manque d'efficience en ce sens qu'elle n'utilise pas tous les prix de transaction disponibles puisqu'elle n'utilise que les informations concernant les unités qui ont été vendues plus d'une fois pendant la période couverte par l'échantillonnage.
- Dans sa version de base, elle ignore la dépréciation (nette) des unités de logement[8].
- Les données sur les ventes répétées peuvent être entachées d'un biais de sélection.
- La méthode ne permet pas d'obtenir des indices de prix séparés pour les terrains et les bâtiments.
- Elle ne peut être utilisée pour obtenir des indices offrant une classification très fine des types de biens immobiliers vendus. Elle peut notamment ne pas permettre d'obtenir des indices mensuels des prix de l'immobilier faute de ventes sur le marché des biens appartenant à une catégorie de biens plus restreinte.
- En principe, les estimations de l'évolution passée des prix par la méthode des ventes répétées devraient être actualisées à mesure que de nouvelles informations sont disponibles sur les transactions. L'indice des prix de l'immobilier reposant sur cette méthode pourrait, de ce fait, être constamment révisé[9].

6.30 Haurin et Hendershott (1991) résument de la façon suivante les inconvénients de la méthode des ventes répétées :

«La méthode fait l'objet de nombreuses critiques : 1) elle ne sépare pas l'évolution du prix des maisons de leur dépréciation, 2) elle ignore les travaux de rénovation effectués entre les ventes, 3) l'échantillon n'est pas représentatif du stock de logements, 4) les prix des attributs peuvent évoluer au cours du temps, et 5) un grand nombre de ventes doit avoir lieu pour pouvoir constituer un échantillon acceptable de ventes répétées.» Donald R. Haurin et Patric H. Hendershott (1991; 260).

La cinquième critique de la méthode figurant dans cette citation — le grand nombre de ventes nécessaire pour obtenir un ensemble de données acceptable — n'avait pas encore été mentionnée. Nous allons, dans la prochaine section, construire un indice élémentaire des ventes répétées reposant sur la méthode des MCO à l'aide des données relatives à la ville de «A», qui a été utilisée plus tôt aux chapitres 4 et 5 pour montrer l'effet résultant de l'étroitesse d'un ensemble de données sur les ventes répétées.

Exemple d'utilisation des données relatives à la ville de «A»

6.31 Souvenons-nous qu'après avoir exclu les maisons de plus de 50 ans au moment de leur vente ainsi que les observations portant sur les terrains de plus de 1.200 m^2, nous n'avions plus que 2.289 ventes pour la période de 14 trimestres couverte par l'échantillon, comprise entre le premier trimestre de 2005 et le deuxième trimestre de 2008, c'est-à-dire que nous avions une moyenne de 163,5 ventes uniques d'unités d'habitation individuelle par trimestre pour la ville néerlandaise de «A». Quelques maisons ont été vendues deux fois au cours d'un même trimestre et nous avons éliminé ces maisons détenues pendant peu de temps pour estimer l'indice des ventes répétées (du fait qu'il pouvait s'agir de ventes effectuées en catastrophe). Nous avons fini par n'avoir que 85 ventes répétées sur la période de 14 trimestres. L'indice P_{RS} des ventes répétées calculé avec la méthode des MCO sur la base de cet ensemble restreint de données est indiqué dans le graphique 6.1 avec l'indice-chaîne moyen P_{FCH} de Fisher à échantillon stratifié décrit au chapitre 4, et l'indice P_{HIF} de Fisher à imputation hédonique décrit au chapitre 5. Ces trois séries de prix sont présentées dans le tableau 6.1.

8 Comme on l'a déjà indiqué, ce problème peut être résolu de diverses façons qui paraissent toutefois trop rudimentaires ou trop complexes pour qu'il soit envisagé d'y recourir pour la compilation de statistiques officielles.

9 Dans la pratique, cela n'est pas nécessairement très problématique. Un problème analogue se pose quand des données mensuelles saisies par lecture optique sont utilisées dans un indice des prix à la consommation (IPC); une fenêtre mobile d'observations peut être utilisée pour construire la composante d'un IPC mensuel en n'utilisant que l'accroissement du taux d'inflation au cours du mois précédent pour actualiser l'indice; voir Ivancic, Diewert et Fox (2011) ainsi que de Haan et van der Grient (2011).

Graphique 6.1. Indice des prix reposant sur la méthode des ventes répétées, indice-chaîne des prix de Fisher à échantillon stratifié et indice des prix de Fisher à imputation hédonique

Source : calculs des auteurs basés sur les données du cadastre néerlandais.

Tableau 6.1. Indice des prix reposant sur la méthode des ventes répétées, indice-chaîne de Fisher des prix moyens à échantillon stratifié et indice des prix de Fisher avec imputation hédonique

Trimestre	P_{RS}	P_{FCH}	P_{HIF}
1	1,00000	1,00000	1,00000
2	1,00650	1,02396	1,04356
3	1,02802	1,07840	1,06746
4	1,02473	1,04081	1,03834
5	1,03995	1,04083	1,04794
6	1,04206	1,05754	1,07553
7	1,08663	1,07340	1,09460
8	1,07095	1,06706	1,06158
9	1,14474	1,08950	1,10174
10	1,15846	1,11476	1,10411
11	1,12709	1,12471	1,11430
12	1,13689	1,10483	1,10888
13	1,14903	1,10450	1,09824
14	1,12463	1,11189	1,11630

Source : calculs des auteurs basés sur les données du cadastre néerlandais.

6.32 L'indice reposant sur la méthode des ventes répétées s'est avéré très instable par rapport aux deux autres indices de prix pendant la deuxième partie de la période couverte par l'échantillon. Au trimestre 14, il indique une baisse des prix alors que l'indice avec imputation hédonique et l'indice moyen reposant sur un échantillon stratifié font apparaître une hausse de prix. Nous ne pouvons certes tirer de conclusions définitives de ce simple exemple, mais celui-ci confirme toutefois que la méthode des ventes répétées nécessite un grand nombre d'observations pour que les indices de prix puissent être estimés avec une précision acceptable.

7 Méthodes basées sur l'estimation

Introduction

7.1 Comme il a été dit dans les chapitres précédents, on ne peut, pour construire des indices des prix des logements, appliquer la méthode du panier fixe, qui consiste à comparer au cours du temps les prix d'éléments identiques, et cela notamment en raison de la faible incidence des reventes et du changement de la qualité des biens vendus. La méthode des ventes répétées, évoquée au chapitre 6, tente de résoudre le problème des changements de qualité en ne retenant que des biens vendus au moins deux fois pendant la période d'échantillonnage. Toutefois, s'en tenir à ces seules observations risque d'être très inefficace, car toutes les observations sur les ventes uniques sont écartées, et pourrait aussi mener à un biais de sélection.

7.2 Dans plusieurs pays, on a des informations sur les valeurs évaluées ou les estimations des biens immobiliers qui pourraient être utiles pour représenter les prix de vente ou, plus généralement, les valeurs sur le marché. Là où les évaluations sont faites à des fins fiscales, on en dispose généralement pour tous les biens pendant une période de référence donnée. Un certain nombre d'études ont utilisé les valeurs évaluées en sus des prix de vente, dans un cadre de ventes répétées, pour atténuer le problème de l'inefficacité et l'éventualité d'un biais de sélection. Ainsi, Gatzlaff et Ling (1994) se sont servis des prix de vente comme première mesure et des évaluations comme deuxième mesure dans une régression sur les «ventes» répétées. Clapp et Giaccotto (1998) ont fait le contraire en se servant des évaluations comme première mesure et des prix de vente comme seconde. Les deux études ont conclu que ces méthodes produisaient des indices de prix similaires à un indice standard de ventes répétées.

7.3 Les méthodes d'évaluation par les ventes répétées mentionnées ci-dessus sont basées sur des rapports de pseudo-prix dans lesquels les valeurs évaluées peuvent être calculées à partir de différentes périodes. Mais, quand on dispose pour tous les biens de valeurs estimées se rapportant à une seule période d'évaluation ou date de référence, il est possible d'employer la méthodologie standard du panier fixe. Pour chaque bien vendu dans une période de comparaison, nous disposons donc d'un prix de vente et nous avions aussi un «prix» de la période de référence : la valeur évaluée. On peut alors construire des rapports de prix ayant une période de référence commune : la période d'évaluation. Ces ratios des évaluations des prix de vente peuvent ensuite être agrégés au moyen d'une formule standard de calcul d'indice, mais il arrive que cela exige un certain changement d'échelle.

7.4 L'emploi d'une formule d'indice de prix utilisant un panier fixe simplifie le calcul de l'indice, car il n'est pas nécessaire d'utiliser les techniques économétriques pour estimer l'indice ou de procéder à des ajustements pour tenir compte du changement de composition qualitative, comme c'est le cas avec les méthodes hédoniques et celle des ventes répétées : voir chapitres 5 et 6. La *méthode rapport prix de vente/évaluation* (RPVE), analysée dans le présent chapitre, a aussi pour caractéristique de ne pas appeler de révisions parce qu'il n'y a ni modélisation, ni mise en commun des données concernées. Ainsi, contrairement à la méthode des ventes répétées et à celle des variables indicatrices temporelles dans le modèle des prix hédoniques, on ne réestime pas les indices de prix déjà calculés lorsque de nouvelles observations sur les ventes deviennent disponibles.

7.5 La méthode RPVE est utilisée en Nouvelle-Zélande depuis le début des années 60 et actuellement dans plusieurs pays européens, notamment le Danemark, les Pays-Bas et la Suède. Comme peu de pays du monde l'emploient, il n'est pas surprenant que les études sur ce sujet soient rares, bien qu'en augmentation. Il semble que Bourassa, Hoesli et Sun (2006) aient été les premiers à en publier une. Selon eux, «du fait de ses avantages et de son nombre relativement limité d'inconvénients, la méthode RPVE est une solution idéale pour les agences administratives qui établissent les indices de prix des logements». Rossini et Kershaw (2006) constatent que la méthode RPVE parvient mieux que plusieurs autres à réduire la volatilité des indices hebdomadaires. Selon de Vries *et al.* (2009), les indices RPVE mensuels utilisés aux Pays-Bas sont plus précis que les indices mensuels de ventes répétées. Shi, Young et Hargreaves (2009), après avoir comparé les indices RPVE et les indices de ventes répétées en Nouvelle-Zélande, ont trouvé une assez faible corrélation mensuelle.

7.6 Quand les biens immobiliers sont réévalués et que de nouvelles estimations deviennent disponibles, l'indice RPVE peut, et doit sans doute, être rebasé. On obtient une série d'indices à long terme en chaînant les séries existantes et les nouvelles. Aux Pays-Bas, les biens sont actuellement réévalués chaque année, ce qui permet de construire un indice-chaîne annuel, la période d'évaluation (janvier) servant de nouvelle base. Selon Shi, Young et Hargreaves (2009), des réévaluations fréquentes peuvent biaiser les statistiques. De Vries *et al.* (2009) n'ont pas trouvé de biais de chaînage, mais font observer que l'erreur-type de l'indice-chaîne RPVE augmente à chaque fois que l'on introduit de nouvelles estimations parce que l'on ajoute une source supplémentaire d'erreurs d'échantillonnage.

Précisions sur la méthode RPVE

7.7 Supposons que nous disposions d'échantillons de biens vendus pour la période de début ou de base 0 et pour des périodes de comparaison t ($t = 1,..., T$). Comme dans les chapitres précédents, les échantillons sont représentés par $S(0)$ et $S(t)$. Dans chaque période, nous connaissons les prix de vente de tous les biens échantillonnés; le prix du bien n pendant la période t est représenté par p_n^t. Comme on l'a mentionné auparavant, les logements vendus au cours de la période t ne l'ayant généralement pas été au cours de la période 0, il y a un manque d'appariement. Supposons toutefois que

nous disposions de valeurs évaluées ou d'estimations pour tous les biens compris dans le stock de logements et qu'elles se réfèrent à une seule période d'évaluation. La période d'évaluation sert de période de référence et l'estimation du bien n est représentée par a_n^0. Ainsi, pour chaque bien inclus dans l'échantillon $S(t)$ de la période t, nous connaissons à la fois le prix de vente p_n^t de la période t et la valeur évaluée a_n^0 de la période de référence. En d'autres termes, pour tous les $n \in S(t)$, nous pouvons établir un rapport de prix — un *ratio de l'estimation des prix de vente* — p_n^t / a_n^0, qui peut être utilisé dans un modèle à panier fixe pour calculer un indice des prix de l'immobilier résidentiel (IPIR).

7.8 Il serait possible d'utiliser des moyennes géométriques pour agréger ces ratios d'estimations de prix, mais nous nous en tiendrons ici aux indices arithmétiques, car ils nous semblent plus appropriés au contexte des biens immobiliers. L'indice arithmétique basé sur ces estimations peut être défini comme suit :

$$P_{AP}^{0t} = \frac{\sum_{n \in S(t)} 1 p_n^t}{\sum_{n \in S(t)} 1 a_n^0} = \sum_{n \in S(t)} w_n^0(t) \left(\frac{p_n^t}{a_n^0} \right) \qquad (7.1)$$

L'expression (7.1) décrit un indice de *type Paasche,* car nous utilisons l'échantillon $S(t)$ de la période de comparaison à la fois au numérateur et au dénominateur. Chaque bien étant fondamentalement unique, les quantités sont égales à 1. La construction d'un indice de prix de type Laspeyres serait difficile, voire impossible : les informations sur les prix en période t des unités de logement appartenant à l'échantillon $S(0)$ de la période de référence ne sont disponibles que pour les quelques unités, s'il y en a, qui ont été revendues en période t. Cela signifie que l'on ne peut pas non plus construire un indice de type Fisher. Comme le montre la seconde expression, la formule (7.1) peut aussi être exprimée comme une moyenne pondérée, par des valeurs, des ratios des estimations des prix de vente p_n^t / a_n^0, où les pondérations $w_n^0(t) = a_n^0 / \sum_{n \in S(t)} a_n^0$ reflètent les parts des valeurs évaluées à la période de référence dans l'échantillon $S(t)$.

7.9 L'indice de type Paasche basé sur les estimations P_{AP}^{0t} données par la formule (7.1) est évidemment un indice à panier fixe. Il n'y a donc pas à tenir compte d'un changement de composition en comparant directement la période t à la période 0. Toutefois, comme il n'y a en général pas de chevauchement, les échantillons $S(t)$ des périodes $(t = 1,..., T)$ sont complètement différents et le changement de composition est présent d'une période à l'autre. Sachant que l'on ne peut procéder à des ajustements pour tenir compte de ces changements de composition des échantillons d'une période à l'autre, il y a une forte probabilité de volatilité à court terme. Cette caractéristique n'est pas propre à l'indice basé sur des estimations; on peut s'attendre à observer plus ou moins le même phénomène avec les indices à imputation hédonique de type Paasche évoqués au chapitre 5. La prochaine section traitera la similitude avec l'indice d'imputation de Paasche.

7.10 L'indice des prix basé sur des estimations de la formule (7.1) ne se sert pas des prix de vente observés pendant la période de référence. Il diffère donc de l'unité dans cette période, ce qui pose problème. Mais ce problème peut facilement être résolu au moyen d'une normalisation des indices consistant à les diviser par la valeur de la période de référence. Nous obtenons alors l'*indice RPVE* arithmétique suivant :

$$P_{SPAR}^{0t} = \frac{\sum_{n \in S(t)} p_n^t}{\sum_{n \in S(t)} a_n^0} \left[\frac{\sum_{n \in S(0)} p_n^0}{\sum_{n \in S(0)} a_n^0} \right]^{-1}$$

$$= \frac{\sum_{n \in S(t)} p_n^t / N(t)}{\sum_{n \in S(0)} p_n^0 / N(0)} \left[\frac{\sum_{n \in S(0)} a_n^0 / N(0)}{\sum_{n \in S(t)} a_n^0 / N(t)} \right] \qquad (7.2)$$

où $N(0)$ et $N(t)$ représentent le nombre de biens vendus au cours des périodes 0 et t (les tailles respectives des échantillons).

7.11 Avec la seconde expression figurant à la droite de (7.2), l'indice RPVE est le produit du ratio d'une moyenne des échantillons et d'un facteur de correction. Bien que la méthode RPVE utilise une méthode de produits appariés, appelée aussi méthode du panier fixe (en ce qui concerne les périodes 0 et t), le facteur de correction ajuste le ratio de la moyenne des échantillons en fonction des changements de composition du panier qui se produisent entre chaque période t et la période de référence 0. Ainsi, bien qu'il y ait lieu de s'attendre à une volatilité à court terme en raison des changements de composition d'une période à l'autre, la méthode RPVE est censée montrer moins de volatilité que le ratio de la moyenne des échantillons.

7.12 L'indice arithmétique RPVE peut être considéré comme représentant un IPIR de Paasche basé sur les ventes[1]. Mais de nombreux pays, dont des États membres de l'UE, cherchent à élaborer un indice Laspeyres plutôt qu'un indice Paasche. On peut recourir à la stratification comme moyen de calculer approximativement cet indice objectif, tout en employant la méthode RPVE. On agrège ensuite les indices RPVE (de Paasche) au niveau des strates en pondérant les parts des dépenses de la période de référence pour obtenir l'indice global «de type Laspeyres». L'IPIR des Pays-Bas est un exemple de la méthode RPVE stratifiée dans lequel les régions et les catégories de logements servent de variables de stratification. L'indice est calculé mensuellement et publié conjointement avec le cadastre néerlandais. La stratification pourrait aussi permettre de tenir compte d'éventuelles différences systématiques entre les estimations et les valeurs de marché selon les régions ou les différentes catégories de logements (de Vries *et al.,* 2009; de Haan, van der Wal et de Vries, 2009).

1 On trouve généralement dans les données administratives, en particulier celles des cadastres, l'ensemble des ventes (à l'exclusion des biens nouvellement construits) de chaque période. Du point de vue des ventes, il n'y a pas d'échantillonnage. Dans cette interprétation, l'indice RPVE ne comporte pas d'erreur d'échantillonnage, mais contient effectivement une erreur due à l'utilisation des estimations qui sont des estimations des «véritables» valeurs marchandes.

7.13 L'indice RPVE peut aussi être interprété comme un estimateur d'échantillon d'un IPIR de stock. Si, pendant chaque période, les biens vendus sont considérés comme des échantillons aléatoires du stock de logements pendant la période de référence, l'indice RPVE est un estimateur de l'IPIR de Laspeyres des stocks. Dans ce cas, il faut exclure les biens vendus qui ont été ajoutés au stock après la période de référence[2]. Comme on l'a dit dans les chapitres précédents, il est possible que l'échantillon de logements vendus ne soit pas représentatif du stock total, ce qui peut créer un biais de sélection des échantillons. La stratification est de nouveau un instrument utile pour atténuer cette difficulté.

Questions méthodologiques et pratiques

Changement de qualité

7.14 Puisque les estimations portent sur la période de référence, les biens sont en général évalués en fonction de leurs caractéristiques pendant cette période. Mais, pour que l'indice RPVE de la formule (7.1) soit un indice de prix à qualité constante, les estimations devraient être faites en fonction des caractéristiques des logements de la période de comparaison. Dès lors, si ces caractéristiques ont changé au fil du temps, la méthode RPVE ne permettra pas de procéder aux ajustements pour tenir compte de ces évolutions, à l'instar des méthodes de ventes répétées. C'est un inconvénient sérieux.

7.15 En pratique, pourtant, il pourrait y avoir un certain ajustement implicite au titre des changements de qualité. À propos de l'indice RPVE de la Nouvelle-Zélande, Bourassa, Hoesli et Sun (2006) notent : «l'estimation de base est ajustée en fonction des améliorations ultérieures du bien qui exigent un permis de construire». Si cela est fait en temps réel, des ajustements en fonction d'améliorations qualitatives importantes seront effectivement opérés. Toutefois, outre que toutes les améliorations d'un bien n'exigent pas un permis de construire, il est improbable que ces ajustements rendent bien compte de l'*effet net* des améliorations et de la dépréciation des bâtiments.

7.16 Aux Pays-Bas également, il peut y avoir un ajustement qualitatif implicite de l'indice RPVE. Les évaluations ont normalement lieu un certain temps après le mois de référence des estimations et peuvent éventuellement tenir compte d'améliorations importantes des biens. En outre, comme on l'a indiqué plus haut, elles sont maintenant effectuées chaque année. Le chaînage annuel peut en lui-même atténuer le problème du changement qualitatif si les estimations actualisées prennent bien en compte l'évolution des caractéristiques. Cela dépend bien sûr de la manière exacte dont les biens sont évalués, que les personnes établissant les indices ignorent éventuellement.

Qualité des informations tirées des évaluations

7.17 Outre les problèmes d'ajustement qualitatif, la méthode RPVE dépend évidemment de la qualité des informations dont on dispose. Il y a trois moyens principaux d'estimer des biens (ne faisant pas l'objet d'échanges) : en utilisant la régression hédonique, en les comparant à des biens similaires faisant l'objet de transactions et en recourant au jugement d'experts. Les méthodes employées diffèrent selon les pays et parfois même au sein d'un pays donné. Dans divers pays, des sociétés privées se livrent à des estimations à grande échelle. Bien que les détails des méthodes utilisées soient souvent non diffusés publiquement, il s'avère que certaines sociétés conjuguent la régression hédonique avec des informations locales ou le jugement d'experts.

7.18 Bourassa, Hoesli et Sun (2006) relèvent que les estimations effectuées en Nouvelle-Zélande sont tirées de régressions hédoniques, mais ils ne présentent malheureusement pas la méthode exacte utilisée. On explique au chapitre 5 qu'il existe différentes méthodes hédoniques et que les prix prédits, en l'occurrence les estimations, dépendent de la nature des données utilisées, du nombre d'observations, de la forme fonctionnelle spécifiée, des variables incluses et d'autres choix. Ainsi, même si la régression hédonique est la moins arbitraire des méthodes susmentionnées, il peut subsister beaucoup d'incertitudes et d'erreurs, ce qui a un effet inconnu sur les ratios des estimations des prix de vente et sur l'indice RPVE qui en résulte.

7.19 L'emploi de biens comparables semble répandu. Chinloy, Cho et Megbolugbe (1997) ont comparé pour les États-Unis un échantillon d'estimations avec les prix de vente. Ils pensaient que se référer à un nombre relativement faible de logements comparables entraîne une volatilité supérieure à celle observée pour l'ensemble des prix de vente du marché. Fait peut-être plus important encore, ils ont constaté que les estimations dépassaient les prix de vente dans à peu près 60 % des cas, d'où un biais moyen de 2 % à la hausse.

7.20 Dans les pays comme les Pays-Bas, où les évaluations officielles ont pour but de taxer l'immobilier, les valeurs obtenues ne sont peut-être pas loin de la réalité parce que l'administration est incitée à évaluer aussi largement que possible pour maximiser les recettes fiscales, alors que les contribuables souhaitent au contraire que les évaluations soient aussi basses que possible. Aux Pays-Bas,

2 Il semblerait que les biens figurant nouvellement dans le stock ne puissent même pas être utilisés faute de disposer des estimations nécessaires. Toutefois, cela dépend du système d'estimation. Si des logements anciennement locatifs ont été vendus et sont donc ajoutés au stock de logements occupés par leur propriétaire, ils ont une valeur estimée pour la période de référence dans le cas où les logements locatifs sont aussi évalués. De plus, si les impôts immobiliers sont uniformément basés sur des évaluations portant sur la période 0 pendant un certain nombre d'années, les autorités ont aussi besoin de ces évaluations pour les logements nouvellement construits. Le problème est évidemment qu'elles doivent «inventer» une valeur évaluative pour un nouveau logement dans la période 0, même s'il n'existait pas pendant cette période. Ces évaluations risquant de poser problème, il faut probablement les exclure du calcul de l'indice.

la responsabilité de l'évaluation incombe aux municipalités et les méthodes utilisées sont variables. Certaines communes comme celle d'Amsterdam, la capitale, utilisent la méthode des logements comparables, alors que d'autres se servent de méthodes apparentées aux régressions hédoniques. Selon de Vries *et al.* (2009), les autorités néerlandaises ont peut-être intérêt à ce que les évaluations ne soient pas trop élevées pour éviter les contentieux, car les ménages qui jugent la valeur estimée trop élevée peuvent engager une action en justice.

Autres questions

7.21 En comparaison avec les méthodes de régression hédoniques, la méthode RPVE a l'avantage de n'exiger des informations que sur quelques caractéristiques des biens : les évaluations (portant sur une période de référence commune), éventuellement certaines variables de stratification, et des adresses pour fusionner les bases de données si les prix de vente et les estimations proviennent de sources différentes. Aux Pays-Bas, par exemple, les prix des opérations et un nombre limité de variables de stratification sont enregistrés au cadastre, alors que les estimations émanent d'une deuxième source administrative. Il est bien connu qu'il peut être difficile de fusionner des fichiers de données par adresse, mais cela ne semble pas constituer un problème majeur aux Pays-Bas.

7.22 L'apurement des données constitue une autre question pratique importante. La méthode RPVE dépend de la qualité des estimations. Certains ratios d'estimation des prix de vente peuvent paraître peu plausibles, peut-être parce que les estimations sont jugées «erronées», et peuvent être sortis de l'ensemble des données[3]. Supprimer les observations erronées, comme les erreurs de saisie évidentes, est une bonne pratique. Toutefois, il faut faire preuve de prudence, car l'exclusion de rapports de prix peut fausser les résultats. On doit au moins formuler explicitement une règle d'exclusion des valeurs aberrantes pour informer les utilisateurs.

Interprétation des imputations basées sur des régressions

7.23 Dans cette section, nous montrons que la méthode RPVE est essentiellement une technique d'imputation dans laquelle les prix manquants de la période de référence sont estimés au moyen d'une régression linéaire des prix de vente sur les estimations. Rappelons d'abord que les prix pour la période de référence des biens appartenant à l'échantillon $S(t)$ de la période t ne peuvent être observés directement puisqu'ils ne faisaient généralement pas l'objet de transactions pendant cette période 0 de base. Nous pouvons tenter d'estimer les prix manquants pour obtenir l'indice de prix de Paasche imputé :

$$P_{IP}^{0t} = \frac{\sum_{n\in S(t)} 1 p_n^t}{\sum_{n\in S(t)} 1 \hat{p}_n^0(t)} \tag{7.3}$$

7.24 La valeur imputée $\hat{p}_n^0(t)$ dans la formule (7.3) devrait estimer le prix de la période 0 pour le bien n, évalué avec ses caractéristiques de la période t. Les (quantités des) caractéristiques doivent rester fixes pour traiter l'ajustement de qualité. L'utilisation de l'imputation *hédonique* a été analysée au chapitre 5. Les modèles de régression hédoniques expliquent le prix de vente d'un bien par un ensemble de caractéristiques du bâtiment et par l'emplacement. Cette section présente un mode d'imputation différent basé sur des régressions.

7.25 Considérons le modèle suivant de régression à deux variables pour la période de référence :

$$p_n^0 = \beta_0 + \beta_1 a_n^0 + \varepsilon_n^0 \tag{7.4}$$

L'équation (7.4) est un simple modèle *descriptif* dans lequel les prix de vente sont régressés sur les estimations. Nous supposons que ce modèle est estimé par les moindres carrés ordinaires (MCO) sur les données de l'échantillon de la période de référence $S(0)$. Les prix prédits pour $n \in S(0)$ sont :

$$\hat{p}_n^0 = \hat{\beta}_0 + \hat{\beta}_1 a_n^0 \tag{7.5}$$

où $\hat{\beta}_0$ est la constante estimée et $\hat{\beta}_1$ le coefficient de pente estimé. Nous nous attendons à trouver $\hat{\beta}_0 \cong 0$ et $\hat{\beta}_1 \cong 1$ si le système d'estimation fonctionne bien[4]. On utilisera ci-après l'équation (7.5) pour prédire les «prix manquants» dans le dénominateur de l'indice d'imputation de Paasche (7.3).

7.26 Par commodité, nous commençons par réécrire la formule (7.3) comme suit :

$$\begin{aligned} P_{IP}^{0t} &= \frac{\sum_{n\in S(t)} p_n^t / N(t)}{\sum_{n\in S(0)} p_n^0 / N(0)} \frac{\sum_{n\in S(0)} p_n^0 / N(0)}{\sum_{n\in S(t)} \hat{p}_n^0(t) / N(t)} \\ &= \frac{\sum_{n\in S(t)} p_n^t / N(t)}{\sum_{n\in S(0)} p_n^0 / N(0)} \frac{\sum_{n\in S(0)} \hat{p}_n^0 / N(0)}{\sum_{n\in S(t)} \hat{p}_n^0(t) / N(t)} \end{aligned} \tag{7.6}$$

Dans la seconde phase de (7.6), nous avons utilisé $\sum_{n\in S(0)} p_n^0 / N(0) = \sum_{n\in S(0)} \hat{p}_n^0 / N(0)$, qui s'avère exact, car la somme des résidus des régressions des MCO est égale à 0. Le premier problème qui se pose est qu'il faudrait maintenir fixes les caractéristiques des logements quand on prédit les prix de la période de référence $\hat{p}_n^0(t)$ pour $n \in S(t)$. C'est manifestement impossible en utilisant l'équation (7.5). Ainsi, la *première hypothèse* est celle d'un changement qualitatif de sorte que nous remplaçons $\hat{p}_n^0(t)$ dans (7.6) par $\hat{p}_n^0(0) = \hat{p}_n^0$.

3 L'exemple de la ville de «A» à la fin de ce chapitre montre que la suppression d'un nombre relativement faible de valeurs aberrantes peut avoir un effet substantiel sur l'indice RPVE.

4 Si les prix de vente étaient utilisés comme évaluations officielles, les valeurs 0 et 1 se vérifieraient bien sûr exactement et nous trouverions un ajustement parfait de (7.4) aux données de la période 0.

En employant (7.5) pour $n \in S(0)$ et $n \in S(t)$, l'équation (7.6) devient :

$$P_{IP}^{0t} = \frac{\sum_{n \in S(t)} p_n^t / N(t)}{\sum_{n \in S(0)} p_n^0 / N(0)} \left[\frac{\hat{\beta}_0 + \hat{\beta}_1 \sum_{n \in S(0)} a_n^0 / N(0)}{\hat{\beta}_0 + \hat{\beta}_1 \sum_{n \in S(t)} a_n^0 / N(t)} \right] \quad (7.7)$$

On remarquera qui si $\hat{\beta}_0 = 0$, c'est-à-dire si la ligne de régression passe par l'origine, (7.7) se simplifie comme l'indice RPVE de la formule (7.2), quel que soit le coefficient de pente $\hat{\beta}_1$. Donc, si l'objectif est d'estimer un indice de Paasche utilisant des imputations, la *deuxième hypothèse* qui sous-tend la méthode RPVE semble être le caractère négligeable de la constante estimée $\hat{\beta}_0$.

7.27 La *troisième hypothèse* est que l'équation (7.5) se vérifie pour $n \in S(t)$: la relation linéaire entre les prix des biens vendus à la période de référence et leurs estimations pour cette même période 0 (pour $n \in S(0)$) est censée se vérifier aussi pour les biens qui n'étaient pas vendus. Mais c'est une hypothèse très restrictive. Alors que le rapport linéaire peut être testé pour $n \in S(0)$[5], il serait difficile, voire impossible, de le tester pour $n \in S(t)$, car les prix de vente sont «manquants». La présence d'un biais d'estimation, signifiant que les estimations sous-estiment ou surestiment les valeurs de marché inconnues (les prix auxquels les biens auraient été vendus), peut fausser l'indice RPVE. Le biais de l'indice se manifeste particulièrement si la «véritable» valeur de β_1 pour $n \in S(t)$ est très différente de β_1 pour $n \in S(0)$.

7.28 Dans cette section, nous nous sommes intéressés à l'indice RPVE en tant qu'IPIR des ventes. De Haan (2010b) présente une technique voisine dans laquelle les estimations servent d'informations auxiliaires dans un cadre de «régression généralisée» (GREG), afin d'estimer un IPIR basé sur les stocks. La méthode GREG utilise les informations tirées d'estimations plutôt que celles qui sont tirées d'échantillons. L'auteur montre que l'indice RPVE est un estimateur direct de l'indice GREG basé sur les stocks qui, quand on l'applique aux données néerlandaises, s'avère presque aussi efficace.

Principaux avantages et inconvénients

7.29 On énumère ci-dessous les principaux avantages de la méthode RPVE :

- Elle s'appuie essentiellement sur la méthodologie ordinaire de modèle d'appariement et est liée à la théorie traditionnelle des indices de prix.
- La méthode donne lieu à des calculs simples.
- Pour l'appliquer, on n'a pas besoin d'informations sur les caractéristiques des logements; les seuls renseignements requis sont les données sur les prix de vente et les estimations. Dans certains pays, on trouve ces données dans des sources administratives comme les cadastres et elles englobent actuellement toutes les opérations (portant sur les biens revendus).
- Comme cette méthode utilise beaucoup plus d'observations que celle des ventes répétées, il y a moins de problèmes dus à la rareté des données. Le biais de sélection devrait notamment être inférieur. De plus, la méthode RPVE ne souffre pas de la révision des chiffres calculés antérieurement lorsque de nouvelles données deviennent disponibles.
- Sous réserve des règles d'apurement des données, la méthode RPVE est reproductible.

7.30 Les principaux inconvénients de cette méthode sont les suivants :

- Elle ne peut traiter adéquatement les changements de qualité (principales réparations ou rénovations et dépréciation) des logements[6].
- La méthode RPVE dépend de la qualité des informations sur les évaluations de la période de référence. Il arrive que leurs modalités exactes ne soient pas claires, ce qui a une incidence inconnue sur les résultats.
- La méthode ne peut distinguer les composantes «terrain» et «bâtiments» de l'indice global des prix des logements[7].

Un exemple de données pour la ville de «A»

7.31 On a calculé un indice RPVE en utilisant la série de données pour la ville de «A» présentée au chapitre 4. On se souvient qu'il s'agissait de données sur des ventes de maisons individuelles pendant 14 trimestres, commençant au premier trimestre de 2005 et finissant au deuxième trimestre de 2008. Après avoir effectué un certain apurement des données, en supprimant en particulier des logements ayant plus de 50 ans à la date de cession, il restait 2.289 ventes.

7.32 Pour calculer les indices RPVE, nous avons aussi besoin des valeurs estimées des biens vendus. Nos données d'estimations portent sur le premier trimestre (janvier) de 2005. Le rapprochement des séries de données de ventes

5 Van der Wal, ter Steege et Kroese (2006) ainsi que de Vries *et al.* (2009) comparent les estimations de l'administration néerlandaise aux prix de vente. Dans la seconde étude, le rapport linéaire (7.4) est testé explicitement (pour les biens faisant l'objet de transactions pendant le mois d'évaluation) pour divers mois d'évaluation. Il s'avère que la constante estimée est très réduite et que le coefficient de pente ne diffère pas sensiblement de 1.

6 Dans les pays où les évaluations donnent des informations distinctes sur la valeur des terrains et des bâtiments, l'indice RPVE peut en principe être ajusté en utilisant des renseignements exogènes sur la dépréciation nette des logements de la catégorie considérée.

7 Si de nouvelles informations sur l'évaluation des biens apparaissent chaque mois ou chaque trimestre, on peut s'en servir pour élaborer des indices de prix distincts des terrains et des bâtiments, à condition que les évaluations décomposent la valeur totale évaluée en ces deux composantes. Malheureusement, les évaluations officielles n'ont généralement lieu qu'une fois par an ou moins encore. Il est toutefois possible d'utiliser ces informations à basse fréquence pour vérifier les indices de prix des terrains et des bâtiments générés par les méthodes de régression hédoniques.

et de données d'estimations a produit de très bons résultats : 99,3 % des prix de vente ont pu être rapprochés des estimations correspondantes. Pour 15 observations seulement, nous n'avons pas trouvé une estimation correspondante et elles ont donc été supprimées. L'indice RPVE obtenu, P_{SPAR}, est représenté au graphique 7.1 et chiffré au tableau 7.1 conjointement à l'indice de Fisher à imputation hédonique, P_{HIF}, décrit au chapitre 5 et à l'indice des ventes répétées, P_{RS}, estimé au chapitre 6. La tendance de P_{SPAR} est très semblable à celle de P_{HIF}, mais P_{SPAR} est un peu plus volatil.

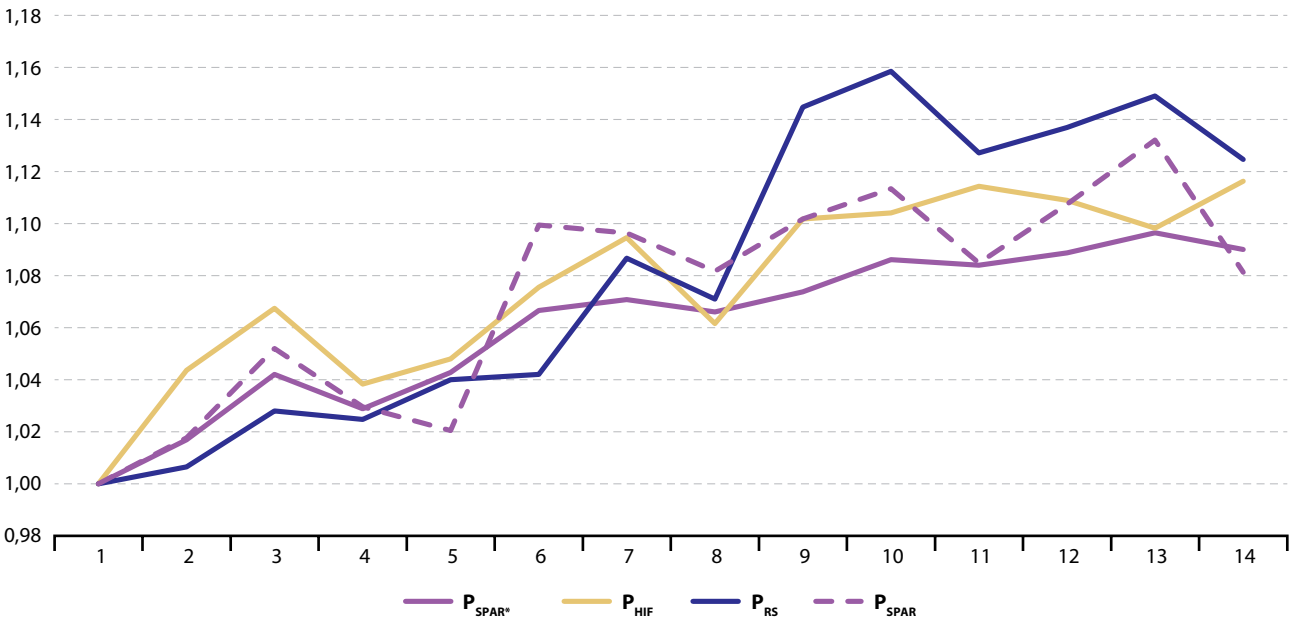
7.33 La méthode RPVE peut présenter un inconvénient : elle dépend entièrement de l'exactitude des données d'estimations. Une inspection de la distribution des ratios d'estimation des prix de vente a montré un certain nombre de valeurs aberrantes. Plus précisément, il y avait plusieurs observations avec des ratios très élevés (jusqu'à 10,5), résultant le plus souvent de valeurs d'estimation inhabituellement basses. Il est très probable qu'une proportion significative de ces valeurs aberrantes correspondait à des erreurs d'enregistrement. Nous avons donc décidé d'exclure les valeurs les plus aberrantes. Conformément aux méthodes d'apurement des données alors en vigueur à l'Office néerlandais de la statistique, basées sur la distribution du logarithme naturel des ratios d'estimation des prix de vente, on a supprimé 26 observations pour lesquelles le log du ratio différait de plus de 5 écarts-types de la moyenne[8]. Il est resté 2.248 observations.

8 La première étape de la procédure d'apurement employée par l'Office néerlandais de la statistique consistait à éliminer tous les biens dont les prix de vente ou les estimations étaient inférieurs à 10.000 euros ou supérieurs à 5.000.000 euros. Toutefois, il n'y en avait pas dans notre série de données. On notera que l'Office néerlandais de la statistique a récemment modifié les procédures de détection et d'exclusion des valeurs aberrantes.

7.34 L'indice RPVE amélioré, qualifié de P_{SPAR*}, calculé avec la série de données nettoyée, apparaît aussi au graphique 7.1 et au tableau 7.1. Comme on peut le constater, l'apurement des données a eu un effet notable sur le résultat : P_{SPAR*} est beaucoup moins volatil que l'indice P_{SPAR} qui avait été calculé avec la série de données initiale. La tendance a aussi été modifiée : P_{SPAR*} est généralement inférieur à P_{SPAR}, parce que la plupart des observations supprimées avaient des ratios d'estimation des prix de vente inhabituellement élevés. Le graphique 7.1 confirme que, si l'on utilise un ensemble de données relativement limité couvrant une courte période, la méthode RPVE donne des résultats plus crédibles que la méthode standard des ventes répétées, surtout après l'apurement des données.

7.35 Une comparaison de P_{SPAR*} avec l'indice de Fisher à imputation hédonique P_{HIF} révèle que pendant plusieurs périodes, par exemple les quatre derniers trimestres, les changements de prix ressortant des deux méthodes sont de sens opposé. En outre, P_{SPAR*} est généralement inférieur à P_{HIF}; à la fin de la période d'observation, au trimestre 14, la différence représente 0,026 point d'indice. À première vue, cela semble indiquer que P_{SPAR*} présente un biais à la baisse. Toutefois, on trouve déjà une différence du même ordre de grandeur (0,027 point) au trimestre 2. Dans ces conditions, si nous avions normalisé les deux séries pour qu'elles soient égales à 1 au trimestre 2, les deux méthodes auraient donné à peu près la même valeur d'indice au trimestre 14. C'est une illustration du *problème général de démarrage* qui se pose quand on compare des séries temporelles volatiles : le choix de la période de départ ou de base influe sur la différence moyenne pendant les périodes d'observation.

Graphique 7.1. Indices RPVE, indice de Fisher à imputation hédonique et indice des ventes répétées



Source : calculs des auteurs basés sur les données du cadastre néerlandais.

Tableau 7.1. Indices RPVE, indice de Fisher à imputation hédonique et indice des ventes répétées

Trimestre	P_{SPAR}	P_{HIF}	P_{RS}	P_{SPAR*}
1	1,00000	1,00000	1,00000	1,00000
2	1,01769	1,04356	1,00650	1,01693
3	1,05196	1,06746	1,02802	1,04204
4	1,02958	1,03834	1,02473	1,02883
5	1,02040	1,04794	1,03995	1,04273
6	1,09938	1,07553	1,04206	1,06655
7	1,09635	1,09460	1,08663	1,07076
8	1,08169	1,06158	1,07095	1,06604
9	1,10173	1,10174	1,14474	1,07378
10	1,11333	1,10411	1,15846	1,08609
11	1,08477	1,11430	1,12709	1,08396
12	1,10742	1,10888	1,13689	1,08869
13	1,13206	1,09824	1,14903	1,09642
14	1,08132	1,11630	1,12463	1,09003

Source : calculs des auteurs basés sur les données du cadastre néerlandais.

Décomposition d'un IPIR en composantes «terrain» et «bâtiments»

8

Introduction

8.1 Au chapitre 3, on a mentionné qu'il serait utile ou nécessaire, dans l'intérêt de la comptabilité nationale et de l'IPC, de décomposer l'indice des prix de l'immobilier résidentiel (IPIR) en deux composantes : un indice des prix des bâtiments, ajusté des variations de qualité et un indice de prix du terrain sur lequel le logement est édifié. Le présent chapitre indique dans les grandes lignes comment la régression hédonique peut servir à calculer cette décomposition. Les méthodes de régression hédonique ont été présentées au chapitre 5.

8.2 Le raisonnement économique permet de formuler un modèle approprié de régression hédonique. Imaginons un promoteur qui envisage de construire un bâtiment sur un espace donné. Il ou elle déterminera probablement le prix de vente du bien après l'achèvement de la construction en calculant d'abord le coût total anticipé. Ce dernier sera égal à la surface de plancher du bâtiment, disons S mètres carrés, multipliée par le coût de construction au mètre carré, par exemple γ, et majoré du coût du terrain, qui sera lui-même égal au coût du mètre carré, disons β, multiplié par la superficie du site, L. Nous suivons une *méthode de coût de production* pour modéliser le prix du bien. En d'autres termes, la forme fonctionnelle de la fonction de prix hédoniques est censée être déterminée par le côté offre du marché, c'est-à-dire par les entrepreneurs indépendants[1].

8.3 Considérons maintenant un échantillon de biens du même type général qui ont les bâtiments de surface S_n^t et les terrains L_n^t dans la période t pour $n = 1,..., N(t)$; les prix p_n^t sont égaux aux coûts de ces éléments majorés des termes d'erreur ε_n^t, lesquels sont supposés avoir une moyenne de 0. Cela aboutit au modèle suivant de régression hédonique dans la période t où β^t et γ^t sont les paramètres à estimer[2] :

$$p_n^t = \beta^t L_n^t + \gamma^t S_n^t + \varepsilon_n^t \qquad (8.1)$$

$$t = 1,..., T;\ n = 1,..., N(t)$$

La quantité de terrain L_n^t et la quantité de bâtiments S_n^t associées à la cession du bien n pendant la période t sont les deux seules caractéristiques immobilières incluses dans ce modèle très simple; les prix correspondants au cours de la période t sont le prix d'un mètre carré de terrain β^t et le prix d'un mètre carré de surface de plancher des bâtiments γ^t. Il est possible d'effectuer des régressions linéaires distinctes avec le modèle (8.1) pour chaque période t de l'échantillon.

8.4 Le «modèle des constructeurs» (8.1) se réfère essentiellement aux nouveaux logements. Pour pouvoir l'appliquer aux logements existants ou revendus, nous devons tenir compte du fait que les bâtiments anciens valent moins que les nouveaux du fait de leur *dépréciation.* Il est donc nécessaire d'obtenir des informations sur l'ancienneté du bâtiment. La prochaine section montre comment la dépréciation peut être intégrée au modèle, de façon similaire à ce qui a été fait dans les exemples concernant la ville de «A» présentés au chapitre 5. On montrera aussi comment on peut inclure des caractéristiques supplémentaires des terrains et des bâtiments en tant que variables explicatives.

Prise en compte de la dépréciation et des autres caractéristiques

La dépréciation

8.5 Supposons qu'outre les informations sur le prix de vente du bien n pendant la période t, p_n^t, sur la surface du terrain où il a été édifié, L_n^t, et sur la surface du bâtiment, S_n^t, on dispose de renseignements sur l'ancienneté du bâtiment au moment t, disons A_n^t. Dans l'hypothèse d'une dépréciation linéaire, le modèle suivant constitue une extension simple de (8.1) pour appréhender les logements «existants» :

$$p_n^t = \beta^t L_n^t + \gamma^t (1 - \delta A_n^t) S_n^t + \varepsilon_n^t \qquad (8.2)$$

$$t = 1,...,T;\ n = 1,..., N(t)$$

où le paramètre δ représente le *taux d'amortissement* (linéaire) quand l'ancienneté du bâtiment augmente d'une période. Si cette ancienneté est mesurée en nombre d'années, δ se situera probablement entre 0,5 % et 2 %. Ce chiffre sous-estimera la dépréciation «réelle», parce qu'il ne prendra pas en compte les rénovations ou adjonctions importantes. Le taux estimé de dépréciation linéaire (8.2) doit être interprété comme un *taux d'amortissement net*, c'est-à-dire un taux d'amortissement brut minoré du taux de rénovation et d'adjonction au bâtiment. Le modèle (8.2) ne fonctionne pas pour les bâtiments très anciens, car, s'ils sont encore utilisés, ils ont probablement été profondément rénovés[3].

8.6 Il convient de noter que (8.2) est un modèle de régression *non linéaire,* alors que (8.1) est un modèle de régression linéaire[4]. Le paramètre de dépréciation δ étant considéré comme fixe dans le temps, (8.2) doit être estimé comme une seule régression non linéaire durant toutes les périodes de l'échantillon, alors que le modèle (8.1) peut fonctionner comme une régression linéaire période par période. Le prix

1 McMillen (2003) évoque un modèle Cobb Douglas du côté de la demande. À propos des questions d'identification dans les modèles de régression hédonique, voir Rosen (1974).

2 À la suite de Muth (1971), Thorsnes (1997; 101) présente un modèle de coût de production du même ordre. Il fait l'hypothèse que la valeur du bien considéré dans la période t, p^t, est égale au prix de la production de logements dans la période t, ρ^t, multiplié par la quantité de cette production $H(L,K)$, la fonction de production H étant une fonction CES. Ainsi, Thorsnes suppose que $p^t = \rho^t H(L,K) = \rho^t [\alpha L\sigma + \beta K\sigma]^{1/\sigma}$ où ρ^t, σ, α et β sont les paramètres, L est la dimension de la parcelle et K le montant du capital en bâtiments (en unités à qualité constante). Notre objection à ce modèle est qu'il n'y a qu'un paramètre de temps indépendant ρ^t , alors que le notre en a deux, β^t et γ^t pour chaque t, ce qui permet aux prix des terrains et des bâtiments de varier librement d'une période à l'autre.

3 Voir par exemple Meese et Wallace (1991; 320) qui ont constaté que, dans leur modèle de régression hédonique, la variable d'ancienneté avait le mauvais signe.

4 Le modèle défini par (8.2) peut être converti en un modèle de régression linéaire.

du terrain en période t dans le modèle (8.2) est l'estimation du paramètre β^t et le prix d'une unité d'un nouveau bâtiment pendant la période t est l'estimation de γ^t. La quantité de terrain en période t pour le bien n est L_n^t, tandis que la quantité de bâtiments en période t pour le bien n, exprimée en unités équivalentes d'un nouveau bâtiment, est $(1-\delta A_n^t)S_n^t$, où S_n^t est la surface de plancher du bien n pendant la période t.

8.7 Comme les biens onéreux donnent probablement lieu à des erreurs relativement importantes en valeur absolue comparativement aux biens non onéreux, il vaudrait peut-être mieux faire l'hypothèse d'erreurs se multipliant plutôt que s'additionnant. Nous préférons toutefois spécifier un modèle additif, parce que l'objectif est de décomposer la valeur totale des logements en la *somme* des composantes «terrain» et «bâtiments»; le recours aux erreurs additives facilite cette décomposition. Quand il y a des signes tangibles d'hétéroscédasticité, il est possible d'envisager des régressions pondérées. Plusieurs chercheurs ont recommandé des modèles de régression hédonique qui permettent des décompositions additives du prix d'un bien en composantes «terrain» et «bâtiments»[5].

8.8 Le modèle des constructeurs susmentionné peut éventuellement poser un problème de multicolinéarité. Les grands bâtiments étant généralement édifiés sur de vastes surfaces, S_n^t et L_n^t pourraient être très colinéaires (autrement dit, les ratios terrains/bâtiments L_n^t/S_n^t pourraient être centrés autour d'une constante); d'où une possible instabilité des estimations des prix β^t et γ^t ajustés de la qualité du terrain et des bâtiments. Comme on le verra dans l'exemple utilisant les données concernant la ville néerlandaise de «A», les problèmes de multicolinéarité et d'instabilité existent effectivement. En général, la multicolinéarité n'est pas une difficulté majeure si l'objectif est de produire un indice global des prix des logements, mais elle est préoccupante si l'on cherche à produire des indices distincts des prix des terrains et des bâtiments. On suggérera plus loin certaines méthodes pouvant permettre de résoudre le problème de la multicolinéarité.

8.9 Le modèle de régression hédonique (8.2) impliquerait une réestimation des paramètres à chaque fois que les données portant sur une nouvelle période deviendraient disponibles. Pour surmonter cette difficulté, on pourrait appliquer une méthode des «fenêtres mobiles». On choisirait une longueur appropriée des intervalles T[6], le modèle défini par (8.2) ou (8.3) serait estimé en utilisant les données des dernières périodes T, et les séries existantes de prix du terrain et des bâtiments seraient actualisées au moyen des facteurs reliés en série β^T/β^{T-1} et γ^T/γ^{T-1}. Cette méthode est illustrée ci-dessous.

5 Voir Clapp (1980) Francke et Vos (2004), Gyourko et Saiz (2004), Bostic, Longhofer et Redfearn (2007), Diewert (2007), Francke (2008), Koev et Santos Silva (2008), Office statistique du Portugal (2009), Diewert, de Haan et Hendriks (2010) (2011) ainsi que Diewert (2010).

6 Le modèle devient un modèle de régression hédonique modifié des périodes adjacentes pour $T = 2$.

Adjonction d'autres caractéristiques

8.10 Le cadre de régression hédonique non linéaire de base présenté ci-dessus peut être généralisé pour englober l'ensemble traditionnel de caractéristiques utilisé dans les régressions hédoniques portant sur l'immobilier. Supposons que l'on puisse associer à chaque bien n faisant l'objet de transactions pendant la période t une liste de K caractéristiques $X_{n1}^t, X_{n2}^t,\ldots, X_{nK}^t$ qui déterminent le prix du terrain sur lequel le bâtiment a été édifié et une liste similaire de M caractéristiques $Y_{n1}^t, Y_{n2}^t,\ldots, Y_{nM}^t$ qui déterminent le prix de ce type de bâtiments. Les équations suivantes généralisent de la façon suivante[7] :

$$p_n^t = \beta^t\left[1+\sum_{k=1}^{K} X_{nk}^t\eta_k\right]L_n^t + \gamma^t(1-\delta A_n^t)\left[1+\sum_{m=1}^{M} Y_{nm}^t\lambda_m\right]S_n^t + \varepsilon_n^t$$
$$t=1,\ldots,T;\ n=1,\ldots,N(t) \qquad (8.3)$$

où les paramètres à estimer sont maintenant la qualité K des paramètres du terrain, $\eta_1,\ldots,\eta_K$ la qualité M des paramètres du bâtiment, $\lambda_1,\ldots,\lambda_M$, le prix qualitativement ajusté pour la période t du terrain β^tet le prix qualitativement ajusté pour la période t du bâtiment γ^t. La superficie qualitativement ajustée L_n^{t*} et la quantité correspondante de bâtiments S_n^{t*} qualitativement ajustée pour le bien n en période t sont définies comme suit :

$$L_n^{t*} \equiv \left[1+\sum_{k=1}^{K} X_{nk}^t\eta_k\right]L_n^t \qquad (8.4)$$
$$S_n^{t*} \equiv \left[1+\sum_{m=1}^{M} Y_{nm}^t\lambda_m\right]S_n^t$$
$$t=1,\ldots,T;\ n=1,\ldots,N(t)$$

8.11 Pour illustrer comment les variables X et Y peuvent être formées, considérons la liste des variables explicatives du modèle de régression hédonique de l'immobilier présentée par Li, Prud'homme et Yu (2006; 23). Les variables suivantes figurant dans cette liste peuvent être considérées comme des variables qui influencent la qualité du bâtiment; il s'agit de variables de type Y : nombre de chambres, nombre de salles de bain, nombre de garages, nombres de cheminées, ancienneté du logement, ancienneté au carré du logement, existence ou non d'une finition en briques, variables auxiliaires pour les nouveaux logements, présence ou non de parquet, chauffage au gaz naturel ou non, existence ou non d'un patio, d'un système centralisé d'aspiration, d'une piscine intérieure ou extérieure, d'un jacuzzi, d'un sauna et d'un système de climatisation. On peut supposer que les variables suivantes ont une incidence sur la qualité du terrain; il s'agit de variables de type X relatives à l'emplacement : le logement se trouve ou non à l'angle de deux rues; il se situe ou non dans un cul-de-sac; il y a ou non un centre commercial à proximité; à cela s'ajoutent diverses variables explicatives relatives à une situation géographique en banlieue.

8.12 Les équations (8.3) et (8.4) montrent comment on calcule les quantités, ajustées de la qualité, du terrain

7 Cette généralisation a été suggérée par Diewert (2007).

et des bâtiments si l'objectif est de calculer des indices de prix pour les *ventes* de biens du type inclus dans le modèle de régression hédonique. Si l'objectif est de construire des indices de prix pour le *stock* de biens du type inclus dans la régression, le choix de pondérations appropriées devient plus complexe. Ces problèmes de pondération sont évoqués dans la prochaine section.

Problèmes d'agrégation et de pondération : différences entre indices des ventes de logements et indices des stocks de logements

8.13 Comme il a été expliqué au chapitre 5, la construction d'un IPIR pour les ventes de biens immobiliers au moyen des techniques classiques de régression hédonique est très simple. Habituellement, une régression hédonique distincte du type défini par (8.3) est opérée pour chaque localité ou région d'un pays[8]. On se souvient qu'après avoir effectué une régression donnée, les prix qualitativement ajustés du terrain en période *t*, P_L^t, et des bâtiments, P_S^t, de la région considérée peuvent être définis comme suit en terme de paramètres estimés :

$$P_L^t \equiv \beta^t \qquad (8.5)$$

$$t = 1,...,T$$

$$P_S^t \equiv \gamma^t \qquad (8.6)$$

$$t = 1,...,T$$

Les quantités correspondantes de terrain et de bâtiments ajustés en fonction de la qualité pour la région, disons Q_L^t et Q_S^t, peuvent aussi être définies comme suit en termes de paramètres estimés en se référant aux définitions (8.4) susmentionnées :

$$Q_L^t \equiv \sum_{n=1}^{N(t)} L_n^{t*} = \sum_{n=1}^{N(t)} \left[1 + \sum_{k=1}^{K} X_{nk}^t \eta_k \right] L_n^t \qquad (8.7)$$

$$t = 1,...,T$$

$$Q_S^t \equiv \sum_{n=1}^{N(t)} S_n^{t*} = \sum_{n=1}^{N(t)} \left[1 + \sum_{m=1}^{M} Y_{nm}^t \lambda_m \right] S_n^t \qquad (8.8)$$

$$t = 1,...,T$$

8.14 Si des régressions hédoniques du type défini par (8.3) ont été effectuées pour les périodes *T* des données, par exemple pour les régions *R*, les opérations algébriques décrites de (8.5) à (8.8) peuvent être répétées pour chaque région *r*. Représentons les prix en résultant et les quantités pour la région *r,* qui sont les équivalents de (8.5)–(8.8), par P_{Lr}^t, P_{Sr}^t, Q_{Lr}^t et Q_{Sr}^t pour $r = 1,...,R$ et $t = 1,...,T$. Maintenant, il est possible de construire des IPIR de Fisher (des ventes) pour les terrains en se servant des données régionales de prix et de quantités du terrain, $P_L^t \equiv [P_{L1}^t,...,P_{LR}^t]$ et $Q_L^t \equiv [Q_{L1}^t,...,Q_{LR}^t]$ pour chaque période *t* ($t = 1,...,T$). De même, il est possible d'établir des IPIR de Fisher (des ventes) à partir des données de prix et de quantités des bâtiments dans chaque période *t*, $P_S^t \equiv [P_{S1}^t,...,P_{SR}^t]$ et $Q_S^t \equiv [Q_{S1}^t,...,Q_{SR}^t]$, pour $t = 1,...,T$[9].

8.15 Comme avec les méthodes de stratification, il faut maintenant se demander comment construire un IPIR pour le *stock* de biens quand on recourt aux techniques de régression hédonique. Les prix des cellules hédoniques de la période *t*, P_{Lr}^t et P_{Sr}^t, définis par les équivalents pour la région *r* de (8.5) et (8.6), peuvent encore être utilisés comme prix de cellules dans le but de construire des indices de prix des stocks du terrain et des bâtiments, mais les quantités équivalentes Q_{Lr}^t et Q_{Sr}^t, définies par (8.7) et (8.8) ne sont plus appropriées; il faut remplacer ces quantités par des estimations qui s'appliquent au *stock total d'unités de logements* de la région (ou à une autre population de référence) pour la régression *r* au moment *t*, par exemple Q_{Lr}^{t*} et Q_{Sr}^{t*}, pour $r = 1,...,R$. Ainsi, les sommations des équivalents de (8.7) et (8.8) appréhendent maintenant la totalité du stock de logements de la région *r* en période *t,* et pas seulement les unités de logement vendues pendant la période *t*. Les informations relatives à la période *t* sur la quantité de terrain L_{nr}^t pour chaque unité *n* de la région relevant du modèle de régression hédonique *m* sont maintenant nécessaires; il en va de même des informations connexes sur les caractéristiques X_{nrk}^t pour chaque caractéristique des terrains *k*, des données sur la quantité des bâtiments, S_{nr}^t, et des informations connexes sur les caractéristiques Y_{nrm}^t pour chaque caractéristique des bâtiments, *m*. Avec ces nouvelles pondérations des quantités de populations, le reste des détails de la construction de l'indice est le même que pour l'IPIR des ventes.

8.16 Pour calculer les pondérations appropriées du stock de populations en période *t*, un pays doit disposer de renseignements tirés d'un recensement du stock de logements avec assez de détails sur chaque unité de logement pour pouvoir calculer les données nécessaires sur la quantité de terrain et de bâtiments, ainsi que sur les caractéristiques connexes. À condition de disposer en temps utile d'informations relatives aux nouveaux bâtiments (complétées par les données requises sur les caractéristiques) et aux démolitions, les renseignements tirés des recensements peuvent être actualisés et des estimations pour la période *t* des quantités de terrain et de bâtiments de qualité constante, Q_{Lr}^{t*} et Q_{Sr}^{t*}

8 Il est aussi possible d'effectuer des régressions hédoniques distinctes pour diverses catégories de biens et différents emplacements. Toutefois, il arrive que les considérations de coût ne permettent pas d'appliquer un système de régressions couvrant tous les biens existants dans un pays et il y a alors seulement un échantillon de régressions hédoniques représentatives. Les problèmes d'agrégation de l'échantillon sont trop complexes pour être traités ici; les détails exacts de l'élaboration d'un indice national dépendraient de la conception de l'échantillonnage.

9 Comme c'était le cas avec les méthodes de stratification, on pourrait construire des indices-chaînes ou à base fixe. Il serait également possible d'effectuer des régressions hédoniques à fenêtres mobiles; la méthode des fenêtres mobiles sera expliquée plus loin.

, peuvent être calculées approximativement en temps utile. Il est donc possible d'établir des IPIR des stocks de terrains et de bâtiments en utilisant les indices de Fisher, comme on l'a fait pour l'IPIR des ventes. Si l'on ne dispose pas à temps de données sur les nouveaux bâtiments et les démolitions, on ne peut que construire des indices de prix à base fixe de type Laspeyres en utilisant les pondérations de quantités du dernier recensement des logements.

8.17 En l'absence de toute information provenant de recensements (ou si les données concernant les caractéristiques des unités de logement manquent), il reste possible de calculer approximativement des IPIR du terrain et des bâtiments au moyen des techniques de régression hédonique. Si les données sur les caractéristiques des biens vendus pendant chaque période sont conservées durablement, on peut construire une distribution approximative par type d'unités de logement. Ces renseignements peuvent ensuite servir à calculer approximativement un IPIR des stocks selon les modalités exposées plus haut.

Principaux avantages et inconvénients

8.18 Cette section résume les principaux avantages et inconvénients du recours aux méthodes de régression hédonique pour construire un IPIR des composantes «terrain» et «bâtiments». Les principaux avantages sont les suivants :

- Si la liste des caractéristiques disponibles des biens est suffisamment détaillée, la méthode permet un ajustement en fonction à la fois des changements de dosage de l'échantillon et des changements qualitatifs de chaque logement.
- On peut établir des indices de prix pour différentes catégories de logements et d'emplacements au moyen d'une stratification appropriée de l'échantillon. La stratification présente un certain nombre d'avantages supplémentaires.
- La méthode de régression hédonique est la plus efficace pour utiliser les données disponibles.
- C'est aussi pratiquement la seule qui permet de décomposer l'indice global des prix en ses composantes «terrain» et «bâtiments».

8.19 Les principaux inconvénients de la régression hédonique sont les suivants :

- Elle exige de disposer de données sur toutes les caractéristiques pertinentes du bien (en particulier l'ancienneté, la catégorie et l'emplacement des biens de l'échantillon ainsi que des informations sur la dimension du bâtiment et du terrain), ce qui rend son application relativement coûteuse.
- Elle risque de ne pas aboutir à des résultats raisonnables à cause de problèmes de multicolinéarité.
- Bien qu'elle soit naturellement reproductible, on peut faire des choix différents à propos de l'ensemble des caractéristiques entrées dans la régression, de la forme fonctionnelle du modèle, de la spécification stochastique, des transformations possibles de la variable de dépendance, etc., qui peuvent aboutir à des estimations diverses de l'évolution globale des prix.
- On comprend aisément l'idée générale sur laquelle repose la méthode hédonique, mais certains aspects techniques ne sont pas simples à expliquer aux utilisateurs.

Application aux données sur la ville de «A» : modalités préliminaires

8.20 Nous allons maintenant illustrer les techniques générales expliquées dans ce chapitre en utilisant l'ensemble des données concernant la ville néerlandaise de «A», qui a été décrit à la fin du chapitre 4. Nous disposons de données sur les ventes de maisons individuelles pour 14 trimestres, à partir du premier trimestre de 2005. Souvenons-nous de la notation mathématique employée ci-dessus ainsi qu'aux chapitres 4 et 5 : $N(t)$ ventes de maisons individuelles ont eu lieu au trimestre t, p_n^t étant le prix de vente de la maison n. On dispose d'informations sur trois caractéristiques : la superficie du terrain en mètres carrés, L_n^t; la surface de plancher du bâtiment en mètres carrés, S_n^t; le nombre de décennies d'ancienneté de la maison n pendant la période t, A_n^t.

Un exemple simple

8.21 Le modèle simple de régression hédonique défini par (8.2), estimé avec cet ensemble de données, est répété ici par souci de commodité :

$$p_n^t = \beta^t L_n^t + \gamma^t (1-\delta A_n^t) S_n^t + \varepsilon_n^t \qquad (8.9)$$
$$t = 1,...,14;\ n = 1,...,N(t)$$

Les paramètres à estimer sont β^t (c'est-à-dire le prix du terrain au trimestre t), γ^t (le prix des bâtiments de qualité constante au trimestre t) et δ (le taux d'amortissement commun à tous les trimestres). Le modèle (8.9) comprend 14 paramètres β^t inconnus, 14 paramètres γ^t inconnus et un paramètre δ inconnu, soit 29 paramètres inconnus au total[10].

8.22 Le R^2 de ce modèle était égal à 0,8847, soit le chiffre le plus élevé à ce jour pour les régressions utilisant l'ensemble

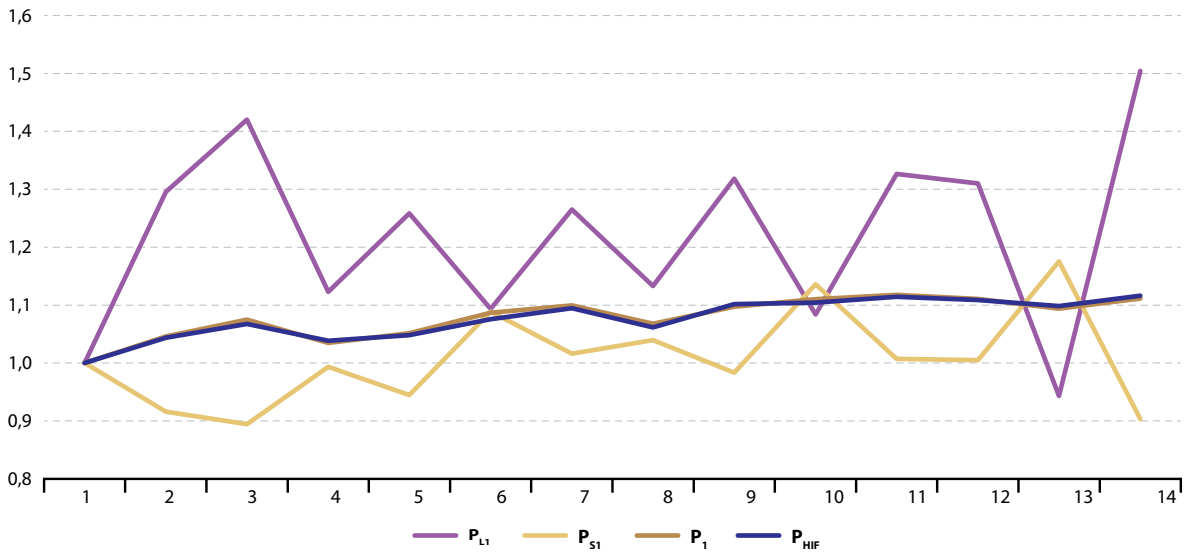
10 La structure du modèle (8.9) est similaire à celle du modèle à imputation hédonique décrit auparavant, sauf que le présent modèle est plus parcimonieux; il n'y a qu'un taux d'amortissement, au lieu de 14 dans le modèle d'imputation défini par les équations (5.25), et il n'y a pas de constante. L'élément important des deux modèles est que les prix du terrain et des structures ajustées en fonction de la qualité peuvent varier librement dans le temps.

de données concernant la ville de «A». Le log de vraisemblance était −10.642,0, ce qui est très supérieur à ceux des deux régressions sur des variables de temps qui utilisaient les prix comme variable de dépendance; on se souvient des résultats des régressions associés à la construction des indices P_{H4} et P_{H5} définis au chapitre 5 où le log de vraisemblance était −10.790,4 et −10.697,8. L'estimation du taux net de dépréciation linéaire décennale était 0,1068 (0,00284).

8.23 La série estimée de prix du terrain $\hat{\beta}^1,...,\hat{\beta}^{14}$ (reformulée pour être égale à 1 au trimestre 1), désignée comme P_{L1}, et la série de prix qualitativement ajustés des bâtiments $\hat{\gamma}^1,...,\hat{\gamma}^{14}$ (également reformulée), désignée comme P_{S1}, sont représentées au graphique 8.1 et figurent au tableau 8.1. En utilisant ces séries de prix et les données de quantités correspondantes pour chaque trimestre t — c'est-à-dire la quantité de terrains ayant fait l'objet de transactions, $L^t \equiv \sum_{n=1}^{N(t)} L_n^t$, et la quantité de bâtiments de qualité constante, $S^{t*} \equiv \sum_{n=1}^{N(t)} (1-\hat{\delta}A_n^t)S_n^t$ — on a construit un indice général des prix de l'immobilier au moyen de la formule de Fisher. Cet indice global, appelé P_1, figure également au graphique 8.1 et au tableau 8.1. Pour permettre des comparaisons, on présente aussi l'indice de Fischer **à** imputation hédonique du chapitre 5, P_{HIF}.

8.24 On peut constater que le nouvel indice global des prix hédoniques, basé sur une approche de la forme fonctionnelle hédonique par le coût de production, P_1, est très proche de l'indice de Fischer à imputation hédonique, P_{HIF}. Toutefois, la série de prix du terrain, P_{L1}, et celle des bâtiments ajustés en fonction de la qualité, P_{S1}, ne sont pas du tout crédibles : les deux présentent d'importantes fluctuations aléatoires. On remarque que, quand le prix du terrain monte brusquement, il y a une baisse correspondante du prix des bâtiments. C'est un signe clair de multicolinéarité entre les variables du terrain et celles des bâtiments ajustés en fonction de la qualité, ce qui entraîne une grande instabilité des estimations des prix des terrains et des bâtiments.

Graphique 8.1. Prix du terrain (P_{L1}), prix des bâtiments ajustés en fonction de la qualité (P_{S1}), indice du coût total de la production de logements (P_1) et indice de Fischer à imputation hédonique pour les logements



Source : calculs des auteurs basés sur les données du cadastre néerlandais.

Tableau 8.1. Prix du terrain (P_{L1}), prix des bâtiments ajustés en fonction de la qualité (P_{S1}), indice du coût total de la production de logements (P_1) et indice de Fischer à imputation hédonique pour les logements

Trimestre	P_{L1}	P_{S1}	P_1	P_{HIF}
1	1,00000	1,00000	1,00000	1,00000
2	1,29547	0,91603	1,04571	1,04356
3	1,42030	0,89444	1,07482	1,06746
4	1,12290	0,99342	1,03483	1,03834
5	1,25820	0,94461	1,05147	1,04794
6	1,09346	1,08879	1,08670	1,07553
7	1,26514	1,01597	1,09941	1,09460
8	1,13276	1,03966	1,06787	1,06158
9	1,31816	0,98347	1,09713	1,10174
10	1,08366	1,13591	1,11006	1,10411
11	1,32624	1,00699	1,11782	1,11430
12	1,30994	1,00502	1,11077	1,10888
13	0,94311	1,17530	1,09373	1,09824
14	1,50445	0,9032	1,11147	1,11630

Source : calculs des auteurs basés sur les données du cadastre néerlandais.

L'utilisation de splines de degré 1 (lignes brisées)

8.25 Le prix du mètre carré du terrain a tendance à diminuer pour les superficies importantes. Afin de tenir compte de ce phénomène, on utilisera une *modélisation par une ligne brisée du prix du terrain*[11]. Pour les surfaces inférieures à 160 m², on suppose que le coût du mètre carré du terrain est β_S^t au trimestre t. Pour les biens situés sur des surfaces comprises entre 160 et 300 m², on suppose que le coût du terrain change pour atteindre β_M^t par mètre carré supplémentaire au trimestre t. Enfin, pour les surfaces supérieures à 300 m², le prix marginal d'une unité supplémentaire de terrain est fixé à β_L^t par mètre carré au trimestre t. Représentons les ensembles de ventes de surfaces réduites, moyennes et importantes respectivement par $S_S(t)$, $S_M(t)$ et $S_L(t)$, pour $t = 1,...,14$. S'agissant des ventes n de biens dans la catégorie des superficies réduites au trimestre t, le modèle de régression hédonique est donné par (8.10); pour la catégorie des superficies moyennes, il est donné par (8.11) et, pour la catégorie des superficies importantes, il est donné par (8.12) :

$$p_n^t = \beta_S^t L_n^t + \gamma^t (1 - \delta A_n^t) S_n^t + \varepsilon_n^t \qquad (8.10)$$

$$t = 1,...,14;\ n \in S_S(t)$$

$$p_n^t = \beta_S^t [160] + \beta_M^t [L_n^t - 160] + \gamma^t (1 - \delta A_n^t) S_n^t + \varepsilon_n^t \qquad (8.11)$$

$$t = 1,...,14;\ n \in S_M(t)$$

$$p_n^t = \beta_S^t [160] + \beta_M^t [140] + \beta_L^t [L_n^t - 300] + \gamma^t (1 - \delta A_n^t) S_n^t + \varepsilon_n^t$$

$$t = 1,...,14;\ n \in S_L(t) \qquad (8.12)$$

8.26 Dans l'estimation du modèle défini par (8.10)–(8.12) sur les données relatives à la ville de «A», le taux estimé de dépréciation décennale était $\hat{\delta} = 0{,}1041$ (0,00419). Le R^2 de ce modèle était 0,8875, soit une hausse par rapport au précédent modèle sans la modélisation par ligne brisée dont le R^2 ressortait à 0,8847. La probabilité logarithmique était −10.614,2 (soit une augmentation de 28 par rapport à celle du modèle précédent). Les valeurs des paramètres de la première période pour les trois prix marginaux du terrain étaient $\hat{\beta}_S^1 = 281{,}4$ (55,9), $\hat{\beta}_M^1 = 380{,}4$ (48,5) et $\hat{\beta}_L^1 = 188{,}9$ (27,5). En d'autres termes, au trimestre 1, le coût marginal au mètre carré est estimé à 281,4 euros pour les petites surfaces, à 380,4 euros/m² pour les surfaces moyennes et à 188,9 euros/m² pour les surfaces importantes. La valeur du paramètre de la première période pour les bâtiments ajustés en fonction de la qualité est $\hat{\gamma}^1 = 978{,}1$ euros/m², avec un écart-type de 82,3. La plus faible statistique t du total des 57 paramètres étant 3,3, tous les coefficients estimés dans ce modèle sont significativement différents de zéro.

8.27 Après l'estimation des paramètres du modèle, la valeur prédite des ventes de terrains de taille réduite, moyenne et grande, respectivement V_{LS}^t, V_{LM}^t et V_{LL}^t, dans chaque trimestre t peut être calculée comme suit en même temps que les quantités correspondantes de terrain : L_{LS}^t, L_{LM}^t et L_{LL}^t :

$$V_{LS}^t \equiv \sum_{n \in S_S(t)} \hat{\beta}_S^t L_n^t \qquad (8.13)$$

$$t = 1,...,14$$

11 Cette méthode s'inspire de celle de Diewert, de Haan et Hendriks (2010) (2011). L'usage de lignes brisées pour modéliser les non-linéarités des prix du terrain comme une fonction de la superficie du terrain est tiré de Francke (2008).

$$V^t_{LM} \equiv \sum_{n \in S_M(t)} \{\hat{\beta}^t_S[160] + \hat{\beta}^t_M[L^t_n - 160]\} \quad (8.14)$$

$t = 1,...,14$

$$V^t_{LL} \equiv \sum_{n \in S_L(t)} \{\hat{\beta}^t_S[160] + \hat{\beta}^t_M[140] + \hat{\beta}^t_L[L^t_n - 300]\} \quad (8.15)$$

$t = 1,...,14$

$$L^t_{LS} \equiv \sum_{n \in S_S(t)} L^t_n \quad (8.16)$$

$t = 1,...,14$

$$L^t_{LM} \equiv \sum_{n \in S_M(t)} L^t_n \quad (8.17)$$

$t = 1,...,14$

$$L^t_{LL} \equiv \sum_{n \in S_L(t)} L^t_n \quad (8.18)$$

$t = 1,...,14$

On définit les *prix trimestriels moyens* correspondants, P^t_{LS}, P^t_{LM} et P^t_{LL}, des trois catégories de surfaces comme les valeurs ci-dessus divisées par les quantités ci-dessus :

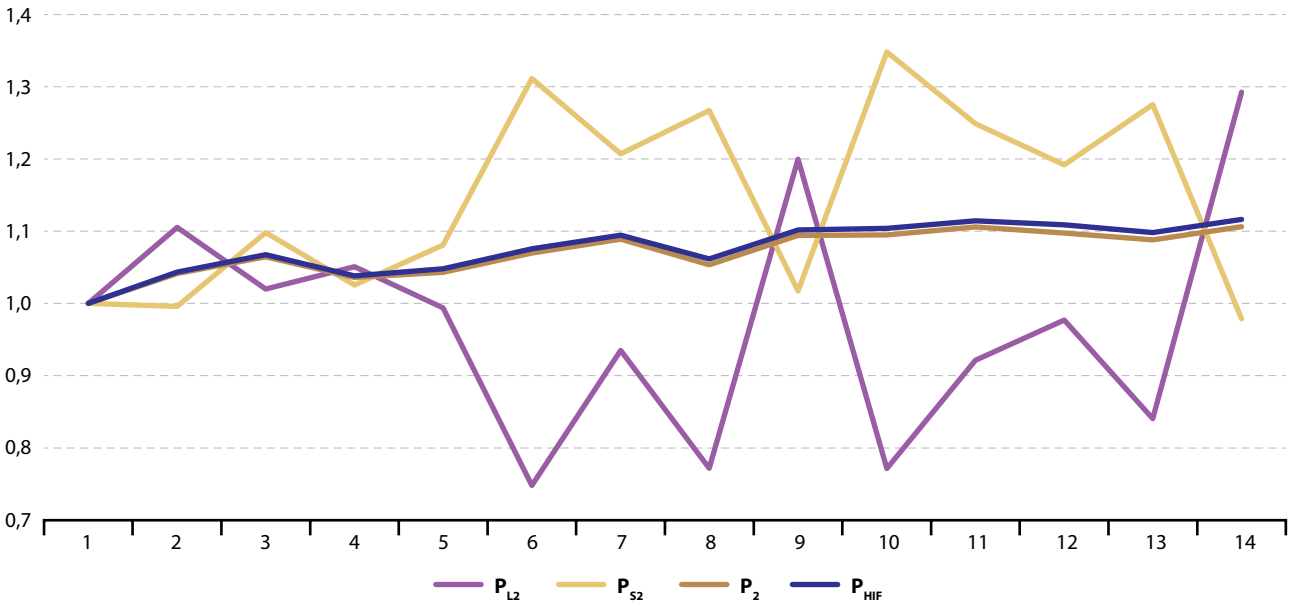
$$P^t_{LS} \equiv V^t_{LS} / L^t_{LS}; \; P^t_{LM} \equiv V^t_{LM} / L^t_{LM}; \; P^t_{LL} \equiv V^t_{LL} / L^t_{LL} \quad (8.19)$$

$t = 1,...,14$

8.28 Les prix moyens du terrain pour les surfaces petites, moyennes et grandes, définis par l'équation (8.19), ainsi que les quantités correspondantes de terrains définies par (8.16)–(8.18) peuvent servir à construire un indice-chaîne de Fischer des prix du terrain représenté par P_{L2}. Cet indice figure au graphique 8.2 et au tableau 8.2. Comme auparavant, le prix estimé au trimestre t du mètre carré de bâtiments ajustés en fonction de la qualité est $\hat{\gamma}^t$, tandis que la quantité de bâtiments à qualité constante est donnée par $S^{t*} \equiv \sum_{n=1}^{N(t)} (1 - \hat{\delta} A^t_n) S^t_n$. Le prix des bâtiments ainsi que les séries de quantités $\hat{\gamma}^t$ et S^{t*} ont été conjugués aux trois séries de prix du terrain et des quantités pour former un indice-chaîne global de Fisher des prix des logements, P_2, qui figure aussi au graphique 8.2 et au tableau 8.2. L'indice des prix des bâtiments à qualité constante P_{S2} (qui est une normalisation de la série $\hat{\gamma}^1,...,\hat{\gamma}^{14}$) est également présenté.

8.29 L'indice global des prix de l'immobilier résultant de la modélisation avec des lignes brisées, P_2, est très proche de l'indice de Fisher à imputation hédonique P_{HIF}. Toutefois, cette modélisation avec des lignes brisées ne génère pas des séries vraisemblables du prix du terrain, P_{L2} et du prix des bâtiments, P_{S2} : les deux séries sont extrêmement volatiles, mais de sens opposé. À l'instar du précédent modèle basé sur

Graphique 8.2. Prix du terrain (P_{L2}), prix des bâtiments (P_{S2}), indice global des prix de l'immobilier utilisant la modélisation par lignes brisées du terrain (P_2) et indice de Fisher à imputation hédonique



Source : calculs des auteurs basés sur les données du cadastre néerlandais.

Tableau 8.2. Prix du terrain (P_{L2}), prix des structures (P_{S2}), indice global des prix de l'immobilier utilisant la modélisation par lignes brisées du terrain (P_2) et indice de Fisher à imputation hédonique

Trimestre	P_{L2}	P_{S2}	P_2	P_{HIF}
1	1,00000	1,00000	1,00000	1,00000
2	1,10534	0,99589	1,04137	1,04356
3	1,02008	1,09803	1,06465	1,06746
4	1,05082	1,02542	1,03608	1,03834
5	0,99379	1,08078	1,04294	1,04794
6	0,74826	1,31122	1,06982	1,07553
7	0,93484	1,20719	1,08912	1,09460
8	0,77202	1,26718	1,05345	1,06158
9	1,19966	1,01724	1,09425	1,10174
10	0,77139	1,34813	1,09472	1,10411
11	0,92119	1,24884	1,10596	1,11430
12	0,97695	1,19188	1,09731	1,10888
13	0,84055	1,27531	1,08811	1,09824
14	1,29261	0,97875	1,10613	1,11630

Source : calculs des auteurs basés sur les données du cadastre néerlandais.

les coûts de production, ce modèle pâtit aussi d'un problème de multicolinéarité.

8.30 En comparant les graphiques 8.1 et 8.2, on s'aperçoit que, dans le premier, l'indice du prix du terrain dépasse la plupart du temps l'indice global et que l'indice du prix des bâtiments lui est inférieur, alors que, dans le second, cette tendance *s'inverse*. Cette instabilité est encore une fois un signe de multicolinéarité. Dans la section suivante, on tente de remédier à ce problème en imposant des restrictions de monotonie aux prix des bâtiments à qualité constante.

Méthode basée sur les restrictions de monotonie

8.31 Il est probable que les coûts de construction néerlandais n'ont pas baissé sensiblement pendant la période de l'échantillon[12]. S'il en est vraiment ainsi, des *restrictions de monotonie* sur les prix trimestriels des bâtiments qualitativement ajustés, $\gamma^1, \gamma^2, \gamma^3, \ldots, \gamma^{14}$, peuvent être imposées au modèle de régression hédonique (8.10)–(8.12) en remplaçant les paramètres γ^t des prix des bâtiments à qualité constante au trimestre t par la séquence suivante de paramètres pour les 14 trimestres : γ^1, $\gamma^1 + (\phi^2)^2$, $\gamma^1 + (\phi^2)^2 + (\phi^3)^2, \ldots$, $\gamma^1 + (\phi^2)^2 + (\phi^3)^2 + \ldots + (\phi^{14})^2$, où $\phi^2, \phi^3, \ldots, \phi^{14}$ sont des paramètres scalaires[13]. Pour chaque trimestre t à partir du trimestre 2, le prix au mètre carré des bâtiments de qualité constante γ^1 est donc égal au prix de la période antérieure γ^{t-1} majoré du carré d'un paramètre ϕ^t, $(\phi^t)^2$. Remplaçons maintenant cette nouvelle modélisation du prix des bâtiments γ^t dans (8.10)–(8.12), de façon à obtenir une modélisation par lignes brisées du prix du terrain avec des restrictions de monotonie sur le prix des structures de qualité constante.

8.32 Quand on a appliqué ce nouveau modèle en utilisant les données concernant la ville néerlandaise de «A», le taux d'amortissement décennale estimé était $\hat{\delta} = 0{,}1031$ (0,00386). Le R^2 de ce modèle était 0,8859, en baisse par rapport à la précédente modélisation par lignes brisées sans restrictions où le R^2 était 0,8875. Le log de vraisemblance était −10.630,5, en baisse de 16,3 par rapport au précédent modèle dépourvu de restrictions. Huit des treize nouveaux paramètres ϕ^t sont égaux à zéro dans cette régression hédonique avec restrictions de monotonie. Les valeurs des paramètres de la première période pour les trois prix marginaux du terrain sont $\hat{\beta}^1_S = 278{,}6$ (37,2), $\hat{\beta}^1_M = 380{,}3$ (41,0) et $\hat{\beta}^1_L = 188{,}0$; ces valeurs sont presque identiques aux estimations correspondantes du précédent modèle sans restrictions. L'estimation du paramètre de la première période pour les bâtiments ajustés en fonction de la qualité est $\hat{\gamma}^1 = 980{,}5$ (49,9) euros/m², ce qui diffère peu de l'estimation antérieure sans restrictions de 978,1 euros/m².

8.33 Une fois que les paramètres du modèle ont été estimés, il faut convertir les paramètres ϕ^t estimés en paramètres $\hat{\gamma}^t$ estimés au moyen des équations récursives suivantes :

$$\hat{\gamma}^t \equiv \hat{\gamma}^{t-1} + (\hat{\phi}^t)^2 \qquad (8.19)$$

$$t = 2,\ldots,14$$

12 La section suivante présentera certaines preuves directes à l'appui de cette affirmation.

13 Cette méthode consistant à imposer des restrictions de monotonie a été employée par Diewert, de Haan et Hendriks (2010) à une différence près : ils ont imposé la monotonie à la fois au prix des bâtiments et à celui du terrain, alors qu'ici les restrictions de monotonie ne portent que sur le prix des bâtiments.

Utilisons maintenant les équations (8.13)–(8.19) de la section précédente pour construire un indice-chaîne de Fischer des prix du terrain qui est représenté par P_{L3}. Cet indice figure au graphique 8.3 et au tableau 8.3. Comme dans les deux modèles antérieurs, le prix estimé en période t d'un mètre carré de bâtiments qualitativement ajustés est $\hat{\gamma}^t$ et la quantité correspondante de bâtiments à qualité constante est $S^{t*} \equiv \sum_{n=1}^{N(t)} (1-\hat{\delta}A_n^t)S_n^t$. Les séries de prix et de quantités $\hat{\gamma}^t$ et S^{t*} ont été conjuguées aux trois séries de prix du terrain et des quantités pour construire un indice-chaîne global de Fischer des prix de l'immobilier P_3, qui figure aussi au graphique 8.3 et au tableau 8.3. On peut aussi trouver au graphique 8.3 et au tableau 8.3 l'indice des prix des bâtiments à qualité constante P_{S3} (une normalisation de la série $\hat{\gamma}^1,...,\hat{\gamma}^{14}$).

8.34 Le nouvel indice global des prix de l'immobilier, P_3, figurant au graphique 8.3, qui a imposé des restrictions de monotonie aux prix des bâtiments ajustés en fonction de la qualité, ne peut guère être distingué du précédent indice global, P_2, qui était basé sur un modèle similaire de régression hédonique, si ce n'est que l'évolution des prix des bâtiments n'était pas soumise à des restrictions. Les fluctuations des prix du terrain et des bâtiments ajustés en fonction de la qualité ne sont plus brutales.

8.35 Bien que les résultats ci-dessus semblent «raisonnables», la rapidité de la hausse des prix des bâtiments au début, puis sa lenteur du trimestre 6 au trimestre 14 ne sont pas très vraisemblables. Dans la section suivante, on essaie donc une nouvelle méthode pour extraire des données sur les ventes immobilières, les composantes distinctes du bâtiment et du terrain.

Méthode basée sur des informations exogènes relatives aux prix des bâtiments

8.36 Il existe dans de nombreux pays de nouveaux indices des prix de la construction disponibles trimestriellement; c'est par exemple le cas aux Pays-Bas[14]. Si l'on veut faire l'hypothèse que les coûts de construction des logements augmentent au même rythme pendant la période étudiée dans toutes les villes de ce pays, les informations relatives à ces coûts peuvent servir à éliminer les problèmes de multicolinéarité rencontrés dans les sections précédentes.

8.37 Revenons aux équations (8.10)–(8.12) présentées plus haut. Ce sont les équations d'estimation du modèle de régression hédonique sans restrictions, basé sur les coûts de production. Dans cette section, les paramètres de prix des bâtiments à qualité constante, γ^t pour $t=2,...,14$ dans (8.10)–(8.12) sont remplacés par les chiffres suivants qui retiennent le seul paramètre inconnu γ^1 [15] :

$$\gamma^t = \gamma^1 \mu^t \qquad (8.20)$$

$$t = 2,...,14$$

où μ^t est l'*indice du coût de la construction* établi par l'agence statistique pour l'emplacement et le type de logements considérés, normalisé pour être égal à 1 au trimestre 1. Le nouveau modèle de régression hédonique est de nouveau défini par les équations (8.10)–(8.12), à l'exception du fait que les 14 paramètres γ^t inconnus sont maintenant définis par (8.20), de sorte qu'il faut seulement estimer γ^1. Le nombre de paramètres à estimer dans ce nouveau modèle assorti de restrictions est 44, alors que l'ancien nombre était 57.

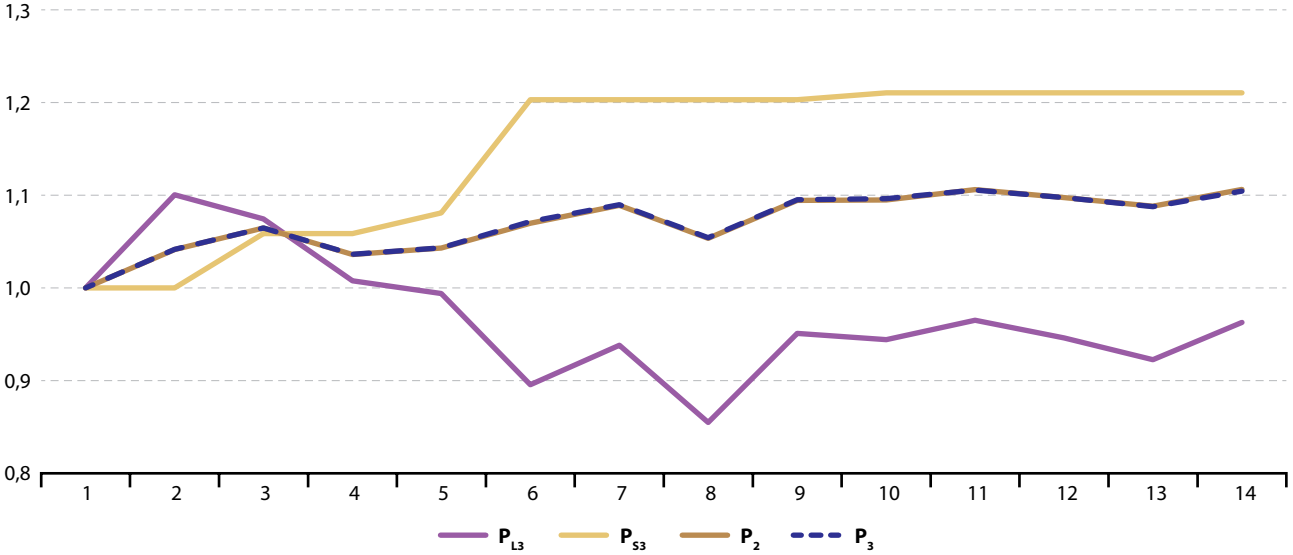
8.38 Quand on a utilisé les données concernant la ville de «A», le taux d'amortissement décennale estimé était $\hat{\delta} = 0{,}1028$ (0,00433). Le R^2 de ce modèle était 0,8849, soit une légère baisse par rapport à la précédente modélisation par lignes brisées soumise à des restrictions, où le R^2 était 0,8859, et une baisse plus marquée par rapport au R^2 de la modélisation par lignes brisées sans restriction de la section 8.5, qui était 0,8875. La probabilité logarithmique s'établissait à −10.640,1, en diminution de 10 par rapport au modèle sujet à des restrictions de monotonie. Les estimations des paramètres de la première période pour les 3 prix marginaux du terrain sont maintenant $\hat{\beta}_S^1 = 215{,}4$ (30,0), $\hat{\beta}_M^1 = 362{,}6$ (46,7) et $\hat{\beta}_L^1 = 176{,}4$ (28,4). Elles diffèrent légèrement des chiffres précédents. L'estimation du paramètre de la première période pour les bâtiments qualitativement ajustés est $\hat{\gamma}^1 = 1.085{,}9$ (22,9) euros/m^2, soit sensiblement plus que l'estimation sans restriction de 980,5 euros/m^2. L'imposition d'un taux d'augmentation (à l'échelle nationale) à la variation du prix des bâtiments qualitativement ajustés a donc eu un certain effet sur les estimations des niveaux de prix du terrain et des bâtiments.

8.39 Comme de coutume, on s'est servi des équations (8.13)–(8.19) pour construire un indice-chaîne de Fisher des prix du terrain qui est représenté par P_{L4}. Cet indice figure au graphique 8.4 et au tableau 8.4. Comme pour les trois modèles précédents, le prix estimé au trimestre t d'un mètre carré de bâtiments qualitativement ajustés est $\hat{\gamma}^t$ (maintenant égal à $\hat{\gamma}^1\mu^t$) et la quantité correspondante est $S^{t*} \equiv \sum_{n=1}^{N(t)} (1-\hat{\delta}A_n^t)S_n^t$. Ces séries de prix des bâtiments et des quantités ont de nouveau été conjuguées aux trois séries de prix du terrain et des quantités pour former un indice global de Fisher des prix de l'immobilier, P_4, qui figure au graphique 8.4 et au tableau 8.4. On présente aussi l'indice de prix des bâtiments à qualité constante P_{S4} (une normalisation de la série $\hat{\gamma}^1,...,\hat{\gamma}^{14}$).

14 Sur le site Internet de l'Office néerlandais de la statistique (2010), Statline, on a téléchargé la série suivante du nouvel indice des prix de la construction de logements pour les 14 trimestres de notre échantillon de ventes : 98,8; 98,1; 100,3; 102,7; 99,5; 100,5; 100,0; 100,3; 102,2; 103,2; 105,6; 107,9; 110,0; 110,0. Cette série a été normalisée à 1 au premier trimestre en divisant chaque entrée par 98,8. La série ainsi obtenue est représentée par μ^1 (=1), $\mu^2,...,\mu^{14}$.

15 La technique suggérée ici de décomposition des prix de l'immobilier en une composante «terrain» et une composante «bâtiments» peut être considérée comme une variante d'une technique employée par Davis et Heathcote (2007) ainsi que par Davis et Palumbo (2008).

Graphique 8.3. Prix du terrain (P_{L3}), prix des bâtiments ajustés en fonction de la qualité (P_{S3}), indice global des prix de l'immobilier avec restrictions de monotonie sur les structures (P_3) et indice global des prix de l'immobilier utilisant la méthodologie des lignes brisées pour le terrain (P_2)



Source : calculs des auteurs basés sur les données du cadastre néerlandais.

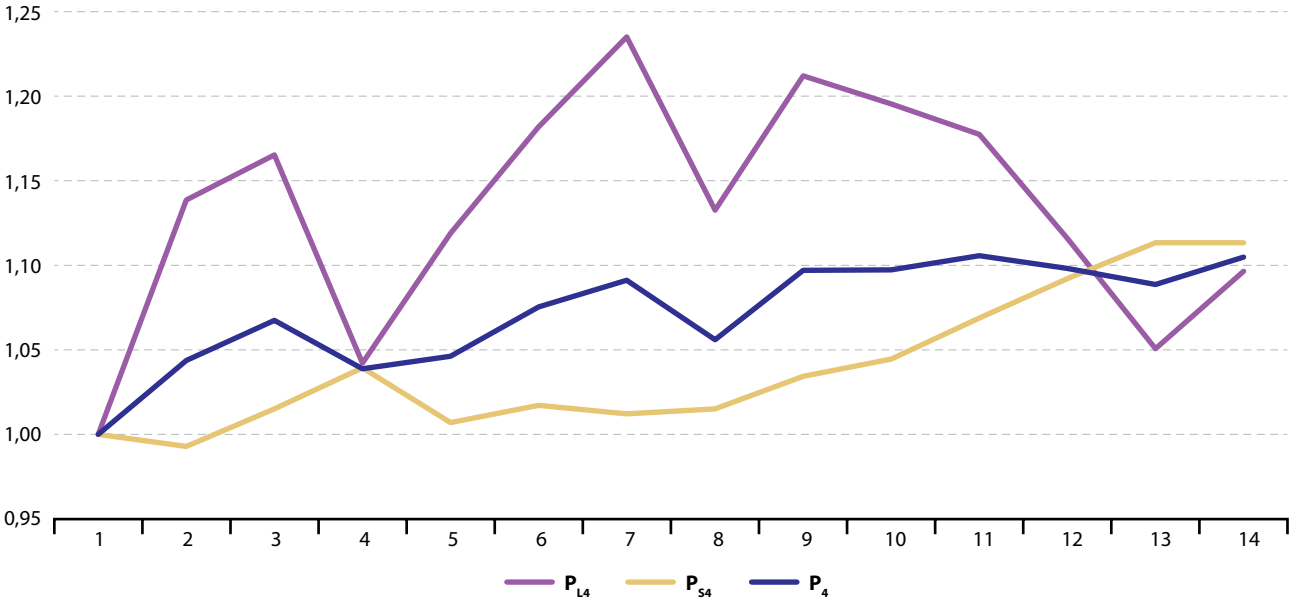
Tableau 8.3. Prix du terrain (P_{L3}), prix des bâtiments ajustés en fonction de la qualité (P_{S3}), indice global des prix de l'immobilier avec restrictions de monotonie sur les structures (P_3) et indice global des prix de l'immobilier utilisant la méthodologie des lignes brisées pour le terrain (P_2)

Trimestre	P_{L3}	P_{S3}	P_3	P_2
1	1,00000	1,00000	1,00000	1,00000
2	1,10047	1,00000	1,04148	1,04137
3	1,07431	1,05849	1,06457	1,06465
4	1,00752	1,05849	1,03627	1,03608
5	0,99388	1,08078	1,04316	1,04294
6	0,89560	1,20300	1,07168	1,06982
7	0,93814	1,20300	1,08961	1,08912
8	0,85490	1,20300	1,05408	1,05345
9	0,95097	1,20300	1,09503	1,09425
10	0,94424	1,21031	1,09625	1,09472
11	0,96514	1,21031	1,10552	1,10596
12	0,94596	1,21031	1,09734	1,09731
13	0,92252	1,21031	1,08752	1,08811
14	0,96262	1,21031	1,10427	1,10613

Source : calculs des auteurs basés sur les données du cadastre néerlandais.

8.40 Une comparaison des graphiques 8.3 et 8.4 montre que l'imposition de taux nationaux d'augmentation aux coûts de construction de logements neufs change la nature des indices de prix des bâtiments et du terrain : au graphique 8.3, la série de prix du terrain se situe au-dessous de la série de prix globaux de l'immobilier pendant la plus grande partie de la période de l'échantillon, alors qu'au graphique 8.4, la tendance est inversée : la série de prix du terrain se situe au-dessus de la série de prix globaux de l'immobilier pendant la plus grande partie de la période de l'échantillon (et vice-versa pour le prix des bâtiments). Mais quel est le meilleur modèle? On peut certes préférer le précédent d'un point de vue statistique, parce que la probabilité logarithmique est quelque peu supérieure; nous préférerions néanmoins le présent modèle qui utilise des informations exogènes sur les prix des bâtiments, parce qu'il en résulte une configuration plus plausible de l'évolution des prix du terrain et des bâtiments.

Graphique 8.4. Prix du terrain (P_{L4}), prix des bâtiments ajustés en fonction de la qualité (P_{S4}) et indice global des prix de l'immobilier utilisant des informations exogènes sur le prix des bâtiments (P_4)



Source : calculs des auteurs basés sur les données du cadastre néerlandais.

Tableau 8.4. Prix du terrain (P_{L4}), prix des bâtiments ajustés en fonction de la qualité (P_{S4}) et indice global des prix de l'immobilier utilisant des informations exogènes sur le prix des bâtiments (P_4)

Trimestre	P_{L4}	P_{S4}	P_4
1	1,00000	1,00000	1,00000
2	1,13864	0,99291	1,04373
3	1,16526	1,01518	1,06752
4	1,04214	1,03947	1,03889
5	1,11893	1,00709	1,04628
6	1,18183	1,01721	1,07541
7	1,23501	1,01215	1,09121
8	1,13257	1,01518	1,05601
9	1,21204	1,03441	1,09701
10	1,19545	1,04453	1,09727
11	1,17747	1,06883	1,10564
12	1,11588	1,09211	1,09815
13	1,05070	1,11336	1,08863
14	1,09648	1,11336	1,10486

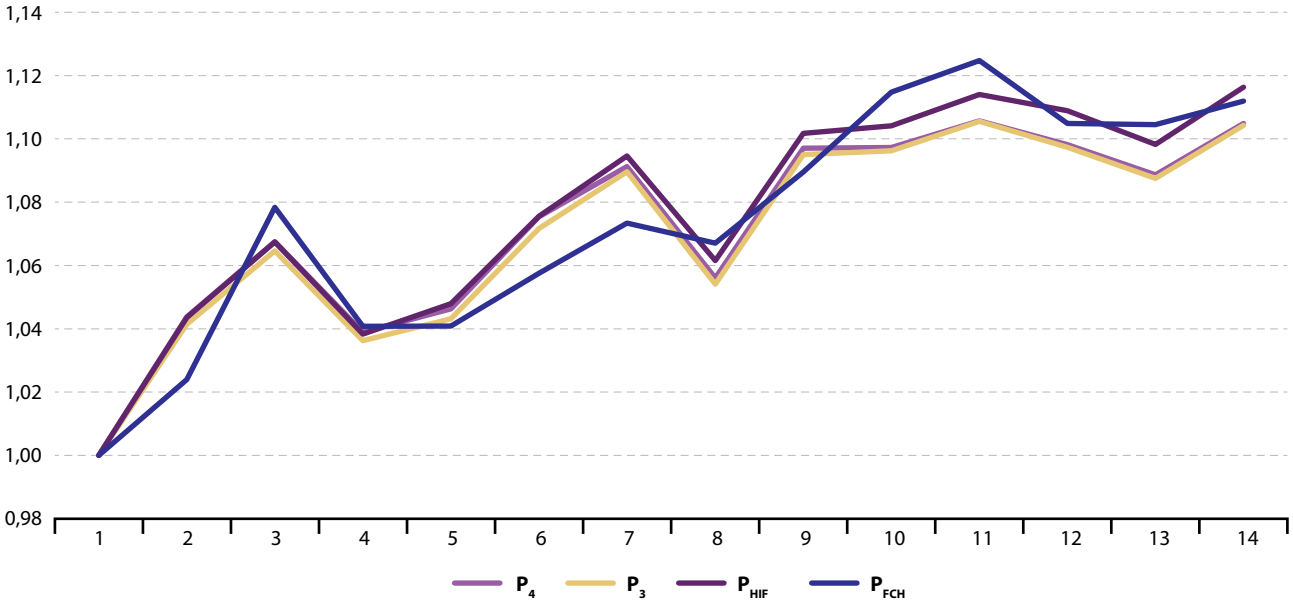
Source : calculs des auteurs basés sur les données du cadastre néerlandais.

Choix du «meilleur» indice global

8.41 On conclut cette section en énumérant et en représentant graphiquement nos quatre «meilleurs» indices *globaux* : l'indice-chaîne de Fisher **à** échantillon stratifié, P_{FCH}, élaboré au chapitre 4; l'indice-chaîne de Fisher **à** imputation hédonique P_{HIF} étudié au chapitre 5; l'indice P_3 résultant du modèle de régression hédonique basé sur les coûts, avec restrictions de monotonie, élaboré précédemment; l'indice P_4, issu du modèle de régression hédonique basé sur les coûts et utilisant des informations exogènes sur le prix des bâtiments étudié dans cette section. Comme on peut le constater à la lecture du graphique 8.5, ces quatre indices donnent des résultats très proches. On remarque que P_3 et P_4 sont quasiment identiques.

8.42 Tout bien considéré, l'indice à imputation hédonique P_{HIF} a notre préférence, car il est sujet à moins de restrictions que les autres et semble le plus proche dans son esprit d'un indice à panier fixe; il est suivi des deux indices hédoniques basés sur les coûts de production, P_4 et P_3, qui sont eux-mêmes suivis de l'indice d'échantillon stratifié P_{FCH}. Ce dernier est probablement affecté d'un certain biais de valeurs unitaires. Les indices hédoniques peuvent également être biaisés (si d'importantes variables explicatives sont omises ou si l'on choisit une forme fonctionnelle «incorrecte»), mais en général nous préférons les méthodes de régression hédonique aux méthodes de stratification. Si l'on a besoin d'indices distincts portant sur le terrain et les bâtiments, nous sommes en faveur du modèle de régression hédonique basé sur les coûts qui utilise des informations exogènes sur le prix des bâtiments.

Graphique 8.5. Indices des prix de l'immobilier utilisant des informations exogènes (P_4) et des restrictions de monotonie (P_3), indice-chaîne de Fisher à imputation hédonique et indice-chaîne de Fisher à échantillon stratifié



Source : calculs des auteurs basés sur les données du cadastre néerlandais.

Tableau 8.5. Indices des prix de l'immobilier utilisant des informations exogènes (P_4) et des restrictions de monotonie (P_3), indice-chaîne de Fisher à imputation hédonique et indice-chaîne de Fisher à échantillon stratifié

Trimestre	P_4	P_3	P_{HIF}	P_{FCH}
1	1,00000	1,00000	1,00000	1,00000
2	1,04373	1,04148	1,04356	1,02396
3	1,06752	1,06457	1,06746	1,07840
4	1,03889	1,03627	1,03834	1,04081
5	1,04628	1,04316	1,04794	1,04083
6	1,07541	1,07168	1,07553	1,05754
7	1,09121	1,08961	1,09460	1,07340
8	1,05601	1,05408	1,06158	1,06706
9	1,09701	1,09503	1,10174	1,08950
10	1,09727	1,09625	1,10411	1,11476
11	1,10564	1,10552	1,11400	1,12471
12	1,09815	1,09734	1,10888	1,10483
13	1,08863	1,08752	1,09824	1,10450
14	1,10486	1,10427	1,11630	1,11189

Source : calculs des auteurs basés sur les données du cadastre néerlandais.

Régressions hédoniques à fenêtres mobiles

8.43 Le chapitre 5 a mentionné un problème soulevé par le modèle de régression hédonique présenté à la section précédente (et par tous les autres modèles hédoniques évoqués dans ce manuel à l'exception des modèles à imputation hédonique) : quand on ajoute de nouvelles données, les indices générés par le modèle changent. Du fait de cette caractéristique des méthodes fondées sur les régressions, les modèles en question ne peuvent pas être employés par les agences statistiques, car leurs utilisateurs attendent des chiffres officiels qu'ils ne soient pas modifiés avec le temps. Ils peuvent certes, tolérer quelques révisions des données récentes, mais n'apprécieraient pas en général des révisions indéfinies des chiffres passés au fur et à mesure que de nouvelles données deviennent disponibles. Il existe toutefois une solution simple à ce problème, qualifiée de *méthode des fenêtres mobiles*. Cette méthode sera précisée et appliquée au modèle de régression hédonique basé sur les coûts qui utilise des informations exogènes sur le prix des bâtiments.

8.44 On commence par choisir un nombre «convenable» de périodes (égal à deux ou plus) dans lesquelles il y a lieu de penser que le modèle hédonique produit des résultats «raisonnables»; ce sera la *longueur de l'intervalle* (disons M périodes) pour la séquence de modèles de régression à estimer. Ensuite, on estime un modèle de régression initial et on calcule les indices appropriés en se servant des données concernant les premières périodes M de l'ensemble de données. Enfin, on estime un second modèle de régression où les données sont les données initiales moins les données de la période 1, mais en ajoutant les données de la période M+1. On calcule les indices de prix appropriés à ce nouveau modèle de régression, en ne retenant toutefois que le taux d'augmentation de l'indice entre la période M et la période M+1, pour actualiser la séquence précédente des valeurs de l'indice M. Cette procédure est de nouveau appliquée à chaque régression successive en sortant les données de la période antérieure la plus proche et en entrant celles de la période suivante, avec adjonction d'un nouveau facteur d'actualisation pour chaque régression. Si la longueur de l'intervalle est d'un an, cette procédure est qualifiée de *modèle de régression hédonique sur année glissante*; pour une longueur de l'intervalle quelconque, on parle de *modèle de régression hédonique à fenêtres mobiles*[16].

8.45 En s'appuyant sur les données concernant la ville de «A», on a appliqué la méthode à fenêtres mobiles avec une longueur d'intervalle de 9 trimestres. Le modèle de régression hédonique défini par les équations (8.10)–(8.12) et (8.20) a d'abord été estimé pour les 9 premiers trimestres. Les indices de prix qui en résultent pour le terrain et les bâtiments à qualité constante ainsi que l'indice global sont représentés par P_{RWL4}, P_{RWS4} et P_{RW4} ; ils figurent aux 9 premières lignes du tableau 8.6[17]. On a ensuite effectué une régression couvrant les trimestres 2 à 10, et les indices des prix du terrain ainsi que des bâtiments en résultant ont servi à actualiser les indices

16 Cette procédure a été récemment employée par Shimizu, Nishimura et Watanabe (2010) ainsi que par Shimizu, Takatsuji, Ono et Nishimura (2010) dans leurs modèles de régression hédonique des prix de l'immobilier à Tokyo. Une procédure analogue a aussi été récemment mise en œuvre par Ivancic, Diewert et Fox (2011) ainsi que par de Haan et van der Grient (2011) dans leur adaptation de la méthode GEKS pour faire des comparaisons internationales avec des données scannées.

17 Comme nous avons imposé les restrictions (33) aux régressions à fenêtres mobiles, l'indice à fenêtres mobiles du prix des bâtiments de qualité constante, P_{RWS4}, est égal à l'indice du prix de ces bâtiments figurant au tableau 8.4, P_{S4}.

initiaux; autrement dit, le prix du terrain au trimestre 10 du tableau 8.6 est égal au prix du terrain au trimestre 9 multiplié par le prix relatif du terrain (indice des prix du terrain au trimestre 10 divisé par le même indice au trimestre 9) obtenu au moyen de la régression couvrant les trimestres 2 à 10, etc. On a procédé à une actualisation similaire pour les quatre trimestres suivants à partir des régressions couvrant les trimestres 3–11, 4–12, 5–13 et 6–14.

8.46 En consultant le tableau 8.6, les indices à fenêtres mobiles peuvent être comparés aux indices correspondants basés sur les données relatives aux 14 trimestres établis à la section précédente. On se souvient que le taux d'amortissement estimé et le prix estimé au trimestre 1 des bâtiments qualitativement ajustés pour le dernier modèle étaient respectivement $\hat{\delta} = 0{,}1028$ et $\hat{\gamma}^1 = 1.085{,}9$. Si par hasard les 6 régressions hédoniques à fenêtres mobiles généraient exactement les mêmes estimations de δ et γ, les indices résultant des régressions à fenêtres mobiles coïncideraient alors avec les indices P_{L4}, P_{S4} et P_4. Les estimations de δ générées par les 6 régressions à fenêtres mobiles sont 0,10124; 0,10805; 0,11601; 0,11103; 0,10857 et 0,10592. Les estimations de γ^1 générées par les 6 régressions à fenêtres mobiles sont 1.089,6; 1.103,9; 1.088,1; 1.101,0; 1.123,5 et 1.100,9. Bien que ces estimations ne soient pas identiques aux estimations correspondantes de 0,1028 et 1.085,9 pour P_4, elles en sont très proches. Nous pouvons donc nous attendre à ce que les indices à fenêtres mobiles soient proches de leurs équivalents du dernier modèle de la section précédente. Les valeurs du R^2 pour les 6 régressions à fenêtres mobiles étaient 0,8803; 0,8813; 0,8825; 0,8852; 0,8811 et 0,8892.

8.47 La série à fenêtres mobiles de prix des bâtiments ajustés en fonction de la qualité, P_{RWS}, ne figure pas au tableau 8.6, car elle est identique à la série P_{S4}[18]. La série de prix à fenêtres mobiles du terrain, P_{RWL}, est extrêmement proche de son équivalent P_{L4}, tandis que la série à fenêtres mobiles du prix global des maisons individuelles dans la ville de «A», P_{RW}, est également proche de son équivalent P_4. Les séries correspondantes du tableau 8.6 sont si proches que nous avons décidé de ne pas les représenter sur un graphique.

8.48 Les régressions hédoniques à fenêtres mobiles utilisant les données concernant la ville de «A» ont produit à peu près les mêmes résultats qu'une régression hédonique utilisant l'ensemble de la période couverte par l'échantillon. Cela confirme notre idée selon laquelle les agences statistiques peuvent recourir à la méthode des fenêtres mobiles pour calculer un IPIR fondé sur des régressions hédoniques, incluant une décomposition en composantes «terrain» et «bâtiments».

18 Par construction, P_{S4} et P_{RWS} sont tous deux égaux à l'indice officiel de prix de la construction de nouveaux logements établi par l'Office néerlandais de la statistique, μ^t/μ^1 for $t = 1,...,14$.

Tableau 8.6. Prix du terrain (P_{L4}), prix des bâtiments ajustés en fonction de la qualité (P_{S4}), indice global des prix de l'immobilier utilisant des informations exogènes sur le prix des bâtiments (P_4) et leurs équivalents à fenêtres mobiles (P_{RWL}) et (P_{RW})

Trimestre	P_{RWL}	P_{L4}	P_{RW}	P_4	P_{S4}
1	1,00000	1,00000	1,00000	1,00000	1,00000
2	1,14073	1,13864	1,04381	1,04373	0,99291
3	1,16756	1,16526	1,06766	1,06752	1,01518
4	1,04280	1,04214	1,03909	1,03889	1,03947
5	1,12055	1,11893	1,04635	1,04628	1,00709
6	1,18392	1,18183	1,07542	1,07541	1,01721
7	1,23783	1,23501	1,09123	1,09121	1,01215
8	1,13408	1,13257	1,05602	1,05601	1,01518
9	1,21417	1,21204	1,09698	1,09701	1,03441
10	1,19772	1,19545	1,09738	1,09727	1,04453
11	1,18523	1,17747	1,10718	1,10564	1,06882
12	1,11889	1,11588	1,09779	1,09815	1,09201
13	1,05191	1,05070	1,08893	1,08863	1,11335
14	1,09605	1,09648	1,10436	1,10486	1,11335

Source : calculs des auteurs basés sur les données du cadastre néerlandais.

Construction des indices de prix du stock de logements

8.49 Cette section montre comment on peut utiliser les modèles de régression hédonique pour élaborer un IPIR approximatif du *stock* de logements. Nous commençons par examiner le modèle à imputation hédonique présenté au chapitre 5 et comparons l'indice qui en résulte à un indice approximatif de stock utilisant la méthode de stratification.

Le modèle à imputation hédonique

8.50 On se souvient que le modèle à imputation hédonique était défini par les équations (5.25), dans lesquelles L_n^t, S_n^t et A_n^t représentaient respectivement le terrain, les bâtiments et l'ancienneté (en décennies) d'un bien n vendu en période t. Pour établir un indice de prix du stock de maisons individuelles de la ville de «A», il faudrait en principe connaître L, S et A pour toutes les maisons individuelles de cette ville pendant une certaine période de référence. Faute de disposer de ces informations, nous pouvons considérer le nombre total de maisons individuelles cédées pendant la période d'échantillonnage comme une *approximation* du stock de cette catégorie de logements[19]. Dans notre ensemble de données, ces transactions étaient au nombre de $N(1)+N(2)+...+N(14)=2.289$[20].

8.51 Les paramètres estimés pour la dimension du terrain, celle du bâtiment et pour la dépréciation au trimestre t sont représentés par $\hat{\beta}^t$, $\hat{\gamma}^t$ et $\hat{\delta}^t$; $\hat{\alpha}^t$ représente la constante. Notre approximation de la *valeur totale du stock de maisons* au trimestre t, V^t, est définie comme suit :

$$V^t \equiv \sum_{s=1}^{14}\sum_{n=1}^{N(s)}[\hat{\alpha}^t + \hat{\beta}^t L_n^s + \hat{\gamma}^t(1-\hat{\delta}^t A_n^s)S_n^s] \qquad (8.21)$$

$$t = 1,...,14$$

V^t est (approximativement) la valeur imputée de toutes les maisons vendues pendant les 14 trimestres de notre échantillon, sachant que les coefficients de régression du modèle à imputation hédonique au trimestre t donné par (5.25) servent à pondérer les caractéristiques de chaque maison. Notre première estimation de l'indice de prix du stock de maisons, $P_{Stock\,1}$, pour la ville de «A» s'obtient en divisant la série V^t par la valeur au trimestre 1, V^1[21]. C'est une forme d'indice de Lowe; on consultera le manuel de l'IPC (2004) pour connaître les propriétés des indices de Lowe. Le tableau 8.7 et le graphique 8.6 comparent cet indice de prix du stock de maisons à l'indice correspondant à imputation hédonique de Fisher basé sur les ventes, P_{HIF}.

8.52 On trouve aussi au graphique 8.6 et au tableau 8.7 un autre indice approximatif de prix du stock de logements fondé sur la *stratification*, $P_{Stock\,2}$. Cet indice utilise les prix des valeurs unitaires pour les strates non vides dans chaque trimestre, comme il est expliqué au chapitre 4, et les prix imputés basés sur les régressions à imputation hédonique du chapitre 5 pour strates vides de chaque trimestre. Le vecteur de quantité employé pour $P_{Stock\,2}$ est le vecteur de la quantité totale (de l'échantillon) par strates, ce qui fait de $P_{Stock\,2}$ un indice de prix de Lowe de substitution. On voit que, si $P_{Stock\,2}$ a la même tendance générale que $P_{Stock\,1}$ et P_{HIF}, il diverge sensiblement de ces indices à imputation hédonique pendant plusieurs trimestres. Les différences sont dues à l'existence d'un certain biais des valeurs unitaires dans ces indices de stratification. Ainsi, bien que l'on puisse construire des indices de cette nature pour le stock de logements d'un certain type et d'un certain emplacement (avec l'aide de l'imputation hédonique pour les cellules vides), les indices de stocks qui en résultent ne s'avèrent pas aussi exacts que ceux qui reposent entièrement sur l'emploi des régressions hédoniques[22].

Utilisation d'informations exogènes sur le prix des bâtiments

8.53 Le même mode de construction d'un indice approximatif de prix de stock peut être appliqué aux autres modèles de régression hédonique présentés dans ce chapitre. Nous montrons ici ce que cela donne pour le modèle fondé sur les coûts qui utilisait des informations exogènes sur le prix des bâtiments. Ce modèle était défini par les équations (8.10)–(8.12) et (8.20). On se souvient qu'en période t les ensembles de ventes de maisons bâties sur des surfaces réduites, moyennes et grandes étaient représentés respectivement par $S_S(t)$, $S_M(t)$ et $S_L(t)$; le nombre total de ventes en période t était représenté par $N(t)$ pour $t=1,...,14$. Les paramètres estimés du modèle sont $\hat{\delta}^t$, $\hat{\gamma}^t$ et $\hat{\beta}^1_S$, $\hat{\beta}^1_M$, et $\hat{\beta}^1_L$ pour $t=1,...,14$. Les estimations des *valeurs en période t de toutes les maisons bâties sur des surfaces petites, moyennes et grandes,* vendues pendant les 14 trimestres, soit respectivement V^t_{LS}, V^t_{LM} et V^t_{LL}, sont définies par (8.22)–(8.24) :

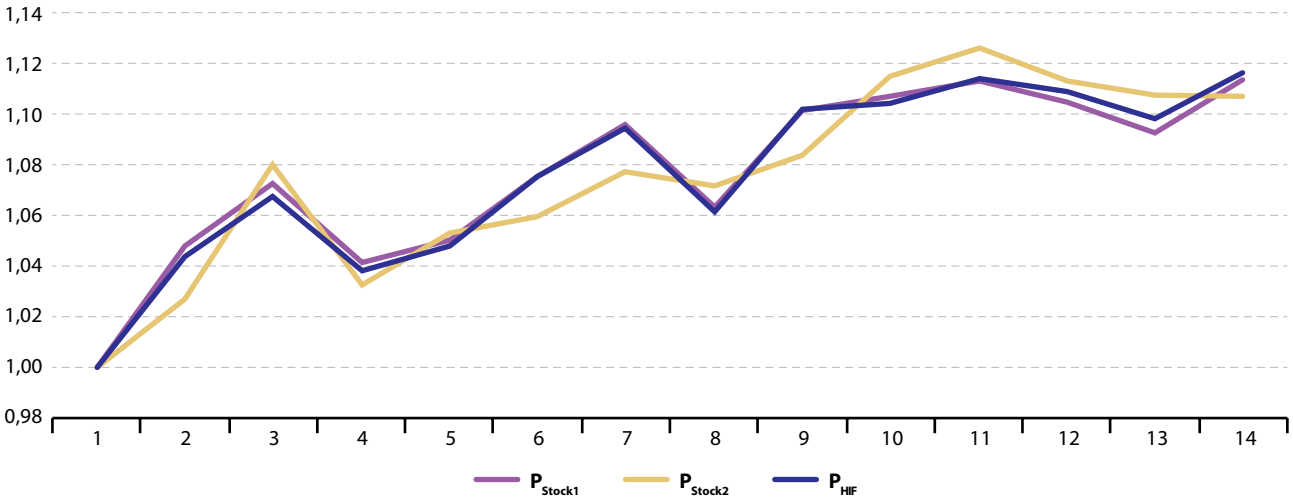
19 Cette approximation conviendrait probablement si la période d'échantillonnage était une décennie ou à peu près. À l'évidence, notre période d'échantillonnage de 14 trimestres est trop courte pour être exacte; il y aussi des problèmes d'autosélection de l'échantillon, les maisons neuves étant surreprésentées. Toutefois, la méthode que nous suggérons ici peut être illustrée par cette approximation grossière.

20 Nous n'avons pas supprimé les observations relatives aux maisons qui ont fait l'objet de multiples transactions pendant les 14 trimestres puisqu'une maison donnée vendue au cours de deux ou plus de ces trimestres n'est pas vraiment la même, du fait de la dépréciation et des rénovations.

21 V^t étant une valeur, elle n'apparaît pas au premier coup d'œil comme une série de prix. Mais, chaque trimestre, le vecteur de quantité sous-jacent à cette valeur est un vecteur unitaire de dimension 2.289, constant pendant les 14 trimestres. V^t peut donc être interprétée comme une série de prix, normalisée pour être égale à 1 au trimestre 1.

22 Si l'on utilise les prix imputés pour chacun des 45 prix de strates de toutes les périodes (et pas seulement pour les strates à transaction nulle, comme c'était le cas pour la construction de $P_{Sctock2}$) et si l'on retient le même vecteur de quantité de la totalité de l'échantillon comme vecteur approximatif des quantités de stocks, l'indice de Lowe qui en résulte se révèle exactement égal à P_{Stock1}. Ces deux modes d'élaboration d'un indice de stock s'avèrent donc équivalents. Le fait que P_{Stock1} ne soit pas égal à P_{Stock2} prouve bien qu'il existe un biais des valeurs unitaires dans les strates : elles ne sont simplement pas définies de façon assez étroite.

Graphique 8.6. Indices approximatifs de prix du stock et indices à imputation hédonique (P_{Stock1}) et à stratification (P_{Stock2}), et indice de Fisher des prix de vente à imputation hédonique



Source : calculs des auteurs basés sur les données du cadastre néerlandais.

Tableau 8.7. Indices approximatifs de prix du stock et indices à imputation hédonique (P_{Stock1}) et à stratification (P_{Stock2}), et indice de Fisher des prix de vente à imputation hédonique

Trimestre	P_{Stock1}	P_{Stock2}	P_{HIF}
1	1,00000	1,00000	1,00000
2	1,04791	1,02712	1,04356
3	1,07255	1,07986	1,06746
4	1,04131	1,03257	1,03834
5	1,05040	1,05290	1,04794
6	1,07549	1,05934	1,07553
7	1,09594	1,07712	1,09460
8	1,06316	1,07172	1,06158
9	1,10137	1,08359	1,10174
10	1,10708	1,11482	1,10411
11	1,11289	1,12616	1,11430
12	1,10462	1,11291	1,10888
13	1,09278	1,10764	1,09824
14	1,11370	1,10686	1,11630

Source : calculs des auteurs basés sur les données du cadastre néerlandais.

$$V_{LS}^t \equiv \sum_{s=1}^{14} \sum_{n \in S_S(s)} \hat{\beta}_S^t L_n^s \qquad (8.22)$$
$$t = 1,...,14$$

$$V_{LM}^t \equiv \sum_{s=1}^{14} \sum_{n \in S_M(s)} \{\hat{\beta}_S^t [160] + \hat{\beta}_M^t [L_n^s - 160]\} \qquad (8.23)$$
$$t = 1,...,14$$

$$V_{LL}^t \equiv \sum_{s=1}^{14} \sum_{n \in S_L(s)} \{\hat{\beta}_S^t [160] + \hat{\beta}_M^t [140] + \hat{\beta}_L^t [L_n^s - 300]\} \qquad (8.24)$$
$$t = 1,...,14$$

La valeur estimée en période t des bâtiments ajustés qualitativement, V_S^t, est définie par

$$V_S^t \equiv \sum_{s=1}^{14} \sum_{n=1}^{N(s)} \hat{\gamma}^1 \mu^t (1 - \hat{\delta} A_n^s) S_n^s \qquad (8.25)$$
$$t = 1,...,14$$

où tous les bâtiments vendus pendant les 14 trimestres sont inclus.

8.54 Les *quantités* qui correspondent aux évaluations ci-dessus pour la période t des trois stocks de terrain et du stock de bâtiments sont définies comme suit[23] :

$$Q_{LS}^t \equiv \sum_{s=1}^{14} \sum_{n \in S_S(s)} L_n^s \qquad (8.26)$$
$$t = 1,...,14$$

$$Q_{LM}^t \equiv \sum_{s=1}^{14} \sum_{n \in S_M(s)} L_n^s \qquad (8.27)$$
$$t = 1,...,14$$

$$Q_{LL}^t \equiv \sum_{s=1}^{14} \sum_{n \in S_L(s)} L_n^s \qquad (8.28)$$
$$t = 1,...,14$$

$$Q_S^t \equiv \sum_{s=1}^{14} \sum_{n=1}^{N(s)} (1 - \hat{\delta} A_n^s) S_n^s \qquad (8.29)$$
$$t = 1,...,14$$

8.55 Les *prix approximatifs des stocks*, P_{LS}^t, P_{LM}^t, P_{LL}^t et P_S^t, qui correspondent aux valeurs et aux quantités définies par (8.22)–(8.29), peuvent être calculés de la façon habituelle :

$$P_{LS}^t \equiv V_{LS}^t / Q_{LS}^t \qquad (8.30)$$
$$P_{LM}^t \equiv V_{LM}^t / Q_{LM}^t$$
$$P_{LL}^t \equiv V_{LL}^t / Q_{LL}^t$$
$$P_S^t \equiv V_S^t / Q_S^t$$
$$t = 1,...,14$$

Au moyen des prix et des quantités ci-dessus, on établit un indice approximatif de stock des prix du terrain, P_{LStock}, en agrégeant les trois catégories de terrains, et on élabore simplement un indice approximatif de prix de stock pour les bâtiments à qualité constante, P_{SStock}, en normalisant la série P_S^t. L'indice approximatif *global* de stock, P_{Stock}, est obtenu en agrégeant les trois catégories de terrains avec les bâtiments à qualité constante (ou, de manière équivalente, en agrégeant P_{LStock} et P_{SStock}). Puisque les quantités sont constantes pendant les 14 trimestres, les indices de prix de Laspeyres, Paasche et Fisher sont tous égaux[24]. Les indices de prix de stock P_{LStock}, P_{SStock} et P_{Stock} figurent au graphique 8.7 et au tableau 8.8. Pour permettre des comparaisons, les indices de prix correspondants basés sur les *ventes* de biens pour le modèle présenté précédemment, P_{L4}, P_{S4} et P_4, figurent aussi au tableau 8.8. Comme on peut le constater sur ce dernier, l'indice approximatif de prix de stock pour les bâtiments, P_{SStock}, coïncide avec l'indice de prix basé sur les ventes pour les bâtiments à qualité constante P_{S4}, de sorte que P_{S4} ne figure pas au graphique 8.7.

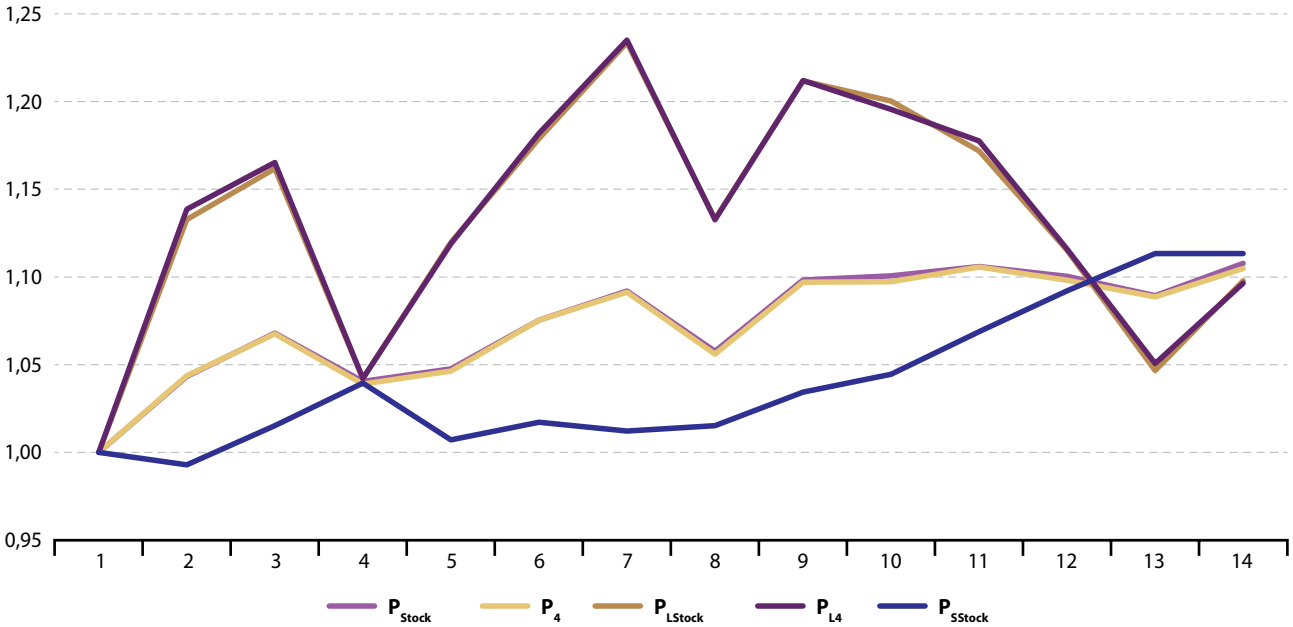
8.56 L'indice approximatif global de prix du stock total de maisons individuelles de la ville de «A» (P_{Stock}) peut difficilement être distingué au graphique 8.7 de l'indice global correspondant du prix des ventes (P_4). De même, on ne peut guère distinguer sur ce graphique l'indice du stock de terrains de «A» (P_{LStock}) de l'indice correspondant du prix des ventes de terrains (P_{L4}). On relève néanmoins au tableau 8.8 de légères différences entre les indices de stock et de ventes.

8.57 Nous concluons que les modèles de régression hédonique pour les ventes de maisons peuvent être facilement adaptés pour calculer des indices de prix de type Lowe portant sur le stock de maisons. L'utilisation de notre ensemble de données ne fait pas apparaître de différences importantes entre les deux catégories d'indices, mais ce résultat pourrait ne pas se vérifier pour d'autres ensembles de données.

23 Les quantités définies par (8.26)–(8.29), constantes pendant les 14 trimestres, sont égales à 77.455, 258.550, 253.590 et 238.476 pour les surfaces petites, moyennes et grandes, respectivement, ainsi que pour les bâtiments.

24 Les indices à base fixe et les indices-chaînes de Laspeyres, Paasche et Fisher sont également égaux dans ces conditions.

Graphique 8.7. Indices approximatifs de prix du stock de maisons (P_{Stock}), du stock de terrains (P_{LStock}), du stock de bâtiments (P_{SStock}) et indices correspondants basés sur les ventes (P_{L4} et P_4)



Source : calculs des auteurs basés sur les données du cadastre néerlandais.

Tableau 8.8. Indices approximatifs de prix du stock de maisons (P_{Stock}), du stock de terrains (P_{LStock}), du stock de bâtiments (P_{SStock}) et indices correspondants basés sur les ventes (P_{L4} et P_4)

Trimestre	P_{Stock}	P_4	P_{LStock}	P_{L4}	P_{SStock}	P_{S4}
1	1,00000	1,00000	1,00000	1,00000	1,00000	1,00000
2	1,04331	1,04373	1,13279	1,13864	0,99291	0,99291
3	1,06798	1,06752	1,16171	1,16526	1,01518	1,01518
4	1,04042	1,03889	1,04209	1,04214	1,03947	1,03947
5	1,04767	1,04628	1,11973	1,11893	1,00709	1,00709
6	1,07540	1,07541	1,17873	1,18183	1,01721	1,01721
7	1,09192	1,09121	1,23357	1,23501	1,01215	1,01215
8	1,05763	1,05601	1,13299	1,13257	1,01518	1,01518
9	1,09829	1,09701	1,21171	1,21204	1,03441	1,03441
10	1,10065	1,09727	1,20029	1,19545	1,04453	1,04453
11	1,10592	1,10564	1,17178	1,17747	1,06883	1,06883
12	1,10038	1,09815	1,11507	1,11588	1,09211	1,09211
13	1,08934	1,08863	1,04668	1,05070	1,11336	1,11336
14	1,10777	1,10486	1,09784	1,09648	1,11336	1,11336

Source : calculs des auteurs basés sur les données du cadastre néerlandais.

9

Sources de données

Introduction

9.1 En pratique, il est très coûteux de mener des enquêtes sur les prix des logements conçues spécifiquement à cet effet. Les offices de statistiques et les autres organismes qui construisent des indices des prix de l'immobilier résidentiel exploitent donc essentiellement des données administratives. Ces données sur les prix de l'immobilier sont généralement produites par les processus juridiques et administratifs des pays lors des achats des logements. Les indices établis avec ces sources dépendent du stade du processus d'achat de logement où le prix est mesuré. Par exemple, on peut retenir comme «prix» de la propriété le prix de transaction final ou l'évaluation antérieure en vue d'obtenir un prêt. De plus, des fichiers de données administratives différents donneront généralement des informations différentes sur les ventes de propriétés. La nature des données disponibles a généralement une incidence sur le choix des méthodes et limite souvent les techniques possibles pour ajuster l'indice des changements de qualité des logements, dont la taille, l'emplacement, etc. C'est pourquoi depuis toujours, les sources de données ont limité la construction des indices.

9.2 Ce chapitre étudie les différentes sources de données employées pour construire les indices des prix de l'immobilier résidentiel. Il s'intéresse principalement aux données sur les prix, mais examine aussi comment le choix de la formule de pondération peut être limité par les informations produites par le processus d'achat du logement. Différentes formules de pondération, en particulier selon qu'un indice est pondéré en fonction du stock ou des ventes, produisent des indices des prix qui mesurent des concepts différents. Il est alors important de bien comprendre ce que l'on souhaite mesurer afin d'évaluer les indices établis en les comparant à la mesure recherchée, de façon à déterminer s'ils répondent aux besoins.

Prix

Le processus d'achat et de vente d'un logement

9.3 Le processus d'achat et de vente d'une propriété prend généralement plusieurs mois, voire davantage. Le stade précis auquel le prix est introduit dans l'indice va dépendre de la source des données et a des conséquences sur ce qui est mesuré et sur la comparabilité des indices. On peut observer les données sur les prix pour un indice des prix de l'immobilier résidentiel aux étapes suivantes :

- Dès que la propriété est sur le marché (prix annoncé ou demandé). Les sources de données généralement utilisées sont les journaux et les agents immobiliers.
- Les demandes de prêt hypothécaire. Les sources de données généralement utilisées sont disponibles dans les institutions de crédit hypothécaire.
- Les approbations de prêt hypothécaire. Les sources de données généralement utilisées sont aussi disponibles dans les institutions de crédit hypothécaire.
- Les signatures d'un contrat contraignant. Les sources de données généralement utilisées sont collectées par les avocats et les notaires.
- Les transactions définitives. Les sources de données généralement utilisées sont le cadastre et proviennent des administrations fiscales.

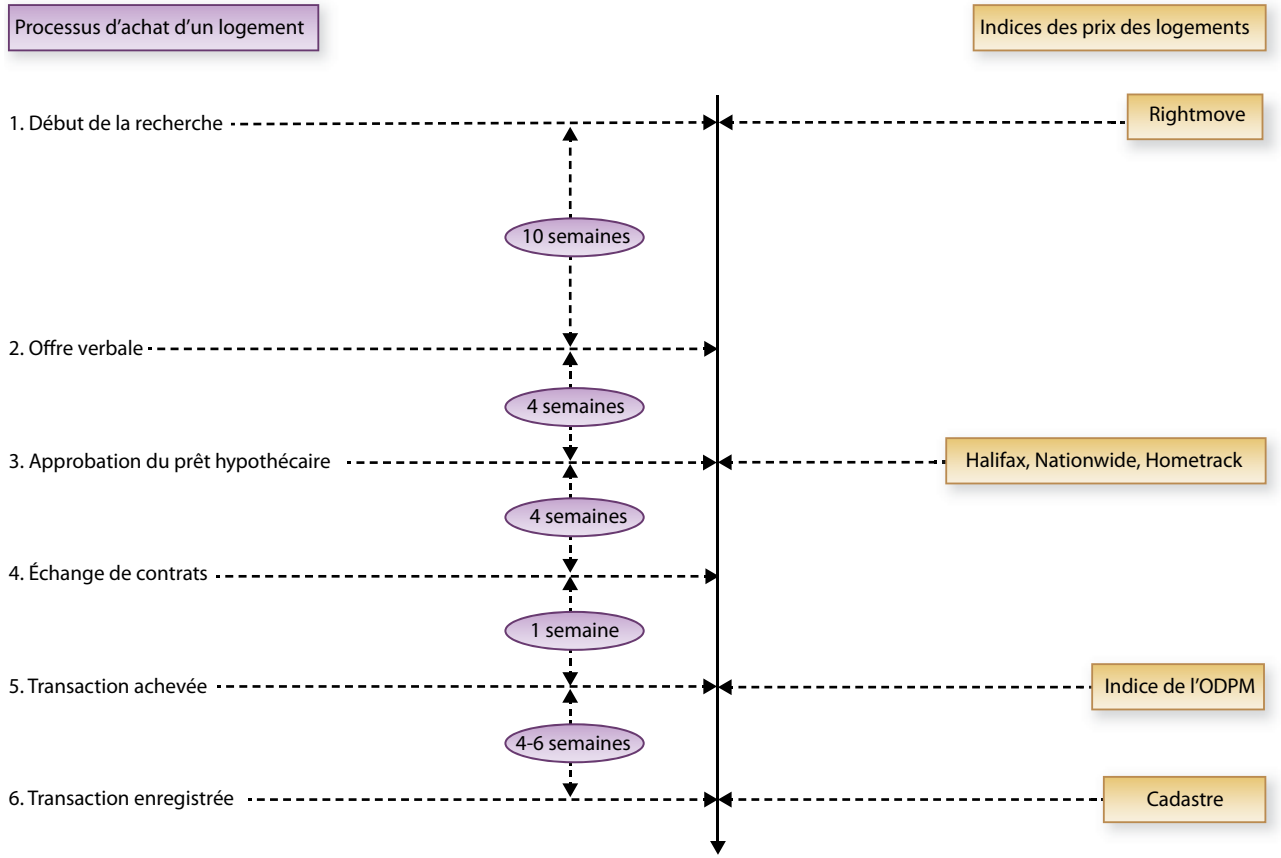
9.4 Chaque source de données sur les prix présente des avantages et des inconvénients. Ainsi, un inconvénient des prix affichés et aussi de ceux indiqués sur les demandes et les approbations de prêts hypothécaires est que tous les prix affichés ne se traduisent pas par une transaction, et qu'ils peuvent être différents des prix finalement négociés. Ces prix sont généralement disponibles avant le prix final de la transaction. Les indices qui mesurent les prix plus tôt dans le processus d'achat peuvent être les premiers à déceler les changements de prix, mais ils comporteront des erreurs, car les prix peuvent être largement négociés avant que le marché ne soit conclu.

9.5 Il faut relever qu'il peut être avantageux de disposer de sources d'information différentes sur les prix à plusieurs stades du processus d'achat et de vente. Ainsi, des changements de la relation entre le prix demandé et le prix de vente peuvent donner une indication précoce d'une évolution du marché du logement. Le diagramme ci-dessous illustre la situation au Royaume-Uni. On verra également l'étude de cas sur ce pays au chapitre 10.

9.6 La plupart des sources de données sont susceptibles de présenter tous les inconvénients que l'on rencontre lorsque l'on recourt aux systèmes administratifs pour établir des statistiques. L'utilisation de données administratives dans les statistiques économiques présente quatre difficultés — les définitions, le champ d'application, la qualité et l'actualité — et oblige à trouver des compromis prévisibles entre ces éléments et le coût d'établissement des données. Les définitions et le champ d'application sont parfois regroupés dans une catégorie unique, «champ d'application», afin de recouvrir les types d'unités couvertes et l'ampleur du champ d'application. Ainsi, on peut enregistrer les ventes au comptant, mais non les propriétés achetées grâce à un prêt hypothécaire, non plus que certaines ventes au comptant, si par exemple elles sont inférieures au seuil d'assujettissement à l'impôt.

9.7 Le problème fondamental vient du fait que les données sont essentiellement enregistrées comme étape du processus administratif, et non pas comme intrants dans un système statistique. Le statisticien ne les maîtrise donc pas. Il faut tenir compte des lacunes inhérentes des données administratives lorsqu'on les utilise et que l'on interprète

Diagramme : Chronologie de l'achat d'un logement et indices des prix des logements



Sources : Banque d'Angleterre et cabinet du vice-premier ministre (ODPM).

les résultats, en particulier si elles servent à remplacer des données statistiques, et non pas à compléter ou accompagner des statistiques conçues pour une finalité précise. On peut combler certaines lacunes en adoptant une méthode adaptée, par exemple en conjuguant plusieurs sources de données complémentaires et, éventuellement, en recourant à une forme de modélisation.

9.8 Plusieurs critères fondamentaux entrent en jeu lorsque l'on détermine dans quelle mesure les différentes sources de données sont adaptées.

- *Définition.* Ce critère est étroitement lié à des questions conceptuelles et au fait de savoir ce que l'indice doit mesurer.
- *Champ d'application.* Les questions relatives au champ d'application seront déterminées par les limites opérationnelles de l'organisme ou de l'entreprise qui fournit les données sur le logement. L'organisme peut, par exemple, couvrir la vente des propriétés dans l'ensemble du pays ou seulement une région particulière, ou bien les transactions couvertes peuvent concerner uniquement les achats au comptant ou les propriétés achetées à l'aide d'un prêt hypothécaire. Dans les organismes publics, les limites opérationnelles seront fonction des réglementations et des processus juridiques qui interviennent dans l'achat des propriétés à usage résidentiel. Lorsque les données proviennent du secteur public et privé, il est inévitable que le champ d'application dépende fortement des moyens à la disposition de l'organisme ou de l'entreprise, et de son efficacité à fournir des données. Les statisticiens qui établissent l'indice ne maîtrisent aucun de ces facteurs, qui peuvent avoir une incidence sur la qualité des données et sur toute divergence entre le champ d'application prévu de l'indice des prix de l'immobilier résidentiel et son champ d'application réel.
- *Qualité.* Lorsque l'on examine la question de la qualité des données, il ne faut pas perdre de vue que l'autorité administrative va vraisemblablement chercher à valider les informations qui ont trait à la vente et à l'accomplissement de sa mission, et qui reflètent les lois et les règlements qu'elle doit respecter. D'autres informations ont peut-être été recueillies qui présentent un intérêt pour l'office de statistiques, mais qui ont une utilité limitée pour l'autorité administrative. C'est le cas par exemple de certaines caractéristiques des logements que l'office de statistiques peut utiliser pour faire des ajustements au titre de la qualité. En fin de compte, la fiabilité des données administratives va varier selon que le fournisseur des données est incité ou non à donner des informations

correctes et complètes. Il peut être mutuellement avantageux pour les deux parties que l'office de statistiques aide l'autorité administrative à améliorer la qualité des données. L'office peut y parvenir en remontant des informations sur la cohérence des données saisies et en donnant un avis sur des lacunes plus générales. Certains offices de statistiques incitent l'autorité administrative à améliorer la collecte des données en établissant des statistiques adaptées aux besoins du fournisseur de données en échange d'un accès aux données brutes.

- *Actualité.* L'actualité des données administratives varie selon le responsable des déclarations à l'autorité administrative et selon son incitation à le faire en temps voulu. Ainsi, un acheteur qui veut obtenir l'approbation d'un prêt immobilier par une institution de crédit hypothécaire et cette institution peuvent être fortement encouragés à obtenir une évaluation précise et à jour, pour que la vente puisse aboutir et que les intérêts de toutes les parties soient protégés, avant qu'un autre acquéreur potentiel ne s'intéresse au bien. En revanche, ils seront moins incités à enregistrer la vente rapidement auprès du cadastre une fois la transaction définitivement acceptée.

L'une des conditions essentielles pour utiliser efficacement les données administratives est d'avoir une connaissance intime et détaillée des processus de collecte des données et des systèmes opérationnels associés.

9.9 On examinera ci-dessous séparément chaque source de données sur les prix. Lorsque le statisticien chargé d'établir l'indice dispose de plus d'une source de données, il a la possibilité d'en vérifier la cohérence et de combiner des données de sources différentes. Il peut par exemple utiliser les évaluations des propriétés réalisées pour approuver les prêts afin d'estimer le prix de transaction final inscrit beaucoup plus tard au cadastre. Ceci dépendra bien entendu de la stabilité de toute corrélation observée entre les deux.

Prix demandé par le vendeur : agents immobiliers, journaux, etc.

9.10 On peut recueillir des informations sur le prix demandé par les vendeurs en enquêtant auprès d'agents immobiliers ou en examinant les annonces publiées dans les journaux, les revues ou en ligne. L'un des principaux avantages des indices construits à partir de ces informations est leur actualité. Les indices construits en retenant les prix demandés peuvent donner une estimation plus à jour des prix des logements que ceux construits à l'aide des transactions ultérieures. Ils présentent également un avantage par rapport aux indices basés sur des informations qui proviennent d'institutions de crédit hypothécaire, car ces derniers se limitent aux transactions réalisées grâce à des prêts hypothécaires. Les indices établis à partir des prix demandés initiaux ont cependant un inconvénient majeur : les logements peuvent être retirés du marché et le prix de vente convenu n'est pas nécessairement égal au prix demandé par le vendeur. Ces indices ne tiennent pas compte des baisses de prix que les vendeurs appliquent ultérieurement, par exemple lorsque le marché immobilier traverse une phase de contraction ou des prix d'offre supérieurs au prix demandé si le marché immobilier est dynamique. Ils peuvent donc présenter des perspectives trop optimistes lorsque le marché immobilier décline et trop pessimistes lorsqu'il se redresse. Le fait que l'on ne puisse pas se fier à ces indices pour présenter une image fidèle du marché de l'immobilier à court terme réduit leur utilité pour la plupart des usagers, en particulier ceux qui souhaitent détecter rapidement les points de retournement du marché ou disposer d'un indicateur avancé de l'orientation future des prix des logements. Il convient de noter que les différences entre les prix initialement demandés et les prix réels des transactions signifient également que le calcul des «estimations du prix moyen des logements» peut être parfois erroné.

9.11 Il n'est pas toujours aisé de vérifier les informations recueillies sur les prix demandés par les vendeurs. Ces informations supposent un échantillon équilibré et représentatif, reposent sur l'honnêteté et les connaissances des personnes interrogées et, lorsqu'elles proviennent d'annonces, en particulier sur des sites Internet, sur leur exactitude. On fait valoir par exemple que les agents immobiliers ont davantage tendance à être optimistes à propos des prix, et ont tout intérêt à une hausse des prix plutôt qu'à une baisse, ce qui peut influer sur les résultats des enquêtes. En revanche, un agent immobilier peut proposer à un vendeur de demander un prix bas, sans rapport avec la réalité, de façon à écouler le bien et obtenir rapidement une commission. D'aucuns disent aussi que les sites Internet ont tendance à favoriser les propriétés dont le prix demandé est concurrentiel pour attirer les acheteurs potentiels. Ce sont bien entendu des conjectures, mais elles n'en rappellent pas moins les difficultés que peuvent poser ces sources.

9.12 Les enquêtes auprès des agents immobiliers présentent certains avantages inhérents par rapport aux enquêtes reposant sur des annonces. Elles peuvent s'appuyer sur un échantillon choisi de façon beaucoup plus scientifique et donner des informations sur un choix représentatif des propriétés sur le marché, en recouvrant celles qui ne font généralement pas l'objet d'annonces. Ces données peuvent également fournir des informations détaillées sur les caractéristiques des propriétés, qui sont extrêmement importantes pour faire des ajustements au titre de la qualité (à l'aide soit de méthodes de régression hédonique, soit de méthodes de stratification, comme on l'a vu dans des chapitres précédents). De plus, le questionnaire d'enquête peut servir à recueillir des informations sur les questions suivantes : quel est en moyenne le délai de vente, quelle a été la différence, récemment, entre les prix demandés et les prix de vente (ont-ils été supérieurs ou inférieurs), ou quel est le nombre d'acquéreurs potentiels s'inscrivant et le nombre de propriétés dans le portefeuille de l'agent immobilier. Ces informations peuvent aider à mettre en contexte les informations sur les prix qui servent à établir l'indice et être utiles pour interpréter les résultats

définitifs. Mais ce type d'enquête ne recense généralement pas le prix demandé d'une propriété précise. Les questionnaires demandent plutôt à l'agent immobilier de donner le «prix moyen demandé» d'une sélection de propriétés représentatives[1]. Il peut s'agir, par exemple, de chacune des catégories habituelles de propriétés (appartements, maisons en rangée, maisons jumelées et maisons individuelles) sur plusieurs lieux différents. Ce sont ces informations qui servent à créer un prix moyen des propriétés pour chaque type de propriété et pour chaque endroit, puis à calculer l'indice des prix correspondant. À l'inverse, l'avantage inhérent d'une enquête portant sur des annonces est qu'elle permet de recueillir le prix réellement demandé de toutes les propriétés affichées.

9.13 En somme, bien que les indices des prix des logements établis à partir d'enquêtes sur les prix demandés soient sans doute plus actuels, il faut être prudent si l'on use ce type d'indice comme baromètre des prix des logements, car il est difficile de déterminer exactement comment ont été établies les informations de l'enquête, et on connaît mal la relation entre le prix demandé et le prix de vente.

Prix d'offre initial accepté par le vendeur : institutions de crédit hypothécaire

9.14 De nombreux pays font appel aux institutions de crédit hypothécaire comme principale source de données pour établir leur indice des prix des logements. Les informations sont stockées dans le système informatique du prêteur et répondent aux besoins commerciaux opérationnels des institutions de crédit hypothécaire. Cette base de données peut comprendre le prix initial offert par l'acquéreur potentiel, le prix de valorisation utilisé pour autoriser l'emprunt, ainsi que, parfois, le prix final de transaction. Les informations obtenues auprès des institutions de crédit hypothécaire présentent tous les inconvénients des données tirées des systèmes administratifs décrits plus haut, mais ces bases de données sont parfois une riche source d'informations actualisées.

9.15 Les données provenant des institutions de crédit hypothécaire présentent cependant un inconvénient majeur : elles ne couvrent pas les achats de logements non financés par un emprunt. Des études ont montré que les acheteurs au comptant représentent un tiers environ du marché au Royaume-Uni et que les propriétés qu'ils acquièrent sont en général soit très bon marché, soit très chères. Ceci ne poserait pas de problème si l'évolution des prix des logements achetés au comptant ne pouvait pas être différente de celle des logements financés par un prêt hypothécaire. Il faut s'attendre à des évolutions différentes lors des retournements du marché. Alors les différentes extrémités du marché peuvent réagir différemment à la situation économique et les acheteurs au comptant peuvent être favorisés. En période de ralentissement par exemple, les personnes situées à l'extrémité supérieure du marché qui envisageaient de vendre leur logement pour dégager des fonds peuvent hésiter à mettre leur logement sur le marché à un prix réduit. L'offre de logements à vendre diminue et provient principalement de propriétaires qui, pour une raison ou une autre, sont très désireux de vendre. Mais, dans le même temps, le nombre d'acheteurs potentiels actifs faisant appel à des prêts hypothécaires peut sensiblement diminuer, car les acheteurs hésitent à contracter des prêts plus importants. Certains auront néanmoins besoin de vendre. Dans ces cas, l'acheteur au comptant d'un logement à l'extrémité supérieure du marché se trouvera en position relativement plus forte pour négocier un prix avantageux que dans un marché plus stable.

Prix de valorisation en vue d'un emprunt : institutions de crédit hypothécaire

9.16 Les institutions de crédit hypothécaire font faire une évaluation indépendante des propriétés avant d'approuver un prêt, mais plusieurs semaines peuvent se passer entre le moment où l'acheteur et le vendeur ont négocié un prix définitif et où l'acheteur a déposé sa demande de prêt et le moment où l'institution donne son évaluation au client. En pratique, ces deux étapes sont séparées par une période de négociation, durant laquelle le prix d'achat convenu du logement peut encore changer. C'est parfois le cas lorsque l'évaluation indépendante est différente du prix convenu entre l'acheteur et le vendeur, ou lorsque l'acheteur a payé une étude détaillée du bien, qui révèle qu'il faut procéder à des réparations importantes. Il est assez courant par exemple qu'un acheteur tente d'obtenir une baisse de prix si l'évaluation de l'institution de crédit hypothécaire se révèle très inférieure au prix précédemment convenu ou si une étude sur l'état de la propriété montre qu'il faut changer la toiture. À l'évidence, si le prix d'offre initial diffère de l'évaluation ultérieure et si un processus de renégociation a lieu ultérieurement, le taux mesuré de l'inflation des prix immobiliers sera différent du taux réel mesuré à partir du prix de la transaction proprement dite.

9.17 La variation des prix des logements mesurée par des indices qui reposent sur les évaluations des institutions de crédit hypothécaire[2] peut être différente de la variation telle qu'elle apparaît à partir du prix d'offre, et les deux peuvent être différentes de la variation calculée à partir des prix de transaction définitifs, même lorsque ceux-ci proviennent du même échantillon d'institutions de crédit hypothécaire. Il est donc important de bien comprendre ce que mesure un indice.

1 Certaines enquêtes demandent également le prix «réaliste» et s'en servent pour construire un indice des prix des logements.

2 Il faut tenir compte du fait que les prix tirés des évaluations des institutions de crédit hypothécaire, comme ceux qui reposent sur n'importe quelle évaluation, dépendent de l'objectivité du processus d'évaluation. C'est pourquoi on a dit que les évaluations des institutions de crédit hypothécaire sont parfois influencées par la politique de crédit de la banque. Ces sources peuvent par conséquent poser des difficultés.

Prix de transaction définitif : institutions de crédit hypothécaire

9.18 Le délai qui sépare la demande de prêt hypothécaire, l'approbation du prêt et les étapes de conclusion de l'achat, ainsi que les différences entre les valeurs correspondantes des prix des logements montrent qu'il faut concilier vitesse de publication et exactitude. Les institutions de crédit hypothécaire ne recensent pas toujours le prix de transaction définitif, que l'on tire souvent des dossiers juridiques comme les inscriptions au cadastre, où figurent également les ventes qui n'ont pas été financées par un prêt hypothécaire. Le délai entre la conclusion de la transaction et l'inscription de la vente au cadastre peut cependant être long. L'un des principaux avantages des données provenant des institutions de crédit hypothécaire est leur fraîcheur. Les prix d'offre initiaux et les évaluations donnent une première indication des prix en vigueur, puisque ces données sont disponibles plus tôt, et on peut obtenir les prix de transaction définitifs auprès des institutions de crédit hypothécaire plus tôt qu'auprès du cadastre. C'est pour cette raison que l'on pourra préférer exploiter les informations sur le prix de transaction définitif qui proviennent des institutions de crédit hypothécaire. On peut facilement vérifier ce prix en le comparant avec les registres du cadastre pour dissiper toute crainte quant à l'exactitude et la crédibilité.

Prix de transaction définitif : données administratives provenant des cadastres et des administrations fiscales

9.19 Dans l'idéal, un indice des prix des logements doit être construit à partir des prix de transaction réels au moment où la propriété est vendue et la vente achevée. La signature du premier contrat contraignant répond le mieux à cette condition en raison de son actualité, mais dans la pratique, il peut y avoir une ambiguïté quant au moment où un contrat est contraignant : est-ce lorsqu'une offre est officiellement acceptée (lorsque les soumissions sous pli cacheté sont ouvertes par exemple) ou lorsque le contrat est signé ou encore quand le contrat est échangé? De même, il peut y avoir une différence entre le moment où un contrat est signé et où se produit la mutation et le moment où celle-ci est inscrite au cadastre ou auprès de l'administration fiscale.

9.20 En théorie, les informations qui proviennent des cadastres ou des administrations fiscales couvrent toutes les propriétés, les achats au comptant tout comme ceux obtenus à l'aide d'un prêt hypothécaire. De toutes les sources à la disposition des statisticiens qui établissent des indices, ce sont ces bases de données qui devraient donc être les plus complètes. En pratique, cependant, on ne peut pas garantir leur exhaustivité, en particulier si le propriétaire est dissuadé d'enregistrer une propriété. Ainsi, lorsque le bien est enregistré essentiellement pour des raisons fiscales, il arrive qu'il ne soit pas du tout enregistré ou qu'il le soit en omettant des détails importants ou en déclarant des détails erronés, comme la surface au plancher, afin d'éviter de payer des impôts ou d'alléger la charge fiscale[3].

Prix de valorisation pour le paiement des impôts et des services locaux : administrations fiscales

9.21 Dans beaucoup de pays, l'administration centrale ou locale peut imposer sur les propriétés à usage résidentiel un impôt mensuel ou annuel, ou une commission de service, afin de financer la prestation de services publics (entretien des routes, services de police et de pompiers ou collecte des ordures, par exemple). Souvent, l'avis d'imposition d'un individu est proportionnel à la valeur fiscale du bien, et cette dernière repose généralement sur une évaluation réalisée par des géomètres experts, soit sous contrat, soit employés directement par l'administration fiscale. Les évaluations doivent prendre en considération les caractéristiques du bien comme l'emplacement et la surface de la parcelle. Elles sont cependant tributaires d'informations exactes sur les biens, ainsi que des évaluations des géomètres experts, qui sont difficiles à vérifier. De plus, il est rare que les évaluations soient actualisées, car cela entraîne des frais sur le terrain. En raison de ces inconvénients, les informations recueillies sont parfois peu utiles pour construire des indices des prix de l'immobilier résidentiel. Des offices de statistiques ont malgré tout exploité cette source d'informations officielles sur les évaluations (voir les données sur la méthode RPVE de construction des indices décrite au chapitre 7).

Informations tirées d'autres expertises : enquêtes menées par des organisations d'agents immobiliers et d'autres organisations professionnelles et leurs membres

9.22 Certains pays réalisent périodiquement des enquêtes auprès d'agents immobiliers, de géomètres experts ou de leurs organisations professionnelles pour les interroger sur les prix et le stock des logements. Ces «enquêtes d'opinion» se limitent généralement à demander aux personnes interrogées si, selon elles, les prix des logements augmentent, diminuent ou restent stables. Elles ne donnent aucune indication sur la valeur des logements, ni sur l'ampleur de la hausse ou de l'augmentation des prix, mais elles peuvent permettre de dresser un tableau général et à jour de l'orientation de la variation des prix sur le marché du logement, pour compléter les derniers chiffres

3 Un problème corollaire se pose : le prix de la transaction n'est pas nécessairement un prix de marché si la transaction, bien que véritable, a lieu entre des parents ou des amis. Des parents peuvent décider par exemple de transmettre la maison familiale à leurs enfants à un prix inférieur au prix de marché.

de l'indice des prix de l'immobilier résidentiel et leur donner plus de crédibilité. Ainsi, une variation significative des proportions respectives d'agents immobiliers qui pensent que les prix augmentent et de ceux qui pensent que les prix baissent peut être le premier indicateur d'une évolution du marché du logement, que n'ont pas encore révélée les statistiques disponibles sur les évaluations des institutions de crédit hypothécaire. Ce type d'information contextuelle apporte une valeur ajoutée et est régulièrement utilisé par les commentateurs, lorsqu'ils interprètent les indices officiels des prix des logements.

Évaluation des sources de données pour déterminer leur conformité aux besoins

9.23 L'utilité globale de ces sources d'information sur les prix de l'immobilier résidentiel dépend pour beaucoup de leur conformité aux besoins des applications particulières pour lesquelles elles sont utilisées. Pour mesurer la conformité aux besoins, il faut évaluer les avantages et les inconvénients intrinsèques de l'indice en le comparant à un ensemble convenu de critères. Autrement dit, il faut l'évaluer en fonction des besoins des usagers.

9.24 Le chapitre 2 a analysé les nombreuses utilisations différentes des indices des prix des logements : indicateur macroéconomique de l'inflation; fixation des objectifs de politique monétaire; mesure de l'évolution de la richesse; indicateur de stabilité financière pour mesurer l'exposition aux risques; déflateur des comptes de la nation; élément permettant à un citoyen de décider s'il convient d'investir dans une propriété à usage résidentiel; élément d'autres indices des prix, en particulier l'indice des prix à la consommation (IPC) et donnée utilisée dans les négociations salariales ou l'indexation des salaires.

9.25 Pour évaluer efficacement les différentes sources de données sur les prix des logements, il faut analyser les exigences des usagers de façon systématique. Les besoins des usagers ont des conséquences importantes sur les décisions relatives au fondement conceptuel d'un indice et sur les exigences statistiques qui s'y rapportent. On pourra poser une série de questions sur les différentes raisons pour lesquelles les usagers cherchent des informations sur le prix des logements. Ainsi, un indice des prix des logements va-t-il servir comme élément d'une série d'indicateurs macroéconomiques généraux, comme donnée pour mesurer l'inflation des prix à la consommation, comme élément dans le calcul du patrimoine des ménages ou comme composante directe de l'analyse de l'exposition d'un prêteur? On pourra alors transformer cette analyse en exigence de l'usager en matière statistique, accompagnée d'un cadre conceptuel, en exprimant les besoins en termes statistiques et en mettant en évidence les liens communs et les relations correspondantes aux niveaux microéconomique et macroéconomique. On évaluera ensuite les sources de données en les comparant aux besoins statistiques.

9.26 La liste suivante des caractéristiques souhaitables d'un indice des prix de l'immobilier résidentiel constitue un ensemble de critères possibles permettant d'évaluer dans quelle mesure d'autres sources de données sont conformes aux besoins de différents usages[4]. Cette liste s'inspire de l'analyse du début de ce chapitre. L'importance relative de chacun des critères dépendra de l'usage et, par essence, constitue une exigence statistique. Il faudra aussi faire les compromis habituels entre la satisfaction de tous les besoins des usagers et le coût de la collecte des données.

Définitions et concept de mesure

9.27 Cette notion couvre également la *cohérence* avec d'autres statistiques produites. Elle représente les exigences des usagers au niveau le plus élémentaire. Considérons les besoins des administrations publiques, des analystes qui étudient les pressions inflationnistes et des investisseurs directs dans l'immobilier. Ces usagers vont sans doute s'intéresser principalement au caractère cyclique des prix et à la capacité des prix de l'immobilier à provoquer des alternances d'expansion et de récession qui déstabilisent l'ensemble de l'économie. À cet effet, ils recherchent toute une gamme d'indicateurs, notamment des indices du volume et du prix des transactions immobilières, ainsi que des indicateurs macroéconomiques permettant de modéliser le cycle économique et de prévoir ses pics et ses creux. Les analystes qui étudient les pressions inflationnistes de la hausse des prix de l'immobilier en les comparant à d'autres hausses de prix peuvent juger intéressant d'inclure dans un IPC les coûts inflationnistes des logements occupés par leur propriétaire à l'aide d'un indice des prix des logements basé sur le coût d'acquisition net, hors terrain.

9.28 Un indice reposant sur tous les achats — au comptant et à l'aide d'un prêt hypothécaire — conviendra aux usagers qui ont besoin d'un indicateur macroéconomique général. Les prix des transactions issus uniquement des données fournies par les institutions de crédit hypothécaire comportent une grave lacune. Sur le plan conceptuel, les données tirées des cadastres sont de meilleure qualité, car elles couvrent en principe l'ensemble des transactions. La difficulté consiste à trouver une source de données sur les prix, immédiatement adaptée ou que l'on puisse manipuler afin de répondre aux exigences des usagers qui souhaitent

4 Voir aussi le chapitre 3, qui présente une liste des besoins des usagers dressée à la suite d'échanges de vues entre des usagers des indices des prix des logements et l'Institut national de la statistique. Il y est indiqué qu'il faut faire un compromis entre le souhait des usagers de disposer d'une famille d'indices plus détaillés (stratifiés en fonction du lieu et du type de logement) et la qualité des indices, des détails plus nombreux produisant inévitablement des indices moins précis.

que les coûts des logements occupés par leur propriétaire soient inclus dans l'IPC sur la base du coût d'acquisition net, c'est-à-dire en excluant le prix du terrain[5].

9.29 À l'inverse, les usagers intéressés par une analyse de la valeur actuelle du portefeuille immobilier qui garantit les prêts hypothécaires à rembourser auront besoin d'un indice des prix des propriétés pour lesquelles des prêts ont été accordés, pondéré par les sommes prêtées pour chaque type de propriété au moment où ils ont été accordés. Dans ces deux cas, la valeur du terrain supportant les bâtiments est tout aussi importante que la valeur des bâtiments proprement dits, et c'est la valeur totale du terrain et des bâtiments qui est intéressante. Les données provenant des institutions de crédit hypothécaire sur les prix des propriétés, le volume des nouveaux prêts hypothécaires et l'encours de la dette répondront aux besoins de ces usagers.

9.30 Examinons maintenant les besoins des employeurs et des syndicats qui négocient des conventions salariales. Ils sont principalement intéressés par les effets des variations des prix sur le niveau de vie des travailleurs. À cet effet, les usagers rechercheront un IPC qui inclut le coût d'un toit au-dessus de leur tête — pour les propriétaires occupants, il s'agira du coût du paiement des intérêts du prêt hypothécaire et des réparations. Pour le mesurer, il faudra donc le montant du prêt hypothécaire au moment de l'achat, et pour déterminer les remboursements ultérieurs, il faudra un indice des prix des logements pondéré par les ventes. Si possible, on exclura les refinancements. On peut mesurer la réparation en calculant l'amortissement. Pour ce faire, c'est un indice des prix des logements lissé et pondéré par les stocks qui convient le mieux. Par ailleurs, la question des terrains se pose puisqu'on soutient souvent que, dans la plupart des cas, un terrain est un investissement qui s'apprécie, et qu'il est abusif de l'inclure dans un calcul de l'amortissement[6]. On doit donc parfois calculer un indice qui exclut le prix des terrains.

9.31 L'usager peut également faire appel aux informations fournies par les institutions de crédit hypothécaire pour calculer la dépense au titre du prêt hypothécaire, mais pas pour estimer l'amortissement, car, dans ce cas aussi, il faut probablement distinguer la valeur du terrain.

9.32 Étudions, avec un dernier exemple, les besoins des comptables nationaux qui recherchent des déflateurs adaptés aux comptes nationaux. Eux aussi auront des besoins particuliers. L'immobilier apparaît dans la comptabilité nationale sous différentes formes (on trouvera des détails complémentaires au chapitre 3) :

- La valeur locative imputée reçue par les propriétaires occupants pour des bâtiments est une composante de la consommation finale des ménages.
- La formation de capital liée aux bâtiments, par opposition aux terrains, est une composante de la formation brute de capital fixe, de l'amortissement et de la mesure du stock de capital fixe.
- La valeur des terrains, qui constitue un élément important du patrimoine national.

Dans chacun des cas, pour calculer les volumes à partir des valeurs, il faut respectivement disposer des indices des prix, du loyer imputé des logements occupés par leur propriétaire, pondéré par le stock des différents types de logements occupés par leur propriétaire; des achats de logements neufs pondérés par les transactions de logements neufs, en excluant la composante terrain, et de l'ensemble du stock de logements, terrains compris, pondéré par le stock de logements.

9.33 On voit que les besoins des usagers vont varier et que, dans certains cas, il faudra produire plus d'une mesure des prix des logements ou de l'inflation de l'immobilier. On voit aussi qu'il est important d'assurer une cohérence des différentes mesures, entre elles et avec d'autres statistiques économiques, et qu'il est particulièrement difficile d'y parvenir, car il est rare que les statisticiens disposent d'une série idéale d'indicateurs des prix.

Champ d'application

9.34 On entend par champ de l'indice pas simplement le fait de savoir si toutes les propriétés sont couvertes, qu'elles soient détenues purement et simplement ou financées par un prêt hypothécaire, mais aussi le fait de savoir si les ventes de propriétés ou les évaluations de l'ensemble du pays sont couvertes, ou uniquement celles d'une région précise, et si toutes les fourchettes de prix sont couvertes. On peut noter que, lorsque l'on a essentiellement besoin d'un indice national, des indices régionaux peuvent servir à des fins analytiques. Il est peu probable que des informations sur les prix des logements provenant d'une seule institution de crédit hypothécaire soient représentatives du pays dans son ensemble, non seulement parce qu'elles ne recouvrent pas les achats au comptant, mais aussi parce que les prêteurs concentrent souvent leurs activités sur des régions particulières.

Qualité

9.35 On entend par qualité l'exactitude et l'exhaustivité des informations. En d'autres termes, il s'agit de savoir si les informations sont exemptes de graves erreurs et correspondent à ce que l'on attend d'elles. Par rapport aux autres données administratives, il y a tout lieu de penser que les informations sur les prix des logements qui proviennent d'un cadastre sont très exactes, car il est légalement obligatoire

5 Dans la plupart des pays et pour la plupart des transactions, le terrain et le bâtiment sont achetés ensemble, en bloc. En règle générale, les deux composantes ne sont donc pas séparées dans les informations tirées des archives sur les transferts de propriété, car il faudrait procéder à une opération supplémentaire pour distinguer les deux prix. On a montré au chapitre 8 comment utiliser une régression hédonique pour décomposer l'indice des prix global entre ses composantes : terrain et bâtiments.

6 D'autres questions plus générales se posent, que nous n'abordons pas ici, qui concernent la mesure de l'amortissement et son inclusion dans l'indice des prix à la consommation.

d'enregistrer les transactions et les changements de propriété. Il est cependant difficile de valider la fiabilité des données de source administrative.

Actualité

9.36 Les indices qui mesurent les prix aux premiers stades du processus d'achat permettent de détecter plus tôt les variations de prix et les changements d'orientation de l'inflation des prix immobiliers. Cela sera sans doute particulièrement important s'ils servent, par exemple, à définir la politique macroéconomique et à fixer des objectifs de croissance monétaire, mais moins important s'ils servent comme déflateurs en comptabilité nationale. Les données des institutions de crédit hypothécaire répondront probablement davantage aux besoins des responsables de la politique macroéconomique et des objectifs de croissance monétaire, même si les achats au comptant ne sont pas couverts, alors que les données tirées des cadastres seront plus adaptées aux besoins des statisticiens qui calculent les déflateurs, par exemple.

Ajustements au titre de la qualité et au titre de la composition de l'échantillon : détails requis

9.37 Deux questions corollaires se posent : dans quelle mesure peut-on ajuster les indices des prix de l'immobilier résidentiel pour tenir compte de l'évolution de la composition de l'échantillon de propriétés vendues et supprimer l'effet de changements de la qualité sur chaque logement? Il faut disposer à cet effet d'informations «en temps réel» sur les caractéristiques qui servent à déterminer les prix comme la surface de la parcelle, la taille du logement, le type de propriété (appartement, maison jumelée ou maison individuelle), l'emplacement, l'état du bien ou la présence d'un chauffage central, d'une cuisine équipée, d'une salle de bains, etc. Il est essentiel de procéder à un ajustement au titre de la qualité (ou à une poststratification) pour construire un indice exact des prix des composantes du logement[7]. Selon toute vraisemblance, aucune des sources de données sur les prix indiquées plus haut ne sera idéale ou ne répondra à tous les besoins. Le volume de données détaillées et de caractéristiques utiles variera en fonction de chaque ensemble de données[8].

Fréquence

9.38 La notion de fréquence désigne essentiellement la fréquence à laquelle on peut calculer un indice (tous les mois ou tous les trimestres par exemple). Il faut concilier fréquence et exactitude. Pour une zone géographique et un type de logement précis, les informations sur les prix courants des logements dans une strate donnée proviendront des ventes de logements neufs et anciens appartenant à cette strate pendant la période donnée. Si l'on choisit une fréquence mensuelle, et non pas trimestrielle, la taille de l'échantillon sera d'environ le tiers de celle de l'échantillon trimestriel. Par conséquent, un indice mensuel des prix des logements construit à partir des ventes de biens dans la strate donnée subira les effets d'une volatilité accrue de l'échantillon (et ne sera donc pas aussi exact) que l'indice trimestriel correspondant. On *peut* réduire la volatilité d'un indice mensuel en «agrandissant» la strate[9], par exemple en regroupant plusieurs quartiers d'un même site général, mais cela oblige à faire un autre compromis entre la finesse de la strate (que beaucoup d'usagers recherchent peut-être) et l'exactitude de l'indice (que recherchent tous les usagers).

9.39 Il est possible de produire des indices des prix des logements mensuels lissés qui sont, par exemple, une moyenne mobile sur trois mois des indices mensuels bruts[10]. L'office de statistiques peut aussi produire à la fois des indices mensuels et trimestriels, et laisser les usagers choisir l'indice qu'ils préfèrent[11]. Il est impossible de se prononcer catégoriquement sur la fréquence à laquelle un indice des prix des logements couvrant une certaine strate doit être publié. La question de la fréquence doit être tranchée par l'office national des statistiques en prenant en considération les besoins des usagers et les données disponibles.

Révisions

9.40 Les révisions peuvent résulter soit d'informations communiquées ultérieurement (c'est alors la série qui est révisée), soit d'autres sources de données plus pertinentes arrivées ultérieurement (une première mesure indicative étant par la suite remplacée par une mesure précise de ce dont on a besoin[12]). Le premier cas se produit par exemple lorsque des

7 Les différentes méthodes permettant de construire des indices des prix des logements ajustés en fonction de la qualité ont été analysées aux chapitres 4 à 8.

8 Lorsque la base de données d'un agent immobilier contient le prix de vente définitif des propriétés répertoriées, accompagné de leurs principales caractéristiques, ces informations seront probablement «optimales» pour la plupart des usages. Il faut cependant comparer l'échantillon des propriétés répertoriées aux propriétés inscrites au cadastre pour veiller à ce que la couverture des propriétés convienne au but recherché. Pour construire des indices des prix du stock de logements, il faut disposer d'informations provenant d'un recensement sur les stocks de logements, ainsi que d'informations postcensitaires sur les démolitions et les constructions de nouvelles habitations.

9 Il n'est pas certain qu'en regroupant plusieurs strates, on réduise la volatilité de l'indice si les prix des logements des différentes microstrates évoluent différemment.

10 L'Australian Bureau of Statistics a souvent recours à cette technique pour toute une gamme de statistiques. Si la fenêtre est de 12 mois, alors l'index lissé ainsi obtenu peut être considéré comme un indice corrigé des variations saisonnières, centré au milieu de la période de 12 mois considérée. On trouvera au chapitre 4 une variante de cette technique de lissage.

11 Certains usagers peuvent être déroutés si plus d'un indice couvre essentiellement la même strate de logements. Aujourd'hui, le Bureau of Labor Statistics publie deux indices des prix à la consommation : l'IPC phare de type Lowe, qui n'est pas révisé, et un second indice qui est une approximation d'un indice superlatif de Törnqvist (qui lui est révisé). Il semble que les usagers des États-Unis aient accepté des indices multiples dans ce contexte.

12 Un problème corollaire se pose : certaines méthodes employées pour construire un indice des prix de l'immobilier résidentiel, comme la méthode hédonique multipériode à variables indicatrices temporelles (chapitre 5) et la méthode des ventes répétées (chapitre 6), sont exposées aux révisions dans le sens où des chiffres déjà calculés changeront lorsque de nouvelles données seront ajoutées à l'échantillon. Dans certains cas, on publie de nouveaux indices et, dans d'autres, on emploie la méthode des fenêtres mobiles avec mises à jour, élaborée par Shimizu, Nishimura et Watanabe (2010) et Shimizu, Takatsuji, Ono et Nishimura (2010). Cette technique ne révise pas l'indice historique jusqu'à la période actuelle.

révisions sont dues à un enregistrement tardif des ventes de propriétés. On se trouve dans le deuxième cas lorsqu'un prix d'offre initial indiqué sur une demande de prêt hypothécaire sert à obtenir une première indication de l'évolution des prix des transactions, mais est ensuite abandonné lorsque les données du cadastre sur les prix réels des transactions (qui prennent en compte toute renégociation des prix avant la clôture des ventes) deviennent disponibles beaucoup plus tard.

9.41 L'ampleur des révisions résultant de la communication ultérieure d'informations dépend en partie du point de référence des données sur les prix, et en partie du moment où l'office de statistiques reçoit l'ensemble de données considéré : plus la période de référence des données se situe tôt dans le cycle d'achat, plus l'ensemble des données considéré est reçu tôt et plus l'indice fera l'objet de révisions. Par conséquent, bien que les informations provenant de l'enregistrement des ventes de propriétés soient bien référencées et constituent une source d'information certaine sur les prix des propriétés, il arrive dans certains pays qu'en raison du délai de transfert de propriété, le cadastre ne soit définitif que douze mois après la vente des biens.

9.42 L'estimation de prix retenue par les administrations fiscales à des fins de taxation et pour le paiement des services locaux, ainsi que le prix de transaction définitif répertorié par les sociétés de crédit hypothécaire sont ceux qui ont le moins de chances de subir des révisions, tandis que le prix de transaction définitive provenant des données administratives des cadastres et des administrations fiscales peut être révisé sur une longue période en fonction des délais des processus d'enregistrement des transferts de propriété.

Comparabilité

9.43 On entend par comparabilité le niveau de *comparabilité* des indices des prix des logements *d'un pays à l'autre*. Cette notion est importante, car il peut être délicat de comparer les prix des logements issus de données nationales non harmonisées. En effet, des différences entre les concepts, la construction de l'indice, la couverture du marché, les procédures d'ajustement au titre de la qualité, etc., peuvent rendre difficiles les comparaisons entre pays. Les différences dans la fréquence, l'actualité et la politique de révisions peuvent également poser des problèmes de comparabilité.

9.44 Ces problèmes se posent au niveau national comme au niveau international :

- Les usagers d'un pays donné peuvent être confrontés soit à un manque de statistiques pertinentes, soit à des statistiques différentes correspondant à des périodes et des délais différents. Ces statistiques peuvent reposer sur des sources de données ou des méthodes de compilation différentes.
- La situation est plus complexe pour les usagers qui souhaitent faire des comparaisons internationales, car la disponibilité des données varie notablement entre les pays. Il est donc difficile d'établir des comparaisons entre données de même nature et d'interpréter des tendances relatives entre les pays. Les indices des prix agrégés qui couvrent des groupes de pays — nécessaires pour coordonner la politique économique et le suivi dans toute une zone économique comme la zone euro[13] — posent une autre difficulté.

Le chapitre 10 montre que les méthodes employées pour établir des indicateurs des prix de l'immobilier résidentiel varient considérablement *entre* les pays, et même entre les différentes sources *au sein* d'un même pays.

Pondérations

9.45 Les sources de données auxquelles on fait appel pour calculer les pondérations d'un indice des prix de l'immobilier résidentiel sont fonction à la fois des besoins de l'indice recherché et de la disponibilité des informations nécessaires. En outre, les besoins de données dépendent non seulement de la base conceptuelle de l'indice, mais aussi des aspects détaillés de sa construction comme la méthode d'ajustement de qualité et de tous les sous-indices dont on a besoin à des fins analytiques et autres. Ainsi, pour construire un indice des prix des propriétés (post)stratifié par catégorie de transactions, il faut connaître suffisamment d'informations sur les ventes durant chaque période pour pouvoir les classer dans des groupes assez homogènes de sorte que les valeurs unitaires puissent être traitées comme des prix. Sur le marché du logement, les problèmes sont amplifiés par le faible volume des ventes de certains types de logements dans des zones géographiques particulières, ce qui fait que de nombreuses cellules peuvent être vides[14].

9.46 Si l'on met de côté ces aspects détaillés de la construction de l'indice, le concept souhaité est le principal facteur qui détermine les besoins de données relatives aux pondérations. Un seul indice des prix ne peut répondre aux besoins divers des usagers. Pour estimer la formation brute de capital par exemple, il faut inclure uniquement les logements neufs, alors que pour estimer l'effet des variations de prix sur les stocks de capital, il faut que l'indice couvre l'ensemble des transactions.

9.47 On peut calculer les pondérations à partir de plusieurs sources, en particulier les données sur les comptes de la nation, les recensements nationaux réalisés périodiquement qui recueillent des informations sur le stock de logements, les informations provenant des banques sur les prêts contractés pour l'achat d'un logement, les statistiques sur le secteur du bâtiment, les registres officiels des propriétés, etc. Ces différentes sources de données peuvent manquer de cohérence, car le processus d'achat et de vente d'un logement est long et souvent complexe, et le prix de valorisation ou

13 Composée des 17 États membres de l'Union européenne qui avaient adopté l'euro en 2012.
14 La méthode de stratification ou poststratification a été analysée au chapitre 4. Dans l'exemple de la ville néerlandaise de «A», beaucoup de cellules étaient effectivement vides. On a proposé de suivre la méthode des produits appariés (panier fixe) pour résoudre ce problème.

le prix d'offre indiqué dans une demande de crédit hypothécaire ne donne pas nécessairement lieu à une vente et à un changement de propriété. D'autres problèmes se posent, comme la distinction entre ce qui est construit en vue d'être vendu et ce qui est construit en vue d'être loué. Il est rare que l'on trouve ces informations facilement à partir d'une seule source statistique. C'est pourquoi la construction des pondérations peut faire appel à une multitude de sources.

Pays en développement, logements traditionnels et marché informel du logement

9.48 Dans de nombreux pays en développement, une grande partie du stock de logements se compose de bâtiments neufs sur des terrains familiaux ou de bâtiments anciens fortement rénovés. Par ailleurs, bon nombre de logements sont construits par les propriétaires eux-mêmes. Leur construction peut durer plusieurs années si bien qu'à tout moment, il existe une part significative de logements qui peuvent être considérés comme non finis. Le recours au financement formel par crédit hypothécaire est souvent très limité, et peut être remplacé par un financement informel. Les constructions de logements vont de baraques édifiées avec des matériaux de récupération sur de la terre battue à de grandes habitations de plusieurs pièces en parpaings sur fondations en béton. Le niveau de confort varie du tout au tout. La mobilité des habitants est généralement très faible, surtout pour les propriétaires qui ont construit leur propre logement. Le marché de l'achat et de la vente est donc limité, avec très peu de passage de l'un à l'autre. En principe, l'établissement d'un indice des prix des logements est le même selon que le logement a été construit par son propriétaire ou par un tiers, mais les problèmes de mesure sont différents et généralement plus difficiles[15].

9.49 Ces complications font que l'on conserve rarement les factures attestant des coûts de construction d'un nouveau logement ou de rénovation d'un ancien, par exemple lorsque l'on apporte l'eau courante ou que l'on ajoute des toilettes à l'intérieur ou des pièces supplémentaires. Parfois, aucun transfert formel de propriété n'a lieu; souvent, aucune valorisation formelle n'est disponible et les méthodes de financement peuvent être informelles (la famille) ou tout simplement non-actées ou sans archivage centralisé. Il est alors impossible de calculer les remboursements d'intérêts d'emprunt (avec ou sans versement d'intérêts notionnels aux proches) ou d'estimer les coûts nets d'acquisition.

9.50 En l'absence de ces informations basiques, l'approche des loyers imputés est souvent la seule qui permette de construire un indice des prix des logements. L'indicateur du prix des loyers fictifs peut être calculé de deux manières : soit à partir d'une série de prix courants de loyers disponibles, pondérés conformément à la composition du stock de logements occupés par leur propriétaire et comparés aux équivalents-loyers pour la période de référence; soit en demandant à un expert de fournir tous les mois les équivalents-loyers pour un échantillon de logements représentatif du stock de logements occupés par leur propriétaire.

9.51 Dans chacun des cas, les stratifications par type de logement (maison individuelle ou appartement), par emplacement (région, zone urbaine ou rurale), ainsi que par d'autres caractéristiques qui influencent le loyer sont importantes pour que les loyers retenus reflètent la composition globale des biens occupés par leurs propriétaires. Il existe d'autres variables de stratification, comme la superficie totale de la parcelle, la surface habitable, le nombre de pièces, l'accès à l'eau courante, la présence de toilettes à l'intérieur, le raccordement au réseau électrique, les matériaux de construction utilisés et la conception (traditionnelle ou non) du bâtiment. Le statisticien des prix devra chercher conseil auprès d'un expert de la location, comme une société d'habitation, pour vérifier les critères déterminants pour le loyer, en veillant à ne pas les multiplier à l'infini. Les informations de pondération peuvent provenir des derniers recensements du logement ou des recensements de la population et du logement. Dans la pratique, ces informations risquent de ne pas être à jour à cause des changements dans le stock des logements occupés par leur propriétaire, toujours possible dans la période qui s'écoule entre deux recensements. Lorsque tel est le cas, il convient de faire des sondages spécifiques ou, notamment dans les zones urbaines y compris dans les bidonvilles, d'utiliser les demandes de permis de construire pour actualiser le dernier recensement.

9.52 Les problèmes de mesure peuvent cependant être de taille. Pour résumer, les logements traditionnels ou informels sont généralement construits par des membres de la famille ou par d'autres personnes non rémunérées. Les murs peuvent être composés de matériaux moins durables (argile séchée, bambou ou treillis) et les toits de roseaux, de paille, de feuilles de palmiers ou de tôle ondulée. Les logements ne sont pas nécessairement équipés d'électricité ou d'eau courante, *a fortiori* d'autres équipements. Les logements traditionnels se trouvent généralement dans les régions rurales. On peut rencontrer d'autres complications lorsque l'on tente d'inclure les coûts des logements occupés par leur propriétaire dans un indice des prix à la consommation :

- Beaucoup de ces logements se situent dans de grandes villes ou dans leur périphérie, dans des bidonvilles par exemple. Ils sont loués ou occupés par leur propriétaire, et il peut être

15 En particulier, les caractéristiques importantes qui servent à déterminer les prix des bâtiments peuvent être assez différentes dans les pays en développement et dans les pays développés. Dans un pays développé, il y aura sans doute moins de variations dans le type de constructions et les matériaux utilisés, tandis que les baraques peuvent comporter des différences plus marquées. De plus, dans les pays en développement, on dispose rarement des titres de propriété foncière, ce qui peut là encore poser des problèmes si l'on veut procéder à un ajustement au titre de la composition de l'échantillon et appliquer des techniques de régression hédonique pour faire des ajustements au titre de la qualité des logements.

difficile d'obtenir des détails sur leur régime de propriété. Il est parfois difficile de mener des enquêtes.

- Dans les régions rurales, beaucoup de ces logements ont pu être construits par une main-d'œuvre familiale sur des terrains familiaux, des terrains non cadastrés ou en propriété «commune».

Dans ces cas, le concept de «propriété» devient flou. La définition d'un logement occupé par son propriétaire et de ce dont une famille est véritablement propriétaire est donc sujette à débat; Même lorsque l'on s'entend sur une définition, il arrive que l'on ne dispose même pas des données les plus élémentaires sur le nombre de ces logements occupés par leur propriétaire et encore moins de détails sur les logements.

9.53 Les caractéristiques suivantes des logements traditionnels et des autres logements sur le marché informel peuvent être utiles pour calculer un indice des prix :

- Approvisionnement en électricité. Il s'agira souvent d'électricité fournie par une société de production ou de distribution. L'électricité peut cependant être produite par le ménage lui-même, grâce à un groupe diesel ou par énergie éolienne par exemple, ou obtenue illégalement auprès du distributeur.
- Eau courante. Les canalisations peuvent arriver jusqu'au logement, ou bien le logement s'approvisionne en eau à une borne collective ou à un puits.
- Toilettes privées ou collectives. Il peut s'agir soit de toilettes à chasse d'eau, soit de toilettes chimiques.

Par ailleurs, comme dans tout logement, la question de l'espace habitable se pose, exprimé en nombre de pièces, en superficie, ou les deux. Il faut à cet effet disposer de définitions adaptées. On doit définir notamment la surface de plancher utile (la surface de plancher du salon, de la cuisine, du couloir, de la salle de bains et de toutes les pièces contiguës, en excluant l'épaisseur des murs et les embrasures des portes et des fenêtres, les escaliers, etc.) et le nombre de pièces (en décidant par exemple s'il convient ou non d'inclure les couloirs).

9.54 Enfin, même si l'on dispose d'informations sur les caractéristiques de ces logements, il n'existe pas nécessairement de logement locatif «équivalent» pour évaluer les services d'un logement occupé par son propriétaire. Il est alors impossible de mesurer les prix *indirectement*. Dans ce cas, les statisticiens peuvent mettre en place un système pour mesurer le prix des intrants (coûts de construction), puis exploiter cette information pour construire une mesure du coût d'usage des services de logement à utiliser comme valeur approchée des prix des services de logement consommés[16]. S'agissant de la consommation pour compte propre, le *Système de comptabilité nationale 1993 (SCN 1993)* admet que l'on ne peut mesurer que le prix des intrants.

Les questions évoquées plus haut sont analysées dans l'étude de cas sur l'établissement des indices des prix de l'immobilier résidentiel en Afrique du Sud, au chapitre 10.

16 On trouvera dans Blades (2009) des éléments complémentaires sur la construction de ces coûts d'usage des logements traditionnels dans les pays en développement.

Méthodes utilisées actuellement

10

Introduction

10.1 En pratique, les méthodes employées pour construire les indices des prix de l'immobilier résidentiel peuvent subir les contraintes liées à la nature des données disponibles. Une fois l'indice cible défini, les données nécessaires à son élaboration ne sont pas toujours disponibles à intervalles réguliers et au moment opportun, voire pas du tout. En outre, même là où il existe des données appropriées à l'établissement d'un indice de prix répondant aux besoins d'une catégorie d'utilisateurs, il est fréquent qu'elles ne correspondent pas à ceux d'une autre. Dans de nombreux pays, la mise en place des infrastructures et des procédures de recueil des données nécessaires à la production d'un indice de prix peut être trop onéreuse. Par ailleurs, les changements de méthodologie et de sources des données peuvent empêcher la construction des séries historiques dont on a souvent besoin pour la modélisation économétrique et les analyses portant sur plusieurs cycles d'évolution du marché immobilier qui permettent d'éclairer les décisions de politique économique. Enfin, et ce n'est pas le moins important, le moment d'obtention et la fréquence des données, lorsqu'elles sont disponibles, ne conviennent pas toujours pour produire le type d'indice des prix du logement que les utilisateurs veulent ou dont ils ont besoin.

10.2 Ces déficiences des données servant à construire les indices de prix de l'immobilier et les indicateurs correspondants ont été parfois dénoncées par les utilisateurs. Ainsi, un ancien gouverneur de la Banque du Canada déclarait dans un discours prononcé à la Conférence des statisticiens européens (Dodge, 2003) : «Sachant que l'investissement immobilier représente une grande part des dépenses des ménages et que la plupart des gens considèrent le logement comme leur actif le plus précieux, il est surprenant que de nombreux pays ne disposent pas de données exhaustives et ajustées en fonction de la qualité sur les prix des logements ou sur les loyers.»

10.3 De plus, les sources de données et les méthodes employées ne sont pas toujours bien documentées, tandis que les enquêtes sur les métadonnées relatives aux prix de l'immobilier résidentiel confirment l'absence d'harmonisation des pratiques. Cet état de fait constitue un autre défi pour les utilisateurs. Il limite notamment la possibilité de procéder à des comparaisons internationales significatives des tendances des prix et rend extrêmement difficile toute étude économique comparative. On peut donc s'interroger sur la crédibilité des résultats.

10.4 Outre la question de la disponibilité des données, les méthodes utilisées par les différents pays pour établir les indices de prix de l'immobilier résidentiel se heurtent aussi à certains problèmes inhérents à ce secteur; en particulier, les biens ayant des caractéristiques uniques, il en résulte une hétérogénéité portant sur divers aspects dont beaucoup sont ainsi difficiles à mesurer objectivement. La faible fréquence des transactions concernant chaque bien est aussi un problème particulier de ce secteur. Ces deux problèmes compliquent sérieusement l'établissement des indices de prix. En outre, du fait que le prix demandé par le vendeur est négociable, le prix effectif de l'opération diffère éventuellement du prix demandé initialement, du prix demandé finalement, du prix proposé par l'acheteur et de l'évaluation à dire d'expert.

10.5 À propos des méthodes les plus répandues de construction des indices des prix de l'immobilier résidentiel, la question se pose de savoir s'il est possible de distinguer une meilleure pratique internationale ou si les techniques adoptées dépendent inévitablement des conditions locales.

10.6 D'autres sections du manuel formulent des recommandations à propos de la meilleure pratique. Ce chapitre présente l'ensemble des indices disponibles dans les différents pays et certaines études de cas. Il se réfère aux métadonnées assemblées par diverses organisations, dont la Banque des règlements internationaux et la Banque centrale européenne, ainsi que, plus récemment, à un exercice exploratoire effectué par Eurostat dans le cadre de l'inclusion des coûts des logements occupés par leur propriétaire dans l'indice harmonisé des prix à la consommation de l'Union européenne; cet exercice a été étendu à certains pays non membres de l'UE. Les métadonnées relatives aux indices des prix de l'immobilier résidentiel publiées par divers pays sont consultables sur le site Internet de la Banque des règlements internationaux (BRI); voir www.bis.org/statistics[1].

Indices existants

10.7 En Europe, Eurostat a commencé en décembre 2010 à diffuser des rapports trimestriels concernant des indices expérimentaux des prix des logements dans l'Union européenne (UE) et dans la zone euro[2]. Ils contiennent, pour ceux des offices statistiques de l'UE qui ont autorisé la publication, des données expérimentales. On trouve dans les annexes à ces rapports tous les liens actuels aux pages Internet des instituts nationaux de la statistique traitant des indices des prix des logements avec des précisions sur leur mode d'élaboration.

10.8 L'examen des métadonnées consultables sur le site Internet de la BRI[3] est révélateur de la forte différenciation des méthodes utilisées pour établir les indices de prix de l'immobilier résidentiel, à la fois *parmi* les pays et même *au sein de* chacun. La différenciation à l'intérieur des pays soulève la question fondamentale du choix de la série correspondant aux besoins particuliers des utilisateurs.

1 Les statistiques des prix de l'immobilier que l'on trouve sur le site Internet de la BRI comprennent des données émanant de 37 pays et sont disponibles avec des périodicités diverses. Les données diffèrent sensiblement d'un pays à l'autre, par exemple en ce qui concerne les sources d'information sur les prix, les types de biens, les domaines couverts, les dates de construction, les unités dont on calcule les prix, les détails des méthodes d'établissement des indices et les techniques de correction des variations saisonnières. Il y a deux raisons à cela. D'abord, les procédures d'acquisition et de cession d'un bien immobilier, et donc les données disponibles, varient selon les pays; ensuite, il n'existe pas actuellement de normes internationales spécifiquement applicables aux prix de l'immobilier.

2 Voir http://epp.eurostat.ec.europa.eu/portal/page/portal/hicp/methodology/owner_occupied_housing_hpi/experimental_house_price_indices.

3 Voir http://bis.org/statistics/pp.htm.

La différenciation entre pays met en cause la validité des comparaisons internationales.

10.9 Les différences entre les indices disponibles des prix des logements portent sur presque tous les aspects de la construction d'un indice de prix. Les précédents chapitres en ont fait état; il s'agit du fondement conceptuel de l'indice (quel est l'indice cible convenant aux besoins de chaque utilisateur?); des sources des données (enregistrement des biens, dossiers fiscaux, demandes de crédit immobilier et leur finalisation, agents immobiliers, médias imprimés comme les journaux, et autres formes de publicité); de la couverture du marché (couverture géographique, type de biens, opérations à crédit et au comptant); de la qualité de l'ajustement (hédonique, mixte) et de la pondération (par les stocks ou par les ventes). Les problèmes posés par ces différents facteurs peuvent être compliqués par une forte hétérogénéité des marchés immobiliers. Ainsi, les prix des biens ne varient pas seulement en fonction de leurs caractéristiques physiques, telles que la surface, et du fait qu'il s'agit de maisons individuelles bâties sur un terrain possédé en propre ou d'appartements situés dans un grand immeuble. Ils peuvent aussi diverger fortement selon, par exemple, la région du pays, le secteur de la ville, ou le caractère rural ou urbain de l'emplacement géographique. Celui-ci influe sur l'attrait, lequel modifie les conditions de la demande, ce qui explique pourquoi deux maisons identiques peuvent avoir un prix différent selon le lieu où elles se trouvent. Par exemple, un bien situé dans une région à PIB par habitant élevé, à faible chômage, dans une localité connue pour la qualité de ses écoles et de l'environnement, aura un prix plus élevé qu'un bien identique mais situé dans une zone pénalisée par un chômage élevé, de bas revenus des ménages, des établissements scolaires médiocres et un taux de criminalité important[4].

10.10 On présente ci-dessous une vue d'ensemble de la situation actuelle. Il convient de noter qu'une évolution se produit au fur et à mesure que davantage de pays améliorent leurs propres indices de prix de l'immobilier résidentiel et réexaminent ceux actuellement publiés. Le lecteur doit se référer aux données consultables sur les sites Internet de la BRI, d'Eurostat et de la BCE pour en savoir plus sur les indices d'un pays en particulier.

Responsabilité de l'établissement des indices

10.11 Dans l'UE, les instituts de statistiques ont œuvré ensemble à la conception et à l'élaboration d'indices de prix de l'immobilier résidentiel basés sur des méthodes statistiques à peu près harmonisées, ce qui a ouvert la voie à la création d'indices internationalement comparables. En outre, plusieurs banques centrales nationales établissent des indicateurs de prix des logements, notamment en Allemagne, en Autriche, en Belgique, à Chypre, en Grèce, en Hongrie, en Italie, au Luxembourg, à Malte, en Pologne et en Slovaquie. En Autriche, la Banque centrale partage cette tâche avec l'Université de technologie de Vienne, alors qu'au Luxembourg l'indice est établi par la Banque centrale à partir de données de l'Institut national de statistique. En Espagne, aux États-Unis, en France, en Irlande et au Royaume-Uni, les indices des prix des logements sont établis par des administrations autres que l'office de la statistique. Parfois, comme au Royaume-Uni, c'est parce que le système statistique est décentralisé, les statisticiens se trouvant dans les ministères et travaillant aux côtés de leurs collègues chargés de l'élaboration des politiques et de la prestation de services. Dans certains cas, la responsabilité de l'établissement de l'indice incombe au ministère qui exerce la responsabilité politique, opérationnelle ou juridique du secteur de l'immobilier. On peut citer l'Agence fédérale du financement du logement aux États-Unis et le cas du Royaume-Uni. Le ministère investi de cette responsabilité est souvent le mieux placé pour obtenir des informations administratives à des fins statistiques; il est aussi censé bien connaître le secteur et peut même disposer d'informations utiles sur le contexte.

Sources des données

10.12 Au Canada, aux États-Unis et dans plusieurs pays européens[5], les données sur les prix de l'immobilier résidentiel sont collectées par les instituts nationaux de la statistique ou les ministères. Au Danemark, en Finlande, à Hong Kong, en Lituanie, en Norvège, aux Pays-Bas, au Royaume-Uni, en Slovénie et en Suède, les données obtenues aux fins d'enregistrement ou par l'administration fiscale constituent la source des indices. En Allemagne, l'Office fédéral de la statistique relève les prix calculés par les commissions locales d'experts chargées d'évaluer les biens immobiliers. En Espagne et en France, les instituts de la statistique calculent les indices de prix à partir d'informations communiquées par les notaires. En Allemagne, en Belgique, en France, en Grèce, en Italie, au Portugal et en Slovaquie, les agences immobilières, les associations pour le logement, les organismes de recherche ou les conseils spécialisés sont à l'origine des informations sur les prix. Des données émanant de la presse ou de sites Internet sont recueillies pour établir les indices de prix de l'immobilier résidentiel dans des pays comme Malte, la Hongrie («Origo») et l'Autriche («Austria Immobilienbörse»). Il est intéressant de remarquer qu'il n'existe qu'un nombre limité de cas d'intégration des différentes sources de données pour produire un meilleur indice, malgré le nombre de pays qui déclarent avoir de multiples sources d'information sur ces prix. En Allemagne, en Irlande et au Royaume-Uni, elles proviennent, entre autres, des entités octroyant des prêts immobiliers. Au Royaume-Uni, l'indice établi par le Ministère des communautés et des collectivités locales est basé sur une enquête effectuée par le Conseil

4 Voir par exemple Chiodo, Hernandez-Murillo et Oryang (2010).

5 En ce qui concerne les sources des données dans les pays de l'UE, voir aussi Eiglsperger (2010).

des organismes de crédit immobilier; la longueur des délais d'enregistrement des transferts de propriétés ne permet pas d'en faire un indicateur disponible en temps utile. En Allemagne, l'Association des banques spécialisées («Pfandbrief») se sert des données fournies par ses membres pour établir un indice des prix de l'immobilier résidentiel.

10.13 La diversité des sources de données que l'on vient de mentionner peut limiter sérieusement la comparabilité des indices : achats à crédit ou achats au comptant; prix en zone urbaine ou prix en zone rurale; prix des biens anciens ou prix des constructions neuves; évaluations ou prix affichés ou prix proposés initialement par les acheteurs ou prix finaux des transactions. Le résultat net est que les indices publiés peuvent en pratique mesurer des aspects très différents de l'évolution des prix des marchés immobiliers. La diversité des sources de données et des techniques d'établissement des indices ainsi que leur mode d'utilisation (c'est-à-dire la finalité de l'indice) expliquent les disparités importantes en matière de délais de publication et de politique de révision.

Méthodologie des indices

10.14 En raison des difficultés inhérentes à la mesure des prix et résultant de la diversité des sources de données utilisées, les pays recourent à une large gamme de méthodologies différentes pour construire les indices des prix des logements.

Ajustement en fonction de la qualité (mixte)

10.15 L'ajustement en fonction de la qualité, destiné à prendre en compte les changements de composition et de qualité de chaque bien, constitue un élément fondamental de la méthodologie des indices. Il garantit que la comparaison des prix porte sur des éléments homogènes et permet d'éviter l'existence de distorsions dans les séries lorsque, par exemple, la qualité du stock s'améliore du fait, entre autres facteurs, de rénovations pouvant prendre des formes diverses comme la modernisation des cuisines et des salles de bains ainsi que l'installation d'une meilleure isolation et du chauffage central ou de l'air conditionné. Les techniques d'ajustement en fonction de la qualité jouent aussi un rôle important dans l'établissement des indices de prix de l'immobilier, car les logements commercialisés changent au fil du temps.

10.16 L'ajustement au titre de la qualité s'applique de différentes façons. En Estonie, par exemple, l'indice des prix de l'immobilier résidentiel est calculé à partir de valeurs unitaires, en l'occurrence les prix moyens des transactions par mètre carré de surface de plancher (dans ce cas particulier, on fait la somme des valeurs de toutes les transactions immobilières et on la divise par la somme des mètres carrés de surface de plancher pour toutes les cessions de biens immobiliers, en excluant les valeurs aberrantes). Mais les indices de valeurs unitaires basés sur le prix du mètre carré de surface de plancher, s'ils opèrent dans chaque période un ajustement au titre de la dimension des logements, ne le font pas pour les différences de qualité de la construction ou pour l'âge de l'immeuble; et il y a peut-être plus important : l'absence d'ajustement pour tenir compte de l'évolution de la moyenne des dimensions des parcelles dans l'échantillon des biens vendus au cours d'une période donnée. D'autres évolutions des caractéristiques d'un logement sont susceptibles de se produire parallèlement aux tendances générales du marché; on les retrouve dans les changements de composition de l'échantillon comme l'emplacement, les équipements, l'environnement, la qualité générale des habitations, etc.

10.17 La principale solution de remplacement de la *stratification*, évoquée au chapitre 4, consiste à utiliser une classification des logements selon ce que l'on considère généralement comme les caractéristiques importantes déterminant les prix pour calculer les indices de prix élémentaire de chaque cellule de la matrice de classification. On calcule ensuite l'indice général en faisant la moyenne pondérée de ces sous-indices. Cette méthode est une forme de stratification particulière. Elle est utilisée notamment par le Australian Bureau of Statistics (ABS) pour tenir compte des variations de la composition des logements dans l'établissement des indices trimestriels pour chacune des huit capitales. Leur méthode stratifie les logements selon deux caractéristiques : le niveau à long terme des prix dans la banlieue où un logement est situé, et les caractéristiques du voisinage, représentées par les indices socioéconomiques des différentes zones calculés par l'ABS (SEIFA)[6]. En pratique, le nombre des caractéristiques incluses dans la classification est souvent limité par celui des observations que l'on peut faire régulièrement pour chaque cellule, c'est-à-dire par l'aptitude à remplir la «base de données des caractéristiques déterminant les prix» à partir des sources de données disponibles et par l'existence d'informations relatives à ces caractéristiques.

10.18 La forme la plus sophistiquée d'ajustement au titre de la qualité qu'emploient les pays est la technique de *régression hédonique* (évoquée au chapitre 5), qui consiste à employer un modèle de régression pour isoler la valeur de chacune des caractéristiques retenues et tenir compte de l'évolution des caractéristiques des biens vendus. Mais cette méthode exige généralement davantage de données. On l'utilise parfois en conjonction avec la stratification par type de bâtiment et par emplacement. Son utilisation pour établir les indices de prix de l'immobilier résidentiel est une innovation très récente. Les pays publiant des indices élaborés de cette façon sont l'Allemagne, l'Autriche, l'Irlande, la Finlande, la France, la Norvège et le Royaume-Uni. Le modèle hédonique servant à produire l'indice norvégien ne comprend que quelques variables explicatives et n'est pas ajusté en fonction des normes des logements et de la date de leur construction[7];

6 Voir http://www.abs.gov.au/ausstats/abs@.nsf/mf/6464.0.

7 Comme on l'a vu dans les chapitres antérieurs utilisant les données pour la ville de «A», l'ancienneté de la construction est une caractéristique importante pour déterminer le prix.

l'ajustement se limite à la dimension et à l'emplacement de l'habitation. L'indice est probablement «biaisé» (sauf si l'âge des bâtiments et les types de logements vendus sont stables). Statistics Norway reconnaît cette déficience.

10.19 Il existe une autre méthode, employée par exemple aux États-Unis et au Canada, celle des *ventes répétées* (décrite au chapitre 6); elle sert à l'établissement de l'indice Case–Shiller des prix immobiliers aux États-Unis et de l'indice Teranet–National Bank™ au Canada. Elle consiste à apparier les ventes de logements identiques au fil du temps. Elle exige une très importante base de données sur les transactions et aucun des producteurs européens d'indices ne l'utilise.

10.20 On remarque que l'un des indices allemands des prix de l'immobilier résidentiel est basé sur des données limitées aux logements «de bonne qualité», ce qui pourrait impliquer que l'on ignore l'ajustement en fonction de la qualité. En pratique, il peut y avoir un problème intrinsèque de mesure, puisque la définition donnée par le marché de la «bonne qualité» n'est sans doute pas indépendante, sur la durée, de la hausse générale des normes du logement. L'indice ainsi obtenu est donc exposé à un biais éventuel à long terme. À cela s'ajoutent les préoccupations à propos de l'échantillonnage, en particulier sur l'aptitude des logements «de bonne qualité» à représenter la tendance des prix de l'ensemble.

10.21 On peut conclure des deux paragraphes ci-dessus que toutes les procédures d'ajustement en fonction de la qualité soulèvent deux questions essentielles : 1) les caractéristiques choisies pour procéder à cet ajustement sont-elles les principaux déterminants des différences de prix? 2) L'application de différentes techniques aux mêmes séries de données produira-t-elle les mêmes résultats? (question de la robustesse statistique). En réalité, s'il est facile de mesurer certaines des caractéristiques des déterminants des prix — comme la dimension de la surface habitable — d'autres facteurs importants, tels que l'emplacement[8] et la qualité de la construction, peuvent être intrinsèquement difficiles à appréhender et à mesurer. En outre, il faut bien voir que l'application aux mêmes séries de données de techniques différentes d'ajustement qualitatif ne produit pas nécessairement les mêmes résultats[9].

La valeur des métadonnées

10.22 Plusieurs organisations ont des sites Internet où l'on trouve des métadonnées sur les indices des prix de l'immobilier résidentiel publiés par différents pays. La Banque des règlements internationaux, notamment, diffuse ces informations (voir la référence antérieure). Elles s'ajoutent à celles communiquées par certains pays, par exemple sur les sites Internet de l'institut national de la statistique ou de la banque centrale.

10.23 Tout en donnant à l'utilisateur des indications sur les forces et les faiblesses d'un indice de prix particulier et sur la façon de l'utiliser, une analyse systématique et plus détaillée des métadonnées sur les statistiques actuellement disponibles et les sources de données peut permettre de faire apparaître :

- les principales insuffisances dans la fourniture des données;
- les différentes solutions pour y remédier efficacement à partir des sources directement disponibles;
- les problèmes de cohérence des données;
- les possibilités d'une intégration plus poussée des données et la nécessité de nouvelles sources.

10.24 Une étude des métadonnées de base révèle aussi les compromis opérés en s'appuyant sur les données directement disponibles et dans les situations où un indice des prix de l'immobilier est utilisé à de nombreuses fins. Ainsi, le principal indice officiel publié par le Ministère britannique des communautés et des collectivités locales s'appuie sur les ventes pondérées; il peut donc, par exemple, être inclus dans un indice des prix à la consommation servant à l'indexation de prestations sociales, mais ne répond pas pleinement aux besoins des utilisateurs désireux de calculer la «richesse», auquel cas les pondérations par les stocks sont plus appropriées que celles par les dépenses. On peut satisfaire à ces besoins soit en repondérant l'indice officiel, soit en se référant à l'un des nombreux indices publiés par les établissements de crédit. Toutefois, ces derniers ont l'inconvénient d'assurer une couverture limitée. La repondération de l'indice officiel peut donc constituer une solution économique pour combler cette lacune spécifique des données.

10.25 Une analyse plus détaillée des lacunes existantes peut orienter vers des estimations synthétiques fondées sur l'intégration de données de sources diverses. Au Royaume-Uni, par exemple, l'indice DCLG mentionné ci-dessus a l'avantage d'être disponible en temps utile et de ne pas être révisé, mais a l'inconvénient d'exclure les achats au comptant.

10.26 Une étude systématique de la construction des indices de prix de l'immobilier résidentiel au Royaume-Uni pourrait conclure à l'opportunité de compléter l'indice officiel par des informations sur les achats au comptant émanant du cadastre. Bien qu'elles soient moins à jour en raison des délais d'enregistrement officiel des transactions, une modélisation de séries temporelles permettrait sans doute de remédier à ce problème. Le cadastre établit un indice de ventes répétées en suivant la progression moyenne des prix immobiliers au travers d'opérations multiples portant sur le même bien, et cela pour tenter de maintenir la qualité constante.

8 On peut mesurer assez précisément la situation géographique d'un bien, mais le problème posé par l'«emplacement» concerne les regroupements de biens. Les méthodes de stratification et de régression hédonique doivent grouper les ventes de biens situés au même endroit, mais il reste à savoir comment déterminer exactement les limites d'un emplacement.

9 Ce point est illustré par les différences entre les indices obtenus en appliquant les diverses méthodes d'ajustement qualitatif décrites précédemment aux chapitres 4 à 8, avec le même ensemble de données que pour la ville de «A». Or, toutes ces méthodes ont fait apparaître des tendances de prix à peu près similaires.

La prochaine section présente une série d'études de cas relatives aux indices de prix de l'immobilier résidentiel de certains pays.

Études de cas

Étude de cas : le Canada

10.27 Au Canada, il existe actuellement quatre indices des prix de l'immobilier. Ce sont l'indice des prix des logements neufs de Statistique Canada, l'indice composite des prix des maisons Teranet–National Bank™, la mesure des prix moyens des logements par l'Association canadienne des agents immobiliers (CREA) et l'enquête sur les prix des maisons de Royal LePage. Nous allons les présenter l'un après l'autre.

L'indice des prix des logements neufs (IPLN)

10.28 C'est un indice mensuel qui mesure l'évolution des prix auxquels les constructeurs vendent les nouveaux logements. Les prix recueillis proviennent d'une enquête effectuée auprès des constructeurs dans diverses régions du pays. Il s'agit d'un indice de prix à qualité constante dans la mesure où les caractéristiques des unités de l'échantillon sont identiques; en d'autres termes, l'IPLN est un indice à panier fixe. Les estimations distinctes formulées par les constructeurs sur la valeur actuelle (évaluée au prix du marché) des parcelles constituent aussi une partie importante de l'enquête. C'est pourquoi Statistique Canada publie aussi une série séparée des indices de prix des terrains excluant les bâtiments. La valeur résiduelle (prix de vente total moins valeur du terrain) donne une indication de la tendance du coût des bâtiments et figure aussi dans une série à part. À l'heure actuelle, les trois variantes de l'IPLN sont publiées pour 21 zones métropolitaines du Canada.

10.29 Les experts du marché de l'immobilier, les universitaires et la population considèrent l'IPLN comme un indicateur disponible en temps utile de la situation passée et actuelle de ce marché. Il sert aussi d'intrant pour l'établissement d'autres statistiques économiques. On l'utilise par exemple pour estimer certaines composantes de l'indice des prix à la consommation relatives au logement. De plus, la comptabilité nationale canadienne s'appuie sur l'IPLN pour estimer la valeur à prix constants des nouvelles constructions à usage résidentiel. En raison de son degré de précision géographique et de sa sensibilité aux variations de l'offre et de la demande, l'IPLN présente un intérêt pour les professionnels de l'immobilier et donne une estimation représentative de l'évolution des valeurs des maisons à la revente. Les informations qu'il contient intéressent aussi les entrepreneurs du bâtiment, les experts de la politique du logement, les fournisseurs et les fabricants de matériel de construction, les compagnies d'assurance, les agences fédérales comme la Société canadienne d'hypothèque et de logement (SCHL), ainsi que les organismes des provinces et des communes responsables de la politique du logement et de la politique sociale.

10.30 Les prix recueillis, qui sont les prix demandés par les constructeurs, excluent la taxe sur les biens et services et les autres remises liées à la fiscalité. Les prix qui ne figurent pas, par exemple à cause de l'absence de ventes d'un constructeur pendant un mois donné, sont imputés à partir de la meilleure estimation que le constructeur ferait si un logement devait être vendu. Tous les types de logement ne sont pas inclus dans l'IPLN. Les immeubles en copropriété sont exclus de l'échantillon, alors que les maisons unifamiliales, même en rangée, sont enquêtées. Comme les constructeurs ne déclarent pas de façon uniforme les prix des terrains à bâtir, les indices des prix des terrains sont peut-être moins exacts et précis que l'IPLN global. On peut faire la même réserve pour les valeurs résiduelles calculées qui servent à établir les indices de prix des bâtiments. Tant les grandes entreprises de construction que les petits entrepreneurs indépendants sont représentés dans l'échantillon utilisé pour élaborer l'IPLN .

10.31 Du fait même de sa nature, l'IPLN canadien ne mesure que les changements des prix des logements neufs et n'est donc pas représentatif des logements revendus au Canada (ou de la plupart des logements neufs construits au cœur des villes couvertes par l'enquête). Les logements couverts se trouvent généralement sur de nouveaux emplacements dans les banlieues des villes faisant l'objet de l'enquête, où le prix des terrains est sensiblement plus bas que dans les centre-villes. L'évolution des prix des terrains dans les banlieues est généralement différente de celle observée dans les quartiers réputés des villes canadiennes. Alors que la part du prix de la construction dans l'indice IPLN est probablement exacte (le coût d'édification du bâtiment d'une habitation est à peu près le même partout), la composante «terrain» sous-estime vraisemblablement de façon significative la hausse des prix des terrains pour le stock de logements existants ces dernières années[10].

L'indice composite des prix des maisons Teranet–National Bank™[11]

10.32 L'indice des prix des maisons Teranet–National Bank™ (IPMTNB) est une estimation indépendante, des taux de variation des prix des logements dans les zones métropolitaines, à savoir Ottawa, Toronto, Calgary, Vancouver, Montréal et Halifax. Les six indices sont ensuite agrégés pour former un indice national composite. Ils sont estimés mensuellement à partir des prix des transactions portant sur les immeubles en copropriété, les maisons en rangée et les maisons unifamiliales dans ces zones métropolitaines.

10.33 L'IPMTNB est calculé avec la méthodologie des ventes répétées. Son estimation repose donc sur le postulat

10 Voir graphique 10.1 pour une comparaison entre l'IPLN et les autres indices canadiens. Ce graphique confirme le biais baissier probable de la composante foncière de l'IPLN.
11 ©Teranet et Banque nationale du Canada, tous droits réservés.

selon lequel les logements achetés plus d'une fois dans les périodes d'échantillon sont de qualité constante. L'IPMTNB tente de tenir compte des changements qualitatifs des unités individuelles de logement en réduisant au minimum ou en éliminant l'influence de modifications éventuelles des caractéristiques physiques (par exemple rénovations, rajouts, etc.). Sachant que l'on ne prend pas en compte la dépréciation nette des biens vendus, l'indice est sans doute légèrement faussé dans le sens de la baisse[12]. Les biens affectés de facteurs endogènes sont exclus du calcul de l'indice des ventes répétées. Ces facteurs sont les suivants : cession entre des parties inégales ou liées; changement de catégorie du bien (par exemple après rénovation); erreur de données et rotation fréquente (semestrielle ou plus).

L'indicateur MLS® des prix moyens de revente des logements

10.34 L'Association canadienne des agents immobiliers (CREA) suit, sur une base mensuelle, le nombre et les prix des biens vendus dans le cadre du Service d'inscriptions multiples® (MLS®) géré par les chambres immobilières. Ceux qui veulent utiliser ces statistiques doivent souscrire un abonnement payant. Bien que l'indicateur porte seulement sur les logements vendus au moyen du MLS®, le système est très utilisé puisque quelque 70 % des logements mis sur le marché l'utilisent. On dispose de ces données pour plus de 25 marchés urbains définis par la CREA, ainsi que pour les provinces et deux territoires; un agrégat national est aussi publié.

10.35 Les indicateurs sont de simples moyennes arithmétiques de la totalité des prix de vente sur le marché considéré, quelle que soit la catégorie de logements. En outre, on ne prend pas en considération les changements de composition de l'échantillon au fil du temps ou la qualité inégale des unités. Dans ces conditions, une variation de l'indicateur de prix peut résulter de bien d'autres facteurs que leur évolution véritable. Ces facteurs vont des différences qualitatives au sein de l'échantillon d'une période à l'autre jusqu'à l'influence de valeurs aberrantes, c'est-à-dire de prix extrêmement élevés ou bas dus à des circonstances particulières. L'Association a récemment publié dans ses rapports mensuels une version pondérée de l'indice national (disponible seulement jusqu'en 2006), les pondérations correspondant à la part des unités de logement en propriété sur les principaux marchés calculée au moyen du recensement de 2006. Mais le prix sur chacun de ces marchés reste une moyenne simple et on ne tente pas de suivre les tendances, éventuellement divergentes, des diverses catégories de logements. Le grand avantage des indices MLS par rapport aux autres indicateurs est qu'ils sortent rapidement, puisque les données sont normalement diffusées deux semaines après le mois de référence.

Enquête Royal LePage sur les prix des maisons — Banque du Canada

10.36 Les prix ressortant de l'enquête Royal LePage reflètent les opinions de cette entité à propos de la «valeur de marché équitable» de sept catégories de biens situés dans un grand nombre de zones géographiques. Les informations obtenues s'appuient sur des données locales et sur la connaissance du marché propre aux courtiers de Royal LePage. La couverture géographique est étendue, comme pour les données MLS, mais le mode de classification des logements est plus sophistiqué. L'enquête porte sur les prix de quatre catégories de maisons individuelles (maisons individuelles de plain-pied, maisons individuelles à deux étages de niveau supérieur et standard et maisons luxueuses), de deux catégories d'appartements situés dans des immeubles en copropriété (niveau standard et de luxe) et d'une maison en rangée. Royal LePage normalise chaque catégorie en se référant à la surface, au nombre de chambres, à celui des salles de bain, au type de garage, aux caractéristiques de l'emplacement, à la configuration du sous-sol et à d'autres critères. En outre, les biens couverts par l'enquête sont situés à une distance moyenne du centre ville proche et ressemblent aux autres maisons du voisinage. À condition que le courtier répondant à l'enquête s'en tienne à ces références, on est assuré d'un certain degré de constance de la qualité. L'inconvénient comparatif des données de prix de Royal LePage est la longueur du délai de publication.

10.37 Cette enquête sert de base à l'un des indicateurs des prix des logements utilisés par la Banque du Canada pour suivre l'évolution du marché de l'immobilier dans ce pays[13]. Malgré l'abondance d'informations sur les prix de nombreux autres types de logements que l'on trouve dans l'enquête Royal LePage, l'indicateur mis au point par la Banque porte seulement sur un sous-ensemble de maisons individuelles qui ont été considérées comme représentatives du marché lors de sa création en 1988[14]. Pour le Canada et pour 11 marchés locaux, l'indicateur de prix de la Banque est la somme pondérée des prix des maisons individuelles de plain-pied individuels (pondération de 0,75) et des maisons individuelles à deux étages de niveau supérieur (pondération de 0,25). Ensuite, le prix de chaque catégorie de logements est la somme pondérée des sous-régions, les pondérations fixées correspondant à la part infrarégionale des unités vendues à une date donnée à la fin des années 80. Les données relatives aux «unités» proviennent du MLS®.

Analyse comparative

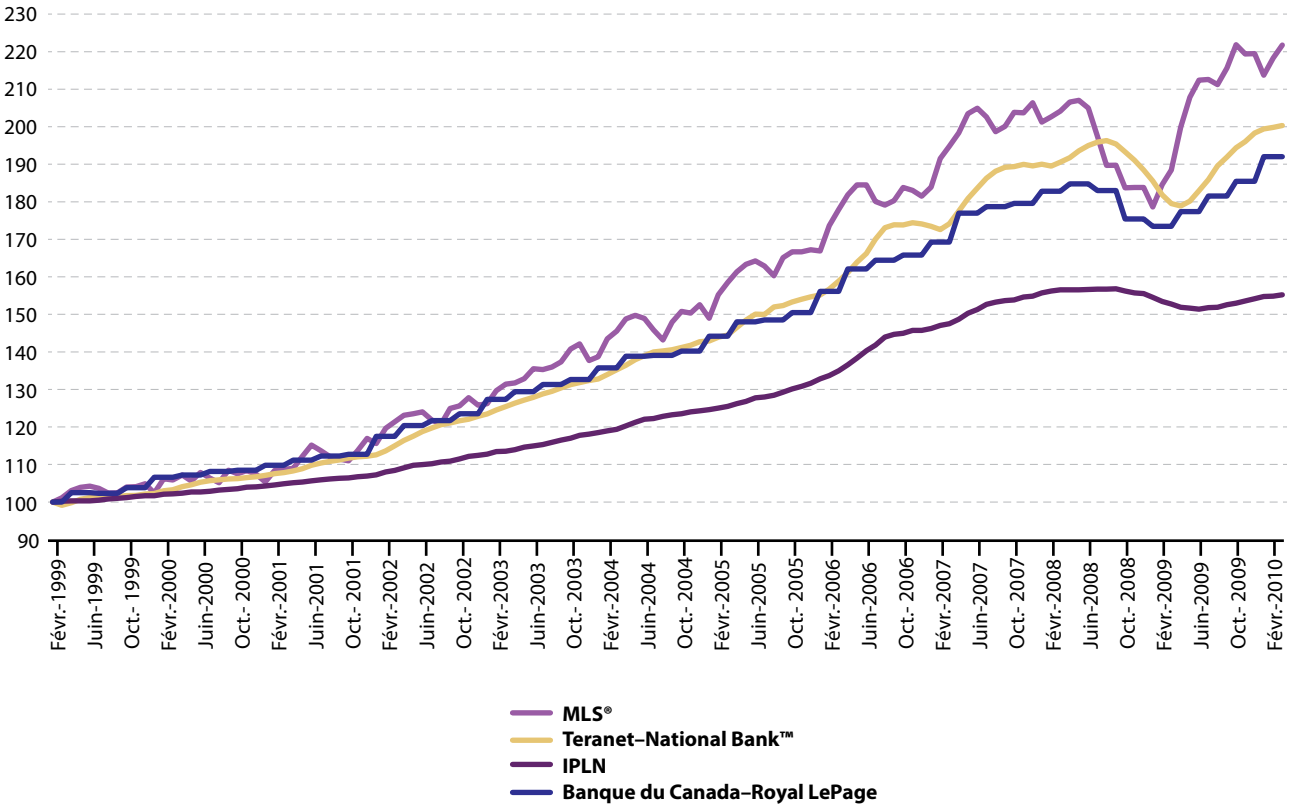
10.38 Le graphique 10.1 compare les quatre indices des prix de l'immobilier qui existent au Canada. Pendant la

12 Ce «biais» baissier n'apparaît pas au graphique 10.1, l'IPMTNB se situant entre les deux indices concurrents qui couvrent le marché de la revente, mais ces derniers ne sont pas ajustés en fonction de la dépréciation nette. Selon certains économistes spécialistes du logement, la méthode des ventes répétées pourrait être faussée dans le sens de la hausse à cause d'un problème de sélection de l'échantillon. Comme les unités de logement vendues plus souvent que la moyenne sont peut-être davantage rénovées et améliorées, la qualité d'une unité à ventes répétées augmente entre deux dates de vente (au lieu de diminuer du fait de la dépréciation).

13 http://www.bankofcanada.ca/en/rates/indinf.html.

14 L'indicateur de la Banque du Canada se limite aux maisons individuelles de plain-pied et aux maisons individuelles à deux étages de niveau supérieur.

Graphique 10.1. Quatre indices des prix de l'immobilier résidentiel au Canada
(Février 1999 = 100)



période couverte, qui va de février 1999 à mars 2010, tous indiquent une tendance à la hausse des prix de l'immobilier résidentiel. Toutefois, les taux de progression diffèrent. L'IPLN affiche la hausse la plus modérée, soit 55 %. En revanche, l'indicateur MLS® fait ressortir une augmentation de 122 %, c'est-à-dire plus du double de la précédente. L'indice Teranet–National Bank™ et l'indicateur Banque du Canada–Royal Lepage progressent respectivement de 100 % et de 92 %.

10.39 Le taux de progression supérieur de l'indicateur MLS® peut s'expliquer, au moins en partie, par la méthodologie des moyennes des prix employée pour le calculer. Il est bien connu qu'elle ne tient pas compte des changements de composition d'une période à l'autre et qu'il peut en résulter un rythme de hausse plus élevé de l'indice en cas d'une évolution des ventes de logements vers le haut de gamme. L'explication de la progression plus lente de l'IPLN est probablement que, si cet indice tient compte dans la durée de la catégorie de logements, il ne le fait pas pour l'emplacement. Les logements neufs sont construits de plus en plus loin des centre-villes, là où le marché de la vente se comporte différemment de ce que l'on observe au cœur des agglomérations ou à proximité.

10.40 Les quatre indices témoignent de la chute des prix intervenue lors de la récession amorcée à la fin de 2008, mais l'indice MLS® commence à baisser juste avant les trois autres et son inflexion est plus marquée. En comparaison avec les trois autres indices, la baisse de l'IPLN commence un peu plus tard et n'est pas aussi prononcée. Les quatre indices commencent à signaler un redressement au début de 2009, mais le changement de tendance de l'indice MLS® se produit plus tôt, alors que le point de retournement de l'indice IPLN se produit en dernier[15]. En ce qui concerne la volatilité, l'indice MLS® est le plus volatil par rapport à sa tendance du fait des changements de composition des échantillons de logements vendus chaque mois. Les trois autres indices, qui sont ajustés dans une certaine mesure pour tenir compte des changements qualitatifs, se montrent moins erratiques au fil du temps.

Étude de cas : l'Allemagne

10.41 En Allemagne, des séries trimestrielles d'indices des prix de l'immobilier résidentiel existent depuis 2000. Avant cette date, on pouvait décrire la situation dans ce pays comme l'existence d'un ensemble non coordonné d'indicateurs différents diffusés par plusieurs instituts privés. «Il manquait à ces indicateurs un fondement méthodologique clair et leur couverture était restreinte. En outre, ils donnaient — dans une certaine mesure — des signaux contradictoires[16].»

15 On trouvera une illustration de l'incidence des différentes méthodologies sur les points de retournement dans Shimizu, Nishimura et Watanabe (2010).
16 Voir Hoffmann et Lorenz (2006).

10.42 L'Office fédéral de la statistique (Destatis) a pris des mesures pour améliorer la situation en s'appuyant sur les sources de données disponibles. Il existait depuis longtemps en Allemagne des statistiques sur les prix de la construction et sur ceux des terrains à bâtir. De plus, au niveau local, les commissions d'experts chargés d'évaluer les biens, qui fonctionnent à l'échelle nationale et sont régies par la législation fédérale, permettaient d'avoir accès à des bases de données exhaustives qui contenaient les prix d'acquisition des terrains à bâtir et des logements, ainsi que les caractéristiques correspondantes sur le plan de l'immobilier. Le principal obstacle à l'exploitation centralisée des données disponibles était l'existence de différences entre les systèmes de recueil des États et des commissions locales. La méthode utilisée par l'Allemagne constitue un exemple intéressant d'intégration des données, c'est-à-dire d'exploitation de sources multiples.

Indices des prix de l'immobilier résidentiel

10.43 On recourt à différentes sources de données et méthodes pour construire des indices de prix portant sur différents compartiments du marché. Ils sont ensuite agrégés afin de calculer un indice des prix de l'immobilier résidentiel couvrant toutes les catégories de biens ainsi que des sous-indices relatifs aux logements existants et neufs. Les pondérations employées pour élaborer un indice des prix des logements existants sont les dépenses liées aux opérations de l'année de base, décomposées en maisons et en appartements et ventilées par État fédéré. Pour les logements clé en main, les pondérations sont calculées à partir des statistiques officielles sur les activités de construction, et pour les constructions effectuées soi-même, on se sert de pondérations provenant de la construction. Les indices sont publiés dans un délai de 90 jours suivant la fin de la communication des données.

Logements neufs clé en main et logements existants

10.44 Les données sont tirées des informations recueillies par les commissions locales d'experts en évaluation des biens. Collectées au moment de la conclusion d'un contrat, elles couvrent la totalité des ventes (au comptant et à crédit) et se composent des prix effectifs des opérations (au comptant et à crédit) et d'un certain nombre de caractéristiques déterminant le prix d'un bien : le type de logement (maison unifamiliale, maison pour deux familles, appartement en pleine propriété); le type de maison (individuelle, en rangée, jumelée); le type de construction (classique, préfabriqué); l'année de construction; la dimension de la parcelle; la surface habitable; l'ameublement/les éléments de confort (cuisine, sauna/piscine, étage grenier); le parking; les caractéristiques du lieu géographique (État, district, commune); la notation générale de l'emplacement (simple/moyen/bon); le nombre de pièces/d'étages. Une évaluation du terrain est aussi fournie.

10.45 On utilise une conjonction de techniques hédoniques et de stratification (une strate pour les maisons destinées à une famille/à deux familles et une pour les appartements situés dans des immeubles), afin de procéder à des ajustements tenant compte de l'incidence des changements qualitatifs des catégories de biens vendus. La méthode de régression hédonique adoptée est la «double imputation», décrite au chapitre 5, avec laquelle on estime les prix à la fois pour la période de référence et pour la période de comparaison. Les valeurs aberrantes sont exclues.

Logements neufs unifamiliaux[17]

10.46 L'élaboration d'un indice de prix pour cette catégorie particulière de logements neufs s'appuie sur des informations tirées des indices officiels nationaux des prix de la construction. On dispose de ces derniers pour diverses formes de structures (par exemple les bâtiments résidentiels/non résidentiels, les routes, les ponts routiers) et pour les travaux d'entretien. Les prix de quelques 190 activités de construction sont recueillis chaque trimestre (y compris pour les matériaux). Au total, à peu près 30.000 prix sont déclarés par 5.000 entreprises environ à chaque collecte. Il s'agit des prix des opérations relatives aux contrats conclus pendant le trimestre, en excluant la taxe sur la valeur ajoutée (TVA), ce qui signifie que l'on prend en compte les bénéfices et l'évolution de la productivité. Pour les logements autoconstruits, on se sert des prix de la construction des «logements résidentiels destinés à une seule famille et édifiés de façon classique». Un modèle à panier fixe est employé pour construire l'indice.

Logements préfabriqués

10.47 L'indice s'appuie sur les statistiques officielles de prix à la production des produits utilisés, en particulier l'indice des prix des maisons préfabriquées pour une famille, sans sous-sol et présentant un ensemble de caractéristiques spécifiques. Ici encore, on calcule l'indice au moyen d'un modèle à panier fixe. Les logements préfabriqués ont pour particularité que les contrats prévoient généralement l'acquisition/la vente de maisons complètes (par exemple celles destinées à une seule famille et dépourvues de cave), dont les caractéristiques ne changent pas sensiblement en courte période.

Terrains à bâtir

10.48 Les indices des prix des logements préfabriqués et des logements «autoconstruits» n'englobent pas le coût des terrains. On établit un indice des prix des terrains à bâtir à partir des chiffres officiels relatifs aux transactions les concernant, qui sont répertoriés à la date de conclusion des contrats. Chaque série de données intègre les aspects suivants : l'emplacement, les caractéristiques de la commune, la date de la vente, la dimension de la parcelle, les détails du

17 Ils sont parfois qualifiés de biens «autoconstruits». Les constructeurs comprennent à la fois les futurs propriétaires qui réalisent eux-mêmes une partie de l'édification et ceux qui font participer un entrepreneur responsable de l'essentiel des travaux (lesquels sont finalisés par le propriétaire).

permis de construire, précisant par exemple s'il s'agit d'une maison ou d'un appartement, et la taille de l'habitation. À la différence de l'IPLN de Statistique Canada, l'indice allemand ne couvre pas seulement les surfaces aménagées, mais il tente de couvrir tous les logements neufs.

10.49 L'indice agrégé des prix des terrains à bâtir aménagés est une moyenne pondérée des indices de valeurs unitaires des sous-agrégats, se référant à la valeur totale des ventes. Ces sous-agrégats sont établis sur la base de la différenciation régionale, principalement par district, par catégorie de zone constructible et par taille des communes au sein des États fédérés. La pondération de ces derniers s'effectue par conjonction des données sur la totalité des prix payés pour l'acquisition de terrains à bâtir aménagés dans les zones de construction résidentielle et dans les zones rurales, sur le chiffre d'affaires des activités de construction et sur le nombre de permis de construire d'habitations résidentielles comportant un ou deux logements.

Étude de cas : le Japon

Informations sur les prix de l'immobilier

10.50 Au Japon, les indices officiels des prix de l'immobilier ne portent que sur les prix des terrains. Les administrations donnent les informations suivantes : publication des prix des terrains (PNLP) assurée par le Ministère des sols, des infrastructures, du transport et du tourisme (MLIT); enquête sur les prix des terrains effectuée par chaque préfecture; évaluation des actifs fonciers pour le calcul des droits de succession opérée par les services nationaux des impôts et évaluation faite par chaque commune pour le calcul des impôts locaux. Toutes ces sources d'information représentent des valeurs estimatives calculées par des spécialistes dotés de licences.

10.51 Les informations servant à établir les indices de prix de l'immobilier résidentiel (bâtiments compris) sont recueillies par des entités privées. La base de données la plus représentative est le Système de réseau d'information immobilière (REINS). Il s'agit d'une base de données conçue sur le modèle du Service d'inscriptions multiples (MLS) qui existe aux États-Unis et au Canada; les renseignements sont fournis par les agents immobiliers. La base de données REINS contient à la fois le prix demandé quand le bien est mis sur le marché et le prix final de la transaction fixé au moment de la conclusion du contrat de vente. Les entreprises qui s'occupent de la publicité des ventes de logements constituent une deuxième source de données, tout à fait spécifique, sur leurs prix. Le secteur privé utilise ces deux sources pour calculer et publier des indices des prix des logements. Toutefois, tous présentent des défauts et ne répondent pas pleinement aux besoins des utilisateurs. MLIT a donc lancé un programme de travail qui devrait aboutir à l'élaboration d'un meilleur indice. Il s'agira du premier indice des prix de l'immobilier résidentiel publié par le secteur public.

Le tableau 10.1 donne un aperçu général des indices de prix de l'immobilier qui existent au Japon. Il s'agit à la fois des indices basés sur l'évaluation des terrains et de ceux portant sur les ventes de biens. C'est à partir de ces dernières que l'on établit les indices des prix de l'immobilier résidentiel.

Prix demandés et prix de vente

10.52 Au Japon, une personne désireuse de vendre un logement passe généralement par un agent immobilier. Elle doit alors choisir entre deux catégories de contrats de mandataire : le mandat exclusif ou le mandat unique. L'autre solution possible est de donner mandat à plusieurs agents. Les contrats sont régis par l'article 34-2 de la loi sur les opérations immobilières.

10.53 Dans le cadre du mandat exclusif, le vendeur peut recevoir un compte rendu de l'agent immobilier au moins une fois par semaine, mais n'a le droit ni de recourir à un autre agent immobilier pour trouver un acheteur, ni d'en chercher un lui-même. Avec le mandat unique, le vendeur ne peut demander à un autre agent immobilier de trouver un acheteur, mais il peut le faire lui-même et l'agent immobilier mandataire doit faire un compte rendu au moins toutes les deux semaines. Avec le mandat général, le vendeur peut chercher lui-même un acheteur et demander à plusieurs agents immobiliers de le faire. En revanche, il ne reçoit pas de compte rendu des agences.

10.54 Sous le régime du contrat exclusif, l'agent immobilier mandataire doit procéder à l'enregistrement dans REINS dans un délai de cinq jours après l'accord d'inscription et il est tenu de procéder à de larges prospections. Avec le contrat donnant mandat unique, l'agence doit enregistrer dans REINS dans les sept jours et également prospecter à grande échelle. Pour avoir le droit d'enregistrer dans REINS, les agents sont obligés de déclarer non seulement le prix demandé au moment de l'enregistrement, mais aussi le prix final de l'opération. Ainsi, pour certaines opérations faisant intervenir les agents immobiliers, tant le prix demandé que le prix définitif sont répertoriés.

Système d'enregistrement public du prix des opérations

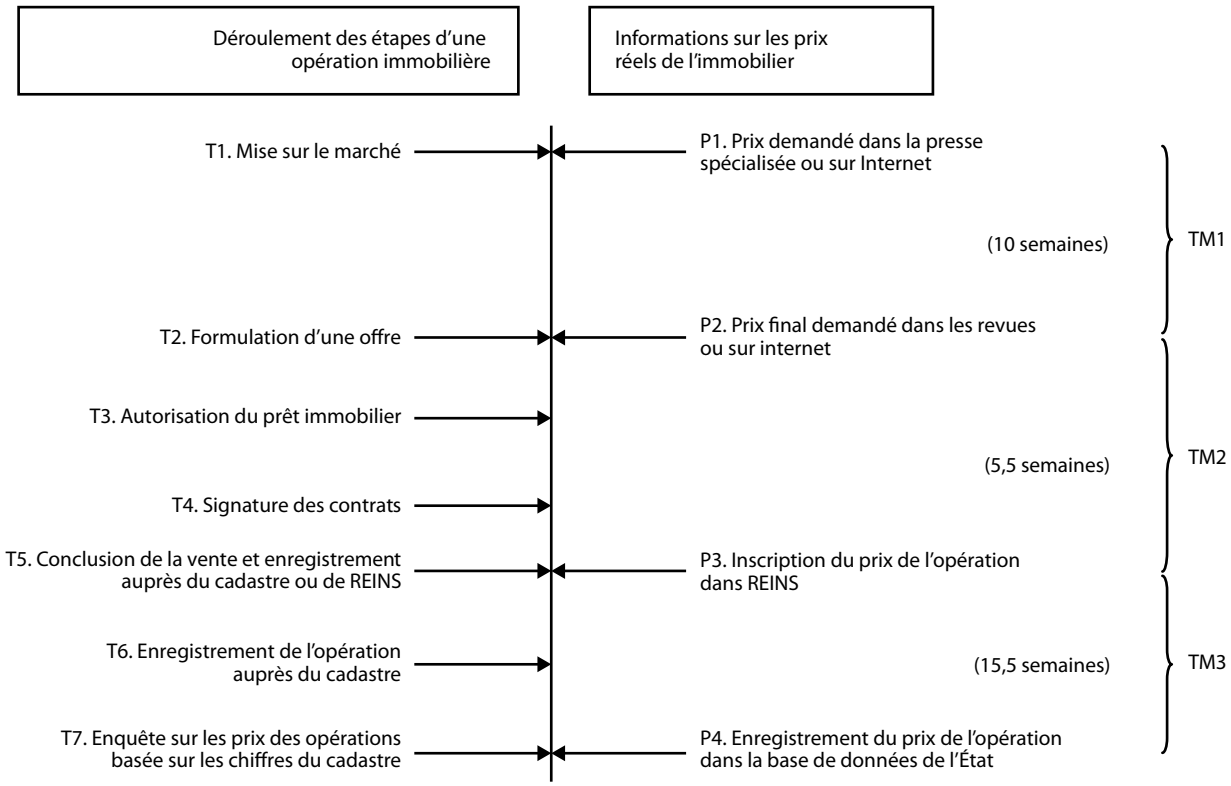
10.55 MLIT recueille et publie des informations sur les prix des opérations immobilières depuis 2005. Ces dernières sont répertoriées par le Bureau des affaires juridiques qui notifie ensuite à MLIT un «changement des données figurant au registre». Sur cette base, MLIT adresse un questionnaire à l'acheteur portant sur les parcelles vacantes, les terrains construits et les immeubles ayant des régimes de propriété distincts (par exemple ceux comportant des bureaux, des commerces et des appartements) en demandant le prix de l'opération. Ensuite, les évaluateurs des actifs immobiliers ou leurs homologues ajoutent des renseignements. Ceux-ci concernent l'usage de l'immeuble, L'emplacement de la parcelle (configuration du terrain), l'accès au réseau routier (largeur

Tableau 10.1. Indices des prix de l'immobilier publiés au Japon

Indice	Échantillon	Méthode	Correction ou non des variations saisonnières (et fréquence)	Méthode de pondération	Étapes de la vente
Indice des taux de variation des prix des terrains (MLIT)	Évaluations figurant dans la publication des prix des terrains par MLIT	Indice de la période précédente multiplié par la volatilité moyenne	Non (annuelle)	Non	Évaluation au 1er janvier de chaque année *(publiée à la fin mars)*
Statistiques de base sur les prix des transactions foncières dans les grandes villes (MLIT)	Prix de vente	Prix unitaire au mètre carré, valeur médiane, écart-type, quartile, etc.	Non (trimestrielle ou annuelle)	Non	Enquête après enregistrement de la vente *(prix de vente)*
Indice des prix des terrains urbains (Institut japonais de l'immobilier)	Évaluations figurant dans la publication des prix des terrains par l'Institut japonais de l'immobilier	Indice de la période précédente multiplié par la volatilité moyenne	Non (semestrielle)	Non	Évaluation chaque année fin mars et septembre
Indice des prix de l'immobilier résidentiel (Institut immobilier Recruit)	Prix demandés définitifs publiés dans des revues ou sur Internet	Régression hédonique avec chevauchement des périodes	Oui (mensuelle)	Quantités	Formulation de l'offre *(prix demandé définitif)*
Indice de l'immobilier résidentiel (Institut japonais de l'immobilier, At home Co., Ltd., Ken Corporation)	Prix demandés et prix de vente	Prix unitaire au mètre carré (ancienneté du logement ajusté par régression hédonique)	Non (semestrielle)	Non	Formulation de l'offre? *(prix demandé ou prix de vente)*
Indice des prix du marché des immeubles en copropriété dans la région de Tokyo (Institut japonais de recherche, Réseau d'information immobilière pour l'Est du Japon)	Prix de vente répertoriés par le réseau d'information immobilière pour l'Est du Japon	Régression hédonique	Non (mensuelle	Non	Conclusion de la vente *(prix de vente)*
Indice des prix des immeubles neufs en copropriété (Tokyo Kantei Co., Ltd.)	Prix demandés	Moyenne mobile	Non (trimestrielle)	Non	*(prix demandé)*

Source : Shimizu, Nishimura et Watanabe

Graphique 10.2. Transmission d'informations immobilières



Source : Bureau national de statistique du Royaume-Uni.

de la voie faisant face à l'immeuble), la distance de la gare la plus proche, d'autres informations pratiques et réglementaires comme le plan d'urbanisme. Le «dossier sur les données de l'opération» ainsi établi est ensuite rendu anonyme, pour que l'on ne puisse pas identifier le bien, et enfin publié sur le site Internet de MLIT en tant qu'information sur les prix des opérations[18]. Comme ni la communication d'informations sur ces prix, ni celle des renseignements demandés aux évaluateurs ne sont obligatoires, on se heurte à des problèmes non-réponses et de retard. Il n'est pas possible de vérifier de façon indépendante les données transmises, notamment le prix d'une opération.

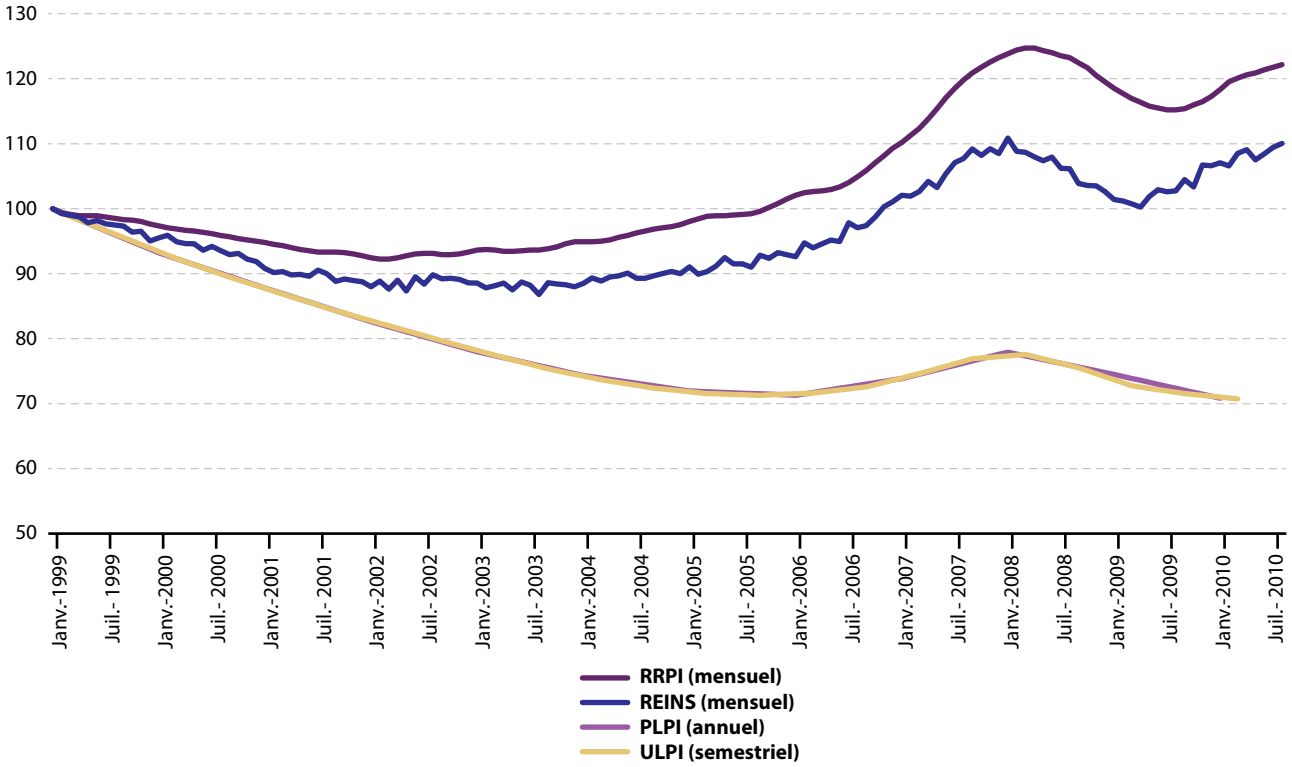
Délais d'acquisition et de vente d'un logement et exactitude du prix

10.56 Le choix des sources de données est important quand on calcule un indice des prix de l'immobilier. Divers problèmes se posent tels que la date du recueil des données sur les prix, l'évolution de ceux-ci (du prix initial demandé au prix final de l'opération) et les délais de diffusion des données. Le graphique 10.2, tiré de Shimizu, Nishimura et Watanabe (2011), inscrit sur une verticale les informations sur les prix de l'immobilier qui sont actuellement disponibles au Japon. À droite, on distingue quatre étapes, les prix allant de P1 à P4. Les intervalles correspondants sont : la période TM1 écoulée entre le début du processus de vente et le moment où l'on trouve un acheteur; la période TM2, qui se situe entre le moment où l'on trouve un acheteur et la finalisation du contrat de vente; la période TM3, située entre la finalisation du contrat de vente et l'enregistrement du prix de vente dans la base de données officielle.

10.57 La période TM1 dure en moyenne 70 jours, ce qui signifie qu'il faut en moyenne 70 jours pour trouver un acheteur après que le vendeur a lancé le processus de cession; la durée maximum constatée a été de 3,72 ans. Le ratio P2/P1 ressort à 0,976 en moyenne; autrement dit, il y a une baisse de 2,4 % entre le prix demandé initialement et le dernier prix demandé. TM2 dure en moyenne 39 jours. Le ratio P3/P2 étant de 0,956 en moyenne, le prix de l'opération est en général inférieur de 4,4 % au dernier prix demandé. TM3 dure en moyenne 109 jours. Cela signifie qu'il y a (pour les données de prix des opérations faisant l'objet de l'enquête) un délai d'enregistrement d'à peu près trois mois dans la base de données de l'État. Les différentiels de prix aux étapes successives du processus de vente peuvent bien sûr varier au fil du temps selon la situation du marché des logements occupés par leur propriétaire.

18 Voir www.land.mlit.go.jp/webland.

Graphique 10.3. Quatre indices des prix de l'immobilier résidentiel au Japon (janvier 1999 = 100)



Source : Shimizu, Takatsuji, Ono et Nishimura (2010).

Analyse comparative des indices des prix des logements dans la zone métropolitaine de Tokyo

10.58 Le graphique 10.3 compare les quatre indices des prix de l'immobilier. Les données de REINS sont utilisées par le réseau d'information immobilière pour l'Est du Japon et par l'Institut japonais de recherche qui établissent ensemble l'indice des prix des immeubles anciens en copropriété pour Tokyo. Cet indice mensuel, publié depuis 1995, est construit au moyen d'une méthode de régression hédonique. L'indice Recruit des prix de l'immobilier résidentiel (RRPI) est aussi un indice de prix hédonique[19], basé sur le prix d'offre final des biens figurant dans la revue Recruit et portant sur les reventes de logements destinés à une seule famille et en copropriété. Également mensuel, il est publié depuis janvier 1986[20], mais n'est largement diffusé dans sa forme actuelle que depuis le début de 2000. On trouve au graphique 10.3 deux indices des prix des terrains, excluant donc ceux des immeubles, l'indice semestriel ULPI et l'indice annuel PNLP. Tous deux reposent sur des estimations[21]. Les indices des prix de l'immobilier qui incluent les bâtiments font clairement ressortir des tendances différentes des indices des prix des terrains. Ainsi, les premiers ont commencé à se redresser peu de temps après la crise financière de 2008, alors que les autres ont continué à fléchir. On notera que l'indice REINS évolue bien en deçà de l'indice Recruit, bien que tous deux soient des indices hédoniques.

Étude de cas : le Royaume-Uni

10.59 C'est probablement au Royaume-Uni qu'il existe le plus d'indices des prix de l'immobilier résidentiel publiés régulièrement. Ce nombre élevé résulte de l'exploitation faite par divers organismes des séries de données générées aux différentes étapes du processus d'achat et de vente. Ce dernier dure souvent plusieurs mois, voire plus, et le moment auquel le prix est fixé et intégré à un indice peut influer sur le taux mesuré de la hausse des prix des logements. Au Royaume-Uni, l'exploitation des données relatives aux prix de l'immobilier intervient à plusieurs étapes :

- Dès que le bien est sur le marché. On a alors le *prix demandé*. Source : les agents immobiliers[22]. Publication par ces derniers, le *Financial Times* et les sites Internet spécialisés.

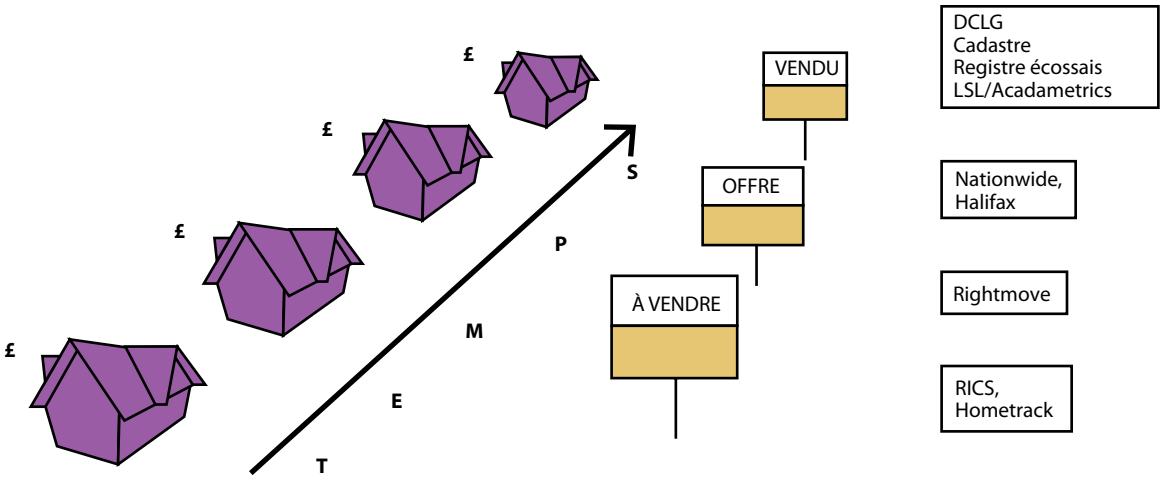
19 L'indice Recruit utilise des variables indicatrices temporelles et fait donc l'objet de révisions (voir chapitre 5).

20 On trouvera des précisions dans Shimizu, Takatsuji, Ono et Nishimura (2010).

21 Shimizu et Nishimura (2006) (2007), qui comparent les valeurs estimées et les prix de vente, signalent les problèmes d'erreurs d'évaluation et de lissage pour les indices basés sur des estimations.

22 Bien que cela soit sans rapport avec la question des délais, l'inconvénient des prix affichés et des autorisations de crédit immobilier est que tous ces prix ne se terminent pas forcément par une transaction finale; si c'est le cas, le prix est souvent supérieur à celui résultant de la négociation finale.

Graphique 10.4. Déroulement de l'achat d'un logement



Source : Bureau national de statistique du Royaume-Uni.

- Autorisation du crédit immobilier. On a alors l'*évaluation par le prêteur*. Source : les organismes octroyant des crédits immobiliers. Publication par plusieurs d'entre eux.
- Finalisation du crédit immobilier. On a alors le *prix fixé lors de la finalisation du crédit*. Source : les organismes accordant des crédits immobiliers. Publication par le Ministère des communautés et des collectivités locales (DCLG).
- Enregistrement de l'opération. On a alors le *prix de l'opération*. Source : le cadastre.

Le graphique 10.4 indique le déroulement de l'achat ou de la vente d'un logement au Royaume-Uni avec les différents stades auxquels on recueille les informations servant à produire un indice des prix.

10.60 Au Royaume-Uni, il existe actuellement deux indices officiels des prix des logements. Le premier, publié mensuellement par le Ministère des communautés et des collectivités locales (DCLG), s'appuie sur des informations émanant des prêteurs, par l'intermédiaire du Conseil des organismes de crédit immobilier, à propos du prix évalué lors de la conclusion d'une vente. Il est publié à peu près six semaines après la date de référence de la vente — soit en moyenne quatre à cinq mois après qu'un logement a été mis en vente. Cet indice porte seulement sur les acquisitions au moyen d'un prêt immobilier. Le second est publié mensuellement par le cadastre à partir des ventes de biens qu'il a enregistrées. La publication intervient un mois après la date de référence, c'est-à-dire un mois après l'enregistrement de la vente; néanmoins, il accuse un certain décalage en raison des délais de notification des transferts de propriété auprès du cadastre par les acheteurs ou leurs mandataires.

10.61 Deux organismes de crédit immobilier, Halifax et Nationwide, publient des indices qui se réfèrent à leur propre évaluation d'un bien au moment où ils accordent un prêt. Ces indices sont établis quelques semaines après la date de référence de l'octroi du prêt et approximativement trois à quatre mois après la mise en vente du bien. Leur publication intervient donc un peu plus tôt que pour l'indice officiel DCLG, mais ils ont une couverture beaucoup plus limitée et on n'est pas certain que les biens ayant donné lieu à leurs opérations de crédit soient représentatifs de toutes les transactions ou de toutes les acquisitions effectuées au moyen d'un crédit immobilier.

10.62 Il existe un autre indice, établi par une société de services appelée Hometrack, qui donne divers renseignements sur le marché de l'immobilier résidentiel à plusieurs entités du secteur, notamment les promoteurs, les associations pour le logement, les investisseurs, les agents immobiliers ainsi que les collectivités locales et l'administration centrale. Hometrack effectue une enquête mensuelle auprès des agents immobiliers, qui recense leur opinion sur les prix de vente réalisables de quatre catégories standards de biens. De tous les indices mentionnés, c'est le plus à jour puisqu'il est publié trois à quatre semaines environ après la période de référence, de sorte qu'il n'y a pas d'autre délai, mais il s'agit d'une enquête d'opinion sur les prix de vente probables des biens mis sur le marché.

Acadametrics, une société de conseil, établit aussi un indice des prix des logements basé sur les données communiquées par le cadastre. L'indice LSL/Acadametrics, publié quelques semaines après la fin de la période de référence, est élaboré au moyen d'une méthode de prévision s'appuyant sur un «indice des indices». L'indice relatif à chaque période est ensuite révisé jusqu'à l'intégration de toutes les opérations. Enfin, un indice fondé sur les prix demandés, figurant sur le site Internet de «Rightmove property», est aussi très utilisé au Royaume-Uni.

Tableau 10.2. Indices des prix de l'immobilier résidentiel publiés au Royaume-Uni

Indice	Échantillon	Méthode	Correction ou non des variations saisonnières	Méthode de pondération	Étapes de la vente
DCLG[1]	Échantillon des organismes de crédit immobilier	Ajustement mixte et régression hédonique	Oui	Dépenses	Finalisation du crédit immobilier *(prix de l'opération figurant dans le dossier de crédit)*
Cadastre (mensuel)	Ventes enregistrées en Angleterre et au Pays de Galles avec vente précédente depuis 1995.	Régression sur des ventes répétées	Oui	Dépenses	Enregistrement de la vente *(prix de l'opération)*
Halifax	Prêts accordés par Halifax pour l'achat de logements	Régression hédonique (ajustement qualitatif)	Oui	Quantités	Autorisation du prêt *(prix évalué)*
Nationwide	Prêts accordés par Nationwide pour l'achat de logements	Régression hédonique (ajustement qualitatif)	Oui	Quantités	Autorisation du prêt *(prix évalué)*
Hometrack	Enquête auprès des agents immobiliers (évaluations)	Ajustement mixte	Non	Dépenses	Prix de vente possible
Rightmove	Prix demandés affichés sur le site Internet	Ajustement mixte	Non	Dépenses	*(prix demandé)*
LSL/ Acadametrics	Ventes enregistrées en Angleterre et au Pays de Galles	Modèle de prévision avec ajustement mixte	Oui	Quantités	Enregistement de la vente *(prix de l'opération)*

[1] Ministère des communautés et des collectivités locales. On peut trouver un examen des indices de prix de l'immobilier effectué par l'Institut officiel de la statistique à l'adresse suivante : http://www.statisticsauthority.gov.uk/national-statistician/ns-guidance-and-reports/national-statistician-s-reports/index.html.

Source : Bureau national de statistique du Royaume-Uni.

10.63 Le tableau 10.2 résume le champ d'application, la définition et les principaux aspects de la méthode d'élaboration des sept indices disponibles au Royaume-Uni qui apparaissent au graphique 10.4. Compte tenu des différences existant sur ces trois points, il n'est pas surprenant que, quand on les réunit, le résultat ne soit pas toujours cohérent.

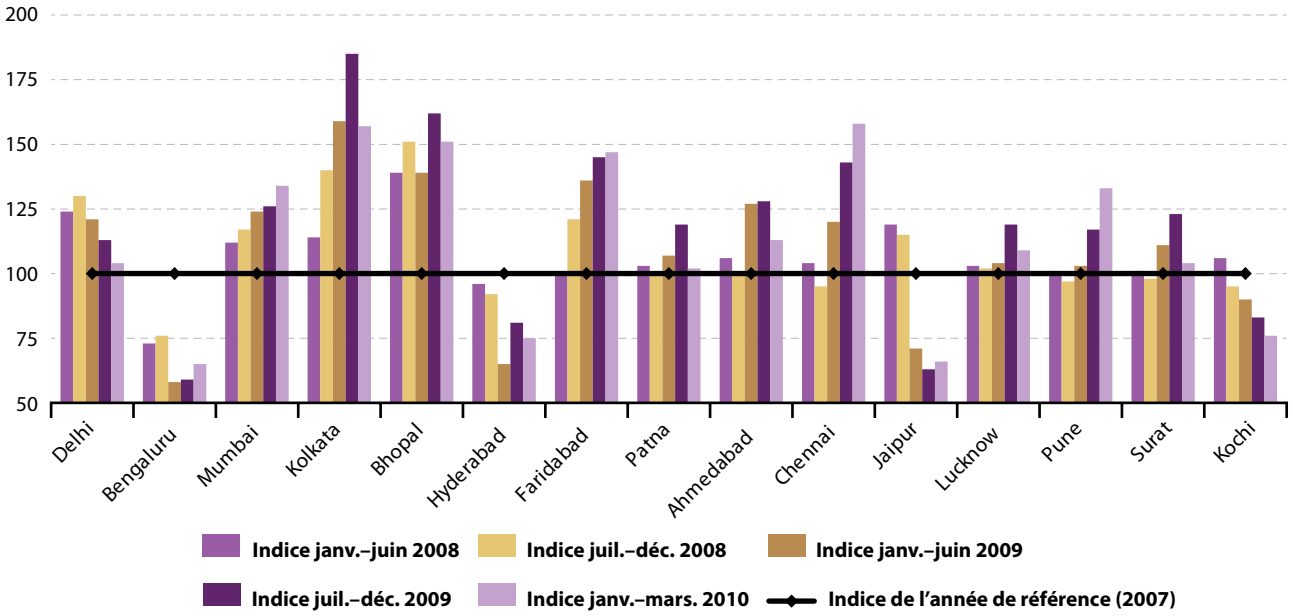
Étude de cas : l'Inde

10.64 En Inde, l'évolution des prix des logements, en particulier de l'immobilier résidentiel, revêt une très grande importance sur le plan macroéconomique et pour chaque ménage. Il n'est pas étonnant que les utilisateurs réclament un indice pertinent et fiable pour suivre cette évolution. Mais le manque de transparence des opérations sur le marché de l'immobilier résidentiel et le caractère limité des informations disponibles en matière de prix rendent difficile de suivre la dynamique des prix réels de l'immobilier.

10.65 En Inde, l'enregistrement de toute opération immobilière est juridiquement obligatoire. En principe, l'autorité qui en est chargée connaît donc les détails de toutes les transactions effectuées pendant une période de référence. Les données sont théoriquement disponibles quotidiennement avec un mois de décalage après la première déclaration d'un changement de propriétaire. Toutefois, il est bien connu que les prix répertoriés sont manifestement sous-estimés en raison du montant très élevé des droits d'enregistrement et du droit de timbre. L'obligation de payer ensuite l'impôt immobilier dissuade aussi les acheteurs (à l'exception des sociétés) de révéler les prix de vente exacts. En outre, la procédure d'enregistrement et la maintenance des dossiers ne sont pas informatisées et ils sont établis dans les différentes langues régionales, ce qui exige un travail supplémentaire pour arriver à une présentation commune.

10.66 Ces raisons empêchent d'exploiter les données administratives concernant l'enregistrement des changements de propriétaire et il a fallu trouver une autre source, émanant du marché. Ce sont les prix des opérations recueillis par le Conseil national de la recherche économique appliquée (NCAER), un organisme national d'étude, par les associations de défense des résidents (RWA), les agents immobiliers et les courtiers. Pour compléter les données sur les prix effectifs des opérations obtenues par enquête, on prend les évaluations liées aux prêts immobiliers distribués par les banques et les sociétés de financement du logement. Elles servent ensuite à construire l'indice RESIDEX de la National Housing Bank.

Graphique 10.5. Indices NHB RESIDEX : indices des prix en zone urbaine en Inde



Source : National Housing Bank, Inde.

L'indice NHB RESIDEX

10.67 NHB RESIDEX est le résultat d'une première tentative de la National Housing Bank (NHB), établissement spécialisé dans le financement du logement et propriété de la Banque centrale, pour mesurer les prix de l'immobilier résidentiel en Inde. On a d'abord étudié cinq villes pilotes : Bangalore, Bhopal, Delhi, Kolkata et Bombay (Mumbai). Le recueil des données s'est heurté à de nombreuses difficultés et leur analyse a aussi posé plusieurs problèmes méthodologiques. Finalement, après beaucoup d'efforts, la NHB a lancé en juillet 2007 son premier indice RESIDEX destiné à suivre les prix de l'immobilier résidentiel en Inde. Il se réfère aux opérations en prenant le prix de vente complété par des évaluations. Les principales données, qui proviennent des agents immobiliers, sont obtenues en recourant aux services d'un organisme d'étude ou de conseil de réputation nationale qui se procure les prix des transactions. Des données sur les prix des logements sont aussi recueillies auprès des sociétés de financement et des banques commerciales. Il s'agit de prix évalués dans le cadre des crédits immobiliers contractés auprès de ces institutions.

10.68 Les principales caractéristiques de l'indice RESIDEX de la NHB sont les suivantes :

- Il couvre toutes les catégories de l'immobilier résidentiel dans 15 villes[23].
- Il est établi trimestriellement avec 2007 comme année de référence[24].
- D'autres séries sont établies à partir de pondérations des opérations et des stocks.
- L'indice couvre les achats au comptant et à crédit.
- Il porte sur les logements neufs et anciens.
- Il est construit à partir de «moyennes pondérées de rapports de prix»[25].
- On n'opère pas actuellement d'ajustement qualitatif au titre de l'emplacement, de la dimension, etc.
- L'indice est révisable pour tenir compte des dernières données disponibles.
- On dispose aussi d'informations sur l'évolution des prix de l'immobilier résidentiel selon l'emplacement, le secteur et la ville; il existe par exemple des indices distincts pour chaque secteur des 15 villes couvertes.

10.69 Pour un pays de la taille de l'Inde, L'emplacement géographique est importante. Ainsi, les indices des prix en zone urbaine, qui figurent au graphique 10.5, facilitent les décisions

23 À terme, en fonction de l'expérience acquise et de la disponibilité des données, il pourrait être étendu à l'immobilier commercial.

24 2001 avait été choisie comme année de référence de l'indice pilote portant sur cinq villes, pour être comparable aux années de référence de l'indice des prix de gros et de l'indice des prix à la consommation. On a pris les variations d'une année sur l'autre au cours de la période 2001–05, puis on a actualisé l'indice avec deux années supplémentaires, c'est-à-dire jusqu'en 2007. Il a ensuite été étendu à dix autres villes — Ahmedabad, Faridabad, Chennai, Kochi, Hyderabad, Jaipur, Patna, Lucknow, Pune et Surat — et au même moment l'année de référence est passée de 2001 à 2007.

25 Il faut voir que, comme il s'agit d'un indice pondéré de Carli, il présente probablement un «biais» dans le sens de la hausse; voir le Manuel de l'IPC (2004), page 421.

des acheteurs, tout en aidant les constructeurs et les promoteurs à prendre des décisions d'investissement pour l'avenir.

10.70 L'indice NHB RESIDEX continue à être développé afin d'en accroître l'intérêt pour les utilisateurs :

- Il va être étendu progressivement pour couvrir les 35 villes indiennes comptant plus d'un million d'habitants selon le recensement de 2001.
- Il est proposé de l'étendre aux 63 villes concernées par la Mission nationale Jawaharlal Nehru de rénovation urbaine (action phare du gouvernement), pour en faire un indice national.
- À terme, en fonction de l'expérience acquise et de la disponibilité des données, il pourrait être étendu à l'immobilier commercial.

Étude de cas : la Colombie

10.71 La banque centrale de Colombie (Banco de la República) établit un indice concernant les prix des logements existants, appelé IPVU. Il existe d'autres indices, produits par l'Institut national des statistiques (DANE), qui portent sur les coûts de la construction et les prix des logements neufs. Il n'y a pas de série fusionnant les informations de ces deux sources pour produire un indice qui couvrirait les ventes de tous les biens relevant de l'immobilier résidentiel dans le pays[26]. Dans le passé, on a envisagé d'exploiter des données administratives, mais cela s'est avéré impossible, car trop complexe.

L'indice IPVU

10.72 Le projet d'élaboration d'un indice des prix des logements existants en Colombie, l'IPVU, a été amorcé en 2003. Auparavant, l'impossibilité d'obtenir des informations de base avait constitué le principal obstacle à l'établissement d'un indice de ce type. Après consultation de plusieurs banques sur l'opportunité de disposer d'une mesure de la valeur des logements existants, le projet a été lancé avec un financement du Banco de la República. Le service des statistiques de cette institution est chargé de produire et de publier l'indice.

10.73 L'indice IPVU ne couvre que les principales zones métropolitaines de la Colombie, c'est-à-dire les villes de Bogotá, Medellín, Cali et Soacha dans la province Cundinamarca, ainsi que Bello, Envigado et Itaguí dans la province d'Antioquia. Il est calculé à partir d'évaluations liées à l'octroi de prêts immobiliers, communiquées par les banques spécialisées Davivienda, BBVA, Av. Villas, Bancolombia, Colmena BCSC et Colpatria. Il ne porte donc que sur les biens acquis à crédit, les achats au comptant étant exclus. Les banques communiquent à la banque centrale les évaluations commerciales et les adresses de tous les biens pour lesquels un crédit immobilier est autorisé. Les prix incorporés à l'indice émanent d'évaluations indépendantes demandées par le prêteur. Ces évaluations sont proches des prix du marché quand le décaissement intervient. L'indice est publié chaque trimestre sur le site Internet de la banque centrale, avec un décalage d'un trimestre, et révisé trimestriellement conformément à la méthode des ventes répétées qui est utilisée (voir ci-dessous). Un indice calculé en moyenne annuelle est également publié et on établit des sous-indices pour les principales zones métropolitaines : Bogotá, Medellín et Cali.

10.74 Les logements sont classés selon qu'ils bénéficient ou non de subventions, d'où l'existence de deux indices — VIS et NOVIS. L'octroi d'une subvention dépend de la valeur et de l'emplacement d'un logement. Le terme «logement pour personnes à bas revenu» (VIS) se réfère aux habitations construites pour garantir le droit au logement des ménages à bas revenu. Pour chaque projet immobilier de ce type, le gouvernement fixe un prix plafond et décide du type de résidence destiné à ces ménages. Il prend en compte, entre autres éléments, leur accès aux marchés de crédit, le montant des fonds pouvant être obtenus auprès du système financier et celui des crédits publics consacrés aux programmes d'aide au logement[27].

10.75 La technique appliquée est similaire à la méthodologie Case–Shiller des ventes répétées. On manque d'informations détaillées sur les caractéristiques des logements nécessaires pour répondre aux exigences constantes d'ajustement de la composition de la méthodologie Case–Shiller qui a recours à la stratification. Toutefois, la situation s'améliore dans la mesure où il est prévu que les informations données par les établissements de crédit immobilier incluront dans le futur une large gamme de données sur les aspects spécifiques des logements. On remédie à l'absence actuelle de précision en révisant les données. Si la variation du prix d'un bien apparaît «anormale», c'est-à-dire si l'on juge qu'il s'agit d'un cas extrême, l'information sur le prix est rejetée au lieu d'être intégrée à l'indice. L'objectif est d'empêcher les maisons complètement rénovées ou négligées d'y figurer. L'indice est révisable conformément à l'une des caractéristiques de la méthodologie des ventes répétées.

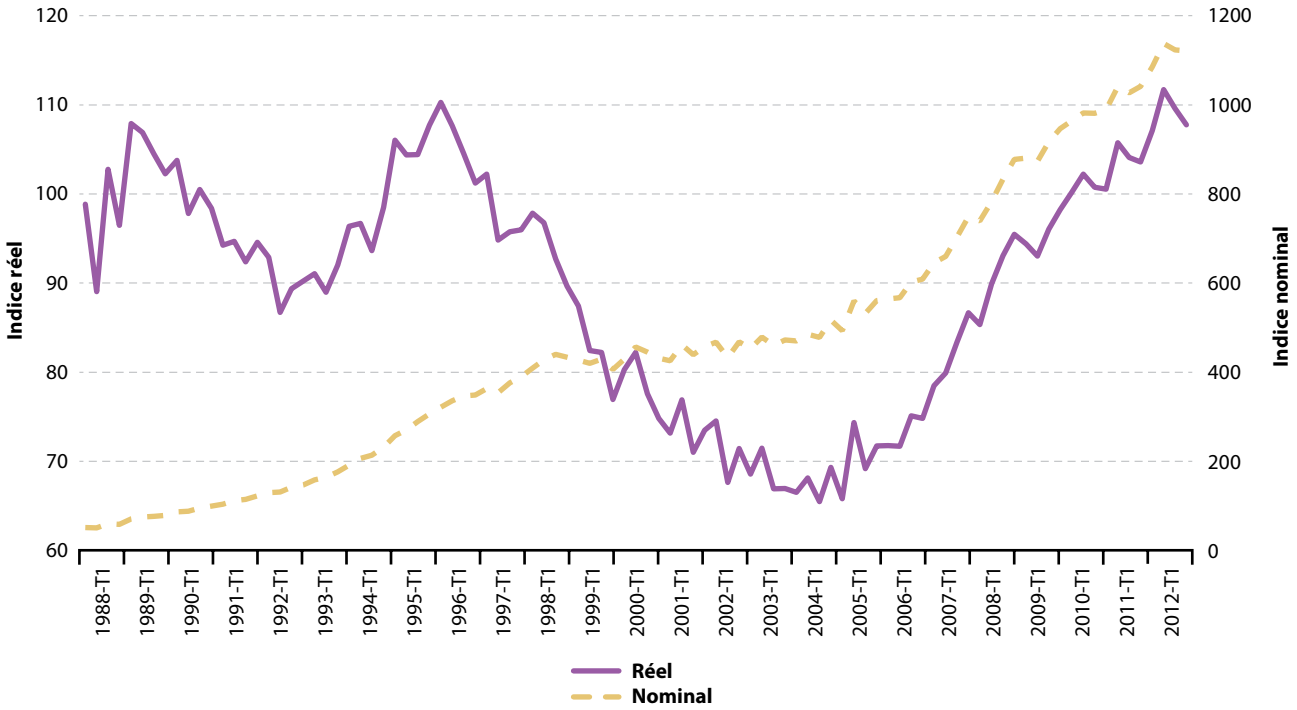
Analyse comparative

10.76 Les sous-indices détaillés qui sont disponibles permettent une analyse plus précise du marché des logements existants. Les graphiques 10.6 à 10.10 montrent l'éventail des statistiques offertes à l'utilisateur. L'«indice nominal» utilise les prix déclarés par les banques et n'est donc pas déflaté; l'«indice réel» est l'indice IPVU déflaté par la variation de l'IPC en moyenne annuelle. L'IPVU trimestriel est déflaté par la variation de l'IPC en moyenne trimestrielle.

26 L'intégration des deux indices poserait des problèmes de cohérence. Ainsi, l'indice IPVU repose sur des évaluations faites indépendamment lors d'une demande de crédit immobilier, tandis que l'indice DANE est basé sur les prix demandés.

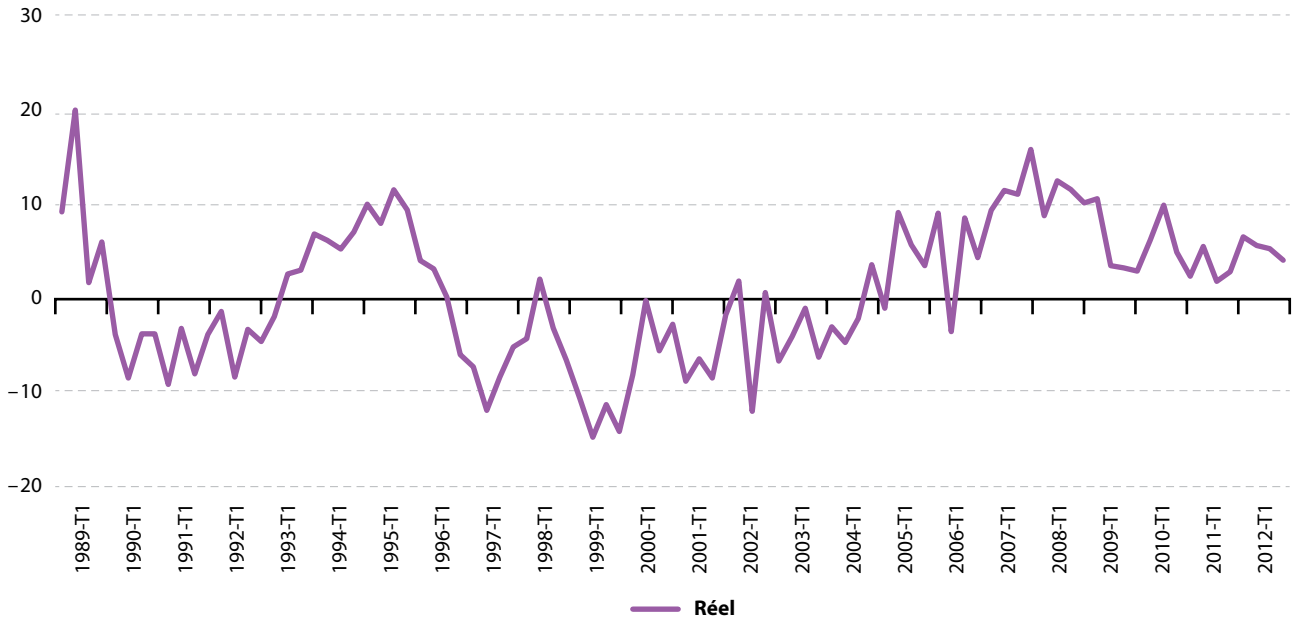
27 Pour plus d'informations sur ce sujet, voir http://www.cijuf.org.co/codian03/junio/c31847.htm.

Graphique 10.6. Indice national trimestriel des prix des logements existants : valeur nominale et réelle (Année de référence 1990 = 100)



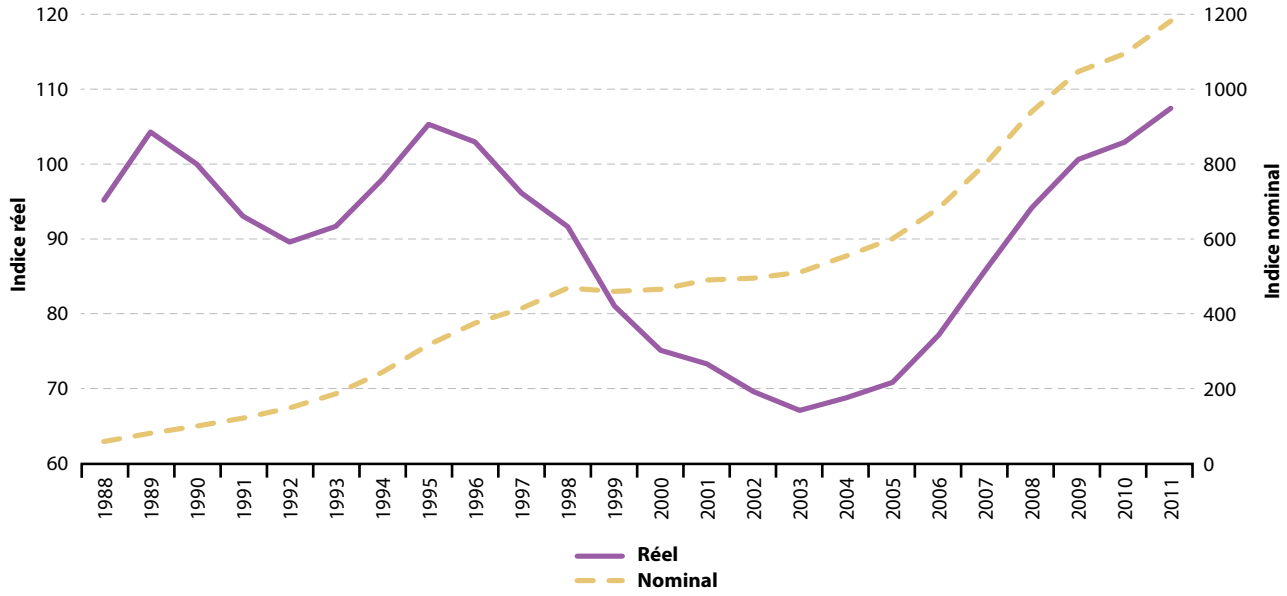
Source : Département de la programmation et de l'inflation du Banco de la República, Colombie.

Graphique 10.7. Indice national trimestriel des prix réels des logements existants : variations annuelles en pourcentage



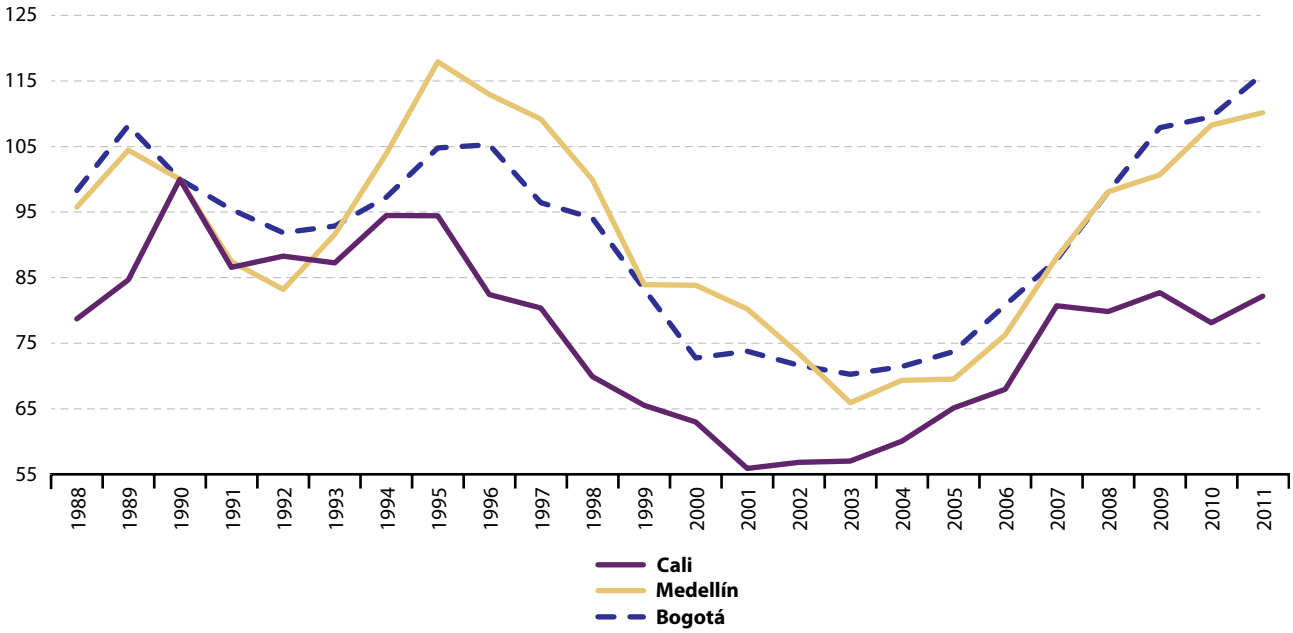
Source : Département de la programmation et de l'inflation du Banco de la República, Colombie.

Graphique 10.8. Indice national annuel des prix des logements existants[1]
(Année de référence 1990 = 100)



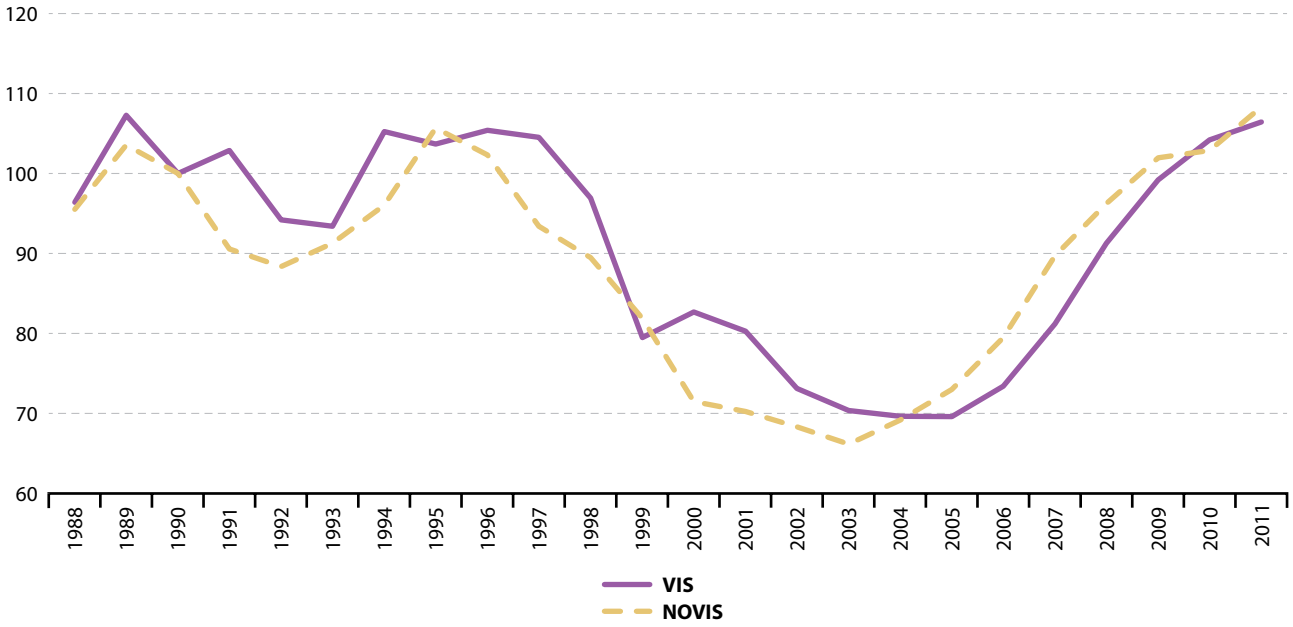
[1] L'IPVU publié annuellement compare le niveau moyen de l'indice pendant une période de 12 mois avec la moyenne des 12 mois précédents.
Source : Département de la programmation et de l'inflation du Banco de la República, Colombie.

Graphique 10.9. Indice annuel des prix réels des logements existants : principales zones métropolitaines
(Année de référence 1990 = 100)



Source : Département de la programmation et de l'inflation du Banco de la República, Colombie.

Graphique 10.10. Indice annuel des prix réels des logements existants : logements subventionnés (VIS) et non subventionnés (NOVIS)
(Année de référence 1990 = 100)



Source : Département de la programmation et de l'inflation du Banco de la República, Colombie.

Étude de cas : l'Afrique du Sud

10.77 L'étude de cas suivante concernant l'Afrique du Sud illustre les obstacles auxquels se heurte l'élaboration d'un indice des prix de l'immobilier résidentiel lorsqu'une proportion significative du stock se compose de logements informels ou traditionnels.

Introduction au marché sud-africain du logement

10.78 En Afrique du Sud, le stock de logements est très différencié; il inclut des éléments formels, informels (y compris de type «backyard»), tribaux, ainsi que des logements dont la propriété est partagée. Les logements formels comprennent les maisons individuelles (subventionnées par l'État ou privées), les maisons en rangée et les appartements; les logements informels, c'est-à-dire ceux dépourvus de permis de construire et non enregistrés officiellement, se composent d'habitations sommaires (généralement en tôle ondulé), de rondavelles et de huttes fabriquées avec des matériaux traditionnels. Les logements de type «backyard» sont situés dans l'enceinte d'une maison principale, tandis qu'il y a partage de propriété quand plusieurs habitations sont construites sur une même parcelle. Le tableau 10.3 décompose le marché immobilier sud-africain. Selon le recensement de la population effectué en 2001, le nombre de logements formels a augmenté de 37,1 % de 1996 à 2001, celui des logements informels de 26,4 % et celui des habitations traditionnelles de 0,6 %. En revanche, le nombre de constructions adjacentes ou avec partage de propriété a diminué de 14,5 %.

10.79 En Afrique du Sud, les entrepreneurs et les promoteurs construisent la totalité des logements, à l'exception des habitations tribales et/ou informelles. S'agissant de la construction de logements formels, elle donne lieu à une transaction monétaire avec financement par les disponibilités de l'acheteur et/ou par un crédit immobilier. Les biens et leur valeur sont répertoriés par la commune et le cadastre. En ce qui concerne les habitations tribales et informelles, il y a très peu de transactions monétaires. Quand il en existe, il s'agit de petites dépenses en numéraire, mais le bien n'est généralement pas répertorié par la commune. Cependant, face à la demande de services de base, les pouvoirs publics ont commencé à recenser le nombre d'habitations implantées sur des emplacements informels et dans les zones rurales, mais la valeur n'est pas répertoriée. Cette situation pose des problèmes très délicats aux producteurs des indices de prix de l'immobilier résidentiel.

Indices des prix de l'immobilier résidentiel en Afrique du Sud

10.80 Il en existe plusieurs, mais ils ne sont pas établis par Statistics South Africa. Les indices publiés sont celui des prix des logements de la First National Bank (FNB), celui de la banque ABSA et l'indice médian des prix des logements de la Standard Bank[28].

10.81 L'indice de la FNB est construit à partir de la valeur moyenne des opérations immobilières financées par cette banque. Pour éliminer les valeurs aberrantes de l'échantillon,

28 L'ABSA, la FNB et la Standard Bank contrôlent la majorité des activités bancaires en Afrique du Sud.

Tableau 10.3. Décomposition par statut du stock de logements en Afrique du Sud (selon le recensement de 2001)

Catégories de logements	Total	Occupation par les propriétaires (%)	En location (%)
Maisons	6.238.454	66,1	45,6
*Maisons subventionnées**	*1.074.028*	*9,6*	–
Appartements	589.109	2,9	16,3
Maisons en rangée	319.868	3,3	4,1
Logements informels	1.836.230	10,3	18,4
Logements traditionnels	1.654.787	15,0	4,1
Logements construits dans un jardin ou à propriété partagée	532.986	2,4	11,5
Total	11.171.434	100,0	100,0

* Estimation du Trésor sud-africain.
Source : Statistics South Africa.

celles qui y sont intégrées doivent être supérieures à 70 % de l'évaluation faite par le service spécialisé de la banque et inférieures à 130 %, tandis que les prix d'achat enregistrés qui dépassent 10 millions de rands sont exclus. Afin de réduire l'incidence sur l'indice des changements soudains et à court terme des pondérations des différents compartiments du marché dus aux variations relatives du nombre des opérations, ces pondérations, liées au nombre de pièces, sont maintenues constantes à leur niveau de pondération moyen sur cinq ans. On applique une fonction de lissage statistique aux données et il arrive qu'elles soient révisées. L'indice FNB est calculé mensuellement.

10.82 L'indice des prix des logements d'ABSA mesure l'évolution nominale d'une année sur l'autre des prix des biens acquis au moyen de prêts immobiliers accordés par cette banque. L'indice se réfère à la totalité des prix d'acquisition des logements d'une surface de 80 m² à 400 m² et d'un montant de 3,1 millions de rands ou moins (améliorations comprises). Les prix sont lissés dans le but d'éliminer l'effet des facteurs saisonniers et des valeurs aberrantes. Cet indice est calculé mensuellement.

10.83 L'indice de la Standard Bank, basé sur le prix médian de toute la gamme des logements, est une moyenne mobile à cinq mois. Les données nationales du service d'enregistrement n'étant disponibles qu'avec un délai pouvant aller jusqu'à neuf mois, l'indice établi par cette banque, qui détient une part de marché de l'ordre de 27,7 % et dont les données sont généralement fortement corrélées à celles du service d'enregistrement, est considéré comme représentatif du marché national. Il est calculé mensuellement.

Obstacles à la construction d'un indice des prix de l'immobilier résidentiel

10.84 Tous les indices mentionnés ci-dessus ne portent que sur les logements formels (c'est-à-dire les maisons ordinaires, les maisons en rangée et les appartements) achetés à crédit — les achats au comptant et ceux d'habitations «informelles» sont exclus. La principale difficulté à laquelle se heurte l'élaboration d'un IPIR en Afrique du Sud est l'absence d'estimations crédibles du stock de logements ainsi que l'absence d'informations sur les prix des habitations informelles et traditionnelles. Comme ces dernières représentent 19,6 % de l'ensemble des bâtiments, elles constituent un secteur important au sein du marché sud-africain.

10.85 Le secteur a aussi ses caractéristiques propres. Comment, par exemple, définit-on un logement informel?

- Des zones habitables où un ensemble de logements ont été construits sur des terrains sur lesquels les habitants n'ont pas de droits reconnus ou qu'ils occupent illégalement.
- Des terrains non aménagés où les habitations ne sont pas conformes à la réglementation foncière et immobilière en vigueur.
- Des habitations informelles dont les murs et le toit sont habituellement en tôle ondulée (sortes de cabanes).
- Les habitants eux-mêmes effectuent l'essentiel de la construction.

En quoi consiste une habitation traditionnelle?

- Il s'agit d'une expression à caractère général qui s'applique aux huttes, aux rondavelles[29], etc. Ces habitations peuvent être séparées ou groupées.
- Elles sont fabriquées avec de l'argile, de la terre, des roseaux ou d'autres matériaux disponibles sur place.

Principaux problèmes posés par la construction d'un indice des prix de l'immobilier résidentiel

10.86 Comme on l'a déjà indiqué dans ce manuel, les deux principaux problèmes posés par l'élaboration d'un indice des prix de l'immobilier résidentiel sont la nature sporadique des transactions et le fait que les caractéristiques qui déterminent

29 Constructions circulaires souvent dotées d'un toit conique en chaume.

Tableau 10.4. Distribution du nombre de pièces dans les habitations informelles

Nombre de pièces	Pourcentage du total des habitations informelles
1	40,0
2	27,2
3	15,1
4	10,5
5 +	7,2

Source : Statistics South Africa.

le prix sont propres à chaque logement. Ces deux facteurs s'appliquent aux logements formels, mais le second a beaucoup moins d'importance pour les logements informels. En Afrique du Sud, exceptionnellement, ces derniers ont des attributs standard, car la plupart sont faits de tôle ondulée et comptent une à quatre pièces. De plus, ils se situent généralement dans des zones similaires. Il est donc peut-être possible d'appliquer sans difficulté le principe du panier fixe. En outre, comme le propriétaire d'une habitation sommaire ne possède pas le terrain sur lequel elle est construite, la décomposition de l'indice en terrains et en bâtiments n'est pas pertinente. Le tableau 10.4 indique la distribution du nombre de pièces ressortant du recensement de 2001.

10.87 La décomposition en terrains et en bâtiments n'est pas non plus pertinente pour les habitations traditionnelles. Une parcelle est allouée à une personne ou à un ménage par le chef tribal, soit gratuitement, soit pour un montant modeste. Toutefois, l'estimation du prix de l'habitation peut s'avérer délicate, si, à la différence des logements formels, la construction se fait surtout à l'aide de matériaux traditionnels.

Pondération des logements non formels

10.88 La pondération des logements non formels (habitations informelles et traditionnelles) sera naturellement complexe pour deux raisons : les propriétaires effectuent eux-mêmes l'essentiel de la construction et les transactions monétaires sont limitées. En outre, la plupart des matériaux utilisés pour la construction d'une habitation informelle sont d'occasion et, pour les habitations traditionnelles, on utilise des matériaux traditionnels; il est difficile d'obtenir des estimations du coût de ces types de matériaux, et il y a des chances qu'ils ne soient pas achetés mais récupérés.

10.89 Le recensement de la population permet de connaître la plupart des caractéristiques des logements, mais il est difficile d'estimer la valeur d'une habitation informelle ou traditionnelle; en effet, il n'y a pas de marché organisé et les prix ne sont pas répertoriés par un service d'enregistrement ou un cadastre. Par ailleurs, la mobilité de l'habitat informel entre différentes implantations risque de compliquer l'estimation du stock de logements. On ne connaît pas le rythme des constructions et des démolitions, faute d'être certain que toutes les habitations démolies soient de nouveau édifiées sur le nouvel emplacement.

Détermination des prix des logements non formels

10.90 Les prix des logements non formels ne dépendent pas des déterminants normaux des prix du marché. En général, la parcelle, l'emplacement, l'ancienneté et les rénovations n'interviennent pas. Les seuls éléments qui influent sur le coût de l'habitation sont la nature des matériaux employés et, sur ce point, la taille du bâtiment joue évidemment un rôle; voir le tableau 10.5.

10.91 Il serait très difficile d'obtenir les prix des logements traditionnels et informels, sachant que le propriétaire effectue le plus souvent lui-même la construction et qu'un logement donne rarement lieu en totalité à une transaction monétaire (l'achat des matériaux se fait normalement en liquide). La seule manière de connaître les prix des nouvelles constructions informelles et traditionnelles est d'effectuer fréquemment des enquêtes, puisque la plupart ne sont pas enregistrées officiellement et, quand elles le sont, la valeur ne figure pas. Une autre solution, restant à explorer, serait d'établir un indice du «coût notionnel de la construction», basé sur l'imputation du prix avec les données quantitatives indiquées au tableau 10.6[30].

Résumé

10.92 En Afrique du Sud, le calcul d'un indice exhaustif des prix de l'immobilier résidentiel serait une tâche très complexe en raison de la diversité du stock de logements de ce pays. Il faudrait recourir à différentes méthodes pour obtenir les prix des diverses catégories de logements. Il serait également difficile d'estimer la pondération de chacune de ces catégories, car elles ont des facteurs de détermination des coûts qui leur sont propres. Enfin, le nombre limité de données disponibles dans chaque catégorie compliquerait encore le problème.

10.93 Les principaux obstacles à la construction d'un indice général des prix de l'immobilier résidentiel en Afrique du Sud sont énumérés au tableau 10.7; ce sont :

- L'absence d'un marché organisé des logements informels et traditionnels.
- L'absence de données fiables sur le coût de ces habitations.
- Le mode de vie nomade. Si une enquête est effectuée, les déplacements des logements informels d'une zone à l'autre posent un problème de mesure de l'évolution des prix de cette catégorie, parce qu'ils sont normalement recueillis à des endroits spécifiques.
- Il n'y a pas d'enregistrement administratif des biens.
- L'obtention ou la construction d'un logement ne donne pas toujours lieu à des transactions monétaires.
- Les prix ne dépendent pas des facteurs habituels, comme les coûts des terrains, de la main-d'œuvre et des matériaux.

30 Voir Blades (2009).

Tableau 10.5. Déterminants des prix

Déterminants des prix	Habitations traditionnelles	Habitations informelles	Habitations formelles
Bâtiment	Non	Non	Oui
Terrain	Non	Non	Oui
Emplacement	Non	Non	Oui
Ancienneté	Non	Non	Oui
Rénovations	Non	Non	Oui
Type de bâtiment	Non	Non	Oui
Type de matériaux	Oui	Oui	Oui
Autres caractéristiques déterminant le prix	Non	Non	Oui

Source : Statistics South Africa.

Tableau 10.6. Décomposition en pourcentage des matériaux utilisés dans la construction des habitations informelles et traditionnelles en Afrique du Sud

Année	2002	2003	2004	2005	2006	2007	2008	2009
Matériaux utilisés pour le toit								
Tôle ondulée/zinc	72,1	72,1	71,6	78,2	79,5	78,6	78,6	83,6
Matériaux organiques	23,2	24,2	23,8	16,8	16,2	17,1	15,8	13,3
Amiante	1,9	1,6	1,4	1,7	1,8	1,2	2,1	0,5
Autres	2,6	2,1	3,1	3,2	2,1	3,1	3,1	2,2
Total	**100,0**	**100,0**	**100,0**	**100,0**	**100,0**	**100,0**	**100,0**	**100,0**
Matériaux utilisés pour les murs								
Briques	2,6	3,3	2,3	1,8	1,4	1,6	2,3	1,9
Ciment/béton	2,9	2,2	2,8	1,9	2,5	2,3	2,4	1,4
Tôle ondulée/zinc	35,1	36,1	33,9	40,0	43,6	43,9	41,4	42,2
Bois	9,8	9,4	8,9	9,6	10,5	10,8	10,1	8,6
Mélange de terre et de ciment	7,0	5,2	6,3	5,0	5,8	6,5	6,7	10,4
Torchis et enduis	1,4	1,1	1,7	1,0	0,5	0,9	1,3	1,2
Terre	38,2	39,8	41,8	37,2	33,7	31,8	32,8	31,8
Autres	2,6	2,9	2,3	2,6	1,8	2,2	2,9	2,5
Total	**100,0**	**100,0**	**100,0**	**100,0**	**100,0**	**100,0**	**100,0**	**100,0**

Source : Statistics South Africa.

Tableau 10.7. Évaluation des obstacles

Problèmes possibles	Logements traditionnels	Logements informels	Logements formels
Marché organisé	Non	Non	Oui
Existence d'estimations fiables du coût des logements	Non	Non	Oui
Mobilité des logements d'une implantation à l'autre	Non	Yes	Non
Enregistrement administratif des biens	Non	Non	Oui
Transaction monétaire à l'établissement de crédit	Non	Non	Oui
Transfert de numéraire pour la construction de la structure	Quelquefois	Quelquefois	Oui
Construction par un promoteur ou un entrepreneur	Non	Non	Oui
Le prix dépend des facteurs habituels	Non	Non	Oui

Source : Statistics South Africa.

11

Exemples empiriques

Introduction

11.1 L'objet de ce chapitre est de donner des exemples empiriques supplémentaires de construction d'indices des prix de l'immobilier résidentiel selon les méthodes des chapitres 5 à 9. Globalement il s'agit de la mesure d'une tendance centrale (moyenne ou médiane), de l'utilisation de régressions hédoniques, des méthodes des ventes répétées et des méthodes basées sur les évaluations des données. Les trois sections suivantes de ce chapitre illustrent comment appliquer les trois premières méthodes à de très petits échantillons. Le but étant qu'à travers ces exemples simplifiés, le lecteur puisse suivre plus facilement les descriptions algébriques, plutôt abstraites, des différentes méthodes décrites aux chapitres 5 à 9.

11.2 La section suivante illustre aussi les différentes méthodes d'agrégation qui peuvent être utilisées pour calculer l'indice d'ensemble de l'immobilier à partir des indices des prix régionaux. Ce sujet n'a pas été traité en détail dans les autres chapitres de ce manuel.

Méthodes de la tendance centrale et de la stratification

11.3 Les estimations de tendances centrales utilisées pour calculer un IPIR, comme la moyenne et la médiane des prix, font partie des méthodes aujourd'hui à la disposition des statisticiens les moins coûteuses en données. La moyenne simple et la médiane n'ont besoin que des prix de vente de l'immobilier d'un endroit donné pour calculer un indice de prix. L'information géographique sera en conséquence nécessaire. De plus, il est habituel de stratifier selon le type d'habitation, et si cela est le cas, cette information sera également requise.

11.4 Le premier exercice consiste à construire un indice en utilisant la moyenne des prix. La moyenne simple des prix relevés d'un échantillon de maisons est calculée pour une période et une zone géographique données. Cet indicateur, qui peut être exprimé en valeur monétaire ou par un indice, mesure uniquement (souvent en pourcentage) la variation du prix moyen de l'échantillon entre deux périodes[1].

11.5 Il est important que l'échantillon de maisons sélectionné pour calculer cet indicateur de prix soit représentatif. Ainsi, il sera peut être nécessaire d'apurer les données, selon les instructions que le statisticien aura fournies à l'équipe de collecte, et selon la volonté et la capacité de cette dernière de les suivre[2]. Par exemple, l'échantillon des prix initialement collectés peut inclure certaines catégories de propriété, comme des terres agricoles, des commerces et des logements collectifs considérés comme hors-champ pour l'indice particulier souhaité. Si cela est le cas, ces observations doivent être exclues de l'échantillon lors de la mesure spécifique souhaitée. Les observations aberrantes doivent aussi être identifiées et retirées de l'échantillon si l'on pense que cela pourrait fausser ou biaiser le résultat.

11.6 Un exemple numérique simple utilisant 5 et 7 observations de prix, respectivement, pour les périodes 1 et 2[3] va servir à illustrer l'approche utilisée pour mesurer l'évolution de la moyenne simple des prix du logement pour une zone géographique, habituellement une ville ou une autre zone bien définie[4].

Période 1, prix du logement et moyenne

(350.000 + 352.000 + 378.000 + 366.000 + 402.000)/5 = 370.000)

Période 2, prix du logement et moyenne

(360.000 + 350.000 + 382.000 + 395.000 + 380.000 + 400.000 + 450.000/7 = 388.000)

Une fois que les prix moyens pour chaque période (un mois, un trimestre ou une année) sont obtenus, il est aisé de calculer l'évolution d'une période à l'autre, habituellement exprimée en pourcentage. Dans l'exemple précédent, entre 370.000 dollars et 388.000 dollars, le prix moyen des maisons a augmenté de 5 % entre les deux périodes.

11.7 L'influence des observations aberrantes est réduite lorsque la médiane des prix de l'échantillon est utilisée à la place de la moyenne. Par exemple, si une ou plusieurs maisons très chères sont vendues pendant une période donnée, la moyenne des prix ne sera probablement pas représentative des prix du marché à cette période. Comme ce qui a été signalé au chapitre 4, l'utilisation de la médiane ne contrôle quand même pas complètement les changements de composition de l'échantillon des maisons vendues d'une période à l'autre. Malgré ce défaut, la médiane est néanmoins un indicateur très utilisé des prix de l'immobilier résidentiel, principalement en raison de sa simplicité, son économie en données et donc la rapidité de sa disponibilité. De plus son interprétation est simple.

11.8 En prenant les mêmes données du calcul de la moyenne de l'exemple précédent, la médiane des prix des périodes 1 et 2 est respectivement de 366.000 dollars et de

1 La variation en pourcentage sera la même, que l'indicateur soit exprimé par une valeur ou par un indice.

2 Bien sûr, des cas particuliers influenceront l'apurement des données. Si le principal utilisateur dirige aussi leur collecte, alors l'enquête sera adaptée à ses besoins et l'apurement sera limité.

3 Comme le nombre de ventes variera vraisemblablement d'une période à l'autre, le nombre d'observations de prix dans l'échantillon pour chaque période variera lui aussi.

4 Notons que la plupart des publications de mesures de tendances centrales ne sont typiquement pas accompagnées d'indicateur de qualité statistique comme l'écart-type.

382.000 dollars. Le prix médian a donc augmenté de 4,4 % entre les deux périodes.

11.9 L'exercice est maintenant répété avec un échantillon plus important de 5.787 prix de ventes effectives de maisons individuelles pendant plusieurs années pour une petite ville[5]. Des statistiques descriptives sont présentées dans le tableau 11.1. Notons que dans ce cas particulier, pour n'importe laquelle des années, la moyenne des prix du logement est supérieure à la médiane. Par exemple, en 2002 la moyenne est de 249.702 dollars contre 236.000 dollars pour la médiane; en 2008 la moyenne et la médiane valent respectivement 365.195 dollars et 340.600 dollars. Comme chaque année, il y a toujours quelques ventes de maisons très chères, cette différence est prévisible. En fait, la distribution des prix présente une asymétrie positive, avec un coefficient d'asymétrie variant selon les années entre 1,44 et 1,87[6]. Le graphique 11.1 illustre la distribution des prix du logement vendues en 2008. Des graphiques similaires pour les autres années de cet exemple produisent des distributions des prix semblables[7].

11.10 La variation annuelle en pourcentage dépend de la mesure de la tendance centrale utilisée[8]. Pour quelques années, la différence numérique entre la médiane et la moyenne est plutôt faible. Par exemple, en 2002, la différence est seulement d'un dixième de point de pourcentage (8,2 % contre 8,1 %) avec une moyenne subissant une augmentation légèrement plus forte. Pour d'autres années, comme en 2008, la différence est plus importante. En 2008, la variation annuelle des prix mesurée par la médiane était de 6,8 % contre 5,2 % pour celle mesurée avec la moyenne.

5 Notons que les données servent à calculer à la fois la moyenne et la médiane. Les étapes sont plutôt aisées et la plupart des logiciels de statistique peuvent effectuer rapidement l'ensemble des calculs avec une intervention réduite de l'utilisateur.

6 Le coefficient d'asymétrie mesure l'asymétrie de la distribution. Quand il est nul, cela veut dire que la distribution est symétrique autour de la moyenne. Un coefficient d'asymétrie positif veut dire qu'un grand nombre d'observations se concentrent sur la gauche du mode de la distribution et vice-versa.

7 Avec ces données particulières, la moyenne était toujours plus grande que la médiane correspondante. Ce résultat n'est pas toujours vrai, en particulier avec les échantillons très petits.

8 D'habitude, la moyenne sera plus grande que la médiane correspondante. Néanmoins, par la suite, il n'y a aucune raison que l'indice formulé à partir de la moyenne augmente plus vite que celui formulé à partir de la médiane.

11.11 Comme il est bien connu, l'emplacement est un facteur déterminant non seulement du prix d'une maison, mais aussi de son évolution dans le temps. Ainsi, pour améliorer la fiabilité d'un indicateur, il est habituellement recommandé de stratifier ou de poststratifier l'estimateur, à condition, bien sûr, que l'information pour segmenter le marché (ou l'échantillon de ventes) soit facilement disponible. La stratification géographique a l'avantage de réduire l'effet sur l'évolution qui est dû au changement de composition de l'échantillon de logements, effet qui affecte les méthodes utilisant la moyenne simple ou la médiane. Une approche répandue pour stratifier le marché du logement est de regrouper les maisons selon un critère géographique, assurant ainsi une certaine homogénéité des logements de la strate. Cette approche réduit aussi les effets des autres critères d'emplacement. La stratification permet aussi de donner aux utilisateurs des informations sur les prix de l'immobilier selon différents segments du marché, comme le type de maison ou le voisinage. Goodman et Thibodeau (2003) ajoutent aussi que la disponibilité des variables géographiques dans les bases de données sur les ventes de logements est aussi une raison pratique pour regrouper les maisons selon leur emplacement général. La stratification est une manière efficace d'utiliser ces données, il faut donc en profiter lorsqu'elles sont disponibles.

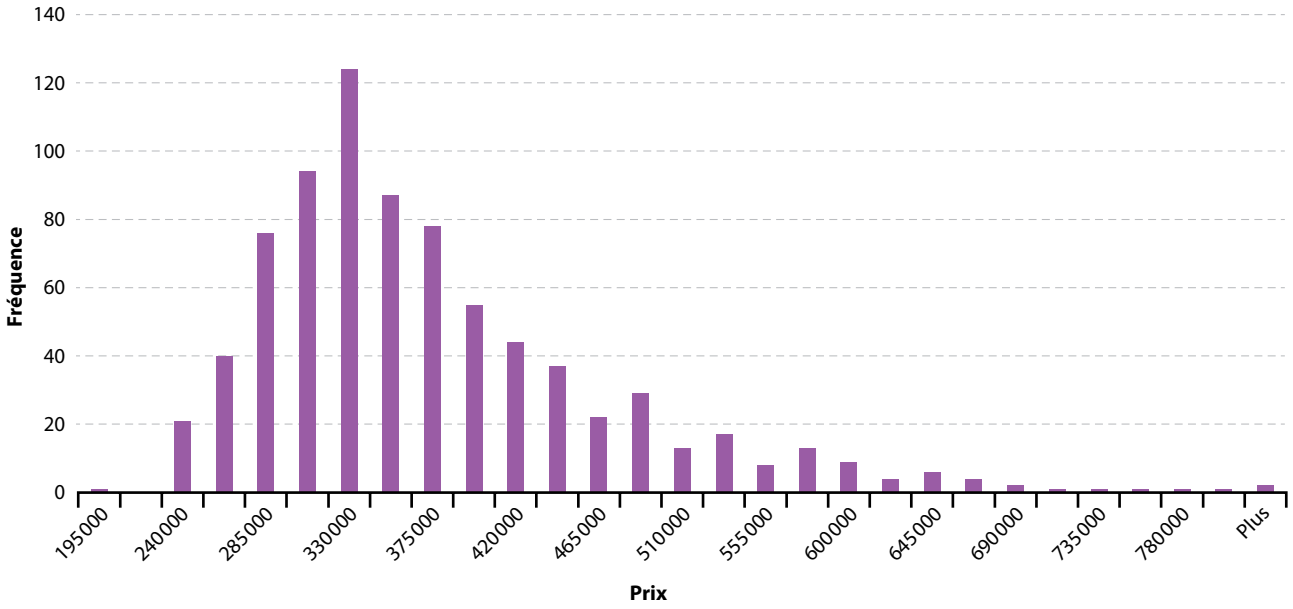
11.12 Quelques pays comme l'Australie (Branson, 2006) ont tiré parti de la relation traditionnellement forte entre les prix et l'emplacement qui caractérise le marché de l'immobilier résidentiel, en stratifiant l'échantillon des résidences selon un critère géographique ou d'autres caractéristiques du marché. Cela peut être une alternative viable, bien qu'imparfaite (ou une solution de compromis) pour mesurer un indice des prix à qualité constante en l'absence des moyens et des données nécessaires pour mettre en œuvre des méthodes plus sophistiquées d'élaboration d'un IPIR telles que les régressions hédoniques. En fait, Prasad et Richards (2008) mesurent un prix médian des maisons de six capitales australiennes, stratifié par les évolutions à long terme des prix. En utilisant une base de données de plus de 3 millions d'observations, les auteurs trouvent que

Tableau 11.1 Moyenne, médiane, variation en pourcentage, écart-type et coefficient d'asymétrie

	2002	2003	2004	2005	2006	2007	2008
Observations	777	804	894	808	834	874	796
Écart-type	64.130	62.042	73.405	76.432	84.587	96.559	96.152
Coefficient d'asymétrie	1,63	1,51	1,71	1,87	1,58	1,46	1,44
Moyenne ($)	249.702	270.174	290.686	299.087	315.099	347.009	365.195
Variation en pourcentage		**8,2%**	**7,6%**	**2,9%**	**5,4%**	**10,1%**	**5,2%**
Médiane ($)	236.000	255.000	273.000	280.000	292.000	319.000	340.600
Variation en pourcentage		**8,1%**	**7,1%**	**2,6%**	**4,3%**	**9,2%**	**6,8%**

Source : calculs des auteurs basés sur les données MLS® pour une ville canadienne.

Graphique 11.1. Distribution des prix du logement en 2008



Source : calculs des auteurs basés sur les données MLS® pour une ville canadienne.

leur approche pour mesurer l'évolution du prix du logement (c'est-à-dire une médiane stratifiée par zones, elles-mêmes définies selon les évolutions de long terme des prix) donne des résultats comparables à celles utilisant des méthodes plus sophistiquées et plus de données, comme les méthodes hédoniques ou les méthodes des ventes répétées.

11.13 La stratification géographique assure probablement que les grappes d'observations au sein de chaque groupe ou strate sont plus homogènes que celles provenant de la population entière. La stratification peut être étendue, en plus des critères géographiques, aux autres facteurs qui déterminent le prix comme le type de la maison et/ou le nombre de pièces. Le regroupement des maisons par critère géographique et selon d'autres critères fournira un échantillon avec de meilleures propriétés d'homogénéité, ce qui est souhaitable pour atténuer les évolutions de l'indice causées par les modifications de composition de l'échantillon qui arrivent au cours du temps. Cependant, un éventuel inconvénient avec cette approche est que le statisticien devra se garder de créer des strates trop étroites, car un petit nombre d'observations dans une strate, à chaque période, pourrait générer certains biais d'échantillonnage. L'objectif est ainsi de définir chaque strate en équilibrant les avantages de l'homogénéité des caractéristiques déterminant le prix avec les avantages d'avoir un échantillon de taille suffisante pour fournir une estimation fiable et représentative de l'évolution du prix du logement.

11.14 Comme nous avons vu précédemment, la construction d'indices par segment du marché (ou stratifié), qui sont ensuite agrégés au niveau général souhaité, utilisera souvent, en pratique, la médiane des prix. La construction d'un indice stratifié consiste dans un premier temps à définir les strates. La deuxième étape est de calculer la médiane par strate des prix des maisons vendues pour la période mesurée. Ensuite, les médianes des prix pour chaque strate doivent être pondérées et additionnées pour former la mesure du prix du marché étudié, souvent une ville, et parfois l'ensemble du pays.

11.15 Le paragraphe suivant propose un exemple simple de la méthode de calcul d'un indice de prix de l'immobilier résidentiel par stratification[9].

- Étape 1 : Définir les strates. Pour cet exemple, les strates seront des sous-divisions géographiques de la ville, telles que les quartiers Ouest ou le centre-ville. Il n'y a pas de règles strictes pour délimiter les strates en question, mais l'emplacement revient souvent, c'est un choix couru qui semble évident, combinable, si les données sont disponibles, avec d'autres caractéristiques du logement, comme le type de maison ou le nombre de chambres, afin de définir plus précisément les strates[10].
- Étape 2 : Calculer le prix médian par strate, comme le quartier, pour la période significative (mois ou trimestre). Il est postulé que la médiane sera représentative du prix de

9 Cet exemple est librement inspiré de celui de McDonald et Smith (2009).

10 Cet exemple utilise le quartier comme sous-strate, mais, en réalité, ce pourrait être n'importe quelle zone géographique pour laquelle le statisticien serait certain d'avoir assez d'observations, actuellement et dans le futur, de façon à calculer un prix fiable et représentatif.

Tableau 11.2. Dépenses régionales, prix et volumes (quantités implicites), utilisant le prix médian comme prix régional

Période	v^t_A	v^t_B	v^t_C	p^t_A	p^t_B	p^t_C	q^t_A	q^t_B	q^t_C
0	1.300	500	675	300	500	200	4,333	1,000	3,375
1	1.725	400	825	300	400	250	5,750	1,000	3,300

Source : calculs des auteurs basés sur les données MLS® pour une ville canadienne.

vente de toute la strate. Cependant, la moyenne des prix peut aussi être une alternative. Cette étape sera répétée pour les périodes futures.

- Étape 3 : Estimer le prix «moyen» des maisons vendues pendant une certaine période en calculant une moyenne pondérée des prix médians par quartier ou par strate[11].

11.16 Supposons que les données pour les ventes de maisons concernant deux périodes (0 et 1) et trois régions géographiques ou quartiers (A, B et C) ont été collectées. Supposons que les prix sont mesurés en milliers de dollars et que 4 ventes ont été enregistrées dans la région A lors de la période 0, dont les prix sont 290, 450, 250 et 310. Ainsi, la moyenne des prix pour cette période est de 325, la médiane de 300 (la moyenne arithmétique des deux prix du milieu 290 et 310) et la dépense totale est de 1.300. Pour la période 1 et la région A, supposons 5 ventes de 300, 500, 250, 400 et 275. Ainsi la moyenne et la médiane pour cette période sont respectivement de 345 et de 300, et la dépense totale de 1.725. Supposons pour la région B une seule vente pour chaque période : 500 pour la période 0, et 400 pour la période 1. Ainsi, la moyenne et la médiane pour la période 0 sont de 500, aussi égales à la dépense de cette période. Pour la période 1 et la région B, la moyenne et la médiane du prix sont de 400, aussi égales à la dépense de cette période. Supposons 3 ventes à chaque période pour la région C. Pour la période 0, supposons que les ventes sont égales à 200, 300 et 175; ainsi la médiane des prix est de 200, la moyenne des prix de 225 et la dépense totale de 675. Pour la période 1, supposons que les ventes de la région C sont égales à 250, 350 et 225; ainsi la médiane des prix est de 250, la moyenne des prix de 275 et la dépense totale de 825. Ce sont les données de base pour l'exemple.

11.17 Supposons que le *prix médian* pour chaque région correspond aux maisons de qualité comparable pour les deux périodes considérées. Comme il est souhaitable d'avoir un produit du prix par le volume égal aux dépenses pour chaque période et chaque région, dès lors qu'un concept de prix à qualité constante a été choisi, le volume correspondant devrait être égal aux dépenses divisées par le prix. En utilisant le prix médian dans chaque région comme un prix à qualité constante pour chaque période, on obtient les dépenses (v^t), les prix (p^t) et les volumes ou quantités implicites $q^t = v^t/p^t$ qui figurent dans le tableau 11.2 ci-dessus.

Notons que les indices régionaux des prix de la période 1 sont égaux à $p^1_A/p^0_A = 1{,}0$, $p^1_B/p^0_B = 0{,}80$ et $p^1_C/p^0_C = 1{,}25$ pour les régions A, B et C, respectivement. Ainsi, il y a de grandes différences d'inflation entre les trois régions pour le prix du logement.

11.18 À ce stade, nous pouvons appliquer la théorie classique des indices de prix pour l'agrégation des indices régionaux dans un taux global de l'inflation du prix du logement. Par exemple des indices généraux de *Laspeyres* et de *Paasche*, P_L et P_P, pour la période 1 peuvent être calculés. Les formules algébriques de ces indices sont les suivantes :

$$P_L \equiv [p^1_A q^0_A + p^1_B q^0_B + p^1_C q^0_C]/[p^0_A q^0_A + p^0_B q^0_B + p^0_C q^0_C] \quad (11.1)$$

$$P_P \equiv [p^1_A q^1_A + p^1_B q^1_B + p^1_C q^1_C]/[p^0_A q^1_A + p^0_B q^1_B + p^0_C q^1_C] \quad (11.2)$$

11.19 Le *Manuel de l'indice des prix à la consommation* (Manuel de l'IPC, 2004) recommande le calcul d'*indices superlatifs* si les données de prix et de quantités sont disponibles pour les périodes considérées, comme c'est le cas pour cet exemple. Les indices de *Fisher*, P_F, et de *Törnqvist-Theil*, P_T sont deux indices superlatifs; ils sont définis pour la période 1 par les formules ci-dessous :

$$P_F \equiv [P_L P_P]^{1/2} \quad (11.3)$$

$$P_T \equiv \exp[0{,}5(s^0_A + s^1_A)\ln(p^1_A/p^0_A) + 0{,}5(s^0_B + s^1_B) \ln(p^1_B/p^0_B) + 0{,}5(s^0_C + s^1_C)\ln(p^1_C/p^0_C)] \quad (11.4)$$

où la *part des ventes pour la période t* dans les régions A, B et C est donnée respectivement par $s^t_A \equiv v^t_A/(v^t_A + v^t_B + v^t_C)$, $s^t_B \equiv v^t_B/(v^t_A + v^t_B + v^t_C)$ et $s^t_C \equiv v^t_C/(v^t_A + v^t_B + v^t_C)$. Notons que l'indice de Fisher (1922) P_F est égal à la moyenne géométrique des indices de Laspeyres et de Paasche, P_L et P_P, et que l'indice de Törnqvist-Theil P_T est égal à la moyenne géométrique pondérée des indices des prix régionaux, p^1_A/p^0_A, p^1_B/p^0_B et p^1_C/p^0_C , où les pondérations sont les moyennes arithmétiques des parts des ventes entre la période 0, s^0_A, s^0_B et s^0_C, et la période 1, s^1_A, s^1_B et s^1_C.

11.20 Les résultats numériques pour les quatre indices définis par les formules (11.1)–(11.4) sont donnés dans le tableau 11.3 ci-dessous. Il faut noter que les deux indices

11 Cela suppose que le statisticien utilise les ventes comme base pour la pondération.

Tableau 11.3. Indice global des prix du logement utilisant la médiane des prix et d'autres formules pour agréger les régions A, B et C

Période	P_F	P_T	P_L	P_P	P_0	P_1	P_A	P_{GL}	P_{GP}
0	1,00000	1,00000	1,00000	1,00000	1,00000	1,00000	1,00000	1,00000	1,00000
1	1,02515	1,02425	1,02778	1,02253	1,02778	1,04280	1,03529	1,01590	1,03267

Source : calculs des auteurs basés sur les données MLS® pour une ville canadienne.

superlatifs, P_F et P_T, sont assez proches entre eux, tandis que l'indice de Laspeyres P_L leur est plus élevé et que l'indice de Paasche leur est plus faible. C'est un résultat empirique typique.

11.21 Les institutions qui calculent des indices de prix de l'immobilier résidentiel ont tendance à utiliser des formules quelque peu différentes pour agréger les régions. Une méthode commune d'agrégation consiste à utiliser une moyenne *pondérée* des indices régionaux des prix pour calculer l'indice d'ensemble, en utilisant les ventes de la période 0 comme pondérations (ou une moyenne de ventes pour des périodes antérieures à la période 0). Notons P_0, l'indice qui utilise les ventes de la période 0 comme pondérations, et P_1, celui qui utilise les ventes de période 1. Les valeurs des indices P_0 et P_1 pour la période 1[12] ainsi que leur moyenne arithmétique, notée P_A, sont déterminées par les formules ci-dessous :

$$P_0 \equiv s_A^0(p_A^1/p_A^0)+s_B^0(p_B^1/p_B^0)+s_C^0(p_C^1/p_C^0) \qquad (11.5)$$

$$P_1 \equiv s_A^1(p_A^1/p_A^0)+s_B^1(p_B^1/p_B^0)+s_C^1(p_C^1/p_C^0) \qquad (11.6)$$

$$P_A \equiv 0{,}5P_0+0{,}5P_1 \qquad (11.7)$$

Les trois indices ci-dessus sont aussi donnés dans le tableau 11.3[13]. Nous constatons que P_0 est égal à P_L et d'environ 0,26 point de pourcentage plus élevé que l'indice de Fisher P_F pour la période 1, alors que P_1 est d'environ 1,77 point de pourcentage plus élevé que P_F. Ce résultat n'est pas inattendu, les indices P_0 et P_1 ne sont en général pas des approximations proches des indices superlatifs et leur usage n'est donc pas recommandé.

11.22 Deux indices supplémentaires sont donnés dans le tableau 11.3 : l'*indice géométrique de Laspeyres* et l'*indice géométrique de Paasche*, respectivement P_{GL} et P_{GP}. Les formules pour la période 1 sont données ci-dessous :

$$P_{GL} \equiv \exp[s_A^0\ln(p_A^1/p_A^0)+s_B^0\ln(p_B^1/p_B^0)+s_C^0\ln(p_C^1/p_C^0)] \qquad (11.8)$$

$$P_{GP} \equiv \exp[s_A^1\ln(p_A^1/p_A^0)+s_B^1\ln(p_B^1/p_B^0)+s_C^1\ln(p_C^1/p_C^0)] \qquad (11.9)$$

Ainsi, en période 1, la valeur pour chacun de ces deux indices est une moyenne géométrique pondérée par des parts de valeur des indices régionaux des prix, p_A^1/p_A^0, p_B^1/p_B^0 et p_C^1/p_C^0, où P_{GL} utilise les parts de valeurs régionales relatives à la période 0, s_A^0, s_B^0 *et* s_C^0, tandis que P_{GP} utilise celles de la période 1, s_A^1, s_B^1 *et* s_C^1. Le tableau 11.3 montre que l'indice géométrique de Laspeyres, P_{GL}, est approximativement 1 point de pourcentage inférieur aux indices superlatifs P_F et P_T tandis que l'indice géométrique de Paasche est approximativement 1 point de pourcentage supérieur à ces indices superlatifs[14]. Par conséquent, l'utilisation des indices géométriques de Laspeyres ou de Paasche ne peut pas être recommandée pour agréger des indices régionaux de prix. Ces formules ne sont vraisemblablement pas de bons estimateurs des indices superlatifs, lesquels peuvent facilement être calculés avec les données régionales sur le prix de vente des maisons.

11.23 Les méthodes utilisées ci-dessus pour agréger les indices régionaux de prix supposent que le prix médian de chaque région provient de maisons de qualité constante pendant les deux périodes qui sont comparées. Supposons maintenant qu'au lieu d'utiliser les prix médians régionaux pour représenter les prix de maisons de qualité constante, nous utilisions la moyenne régionale des prix. De nouveau, comme il est souhaitable que le produit des prix et des quantités soit égal aux dépenses — et ce pour chaque période et pour toutes les régions —, une fois qu'il est décidé d'utiliser la moyenne des prix pour suivre les prix à qualité constante, le volume correspondant devra être égal aux dépenses divisées par les prix. Ainsi, en se servant de la moyenne des prix comme un prix à qualité constante dans chaque région et pour chaque période, on obtient les données régionales de dépenses (v^t), de prix (p^t) et de volumes (ou quantités implicites $q^t = v^t/p^t$) présentées dans le tableau 11.4 ci-dessous.

12 Les valeurs pour la période 0 de tous les indices définis dans cette section sont fixées à 1.

13 Fisher (1922; 466) a montré que P_0 défini par (11.5) est égal à l'indice de Laspeyres P_L défini par (11.1). Fisher a aussi appelé indice de Palgrave, l'indice P_1 défini par (11.6).

14 Il peut être vérifié que la moyenne géométrique de P_{GL} et de P_{GP} est exactement égale à P_T. Ainsi si P_{GL} est plus petit que P_T, alors P_{GP} sera automatiquement plus grand que P_T.

Tableau 11.4. Dépenses régionales, prix et volumes (quantités implicites), utilisant la moyenne pour calculer les prix régionaux

Période	v^t_A	v^t_B	v^t_C	p^t_A	p^t_B	p^t_C	q^t_A	q^t_B	q^t_C
0	1.300	500	675	325	500	225	4	1	3
1	1.725	400	825	345	400	275	5	1	3

Source : calculs des auteurs basés sur les données MLS® pour une ville canadienne.

Tableau 11.5. Indice global des prix du logement utilisant la moyenne des prix et d'autres formules pour agréger les régions A, B et C

Période	P_F	P_T	P_L	P_P	P_0	P_1	P_A	P_{GL}	P_{GP}
0	1,00000	1,00000	1,00000	1,00000	1,00000	1,00000	1,00000	1,00000	1,00000
1	1,05305	1,05222	1,05253	1,05357	1,05253	1,07101	1,06177	1,04187	1,06267

Source : calculs des auteurs basés sur les données MLS® pour une ville canadienne.

11.24 Utiliser la moyenne à la place de la médiane comme prix à qualité constante pour chaque région change l'indice régional des prix. Les indices régionaux des prix utilisant la moyenne, en période 1 pour les régions A, B, et C, sont respectivement égaux à $p^1_A / p^0_A = 345 / 325 = 1{,}06154$, $p^1_B / p^0_B = 400 / 500 = 0{,}80$, et $p^1_C / p^0_C = 275 / 225 = 1{,}2$. Une fois de plus, il y a de grandes différences dans les taux d'inflation des prix du logement dans les trois régions quand la moyenne des prix est utilisée à la place de la médiane.

11.25 En utilisant la moyenne à la place de la médiane, il est possible de calculer les différents indices d'ensemble de prix définis par les formules (11.1) à (11.9). Le tableau connexe du tableau 11.3 présenté ci-dessus est obtenu en appliquant ces formules aux données du tableau 11.4.

On peut constater que l'utilisation de la moyenne régionale des prix à la place de la médiane a donné des indices très différents; les indices superlatifs P_F et P_T sont ainsi plus élevés d'environ 3 points de pourcentage pour la période 1. Cependant, l'utilisation de la moyenne des prix a donné des indices de Laspeyres et de Paasche, P_L et P_P, qui sont passablement proches de leurs équivalents superlatifs. Comme l'indice pondéré par les parts de valeur de la période de référence, P_0, est numériquement égal à P_L, P_0 est aussi plutôt proche de P_F et P_T. Cependant, les deux autres indices pondérés par des parts de valeur, P_1 et P_A, sont bien au-dessus des indices superlatifs. Pour finir, l'indice géométrique de Laspeyres, P_{GL}, est largement en dessous de P_T et l'indice géométrique de Paasche, P_{GP}, est largement au-dessus de P_T. En tout cas, l'utilisation de la moyenne des prix dans le contexte du logement n'est pas recommandée, car la moyenne du prix d'une maison dans une région donnée ne s'accompagnera vraisemblablement pas d'un niveau constant de qualité dans la durée.

Méthodes de régression hédonique

11.26 Le chapitre 5 traite de l'utilisation des méthodes hédoniques pour calculer un indice du prix du logement. Il y a plusieurs façons d'appliquer ces techniques pour calculer un indice des prix en général et un indice des prix de l'immobilier résidentiel en particulier. Le manuel présente trois variantes de la méthode hédonique : la méthode des variables indicatrices temporelles, l'estimation du prix des caractéristiques (l'imputation) et la méthode hédonique stratifiée. Comparées aux autres approches, ces méthodes hédoniques sont normalement plus gourmandes en données et requièrent souvent plus d'informations pour construire un indice à qualité constante du prix du logement. En effet, lorsqu'on applique une méthode hédonique, il faut connaître pour chaque observation utilisée pour la régression des informations sur les caractéristiques pertinentes (structurelles et d'environnement) en plus des données de prix. En principe, plus les données sur les caractéristiques sont détaillées et plus l'échantillon de logement est large, plus l'indice des prix en résultant sera fiable et précis[15].

15 Bien que dans les ouvrages spécialisés la plupart des régressions hédoniques sur les prix du logement utilisent souvent bien plus de variables explicatives, quelques études et les exemples du chapitre 5 montrent que des indices de prix hédoniques fiables peuvent être obtenus avec seulement quatre variables indépendantes.

Tableau 11.6. Résultats du modèle log-linéaire pour un exemple simple

Source |	SS	df	MS			Nombre d'obs. =	796
						F(4, 791) =	156,02
Modèle |	20,0634692	4	5,0158673			Prob > F =	0,0000
Résidu |	25,4293063	791	0,032148301			R^2 =	0,4410
						R^2 ajusté =	0,4382
Total |	45,4927755	795	0,057223617			Racine carrée MSE =	0,1793

lprix |	Coeff.	Erreur-type	t	P>|t|	[Intervalle de conf. 95 %]	
chambres |	0,1156791	0,0098159	11,78	0,000	0,0964108	0,1349473
salles de bain |	0,0999522	0,0095996	10,41	0,000	0,0811086	0,1187958
ancienneté |	-0,002561	0,0004173	-6,14	0,000	-0,0033801	-0,001742
terrain |	9,39e-06	1,28e-06	7,31	0,000	6,87e-06	0,0000119
const. |	12,0647	0,0383342	314,72	0,000	11,98945	12,13995

Source : calculs des auteurs basés sur les données MLS® pour une ville canadienne.

11.27 Un modèle hédonique exprime le prix d'un bien comme une fonction des caractéristiques (ou attributs) qui le déterminent. Le chapitre 5 présente deux formes fonctionnelles couramment utilisées : le modèle linéaire et le modèle logarithmique-linéaire (semi-log). Bien que les ouvrages en la matière évoquent souvent d'autres méthodes (voir par exemple la méthode Box-Cox), elles ne seront pas traitées ici. La forme semi-log est commode, car l'interprétation des coefficients de régression est directe : une fois multiplié par 100, un coefficient fournit, en pourcentage, la variation du prix du logement qui résulte d'une modification d'une unité de la variable explicative qui lui est associée.

11.28 Pour illustrer aussi simplement que possible comment les différents indices de prix hédoniques du logement sont construits, la version intégrale de la série de données utilisée pour calculer ci-dessus les moyennes et les médianes sera reprise dans les exemples suivants. Pour simplifier la présentation, le nombre des caractéristiques déterminant le prix sera limité à quatre variables (continues) : la superficie du terrain (land), le nombre de chambres (rooms), le nombre de salles de bains (bath) et l'ancienneté (age). Le résultat initial de la régression MCO avec un modèle semi-log pour une seule année (2008) est présenté dans le tableau 11.6.

11.29 À partir de la régression sur un échantillon de 796 observations de prix, on trouve que toutes les variables explicatives ont un effet du signe attendu et qu'elles sont toutes significativement non nulles (avec le test de Student). Le R^2 ajusté (ou coefficient de corrélation) est de 44 %, c'est-à-dire que les variations de la surface du terrain, de l'ancienneté de l'habitation, du nombre de chambres et de salles de bain explique 44 % de la variance du prix des logements. Avec plus de variables explicatrices dans la régression, le R^2 devrait augmenter. En fait, avec trois variables indépendantes additionnelles (présence d'une cheminée, d'un garage et l'ancienneté au carré pour tenir compte de la non-linéarité associée à cette variable), le R^2 ajusté augmente à 54 %.

11.30 Les résultats des régressions peuvent s'interpréter comme suit :

- Un pied carré de terrain en plus augmente le prix du logement de 0,000939 %, toutes choses étant égales par ailleurs.
- Chaque chambre en plus augmente le prix du logement de 11,6 %, toutes choses étant égales par ailleurs.
- Un logement avec une salle de bain supplémentaire coûte presque 10 % plus cher qu'une maison sans salle de bain supplémentaire, toutes choses étant égales par ailleurs.
- Une année d'ancienneté de plus fait baisser le prix d'un logement (ou le logement se déprécie) de 0,2 %, toutes choses étant égales par ailleurs.

«Toutes choses étant égales par ailleurs» signifie que toutes les autres variables que celles étudiées sont supposées être constantes. À propos de la variable «nombre de chambres», par exemple, on ne peut pas dire qu'une maison avec plus de chambres coûte toujours plus, car d'autres facteurs peuvent aussi influencer le prix de la maison, comme son emplacement, son ancienneté et la qualité générale de la construction. Dans ce cas, «toutes choses étant égales par ailleurs» signifie que si deux maisons ne diffèrent que par le nombre de chambres (c'est-à-dire qu'elles se valent par ailleurs), celle qui en a le plus coûtera le plus cher.

11.31 Ce qui suit sont des exemples simplifiés des différentes méthodes, décrites au chapitre 5, pour calculer des indices de prix hédoniques. La méthode des variables indicatrices temporelles est présentée en premier. Tous les exemples utilisent la régression des MCO.

La méthode des variables indicatrices temporelles

11.32 La méthode des variables indicatrices temporelles est fondée sur une régression hédonique semi-linéaire qui prend

en compte ensemble toutes les périodes. Le modèle est donné par l'équation (6.5) et il est répété ici pour plus de facilité :

$$\ln p_n^t = \beta_0 + \sum_{\tau=1}^{T} \delta^\tau D_n^\tau + \sum_{k=1}^{K} \beta_k z_{nk}^t + \varepsilon_n^t \qquad (11.10)$$

où D_n^τ représente l'indicatrice qui est égale à un si l'observation provient de la période τ ($\tau = 1,...,T$) et sinon zéro. La variable indicatrice temporelle pour la période de référence 0 — soit la période initiale à partir de laquelle les évolutions de prix vont être mesurées — est laissée de côté pour éviter que la constante β_0 soit la parfaite combinaison linéaire de toutes les indicatrices avec la constante, ce qui connu comme «le piège des variables indicatrices». Avec la méthode des variables indicatrices temporelles, la période de référence et la période de comparaison, $t = 1,...,T$, se réfèrent à la même durée, c'est-à-dire un mois, un trimestre ou une année, selon le contexte particulier, comme les préférences des utilisateurs ou la disponibilité des données.

11.33 L'exponentielle (ou l'anti-logarithme) du coefficient estimé par la régression $\hat{\delta}^\tau$ mesure la variation en pourcentage «à qualité constante» du prix des propriétés entre la période de référence et la période t. Les étapes suivantes montrent pourquoi $(\hat{\delta}^\tau)$ est une mesure d'une évolution pure de prix, ajustés de la qualité. L'estimation du logarithme du prix à la période 0 pour la propriété i, avec les caractéristiques de la période de référence, z_{nk}^0 $(k = 1,...,K)$, est :

$$\ln \hat{p}_n^0 = \hat{\beta}_0 + \sum_{k=1}^{K} \hat{\beta}_k z_{nk}^0 \qquad (11.11)$$

À la période 1, le logarithme du prix doit être estimé avec les *caractéristiques de la période de référence* pour conserver une qualité constante, soit

$$\ln \hat{p}_n^{1*} = \hat{\beta}_0 + \hat{\delta}^1 + \sum_{k=1}^{K} \hat{\beta}_k z_{nk}^0 \qquad (11.12)$$

Prenons la différence entre les estimations de ces deux périodes,

$$\ln \hat{p}_n^{1*} - \ln \hat{p}_n^0 = \ln(\hat{p}_n^{1*} / \hat{p}_n^0) = \hat{\delta}^1 \qquad (11.13)$$

L'expression (11.13) ne dépend pas de n. Ce résultat est valable pour toutes les maisons de l'échantillon. Comme il est montré par Berndt (1991), l'estimation de δ^t peut être interprétée comme la variation du logarithme du prix due au passage du temps, les autres variables étant constantes. L'exponentielle de δ^1 donne l'estimation de l'indice des prix de la période 1 :

$$P_{TD}^{01} = \exp(\hat{\delta}^1) \qquad (11.14)$$

La même démonstration peut être menée pour toutes les autres périodes. Ainsi, l'indice des prix par la méthode des variables indicatrices temporelles, de la période de référence à la période de comparaison est :

$$P_{TD}^{0t} = \exp(\hat{\delta}^t) \qquad (11.15)$$

Manifestement, l'indice des prix pour la période de référence obtenu par la méthode des variables indicatrices temporelles est égal à 1.

11.34 L'exemple suivant illustre les étapes pour calculer un indice des prix par la méthode des variables indicatrices temporelles. Supposons que les informations détaillées sur les ventes de maisons pendant deux années (t = 2006 a t = 2007) sont disponibles. En utilisant les mêmes informations que dans l'exemple simplifié ci-dessus, les données pour toutes les périodes sont introduites dans la même régression groupée :

$$\ln p_n^t = \beta_0 + \beta_1 Taille\ du\ terrain_n + \beta_2 Chambre_n + \beta_3 Bain_n + \beta_4 Ancienneté_n + \delta^1 D_n^1 + \varepsilon_n^t \qquad (11.16)$$

Le côté gauche de l'équation (11.16) est le logarithme du prix de la maison n pour les années t (2006 or 2007) comme variable dépendante. Le côté droit de l'équation a les mêmes variables explicatives (sauf la variable indicatrice temporelle) que l'on trouverait dans une régression hédonique pour une période. Dans ce cas particulier, les variables explicatives sont : la taille du terrain (lot size), le nombre de chambres (bedroom), le nombre de salles de bain (bathroom) et l'ancienneté (age); les paramètres respectifs vont de β_1 à β_4. Comme il s'agit d'une régression regroupant deux années, les paramètres estimés (ou les coefficients de la régression) seront contraints pour les années où les données sont utilisées dans la régression. Le terme d'erreur ε_n^t indique si une valeur observée est au-dessus ou au-dessous de la droite de régression. Aussi du côté droit de l'équation, on trouve la constante, β_0.

11.35 Les résultats de la régression utilisant l'ensemble des données sont fournis dans le tableau 11.7. Le coefficient d'intérêt est celui associé à l'année 2007, $\hat{\delta}^{07}$. Il est de 0,0781548. Ce coefficient est ensuite transformé pour estimer l'indice des prix (ou la variation en pourcentage des prix) pour les maisons entre les années 2006 et 2007. Cette transformation consiste à prendre l'exponentielle du coefficient $\hat{\delta}^{07}$: $P_{TD}^{07/06} = \exp(0{,}0781548) = 1{,}08129$. Ainsi, en gardant constantes toutes les caractéristiques, la variation en pourcentage du prix des maisons entre 2006 et 2007 est de 8,1 %. Notons que la moyenne et la médiane ont conduit à des augmentations respectives de 10,1 % et de 9,2 %, pour la même période.

11.36 En ajoutant une troisième période (année 2008), l'équation de la régression hédonique devient :

$$\ln p_n^t = \beta_0 + \beta_1 Taille\ du\ terrain_n + \beta_2 Chambre_n + \beta_3 Bain_n + \beta_4 Ancienneté_n + \delta^1 D_n^1 + \delta^2 D_n^2 + \varepsilon_n^t \qquad (11.17)$$

Le tableau 11.8 contient les résultats de la régression. La valeur du coefficient de l'indicatrice temporelle pour l'année 2008 est de 0,1332734. Son exponentielle $e^{0{,}1332734}$ = 1,14 montre une augmentation à qualité constante de l'indice du prix des maisons de 14 % entre 2006, l'année de référence, et 2008, l'année la plus récente. En revanche, les augmentations des

prix données par la moyenne et la médiane pour la même période étaient respectivement de 16 % et de 17 %.

11.37 Cette technique peut être généralisée à plus que trois périodes, au fur et à mesure que d'autres périodes deviennent disponibles. Cela consiste à combiner les données de plus de périodes et à ajouter plus de variables indicatrices temporelles. Cependant, des régressions groupées avec plusieurs périodes ne sont pas nécessairement idéales pour construire des séries chronologiques, car l'ajout d'une nouvelle période modifiera vraisemblablement les résultats obtenus pour les périodes antérieures. En effet, dans l'exemple précédent, quand l'année 2008 est ajoutée à la régression groupée antérieure, le coefficient de l'année 2007 devient 0,0781257, alors qu'il était de 0,0781548 (tableau 11.7). Dans cet exemple particulier, la modification du coefficient affecté à l'année 2007 est légère. Néanmoins, la stabilité des coefficients d'une régression groupée peut devenir problématique quand le nombre de périodes augmente.

11.38 Une autre approche mentionnée dans le chapitre 5 est d'utiliser la méthode des variables indicatrices temporelles entre deux périodes consécutives. Si la régression hédonique est fondée sur deux périodes consécutives τ et $\tau + 1$, l'équation devient :

$$\ln p_n^t = \beta_0 + \delta^{\tau+1} D_n^{\tau+1} + \sum_{k=1}^{K} \beta_k z_{nk}^t + \varepsilon_n^t \qquad (11.18)$$

Dans le contexte des trois périodes de données utilisées dans les exemples précédents, une première régression hédonique est calculée pour les périodes consécutives 0 et 1 et ensuite une seconde régression est estimée pour les périodes 1 et 2 en utilisant les quatre variables caractéristiques. Les résultats de la première régression sont évidemment identiques à

Tableau 11.7. Résultats de la régression groupée pour les années 2006 et 2007

Source |	SS	df	MS			Nombre d'obs. =	1.708
						F(5, 1702) =	286,64
Modèle |	48,4501865	5	9,6900373			Prob > F =	0,0000
Résidu |	57,5372376	1.702	0,033805663			R^2 =	0,4571
						R^2 ajusté =	0,4555
Total |	105,987424	1.707	0,062089879			Racine carrée MSE =	0,18386

lprix |	Coeff.	Erreur-type	t	P>|t|	[Intervalle de conf. 95 %]	
chambres |	0,0840483	0,0069071	12,17	0,000	0,0705009	0,0975957
salles de bain |	0,121815	0,0071529	17,03	0,000	0,1077855	0,1358444
ancienneté |	-0,0029137	0,0003183	-9,15	0,000	-0,0035381	-0,0022894
terrain |	0,0000137	9,24e-07	14,78	0,000	0,0000119	0,0000155
d2007 |	0,0781548	0,0089128	8,77	0,000	0,0606736	0,095636
const. |	11,96531	0,0273032	438,24	0,000	11,91176	12,01886

Source : calculs des auteurs basés sur les données MLS® pour une ville canadienne.

Tableau 11.8. Résultats de la régression groupée pour les années 2006 à 2008

Source |	SS	df	MS			Nombre d'obs. =	2.504
						F(6, 2497) =	366,64
Modèle |	73,4886776	6	12,2481129			Prob > F =	0,0000
Résidu |	83,4154327	2.497	0,033406261			R^2 =	0,4684
						R^2 ajusté =	0,4671
Total |	156,90411	2.503	0,06268642			Racine carrée MSE =	0,18277

lprix |	Coef.	Erreur-type	t	P>|t|	[Intervalle de conf. 95 %]	
chambres |	0,0942001	0,0056566	16,65	0,000	0,083108	0,1052923
salles de bain |	0,1139931	0,0057443	19,84	0,000	0,102729	0,1252572
ancienneté |	-0,0028112	0,0002538	-11,08	0,000	-0,0033089	-0,0023135
terrain |	0,0000122	7,51e-07	16,28	0,000	0,0000108	0,0000137
d2007 |	0,0781257	0,008856	8,82	0,000	0,0607598	0,0954916
d2008 |	0,1332734	0,0090681	14,70	0,000	0,1154916	0,1510552
const. |	11,95724	0,0225891	529,34	0,000	11,91295	12,00154

Source : calculs des auteurs basés sur les données MLS® pour une ville canadienne.

Tableau 11.9. Résultats de la régression groupée pour les années 2007 et 2008

Source |	SS	df	MS			Nombre d'obs. =	1.670
						F(5, 1664) =	271,91
Modèle |	45,441478	5	9,0882956			Prob > F =	0,0000
Résidu |	55,6172267	1.664	0,033423814			R^2 =	0,4497
						R^2 ajusté =	0,4480
Total |	101,058705	1.669	0,060550452			Racine carrée MSE =	0,18282

lprix |	Coeff.	Erreur-type	t	P>|t|	[Intervalle de conf. 95 %]	
chambres |	0,1041401	0,0068861	15,12	0,000	0,0906337	0,1176465
salles de bain |	0,1070142	0,0068881	15,54	0,000	0,093504	0,1205244
ancienneté |	0,0026926	0,0003045	-8,84	0,000	-0,0032899	-0,0020953
terrain |	0,0000117	9,42e-07	12,42	0,000	9,85e-06	0,0000135
d2008 |	0,0555370	0,0089625	6,20	0,000	0,073116	0,037958
const. |	12,07482	0,026871	449,36	0,000	12,02212	12,12753

Source : calculs des auteurs basés sur les données MLS® pour une ville canadienne.

ceux présentés dans le tableau 11.7 et l'indice des prix en résultant est estimé à 108,1. Le tableau 11.9 montre le résultat de la régression pour les années consécutives 2007 et 2008.

11.39 L'indice des prix à qualité constante est calculé comme l'exponentielle du coefficient de l'année 2008, (0,0555370). L'indice est ainsi : exp(0,0555370) = 1,057. Souvenons-nous que c'est l'évolution des prix depuis 2007, et non pas celle depuis l'année de référence 2006. À partir de ces résultats, il est possible de construire une série chronologique par *chaînage* des indices de deux périodes consécutives (en démarrant à 1 pour la période de référence) : $P_{TD}^{07/06} = 1{,}081$; $P_{TD,chain}^{08/06} = 1{,}081 \times 1{,}057 = 1{,}143$ Ce résultat est seulement légèrement différent de celui obtenu avec la régression groupée (voir le tableau 11.8), où la variation des prix était estimée à 14,0 % sur l'ensemble de la période. Maintenant, en chaînant les indices des variables indicatrices temporelles sur les deux périodes temporelles consécutives, la variation estimée des prix est de 14,3 %.

Prix des caractéristiques ou méthode d'imputation

11.40 La méthode hédonique suivante, présentée dans le chapitre 5, est celle du prix des caractéristiques ou méthode d'imputation, dorénavant appelée simplement la méthode du prix des caractéristiques. En appliquant cette méthode aux mêmes données utilisées précédemment, un indice des prix à qualité constante est estimé. Par facilité de présentation, on fournira les résultats de l'estimation utilisant le modèle *linéaire* pour la régression[16].

11.41 La méthode du prix des caractéristiques utilise les prix implicites des caractéristiques dans le modèle (les coefficients de la régression) comme point de départ pour construire l'indice des prix, selon une formule classique d'indice, mais dans laquelle les coefficients de la régression sont les prix et les quantités sont, pour chaque caractéristique, le nombre d'unités mesurées. Ainsi, la régression hédonique est estimée séparément pour chaque période. Les modèles linéaires pour la période de référence 0 (2006) et pour la période 1 (2007) sont :

$$p_n^0 = \beta_0^0 + \beta_1^0 Taille\ du\ terrain_n + \beta_2^0 Chambre_n + \beta_3^0 Bain_n + \beta_4^0 Ancienneté_n + \varepsilon_n^0 \quad (11.19)$$

$$p_n^1 = \beta_0^1 + \beta_1^1 Taille\ du\ terrain_n + \beta_2^1 Chambre_n + \beta_3^1 Bain_n + \beta_4^1 Ancienneté_n + \varepsilon_n^1 \quad (11.20)$$

11.42 L'estimation de ces équations sur l'échantillon de données respectivement pour 2006 et 2007 par les MCO donne les résultats présentés dans les tableaux 11.10 et 11.11. Dans cet exemple, le prix implicite d'une chambre supplémentaire en 2006 est de 24.329 dollars tandis que chaque salle de bain supplémentaire ajoutera 43.190 dollars au prix de la maison. Dans cet exemple hautement simplifié, les résultats pour 2007 sont bien entendu différents de ceux pour 2006; En 2007, une chambre supplémentaire semble augmenter le prix de 35.147 dollars, tandis que le prix pour une salle de bain supplémentaire est alors estimé à 43.463 dollars[17].

16 Il n'y a rien qui s'oppose à l'utilisation des modèles semi-log ou logarithmiques. Les deux peuvent être utilisés avec cette méthode hédonique.

17 Notons que les coefficients pour le nombre de chambres sont quelque peu volatils entre les deux années. Cela est attendu, car les régressions hédoniques sont souvent caractérisées par la colinéarité entre ces deux variables explicatives. Il faut souligner que la colinéarité en elle-même n'affecte pas l'exactitude de l'indice d'ensemble. Elle n'est un problème que si une valeur monétaire précise d'une chambre et/ou d'une salle de bain est nécessaire, comme ce serait le cas pour estimer la valeur d'une propriété. Il faut aussi ajouter que, pour cet exercice simplifié, l'échantillon est relativement petit. Ceci explique aussi pourquoi, parfois, les résultats ne sont pas tout à fait aussi robustes qu'avec des échantillons importants.

Tableau 11.10. Résultats de la régression pour 2006

Source |	SS	df	MS				
Modèle |	2,4182e+12	4	6,0454e+11			Nombre d'obs. =	834
Résidu |	3,5420e+12	829	4,2726e+09			F(4, 829) =	141,49
						Prob > F =	0,0000
Total |	5,9601e+12	833	7,1550e+09			R^2 =	0,4057
						R^2 ajusté =	0,4029
						Racine carrée MSE =	65365
prix |	Coeff.	Erreur-type		t	P>|t|	[Intervalle de conf. 95 %]	
chambres |	24329,78	3557,79		6,84	0,000	17346,45	31313,12
salles de bain |	43190,01	3734,288		11,57	0,000	35860,24	50519,79
ancienneté |	-1083,309	164,5957		-6,58	0,000	-1406,382	-760,2357
terrain |	5,168582	0,4474175		11,55	0,000	4,290378	6,046787
const. |	98333,45	14450,86		6,80	0,000	69968,88	126698

Source : calculs des auteurs basés sur les données MLS® pour une ville canadienne.

Tableau 11.11. Résultats de la régression pour 2007

Source |	SS	df	MS				
Modèle |	3,5694e+12	4	8,9236e+11			Nombre d'obs. =	874
Résidu |	4,5702e+12	869	5,2592e+09			F(4, 869) =	169,68
						Prob > F =	0,0000
Total |	8,1397e+12	873	9,3238e+09			R^2 =	0,4385
						R^2 ajusté =	0,4359
						Racine carrée MSE =	72520
prix |	Coeff.	Erreur-type		t	P>|t|	[Intervalle de conf. 95 %]	
chambres |	35147,31	3777,91		9,30	0,000	27732,41	42562,2
salles de bain |	43463,76	3858,683		11,26	0,000	35890,33	51037,19
ancienneté |	-1059,767	173,0922		-6,12	0,000	-1399,495	-720,0394
terrain |	5,829323	0,5388036		10,82	0,000	4,771814	6,886831
const. |	79248,85	14337,87		5,53	0,000	51107,95	107389,7

Source : calculs des auteurs basés sur les données MLS® pour une ville canadienne.

11.43 L'étape suivante est de calculer un indice hédonique des prix à partir des résultats de la régression. Un indice pour 2007, comparé à 2006, peut s'exprimer par :

$$P^{01} = \frac{\hat{\beta}_0^1 + \hat{\beta}_1^1\bar{z}_1^0 + \hat{\beta}_2^1\bar{z}_2^0 + \hat{\beta}_3^1\bar{z}_3^0 + \hat{\beta}_4^1\bar{z}_4^0}{\hat{\beta}_0^0 + \hat{\beta}_1^0\bar{z}_1^0 + \hat{\beta}_2^0\bar{z}_2^0 + \hat{\beta}_3^1\bar{z}_3^0 + \hat{\beta}_4^1\bar{z}_4^0} = \frac{\sum_{k=0}^{K}\hat{\beta}_k^1\bar{z}_k^0}{\sum_{k=0}^{K}\hat{\beta}_k^0\bar{z}_k^0} \quad (11.21)$$

Où $\bar{z}_k^0$ est la moyenne de l'échantillon des valeurs de la k-^{e} caractéristique pour la période de référence; on prend aussi $\bar{z}_0^0 = 1$. Les statisticiens qui établissent des indices de prix reconnaîtront que l'indice de la formule (11.21) est un indice de type Laspeyres : le prix estimé des caractéristiques pour la période 0 (2006) et la période 1 (2007), $\hat{\beta}_k^0$ et $\hat{\beta}_k^1$, sont pondérés par la moyenne des quantités de caractéristiques pour la période de référence. Autrement dit, la moyenne pour la période de référence des quantités pour chaque caractéristique est valorisée par le prix implicite de la caractéristique pour la période de référence et la période courante. Le tableau 11.12 présente les moyennes simples de l'échantillon des valeurs des caractéristiques de cet exemple. Avec ces valeurs et les coefficients des tableaux 11.10 et 11.11, l'indice hédonique du type Laspeyres entre l'année de référence (2006) et 2007 est calculé par :

$$P^{07/06} = \frac{79248 + (35147\times 3{,}63) + (43463\times 2{,}76) + }{98333 + (24329\times 3{,}63) + (43190\times 2{,}76) + }$$
$$\frac{+(-1059\times 23{,}89) + (5{,}829323\times 6719)}{+(-1083\times 23{,}89) + (5{,}168582\times 6719)} = 1{,}082$$

L'augmentation de 8,2 % des prix ainsi calculée est assez proche, dans ce cas particulier, des 8,1 % obtenus en utilisant l'approche par les variables indicatrices temporelles du tableau 11.7.

11.44 Pour les périodes suivantes, le statisticien a une décision à prendre. Il ou elle peut utiliser la même base annuelle, en quantité, pour calculer les indices suivants en utilisant la formule de Laspeyres, mais en remplaçant les prix implicites du numérateur par les prix pertinents.

Tableau 11.12. Valeurs moyennes des caractéristiques pour la période de référence (2006)

	Moyenne	Erreur-type	[Intervalle de conf. 95 %]	
chambres	3,633094	0,0244034	3,585194	3,680993
salles de bain	2,767386	0,0269044	2,714578	2,820195
ancienneté	23,88969	0,5693338	22,77219	25,00719
terrain	6719,492	184,8605	6356,644	7082,339

Source : calculs des auteurs basés sur les données MLS® pour une ville canadienne.

Autrement, les quantités (moyenne des caractéristiques) de la précédente période peuvent être utilisées pour générer des indices basés sur des périodes consécutives. Ces indices bilatéraux sont ensuite chaînés pour créer une série chronologique, un indice chaîne. D'autres options sont possibles et elles sont présentées dans le chapitre 5, mais les mécanismes de construction des indices restent essentiellement les mêmes que ceux présentés ici.

La méthode des ventes répétées

11.45 Le problème le plus important dans l'utilisation des moyennes ou des médianes (non stratifiées) de prix de transaction pour mesurer l'évolution du prix des maisons est que la modification d'une période à l'autre de la composition de l'échantillon des propriétés vendues n'est pas correctement prise en compte. Ce problème peut être partiellement contourné en construisant un IPIR selon la méthode des ventes répétées (voir chapitre 6). En fait, un indice très répandu et suivi de près du prix des logements aux États-Unis, l'indice de Case–Shiller, est fondé sur la méthode des ventes répétées.

11.46 La stratégie pour construire un indice des ventes répétées est assez simple. Elle consiste à comparer les prix d'une même propriété vendue à deux moments différents. En d'autres termes, elle utilise la méthode du panier fixe pour sélectionner l'échantillon qui servira à calculer l'indice. Pour que la méthode des ventes répétées soit facile à mettre en œuvre, il faut avoir accès à une importante base de données couvrant une assez longue période. Pour les autres aspects, le besoin d'informations est relativement modeste; avec la méthode des ventes répétées, en plus du prix de vente et de la date de la vente, il ne faut que l'adresse du logement (ou une autre manière d'identifier l'emplacement)[18].

11.47 Un exemple simple peut illustrer l'application de la méthode des ventes répétées[19].On suppose que l'objectif est d'estimer un indice annuel de la variation des prix entre 2008 et 2010; le tableau 11.13 fournit des données sur quelques transactions. La propriété A vendue en 2008 pour 100.000 dollars et revendue en 2009 pour 120.000 dollars; la propriété B vendue en 2008 pour 175.000 dollars et revendue en 2010 pour 220.000 dollars; la propriété C vendue en 2009 pour 180.000 dollars et revendue en 2010 au même prix.

Dans une première étape, la variation de prix entre 2008 et 2010 est estimée en utilisant la moyenne des prix. Les moyennes annuelles des prix entre 2008 et 2010 sont respectivement 137.000 dollars, 150.000 dollars et 200.000 dollars. Les variations en glissement annuel correspondant à ces moyennes sont 9,1 % et 33,3 % pour les périodes 2009/2008 et 2010/2009.

11.48 Ces résultats sont maintenant comparés avec ceux obtenus en utilisant la méthode des ventes répétées. Appelons *P* le rapport du prix de la maison entre la deuxième et la

18 Une hypothèse est que la qualité de la maison n'a pas changé au cours de la période entre les deux ventes. Si des informations sur les caractéristiques de la propriété sont accessibles au statisticien, alors il sera possible d'exclure des calculs celles qui ont notablement changé au cours de la période considérée et sont susceptibles d'affecter le prix et donc de biaiser l'indice. De plus, puisque la multiplication des reventes rapides d'une propriété indique souvent la présence de caractéristiques indésirables, de telles propriétés peuvent aussi être exclues des calculs. Il faut également signaler qu'un indice des ventes répétées n'est pas strictement un indice à qualité constante, car les maisons sont souvent sujettes à une dépréciation temporelle. En conséquence, un indice des prix des ventes répétées sous-estime généralement la vraie inflation du prix des maisons, à moins d'effectuer des ajustements correctifs pendant l'estimation. Si l'objectif d'un indice est d'être un indicateur de court ou de moyen terme du prix des maisons, alors le problème de la dépréciation qui accompagne la méthode des ventes répétées pourra éventuellement être négligé.

19 L'exemple provient en partie de la documentation de l'indice des prix canadien des ventes répétées de Teranet–Banque Nationale® : http://www.housepriceindex.ca/Default.aspx.

Tableau 11.13. Données de ventes répétées

	2008	2009	2010
Propriété A	100.000 $	120.000 $	Pas de vente
Propriété B	75.000 $	Pas de vente	220.000 $
Propriété C	Pas de vente	180.000 $	180.000 $
Moyenne	*$137.500 $*	*150.000 $*	*200.000 $*

première vente pour chaque transaction[20] définitive entre 2008 et 2010. Le logarithme de *P* sera la *variable dépendante* de la régression des ventes répétées. Trois ventes répétées sont identifiées dans le tableau 11.13 entre 2008 et 2010. Pour la première vente répétée (propriété A), *P* vaut 1,200, soit le rapport de prix entre les ventes de 2009 et 2008; pour la deuxième vente répétée (propriété B), *P* vaut 1,257, soit le rapport de prix entre les ventes de 2010 et 2008; pour la troisième vente répétée (propriété C), *P* vaut 1, car le prix de cette propriété n'a pas changé entre 2009 et 2010.

11.49 Les *variables indépendantes* pour une régression des ventes répétées sont des indicatrices, qui valent –1 pour l'année de la première vente, +1 pendant l'année de la deuxième fois vente et 0 pour les autres années. Les coefficients des variables indicatrices, estimés par la régression, servent à calculer l'indice des ventes répétées. Le tableau 11.14 présente les valeurs prises par les variables indicatrices pour les propriétés A à C. Par exemple, comme la propriété A est vendue une deuxième fois en 2009, la variable indicatrice D2009 prend la valeur 1, mais D2010 prend la valeur 0, car la propriété A n'est plus vendue après 2009. Un raisonnement similaire est appliqué aux autres propriétés et aux autres années. Notons que, pour éviter une parfaite colinéarité, la première période (2008) est exclue des variables explicatives et de la régression. En d'autres termes, si la première vente a lieu l'année de référence, il n'y a pas de variable indicatrice pour cette période.

11.50 Avec ces données de ventes répétées, l'équation de régression — qui n'a pas de constante — peut être exprimée par (voir aussi l'équation (6.3)) :

$$\ln P_n^t = \gamma^{2009} D_n^{2009} + \gamma^{2010} D_n^{2010} + \varepsilon_n^t \qquad (11.22)$$

où ε_n^t est le terme d'erreur («bruit blanc»). Les exponentielles des paramètres estimés, soit $\exp(\hat{\gamma}^{2009})$ et $\exp(\hat{\gamma}^{2010})$, représentent les indices des prix du logement pour chaque période comparée à l'année de référence 2008. En utilisant les moindres carrés ordinaires (MCO) pour estimer l'équation (11.22) à partir des données du tableau 11.14, les indices des prix des ventes répétées sont respectivement 1,219 et 1,238 pour 2009 et 2010. Les glissements annuels de cet exemple, 21,9 % et 23,8 %, sont plutôt différents de ceux obtenus avec la méthode fondée sur la moyenne simple, qui étaient de 9,1 % et de 33,3 %[21].

20 Geltner et Pollakowski (2006) utilisent le terme «aller et retour».
21 Il n'y a que très peu d'observations, donc aucune conclusion définitive ne doit être tirée de cet exemple simplifié. Il ne devrait être utilisé qu'à titre pédagogique.

11.51 Le modèle simple des ventes répétées peut être amélioré. Une façon de le faire serait de réduire la variance de l'indice estimé. Comme Geltner et Pollakowski (2006) l'ont indiqué, la source de la variance (du bruit) de l'indice estimé des prix des propriétés réside dans le fait que les prix des transactions sont distribués aléatoirement autour de la «véritable» mais inobservable valeur sur le marché. Les auteurs ajoutent que ce bruit existe pour tous les indices de prix du logement, quelle que soit la méthode d'estimation de l'indice. Pour atténuer la variance, l'échantillon des ventes répétées pourra être élargi, si les données sont disponibles.

11.52 Comme il l'a été montré précédemment, une régression par les MCO peut être utilisée pour obtenir les variations de prix. Le modèle de Bailey, Muth et Nourse (1963) est un exemple classique des MCO appliqués aux ventes répétées en utilisant la méthode décrite ci-dessus. Cependant, des recherches ultérieures ont montré que la méthode de base des ventes répétées avec régression par les MCO peut être améliorée en utilisant une régression par les moindres carrés pondérés (MCP). En un mot, cette méthode consiste à donner plus de poids dans la régression aux observations considérées comme étant plus précises. Dans le contexte de la méthode des ventes répétées, on peut donner moins de poids aux observations des propriétés dont les ventes sont les plus espacées dans le temps et *vice versa*. Cette méthode corrigera le problème intrinsèque mieux connu sous le nom d'hétéroscédasticité.

11.53 Case et Shiller (1987) suggèrent l'approche suivante en trois étapes :

1. Estimer le modèle (11.22) avec la régression par les MCO et récupérer le vecteur des résidus de la régression.

2. Régresser ces résidus, mis au carré, par les MCO sur les intervalles de temps entre les ventes, en incluant une constante dans le modèle.

3. Estimer avec une régression par les MCO le modèle (11.22), dans lequel chaque observation est divisée par la racine carrée de la valeur estimée par la deuxième étape.

La troisième étape est une régression par les moindres carrés pondérés du modèle (11.22) qui tient compte de l'hypothèse d'hétéroscédasticité.

11.54 Deux modèles des ventes répétées sont à présent simulés en adoptant l'ensemble de données sur les maisons individuelles plus grand et plus réaliste retenu pour la

Tableau 11.14. Variables indicatrices des ventes répétées

	P	D2009	D2010
Propriété A	1,200	1	0
Propriété B	1,257	0	1
Propriété C	1,000	-1	1

Tableau 11.15. Régression non pondérée des ventes répétées

Source |	SS	df	MS				
Modèle |	32,5127473	6	5,41879122			Nombre d'obs.	= 1.186
Résidu |	16,8531146	1.180	0,014282301			F(6, 1180)	= 379,41
						Prob > F	= 0,0000
Total |	49,365862	1.186	0,04162383			R^2	= 0,6586
						R^2 ajusté	= 0,6569
						Racine carrée MSE	= 0,11951
diflnprix |	Coeff.		Erreur-type	t	P>|t|	[Intervalle de conf. 95 %]	
dy2003 |	0,0613539		0,0086332	7,11	0,000	0,0444157	0,0782921
dy2004 |	0,1198942		0,0082047	14,61	0,000	0,1037969	0,1359915
dy2005 |	0,1431862		0,008343	17,16	0,000	0,1268173	0,159555
dy2006 |	0,1845885		0,0084578	21,82	0,000	0,1679945	0,2011826
dy2007 |	0,2658241		0,0083474	31,85	0,000	0,2494468	0,2822015
dy2008 |	0,3438869		0,0087587	39,26	0,000	0,3267025	0,3610713

Source : calculs des auteurs basés sur les données MLS® pour une ville canadienne.

Tableau 11.16. Régression pondérée des ventes répétées

Source |	SS	df	MS				
Modèle |	2098,21619	6	349,702699			Nombre d'obs.	= 1.186
Résidu |	1182,72363	1.180	1,00230816			F(6, 1180)	= 348.90
						Prob > F	= 0,0000
Total |	3280,93982	1.186	2,76639108			R^2	= 0,6395
						R^2 ajusté	= 0,6377
						Racine carrée MSE	= 1,0012
ndifprix |	Coeff.		Erreur-type	t	P>|t|	[Intervalle de conf. 95 %]	
ndy2003 |	0,0635307		0,0085609	7,42	0,000	0,0467345	0,0803269
ndy2004 |	0,1211754		0,0081162	14,93	0,000	0,1052516	0,1370992
ndy2005 |	0,1437457		0,0082962	17,33	0,000	0,1274688	0,1600226
ndy2006 |	0,1864151		0,0084621	22,03	0,000	0,1698127	0,2030175
ndy2007 |	0,2689894		0,0084844	31,70	0,000	0,2523433	0,2856356
ndy2008 |	0,3491619		0,0091085	38,33	0,000	0,3312913	0,3670325

Source : calculs des auteurs basés sur les données MLS® pour une ville canadienne.

plupart des exemples précédents de ce chapitre. Les résultats sont d'abord utilisés dans la régression non pondérée des ventes répétées et sont présentés dans le tableau 11.15. Le tableau 11.16 présente les résultats de la version pondérée de la régression des ventes répétées. Notons que pour cet ensemble particulier de données, tous les coefficients sont significativement différents de 0 et qu'aucune constante n'est utilisée pour les régressions de la méthode des ventes répétées. Un inconvénient souvent cité de la méthode des ventes répétées est le gaspillage de données. Cet exemple le confirme. Sur les 5.787 observations de la base de données initiale, seulement 1.186 (environ 20 %) correspondent à des propriétés vendues plus d'une fois en 6 ans environ.

11.55 Comme avec le modèle hédonique à variables indicatrices temporelles présenté précédemment, l'indice des prix correspondant est obtenu avec l'exponentielle du coefficient estimé, car la variable dépendante est le logarithme du prix. Par exemple, la régression pour la méthode non pondérée des ventes répétées donne pour 2007 un coefficient de 0,2658241; En prenant son exponentielle, on obtient exp(0,2658241) = 1,3045 (ou 130,5 une fois arrondi et multiplié par 100). Les indices de toute la période 2002–08 sont disponibles dans le tableau 11.17. Notons que ces indices sont plutôt semblables, qu'il s'agisse de la version pondérée ou non pondérée. Un tel résultat est propre à cet ensemble de données en particulier et ne s'applique pas nécessairement à des indices de prix du logement estimés à partir d'autres sources.

11.56 Le tableau 11.18 résume les résultats des indices obtenus selon les différentes méthodes présentées ici, en utilisant la base de données élargie pour l'année 2007. La moyenne simple enregistre le taux de croissance le plus élevé parmi tous les indices estimés avec 10,1 % tandis que celui de la médiane est légèrement plus faible à 9,2 %.

Tableau 11.17. Indice des ventes répétées (2002 = 100)

Année	Non pondéré	Variation en pourcentage	Pondéré	Variation en pourcentage
2002	100,0		100,0	
2003	106,3	6,3	106,6	6,6
2004	112,7	6,0	112,9	5,9
2005	115,4	2,4	115,5	2,3
2006	120,3	4,2	120,5	4,4
2007	130,5	8,5	130,9	8,6
2008	141,0	8,1	141,8	8,3

Source : calculs des auteurs basés sur les données MLS® pour une ville canadienne.

L'indice de la méthode hédonique augmente de 5,7 % et ceux des méthodes des périodes temporelles consécutives, des régressions groupées et du prix des caractéristiques augmentent de 5,9 % (les calculs ne sont pas développés). En revanche, la méthode des ventes répétées donne respectivement 8,5 % et 8,6 % pour ses versions pondérée et non pondérée. Si l'échantillon est quelque peu réduit pour se permettre de généraliser, une remarque est toutefois importante. Les indicateurs qui ne sont pas à qualité constante, c'est-à-dire la moyenne et la médiane, donnent les augmentations les plus fortes, alors que les méthodes hédoniques enregistrent les plus faibles. L'approche par les ventes répétées, bien qu'elle fixe beaucoup d'éléments de qualité, n'ajuste pas l'ancienneté. Il n'est donc pas surprenant que l'augmentation des prix ainsi obtenue soit plus forte que celle obtenue avec les méthodes hédoniques.

Tableau 11.18. Taux de croissance en pourcentage des différents indices de prix du logement (2007)

Moyenne	Médiane	Hédonique regroupée	Hédonique Caractéristiques des prix	Ventes répétées non pondérées	Ventes répétées pondérées
10,1	9,2	5,7	5,9	8,5	8,6

Source : calculs des auteurs basés sur les données MLS® pour une ville canadienne.

12 Recommandations

12.1 Ce manuel apporte des informations détaillées et exhaustives sur l'élaboration d'indices des prix de l'immobilier résidentiel (IPIR). Il donne un aperçu des problèmes conceptuels et théoriques susceptibles de se poser, explique les différentes finalités auxquelles ces indices peuvent répondre et fournit des conseils sur la manière de résoudre les difficultés pratiques que les organismes statistiques peuvent rencontrer dans l'établissement de ces indices. Les chapitres précédents couvrent tous les aspects pertinents de la question, notamment une description des différentes méthodes actuellement utilisées et des conseils sur les autres techniques dont on dispose pour élaborer les indices, avec leurs avantages et leurs inconvénients respectifs. Ce chapitre fait la synthèse de toutes ces informations et contient des recommandations sur la meilleure manière de construire des indices des prix de l'immobilier résidentiel, y compris s'agissant d'en améliorer la comparabilité internationale. Vu qu'elles tiennent nécessairement compte de la diversité des situations nationales sur le plan de la disponibilité des données, ces recommandations ne peuvent pas être trop prescriptives.

12.2 Le présent ouvrage s'adresse également aux utilisateurs d'IPIR. Il apporte des informations, non seulement sur les différentes méthodes employées (ou susceptibles de l'être) pour construire les indices, mais également sur les limites statistiques de ce qu'il est possible de mesurer. Les utilisateurs devront conserver à l'esprit ce dernier point afin de pouvoir interpréter comme il se doit les résultats d'un indice donné. Tout ensemble de recommandations doit commencer par une présentation du concept fondamental qui sous-tend l'indice cible; en d'autres termes, il faut savoir précisément ce qu'un indice des prix de l'immobilier résidentiel vise à mesurer. La réponse à cette question dépendra, bien entendu, des besoins de l'utilisateur et de la finalité de l'indice.

12.3 Les recommandations ci-après suivent le même ordre que les chapitres 3 à 8. Le chapitre 3 décrit les principaux éléments du cadre conceptuel des IPIR, tandis que les chapitres 4 à 8 développent les principales méthodes statistiques pouvant servir à construire ces indices. Les différentes méthodes correspondent essentiellement aux solutions que l'on peut apporter au problème de la variation de qualité, à savoir comment corriger un IPIR en fonction d'éventuelles variations de la composition de l'échantillon de biens vendus et de la qualité de ces biens (effet net des rénovations, des agrandissements et de la dépréciation), pour les logements individuels.

Aspects conceptuels

Indice cible et fondement conceptuel

12.4 En principe, l'indice cible, en d'autres termes le type d'indice que l'on cherche à construire, dépendra de l'objectif visé. C'est le *Système de comptabilité nationale 2008* qui doit servir de cadre conceptuel à l'élaboration des IPIR.

Pondération

12.5 Un indice des prix censé mesurer la *richesse* liée au fait d'être propriétaire d'un bien immobilier résidentiel doit être *pondéré en fonction des stocks*. Il faut également un indice *pondéré en fonction des stocks* pour un indicateur de stabilité financière, en particulier si cet indice doit servir à repérer les bulles spéculatives des prix des biens immobiliers.

12.6 Un indice des prix qui doit mesurer la *production réelle* du secteur de la construction de biens immobiliers résidentiels doit être *pondéré en fonction des ventes*. Il faut également un indice pondéré en fonction des ventes pour un indice des prix à la consommation (IPC) basé sur des coûts d'acquisition.

Le champ de l'indice

12.7 Un indice des prix qui doit mesurer la *richesse* liée à un bien résidentiel doit couvrir tous les biens immobiliers résidentiels, c'est-à-dire tant les biens existants que ceux de construction récente[1]. Il en va de même pour un indice servant d'indicateur de stabilité financière.

12.8 Un indice des prix qui doit mesurer l'investissement réel dans le secteur de l'immobilier résidentiel doit couvrir les ventes de biens immobiliers résidentiels neufs[2]. L'élément «construction» des logements neufs fait partie de la *formation brute de capital fixe*. Le coût du terrain, hormis la valeur de toute amélioration apportée à cet élément, doit être exclu à cette fin. Toutefois, comme nous l'avons expliqué au chapitre 3, il faut que l'indice des prix englobe à la fois les ventes d'habitations neuves et les bâtiments existants si l'on veut élaborer des mesures de la production réelle pour les activités des agents immobiliers qui vendent aux acheteurs des biens neufs ou anciens. Le champ de l'indice pour cette application doit couvrir à la fois la valeur des bâtiments et des terrains lors des ventes de biens immobiliers résidentiels.

12.9 Il faut également un indice des prix limité aux seuls biens immobiliers neufs dès lors qu'un indice des prix de l'immobilier résidentiel constitue l'un des intrants de l'IPC pour la mesure du coût des logements occupés par leur propriétaire sur la base du coût d'acquisition net, c'est-à-dire lorsque l'IPC couvre le coût d'acquisition de biens immobiliers neufs sur le marché des logements occupés par leur propriétaire. Cette démarche, qui est l'une des possibilités exposées au chapitre 3, traite l'achat d'un logement au même titre que l'achat de n'importe quel autre bien de consommation[3].

1 On inclut ici les conversions de biens existants, par exemple lorsqu'un entrepôt est converti en immeuble à usage d'habitations ou lorsqu'un bien existant a été sous-divisé.
2 Les rénovations de logements existants font également partie de l'investissement dans la construction de logements.
3 C'est le concept des acquisitions nettes qui est privilégié étant donné qu'il est celui qui se rapproche le plus du concept de l'«acquisition» traditionnellement adopté pour d'autres volets d'un IPC et qui est plus adapté pour un IPC devant servir d'indicateur général des conditions économiques actuelles. Mais la méthode peut susciter les critiques de ceux qui ont besoin d'un IPC à titre d'indice de compensation, puisque ni la pondération ni l'indicateur de prix ne reflètent adéquatement les services fournis par les logements des propriétaires occupants. Ainsi, une hausse des taux d'intérêt ne serait pas reflétée dans un indice calculé sur la base des coûts d'acquisition nets. Voir le Manuel de l'IPC (2004) et le *Guide pratique de l'élaboration des indices des prix à la consommation* (Nations Unies, 2009).

Qualité constante

12.10 Indépendamment des différentes finalités de l'IPIR, un indice des prix de l'immobilier résidentiel doit permettre de comparer la valeur des ventes ou du stock de biens immobiliers résidentiels d'une période à une autre, étant entendu que l'on doit tenir compte des variations éventuelles des caractéristiques des biens. C'est la raison pour laquelle il faut pouvoir distinguer, parmi les variations de prix, celles qui relèvent d'une évolution des attributs du logement des autres, qui s'apparentent à des «variations pures de prix».

12.11 Un indice des prix à qualité constante convient à la fois pour un indice des prix pondéré par les stocks et pour un indice des prix pondéré par les ventes. Il existe un certain nombre de méthodologies pratiques permettant de construire un tel indice. On trouvera ci-après quelques recommandations sur les méthodes à privilégier en fonction des circonstances.

Décomposition entre les composantes «bâtiments» et «terrains»

12.12 Il peut s'avérer nécessaire de décomposer l'IPIR en bâtiments et terrains, surtout si cette distinction est faite dans les comptes de patrimoine du Système de comptabilité nationale. Cette décomposition peut également être nécessaire lorsqu'un indice des prix de l'immobilier résidentiel est l'un des intrants dans un IPC pour la mesure des logements occupés par leur propriétaire sur la base du coût d'acquisition net.

Méthodes statistiques permettant d'élaborer des indices à qualité constante

12.13 Les méthodes adoptées par les organismes statistiques pour construire des IPIR à qualité constante varient d'un pays à l'autre et sont dictées en grande partie par la disponibilité des données issues des processus d'achat et de vente d'un bien immobilier. Les difficultés inhérentes à l'élaboration d'indices des prix de l'immobilier résidentiel à qualité constante sont essentiellement les trois suivantes :

- Les biens résidentiels sont notoirement hétérogènes. Il n'existe pas deux biens identiques.
- Les prix sont souvent négociés. Le prix demandé pour un bien n'est pas fixe et peut évoluer au cours du processus de transaction jusqu'à ce que le prix soit finalisé. En d'autres termes, la valeur de marché d'un bien ne peut être connue avec certitude qu'une fois que celui-ci a été effectivement vendu[4].
- Les ventes de biens immobiliers ne sont pas fréquentes. Dans bien des pays, moins de 10 % du stock de logements change de mains chaque année; une maison est donc susceptible d'être revendue environ une fois tous les dix ans.

12.14 Les différentes méthodes d'élaboration d'indices utilisées par un organisme statistique reflètent les solutions adoptées pour surmonter ces difficultés. Quatre méthodes ont été passées en revue dans le présent manuel : la stratification ou l'ajustement au titre de la composition de l'échantillon, les méthodes de régression hédonique, les ventes répétées et les méthodes fondées sur l'estimation (c'est-à-dire la méthode RPVE). On trouvera ci-après des recommandations relatives à chacune de ces méthodes. Chaque méthode tente de tenir compte d'éventuelles variations des «paramètres de qualité» des habitations dont les prix sont observés et combinés pour construire l'indice. Certaines méthodes, toutefois, ne permettent pas de neutraliser l'effet des variations de qualité des biens immobiliers pris individuellement, c'est-à-dire l'effet net de la dépréciation des bâtiments et des rénovations et des agrandissements. Quand on utilise pour élaborer l'indice des données issues des processus administratifs d'achat et de vente d'un bien résidentiel, le prix sera généralement lié soit au prix demandé, soit au prix de vente — l'un pouvant différer de l'autre.

12.15 Les recommandations n'apportent pas de solution au problème que pose l'élaboration d'un IPIR dans les pays où une proportion significative du stock de logements est composée de logements informels ou traditionnels. On trouvera au chapitre 10 un exemple de calcul d'IPIR dans ces circonstances, qui repose sur le cas de l'Afrique du Sud. Dans ces conditions, il n'est pas possible d'être très prescriptif en termes de recommandations, puisque la situation varie considérablement d'un pays à l'autre, et qu'il n'existe pas de solution idéale permettant d'obtenir un indice des prix de l'immobilier résidentiel pur sur le plan conceptuel et dénué de difficultés pratiques. Le responsable devra en fait tirer parti des meilleures sources d'information disponibles et sera sans doute contraint à des compromis conceptuels et méthodologiques dans le calcul de l'indice. Dans ces conditions, il est particulièrement important que les organismes statistiques fournissent des appréciations des indices des prix ainsi élaborés et guident leurs utilisateurs quant à leurs utilisations potentielles.

Stratification ou ajustement au titre de la composition de l'échantillon

12.16 La stratification ou l'ajustement au titre de la composition de l'échantillon constitue le moyen le plus simple de tenir compte des variations intervenues dans la composition ou dans les paramètres de qualité des biens immobiliers vendus constituant l'échantillon. Cette méthode répond également aux besoins des utilisateurs qui recherchent des sous-indices afférents à différents segments

4 Dans certains cas, même les prix de vente ne reflètent pas les «vraies» valeurs de marché, par exemple lorsqu'il s'agit de ventes en catastrophe consécutives à un divorce, etc.

du marché du logement. L'efficacité de la stratification dépendra des variables de stratification utilisées, car une mesure ajustée en fonction des pondérations ne tient compte que des changements de composition à l'échelle des divers groupes — un indice ajusté en fonction des composantes ne tient pas compte des changements intervenus dans la composition de l'échantillon de biens vendus *au sein* de chaque sous-groupe ou strate.

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

12.18 Les principaux avantages de la méthode de la stratification ou poststratification sont les suivants :

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

12.19 Les principaux inconvénients de la méthode de la stratification ou poststratification sont les suivants :

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

12.20 La stratification ou poststratification convient dès lors que :

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

12.21 *La stratification ou poststratification est recommandée dès lors que le volume des ventes est suffisamment important et que les informations sur les caractéristiques du logement sont suffisamment détaillées pour permettre une classification détaillée des biens*[5].

La régression hédonique

12.22 Depuis quelques années, le recours aux techniques hédoniques pour faire des ajustements au titre de la qualité et calculer des indices des prix a apporté une importante contribution à l'évolution méthodologique des indices des prix et devient rapidement une méthode de choix pour l'élaboration d'indices des prix de l'immobilier résidentiel à qualité constante[6]. Il n'existe aucune uniformité dans l'application pratique de la régression hédonique, mais l'idée qui sous-tend ces techniques est assez simple. La régression hédonique est une technique statistique qui mesure la relation entre les caractéristiques observables d'un produit ou d'un service, d'une part, et son prix ou sa valeur, d'autre part. Dans le contexte des indices des prix de l'immobilier résidentiel, la «meilleure» forme de la fonction hédonique peut être linéaire plutôt que log-linéaire afin de refléter le fait que la valeur d'un bien est généralement égale à la somme du prix des bâtiments et du prix du terrain.

12.23 Il existe essentiellement deux méthodes d'application des techniques hédoniques à l'immobilier résidentiel :

- La méthode des variables indicatrices temporelles. Cette méthode repose généralement sur le recours à une régression simple, avec des variables indicatrices temporelles et

5 Une stratification grossière, disons, par grande ville et par type de bien immobilier, où ce dernier n'est décrit qu'en termes de «récemment construits» ou «anciens», n'est pas recommandée.

6 Si on regarde les indices harmonisés du prix du logement produits par les instituts statistiques européens, en 2011, plus de la moitié des pays utilisaient des techniques hédoniques pour tenir compte des variations de la qualité. Pour plus de détails, voir Marola *et al.* (2012).

des coefficients de caractéristiques fixes, qui couvre toutes les périodes et qui est refaite à chaque fois que l'indice des prix est calculé. Les (exponentielles des) coefficients de variables indicatrices temporelles représentent les variations de prix d'une période à l'autre hors variations de composition. Cette méthode a l'avantage de la simplicité. L'un de ses inconvénients tient au fait qu'elle soulève le problème de la révision de l'indice, car les coefficients des variables indicatrices temporelles seront actualisés chaque fois que de nouvelles périodes seront ajoutées et que la régression sera effectuée.

Toutefois, il existe une variante de la méthode de la variable indicatrice temporelle, dite «des fenêtres mobiles», qui peut fonctionner correctement dans la pratique et résout le problème de la révision. On effectue une régression hédonique avec les données pour les dernières *N* périodes, et la dernière variable indicatrice temporelle sert de maillon de chaînage pour actualiser l'indice pour la période précédente. Pour des références bibliographiques sur cette méthode et un exemple, voir le chapitre 5.

- La méthode de l'imputation hédonique. Une régression hédonique séparée est réalisée au cours de chaque période et les prix «manquants» de la période courante pour les biens vendus durant la période de référence sont imputés à l'aide des prix prévus à partir de l'équation hédonique estimée. Une approche symétrique est possible en imputant également les prix «manquants» pour la période de référence pour les biens vendus au cours de la période courante et en prenant ensuite la moyenne géométrique des deux indices à imputation hédonique.

12.24 Ces deux méthodes de régression hédoniques peuvent être exposées au biais dû à l'omission de variables si l'une des caractéristiques importantes déterminant le prix est omise de l'équation de régression. La multicolinéarité peut être un problème pratique, surtout quand il faut décomposer l'indice en bâtiments et en terrains. La méthode de la variable indicatrice temporelle a souvent été utilisée par les universitaires, en partie du fait de sa simplicité, mais la méthode de l'imputation hédonique est plus souple — elle permet aux prix des caractéristiques d'évoluer de manière indépendante dans le temps, alors que la méthode de la variable indicatrice temporelle force les prix des caractéristiques à bouger de manière proportionnelle — et elle est essentiellement semblable à la méthodologie traditionnelle du panier fixe pour calculer des indices des prix.

12.25 Les méthodes de régression hédonique peuvent être utilisées en parallèle avec la stratification pour traiter toute variation résiduelle de la qualité des échantillons subsistant au sein des strates, ce qui permet aussi de tenir compte du fait que différentes spécifications de modèles peuvent être nécessaires pour différents segments du marché du logement ou que la «valeur» de certaines caractéristiques variera d'un segment à l'autre du marché.

12.26 Les principaux avantages des techniques hédoniques sont les suivants :

- Si la liste des caractéristiques des biens est suffisamment détaillée, la méthode permet de tenir compte à la fois des variations de composition au sein des échantillons et des variations de qualité (dépréciation et rénovation) des biens pris individuellement.
- Des indices de prix peuvent être construits pour différents types de logements et d'emplacements par le biais de la stratification et l'application de techniques hédoniques à chaque strate individuelle.
- Les indices des prix stratifiés fondés sur des régressions hédoniques afin de prendre en compte les variations dans la composition des échantillons au sein des strates permettent d'utiliser des valeurs relatives du stock de logements pour pondérer les indices des strates corrigées des variations de composition (dans un IPIR pondéré par les stocks).
- La méthode permet de faire une utilisation optimale des données disponibles.
- Elle peut en principe être utilisée pour décomposer l'indice global des prix en ses composantes «terrain» et «bâtiments», sous réserve de disponibilité des données.

12.27 Les principaux inconvénients de la régression hédonique sont les suivants :

- La méthode est souvent considérée comme nécessitant de grandes quantités de données, surtout s'agissant des caractéristiques du logement devant servir de variables explicatives[7].
- Il peut être difficile de neutraliser suffisamment l'effet de l'emplacement si les prix des biens et les tendances d'évolution des prix varient d'une région à l'autre.
- La méthode peut être sensible aux variables utilisées dans la régression et à la forme fonctionnelle du modèle.
- La méthode n'est pas particulièrement facile à expliquer aux utilisateurs et, de leur point vue, peut manquer de transparence.

12.28 *Sous réserve que l'on dispose de données sur les principales caractéristiques du logement, la méthode de la régression hédonique est généralement la meilleure technique pour élaborer un indice des prix des biens résidentiels à qualité constante. En ce qui concerne le recours à des techniques de régression pour tenir compte de la qualité, la méthode des imputations comporte un certain nombre d'avantages sur celle de l'approche de la variable indicatrice temporelle. Les indices hédoniques stratifiés sont préférables à une application directe de la régression hédonique à l'ensemble du jeu de données.*

7 Toutefois, comme on l'a vu dans les chapitres précédents, dans certains cas des résultats satisfaisants peuvent être obtenus avec des méthodes de régression hédonique n'utilisant que trois ou quatre caractéristiques du logement.

Ventes répétées

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

12.30 Les principaux avantages de la méthode des ventes répétées sont les suivants :

- Dans sa forme élémentaire, cette méthode ne nécessite aucune information sur les caractéristiques des logements, hormis les adresses des biens vendus. Les données utilisées sont souvent disponibles dans les dossiers administratifs.
- Elle suit la méthodologie dite du panier fixe, partant de l'hypothèse que la dépréciation et les rénovations n'ont pas modifié le logement entre les ventes successives.
- Un grand nombre des caractéristiques liées à l'emplacement et les autres caractéristiques déterminant le prix qui sont difficiles à mesurer sont susceptibles d'être automatiquement incluses.
- Les régressions courantes dans le cadre de ventes répétées sont faciles à effectuer et les indices des prix qui en résultent sont faciles à élaborer.
- Aucune imputation n'est nécessaire. Par construction, l'emplacement est automatiquement pris en compte.
- Les résultats sont, en principe, reproductibles.

12.31 Les principaux désavantages de la méthode des ventes répétées sont les suivants :

- La méthode n'utilise pas tous les prix de vente disponibles; elle n'utilise que les informations disponibles sur les biens ayant été vendus plus d'une fois pendant la période d'échantillonnage.
- La version courante de la méthode ne tient pas compte de la dépréciation (nette) du logement.
- Il peut en résulter un biais de sélection du fait que l'on se limite aux seuls biens ayant été vendus plus d'une fois pendant la période d'échantillonnage.
- La méthode ne permet pas de produire des indices des prix séparés pour les bâtiments et pour les terrains.
- Du fait du recours à la technique des ventes répétées, on peut manquer de données pour calculer des indices mensuels des prix des biens résidentiels pour des catégories restreintes de biens.
- L'échantillon est mis à jour au fur et à mesure que l'on dispose d'informations nouvelles sur les transactions. L'indice des prix sur les biens résidentiels élaboré à partir des ventes répétées peut donc faire l'objet de révisions rétrospectives sur une longue période[8].
- Étant donné qu'une maison doit être vendue au moins deux fois dans un indice de ventes répétées, les logements récents en sont donc exclus.

12.32 *Bien qu'elle constitue un point de départ naturel pour l'élaboration d'un indice, la méthode des ventes répétées ne doit pas être préférée à la méthode hédonique (stratifiée) pour la construction d'un indice des prix des biens résidentiels à qualité constante. Toutefois, elle peut être une solution possible dès lors qu'il n'existe que des informations limitées, voire inexistantes, sur les caractéristiques des logements, que l'on dispose d'un nombre relativement élevé de transactions successives fournissant assez d'observations pour les catégories de résidences pertinentes et que le biais de sélection de l'échantillon n'est pas considéré comme posant un problème. Elle n'est pas recommandée quand une distinction doit être faite entre le prix des bâtiments et le prix des terrains.*

Méthodes basées sur l'estimation

12.33 Les méthodes basées sur l'estimation utilisent des valeurs «évaluées», comme les évaluations à des fins fiscales ou des valorisations émanant d'enquêtes spécialement commandées auprès d'agents immobiliers, souvent menées en référence à des biens similaires ayant été vendus, afin de contourner les deux principaux écueils de la méthodologie des ventes répétées — le nombre relativement restreint de prix observés et la sensibilité au biais de sélection de l'échantillon. Lorsque les valorisations se rapportent toutes à une période de référence type, le modèle d'appariement qui sous-tend les méthodes basées sur une estimation offre également l'avantage de pouvoir être appliqué directement et sans nécessiter de recourir à l'économétrie pour tenir compte des variations de composition. Toutefois, comme la méthode des ventes répétées, les méthodes basées sur l'estimation ne permettent généralement pas de neutraliser les variations de qualité des habitations prises individuellement. Par ailleurs, elles reposent le plus souvent sur l'opinion d'un expert quant à la somme pour laquelle un bien se vendrait plutôt que sur un prix effectif de transaction. Ainsi, on peut faire valoir, à l'extrême, que les méthodes basées sur l'estimation sont influencées par des jugements ou des opinions, même si ceux-ci font autorité et sont objectifs.

8 Dans la pratique, le lien pour les deux dernières périodes dans la régression courante des ventes répétées peut être utilisé pour mettre à jour l'indice courant.

12.34 La méthode RPVE (rapport prix de vente/évaluation) utilise les évaluations avec une période de référence commune comme les prix d'une période de référence dans un modèle d'appariement (même si les résultats sont normalisés pour obtenir un indice égal à 1 (ou 100) dans la période de référence). L'expérience des quelques pays ayant calculé un indice RPVE[9] sont généralement positifs, bien que certains chercheurs aient fait état d'un biais découlant de réévaluations fréquentes et d'une précision réduite dans le temps du fait de nouvelles évaluations.

12.35 Les principaux avantages de la méthode RPVE sont les suivants :

- Reposant sur la méthode standard du panier fixe, elle est conforme à la théorie traditionnelle des indices.
- Elle est facile à calculer.
- La méthode bénéficie d'un nombre beaucoup plus important d'observations que la méthode des ventes répétées et elle est donc moins sujette aux difficultés inhérentes au fait de ne disposer que d'un nombre relativement réduit de prix observés.
- Elle est moins sujette au biais d'autosélection de l'échantillon que la méthode des ventes répétées.
- Elle ne souffre pas d'éventuelles révisions apportées aux chiffres précédemment calculés.
- Elle est reproductible.

12.36 Les principaux désavantages de la méthode RPVE sont les suivants :

- Elle ne peut pas traiter adéquatement les variations de qualité (dépréciation et rénovations) des logements[10].
- Il faut disposer de données sur les évaluations effectuées à chaque adresse et pour tous les biens immobiliers.
- La méthode dépend de la qualité des évaluations.
- Elle ne peut pas être utilisée pour décomposer l'indice global des prix des biens immobiliers en terrains et en bâtiments[11].

12.37 *La méthodologie RPVE remédie à quelques-unes des faiblesses de la méthodologie des ventes répétées et doit lui être préférée si l'on dispose de données d'évaluation de qualité suffisante et si le biais de sélection est considéré comme étant un élément sérieux de l'application de la méthodologie des ventes répétées. La méthodologie RPVE a effectivement ses inconvénients, mais elle est recommandée quand l'usage de techniques hédoniques n'est pas possible. Les résultats de la méthode RPVE sont encore meilleurs en conjonction avec la stratification.*

Ajustement saisonnier

12.38 Si la série initiale de prix de biens immobiliers fait état de certaines fluctuations saisonnières, on pourra recourir à des techniques ordinaires de correction des fluctuations saisonnières pour corriger la série initiale des variations saisonnières. Toutefois, si l'imputation hédonique ou la méthode de stratification sont utilisées pour construire l'indice initial, certaines recommandations spécifiques supplémentaires s'appliquent; elles sont présentées ci-dessous.

12.39 Si la méthode de stratification est utilisée pour construire l'indice initial et que celui-ci montre une certaine saisonnalité, alors la méthode de la période annuelle mobile exposée au chapitre 5 peut être appliquée pour corriger la série des variations saisonnières sans se reporter à des méthodes économétriques.

12.40 Si la méthode de l'imputation hédonique est utilisée pour construire l'indice des prix initial et que ce dernier montre une certaine saisonnalité, il peut s'avérer utile, pour obtenir une série corrigée des variations saisonnières, de construire dans un premier temps une série mensuelle ou trimestrielle d'une année sur l'autre. Cette série initiale peut alors être agrégée à l'aide de la méthode de l'année glissante au sein d'une série lissée, corrigée des variations saisonnières.

9 En Europe, le Danemark, la Suède et les Pays-Bas utilisent la méthode RPVE.

10 Comme avec la méthode des ventes répétées, l'indice des prix obtenu grâce à la méthode RPVE peut en principe être ajusté en utilisant des informations exogènes sur la dépréciation nette des biens dans la catégorie considérée.

11 Quand on dispose de décompositions officielles de la valeur évaluée totale du bien en terrains et bâtiments, celles-ci peuvent servir à vérifier les indices des prix des terrains et des bâtiments découlant de l'application de méthodes de régression hédonique.

Glossaire

Actualisation

Division de la valeur courante d'un agrégat donné par un indice des prix (appelé le *déflateur* dans ce contexte) pour réévaluer les quantités aux prix de la période de référence des prix.

Agrégat

Ensemble de transactions (ou leur valeur totale), tel que le total des achats d'immobilier résidentiel effectués par les ménages, au cours d'une période déterminée.

Agrégat élémentaire

Habituellement défini comme le plus petit agrégat disponible pour les données sur les dépenses servant à la construction d'indices. Les agrégats élémentaires servent aussi de strates pour échantillonner les produits dont les prix sont suivis. Les valeurs des agrégats élémentaires sont utilisées pour pondérer les indices des prix des agrégats élémentaires et ainsi obtenir des indices plus agrégés.

Dans le contexte des indices des prix de l'immobilier résidentiel reposant sur les ventes, le terme d'agrégat élémentaire est moins approprié. Comme chaque propriété est fondamentalement unique, les quantités sont égales à 1; ainsi les pondérations sont disponibles au niveau le plus détaillé.

Agrégation

Action de combiner ou d'ajouter différents ensembles de transactions pour obtenir un ensemble de transactions plus large. On dit de l'ensemble plus large ainsi obtenu qu'il a un niveau d'agrégation plus élevé que les (sous-)ensembles dont il est composé. Le terme «agrégation» est aussi utilisé pour désigner l'action d'ajouter les valeurs des agrégats de niveau inférieur pour obtenir des agrégats d'un niveau plus élevé. Dans le cas des indices de prix, ce terme désigne le calcul de la moyenne des indices de prix des agrégats de niveau inférieur de manière à obtenir des indices de prix pour les agrégats d'un niveau plus élevé.

Ajustement de qualité

Ajustement pratiqué sur la variation du prix d'une propriété dont les caractéristiques ont changé dans le temps, afin d'éliminer l'effet de ce changement sur la variation de prix observée. En pratique, l'ajustement nécessaire ne peut être qu'une estimation. Différentes méthodes d'estimation, dont les méthodes hédoniques, peuvent être utilisées en fonction des circonstances. Ces méthodes peuvent aussi être utilisées pour maîtriser les changements de composition de l'échantillon des propriétés vendues qui surviennent avec le temps.

Approche axiomatique

Manière d'aborder la théorie de l'indice selon laquelle le choix de la formule de l'indice est déterminé en fonction des propriétés mathématiques de celui-ci. Une liste de tests est établie, dans lesquels l'indice doit à chaque fois posséder une certaine propriété ou être conforme à un certain axiome. Un indice peut alors être choisi en fonction du nombre de tests auxquels il satisfait. Les tests n'ont pas forcément tous la même importance; il suffit parfois, pour rejeter un indice, qu'il ne satisfasse pas à un ou deux tests jugés fondamentaux.

Approche économique

Cette approche de la théorie de l'indice part du principe que les quantités sont des fonctions des prix, les données observées étant produites comme solutions à divers problèmes d'optimisation économique. Bien que cette approche soit très pertinente pour un IPC en tant qu'approximation d'un indice du coût de la vie, elle paraît moins adaptée pour un indice des prix de l'immobilier résidentiel. Voir aussi *approche axiomatique*.

Approche par les dépenses engagées ou méthode des paiements

L'une des trois manières principales d'inclure les logements des propriétaires occupants dans l'IPC. Dans la méthode des dépenses engagées, les frais divers relatifs à la propriété du logement sont simplement additionnés.

Approche par les paiements

Voir *approche par les dépenses engagées*.

Apurement des données

Procédures, souvent automatisées, utilisées pour supprimer des fichiers les erreurs de saisies, les observations jugées peu probables ou les données aberrantes.

Base de sondage

Liste des unités statistiques de l'univers dans lequel un échantillon peut être choisi. Cette liste peut contenir des renseignements sur les unités statistiques pouvant être utilisés à des fins d'échantillonnage. Ces listes peuvent ne pas couvrir toutes les unités de l'univers considéré et peuvent aussi inclure des unités qui n'en font pas partie.

Biais

Tendance systématique de l'IPIR calculé à s'écarter d'un indice idéal ou préféré, qui résulte de la méthode utilisée pour collecter ou traiter les données, ou de la formule employée pour calculer l'indice. Voir aussi *biais d'échantillonnage*.

Biais d'échantillonnage

Biais d'un indice qui peut survenir quand l'échantillon représente mal la population. Dans le contexte du logement, l'échantillon de propriétés peut ne pas être représentatif de l'ensemble des ventes (particulièrement pertinent pour un indice des prix de vente) ou ne pas être représentatif du stock de logements (pertinent pour un indice du prix du stock). Si toutes les ventes sont observées, il n'y a pas de biais d'échantillonnage pour un indice des prix de vente des propriétés.

Bien

Objet physique pour lequel il existe une demande, pour lequel les droits de propriété peuvent être établis et dont la propriété peut être transmise entre unités par le biais de transactions sur un marché.

Bien durable de consommation

Bien de consommation qui peut être utilisé de manière répétée ou continue à des fins de consommation pendant une longue période, en général plusieurs années. Une maison est une forme extrême de bien durable en raison de sa très longue durée de vie espérée. Ceci a conduit à différentes approches pour la prise en compte des *propriétaires occupants* dans les statistiques économiques.

Biens saisonniers

Biens qui ne sont pas disponibles sur le marché pendant certaines saisons ou certaines périodes de l'année, ou qui sont disponibles pendant toute l'année, mais dont les quantités et les prix fluctuent régulièrement en fonction de la saison ou de la période de l'année.

Caractéristiques

Qualités physiques et économiques d'un bien ou d'un service qui sert à l'identifier et permet de le classer. Pour les propriétés résidentielles, elles concernent à la fois le bâtiment (l'immeuble) et l'emplacement/le terrain.

Caractéristiques techniques

Description ou liste des caractéristiques à utiliser pour identifier un logement particulier faisant l'objet de l'enquête sur les prix.

Chaînage

Action d'associer deux séquences consécutives d'observations de prix, ou indices de prix, à cheval sur une ou plusieurs périodes pour constituer une seule série. Si les deux séquences se chevauchent pendant une seule période, le procédé habituel consiste à simplement rééchelonner l'une d'elles de telle sorte que la valeur pour la période de chevauchement soit la même dans les deux séquences.

Champ de l'indice

L'ensemble des produits pour lesquels l'indice doit mesurer les variations de prix. La couverture d'un indice désigne l'ensemble effectif de produits inclus, par opposition au champ souhaité de l'indice.

Changement de qualité

Une modification dans une caractéristique (qui détermine la qualité) d'un bien ou d'un service. Dans le cas d'une propriété résidentielle, cela comprend à la fois la dépréciation du bâtiment et les rénovations, comme la modernisation des cuisines et des salles de bains, l'amélioration de l'isolation ou l'installation d'un chauffage central ou d'une climatisation.

Composante

Ensemble des biens et services qui constituent un agrégat donné. Aussi utilisée dans le contexte de la décomposition du prix des propriétés (indice) en composantes terrain et bâtiments.

Concept d'acquisition

Concept pour lequel la consommation est définie comme l'acquisition par un ménage, sur une période déterminée, de biens et de services de consommation (par opposition à leur utilisation totale ou partielle à des fins de consommation). Voir aussi *concept d'acquisitions nettes*.

Concept d'acquisitions nettes

L'une des trois manières principales d'inclure les logements des propriétaires occupants dans l'IPC. Les habitations ajoutées au stock des logements occupés par leur propriétaire (en général, principalement des habitations neuves) font partie de la couverture de l'indice; les habitations existantes sont exclues. Voir aussi le *concept d'acquisition*.

Concept d'utilisation

Concept sur lequel se fonde le calcul de l'IPC lorsque la consommation est définie comme l'utilisation effective par un ménage, sur une période déterminée, de biens et de services de consommation afin de répondre à leurs besoins et désirs (par opposition à l'acquisition de biens et de services de consommation). Lorsque le calcul de l'IPC se fonde sur ce concept, la consommation de biens de consommation durables sur une période déterminée se calcule à l'aide de la valeur des flux de services que procure le stock de biens durables détenus par les ménages. Cette valeur peut être estimée au moyen du *coût d'usage*.

Concordance de l'agrégation

L'agrégation est dite concordante lorsque l'indice calculé pour un agrégat a la même valeur, qu'il ait été obtenu en une seule opération par calcul direct, sans distinction entre ses composantes, ou en plusieurs étapes, par le calcul d'indices distincts, appelés sous-indices, pour ses différentes composantes, qui sont ensuite agrégées, la même formule étant utilisée à chaque étape.

Coût d'usage

Coût résultant, pour le détenteur d'un bien de capital fixe ou de consommation durable, de l'utilisation de ce bien pour fournir un flux de capital ou des services de consommation. Le coût d'usage se partage essentiellement entre l'amortissement du bien en question (mesuré aux prix courants, et non au coût historique) et le coût en capital ou en intérêts.

Couverture

Ensemble des biens et services dont les prix sont effectivement pris en compte dans l'indice. Pour des raisons pratiques, il se peut que la couverture soit moins étendue que le champ idéal de l'indice. En effet, les types de propriétés dont les

prix sont suivis peuvent ne pas couvrir tous les types de propriétés vendues ou appartenant au stock de logements.

Dépréciation

La baisse graduelle et permanente de la valeur économique d'un bâtiment ou d'un stock de logements causé par l'obsolescence et la dégradation au cours du temps.

Domaine

Autre terme désignant le champ d'un indice.

Échantillon

Une sélection (aléatoire ou non) d'éléments d'une population finie. Dans le contexte du logement, les propriétés vendues pendant un certain temps peuvent être considérées comme un échantillon du stock de logements. Cet échantillonnage est particulièrement pertinent pour un indice des prix du stock de l'immobilier résidentiel.

Habitations existantes

Le terme «habitations existantes» est parfois utilisé pour distinguer ces habitations des habitations neuves (et ajoutées au stock de logements).

Imputation hédonique

Une façon d'estimer un indice des prix de l'immobilier résidentiel ajustés de la qualité où les prix «manquants» sont imputés avec une modélisation utilisant une régression hédonique. Les paramètres du modèle sont estimés à chaque période, ce qui rend cette approche plus flexible que la méthode hédonique reposant sur les variables indicatrices temporelles.

Indice à pondérations fixes

Description abrégée d'une série de moyennes arithmétiques pondérées de rapports de prix d'indices où les pondérations sont fixes au cours du temps. Dans le contexte de l'indice des prix de l'immobilier résidentiel, les pondérations peuvent se rapporter aux ventes (dépenses) ou aux stocks.

Indice-chaîne

Série d'indices relative à une longue séquence de périodes, obtenue par l'enchaînement d'indices relatifs à des séquences plus courtes de périodes. Un indice-chaîne, calculé selon une certaine formule d'indice (comme celle de Fisher) est le produit d'indices calculés selon cette formule à chaque période consécutive. Voir aussi *Chaînage*.

Indice de la médiane

Un indice des prix qui retrace la variation de la médiane des prix des propriétés au cours du temps. La médiane est le milieu de la distribution (de l'échantillon) : la moitié des prix se situe au-dessus et l'autre moitié au-dessous. La médiane est moins sensible aux valeurs extrêmes que la moyenne et lui est souvent préférée comme mesure de tendance centrale pour les distributions fortement asymétriques.

Indice de la moyenne

Un indice des prix qui est calculé comme le rapport des moyennes de l'échantillon des prix *(valeurs unitaires)* des propriétés vendues pendant deux périodes.

Indice de Laspeyres géométrique

Moyenne géométrique pondérée des rapports de prix utilisant comme pondérations les parts de dépenses de la période de référence des prix.

Indice de niveau inférieur

Sous-indice, par opposition à un indice d'agrégat.

Indice des prix à la consommation (IPC)

Indice mensuel ou trimestriel calculé et publié par un organisme officiel de statistique, qui mesure l'évolution des prix des biens et des services de consommation acquis ou utilisés par les ménages. Sa définition, y compris le traitement des *propriétaires occupants*, peut varier d'un pays à l'autre. En Europe, l'indice harmonisé des prix à la consommation (IHPC) exclut actuellement les propriétaires occupants.

Indice des prix de Fisher

Moyenne géométrique de l'indice des prix de Laspeyres et de l'indice des prix de Paasche. L'indice de Fisher est *symétrique* et *superlatif*. Un indice des prix de l'immobilier reposant sur les ventes peut toujours être calculé en utilisant la formule de Fisher, car les quantités sont égales à 1 (comme chaque habitation est fondamentalement un bien unique).

Indice des prix de Jevons

Indice des prix d'agrégat élémentaire défini comme la moyenne géométrique non pondérée des rapports de prix de l'échantillon.

Indice des prix de Laspeyres

Un indice des prix pour lequel les quantités de biens et de services se rapportent à la plus ancienne des deux périodes comparées : la période de référence des prix. L'indice de Laspeyres peut aussi être exprimé comme une moyenne arithmétique pondérée des rapports de prix avec les parts de valeur de la période la plus ancienne comme pondérations. La période la plus ancienne est à la fois la période de référence des pondérations et la période de référence des prix.

Indice des prix de Lowe

Indice des prix qui mesure la variation entre la période 0 et la période t de la valeur totale d'un panier donné de biens et de services dont les quantités restent fixes. Il n'est pas nécessaire que le panier soit composé des quantités effectives d'une période particulière. Cette définition couvre une catégorie d'indices très large qui inclut, par exemple, moyennant une définition appropriée des quantités, les indices de Laspeyres et de Paasche.

Indice des prix de Paasche

Indice des prix pour lequel les quantités de biens et de services se rapportent à la plus récente des deux périodes

comparées. Cette période sert de période de référence pour les pondérations, et la période antérieure, de période de référence pour les prix. Cet indice peut aussi être exprimé sous forme d'une moyenne harmonique pondérée des rapports de prix qui utilise les parts de dépenses afférentes à la période la plus récente comme pondérations.

Indice superlatif

Les indices superlatifs sont en général des indices symétriques et ont de bonnes propriétés théoriques. Les indices de Fisher et de Törnqvist sont des exemples d'indices superlatifs.

Indice symétrique

Indice qui traite les deux périodes de façon symétrique en assignant une importance égale aux données sur les prix et les dépenses dans les deux périodes. Les données sur les prix et les dépenses des deux périodes entrent dans la formule de l'indice de façon symétrique.

Logement occupé par son propriétaire (propriétaire occupant)

Logement appartenant au ménage qui y habite. Le logement est un actif immobilisé que le propriétaire utilise pour se fournir un service de logement qu'il consomme lui-même. Ce service est généralement inclus dans le champ de l'IPC. Les loyers peuvent être imputés par référence aux loyers pratiqués sur le marché pour un logement équivalent ou par référence aux coûts d'usage. Voir aussi *méthode des loyers imputés* et *coût d'usage*.

Marché du logement informel

Zones d'habitations où un ensemble de logements a été construit sur un terrain dont les habitants n'ont pas de titre de propriété ou qu'ils occupent illégalement, et zones où les habitations ont été édifiées sans permis de construire ou ne sont pas conformes aux règles d'aménagement et de construction.

Ménage

Personne vivant seule ou groupe de personnes vivant ensemble qui pourvoient en commun à leur subsistance et à leurs autres besoins essentiels. La plupart des pays excluent du champ de l'IPC les groupes de personnes vivant dans les grandes collectivités institutionnelles (casernes, maisons de retraite, etc.).

Méthode des fenêtres mobiles

Une approche où une «fenêtre» mobile constituée d'un nombre fixe de périodes est utilisée pour calculer l'indice initial des prix (de l'immobilier résidentiel). La série temporelle est mise à jour en avançant la fenêtre d'une période et en chaînant l'évolution de l'indice pendant la dernière période à la série temporelle.

Méthode des loyers imputés

L'une des trois principales approches pour prendre en compte les propriétaires occupants dans l'IPC. Le prix imputé du service de logement est égal au prix auquel l'habitation serait louée.

Méthode des produits appariés ou du panier fixe

Procédé consistant à attribuer un prix à un même produit ou modèle exactement pour deux périodes consécutives ou plus. Ce procédé vise à garantir que les variations de prix observées ne sont pas influencées par un changement de qualité. La variation de prix entre deux produits parfaitement appariés est parfois qualifiée de variation pure de prix.

Méthode des ventes répétées

Une méthode pour calculer l'indice des prix de l'immobilier résidentiel qui compare les propriétés vendues au moins deux fois dans la base de données dont on dispose. C'est une méthode basée sur une régression qui n'inclut que des variables indicatrices temporelles.

Méthode hédonique des variables indicatrices temporelles

L'une des principales façons d'utiliser une régression hédonique pour construire un indice des prix (de l'immobilier résidentiel). Dans le modèle log-linéaire à variables indicatrices temporelles classique, les coefficients des caractéristiques restent constants au cours du temps et l'indice des prix peut être calculé directement avec l'exponentielle des coefficients des indicatrices temporelles.

Méthode hédonique du prix des caractéristiques

Une méthode de régression hédonique pour laquelle la modification de la valeur estimée du paramètre associé à une caractéristique des propriétés vendues (moyenne) — c'est-à-dire le prix implicite de la caractéristique — détermine l'indice des prix de l'immobilier résidentiel. Sous certaines hypothèses, cette approche est équivalente à celle de l'*imputation hédonique*.

Méthode RPVE ou SPAR

La méthode rapport prix de vente/évaluation (RPVE ou SPAR en anglais) permet de construire un indice des prix de l'immobilier résidentiel qui combine les prix de vente de la période actuelle avec les évaluations (valeurs estimées) obtenues pour une période de référence donnée.

Mise à jour des pondérations

Remplacement des pondérations utilisées dans un indice par un nouvel ensemble de pondérations.

Modèles hybrides (ventes répétées)

Une méthode utilisant une régression pour estimer les indices de l'immobilier résidentiel qui combine la *méthode des ventes répétées* et les *méthodes hédoniques*.

Non-unicité

On dit d'un indice en chaîne qu'il ne satisfait pas la propriété d'unicité s'il dérive et ne revient pas à l'unité lorsque les prix de la période en cours reviennent à leur niveau de la période de référence. Les indices en chaîne peuvent dériver lorsque les prix fluctuent pendant les périodes auxquelles ils se rapportent.

Période de référence

En principe, la période qui sert de base de comparaison à toutes les autres périodes. Ce terme peut toutefois avoir différentes significations selon le contexte. On distingue trois types de période de référence :

- la *période de référence des prix* — période à laquelle se rapportent les prix qui servent de base de comparaison aux prix d'autres périodes. Les prix afférents à la période de référence figurent au dénominateur des rapports de prix utilisés pour calculer l'indice;
- la *période de référence des pondérations* — période dont les dépenses servent de pondérations pour l'indice. Si les dépenses sont hybrides (c'est-à-dire si les quantités afférentes à une période sont évaluées aux prix d'une autre période), la période de référence des pondérations est la période à laquelle se rapportent les quantités;
- la *période de référence de l'indice* — période pour laquelle la valeur de l'indice est fixée à 100.

On notera qu'en pratique, dans le cas de l'IPIR, la période de référence des pondérations est souvent d'une année, alors que l'indice est généralement mensuel ou trimestriel, la durée de la période de référence des prix étant d'un mois ou d'un trimestre. En pratique, les périodes de référence des prix et des pondérations peuvent ne pas coïncider, tout au moins lorsqu'un IPIR est calculé pour la première fois, mais les périodes de référence des prix et de l'indice coïncident fréquemment.

Période de référence de l'indice

Période pour laquelle la valeur de l'indice est fixée à 100 (ou à 1).

Période de référence des pondérations ou base des pondérations

Période dont les parts de dépenses servent de pondérations ou dont les quantités composent l'ensemble des propriétés utilisées pour le calcul de l'indice de Lowe. Il est possible qu'il n'y ait pas de période de référence des pondérations, par exemple lorsqu'une moyenne est calculée pour les parts de dépenses des deux périodes, comme dans le cas de l'indice de Törnqvist, ou pour les quantités, comme dans le cas de l'indice de Walsh. Voir aussi *période de référence*.

Période de référence des prix

Période pour laquelle les prix figurent au dénominateur des rapports de prix. Voir aussi *période de référence*.

Période en cours ou période de comparaison

Désigne en principe la période la plus récente pour laquelle l'indice a été ou est calculé. Cependant, ce terme est largement utilisé pour désigner la période de comparaison, c'est-à-dire la période comparée à la période de référence, qui est habituellement la période de référence des prix ou la période de référence de l'indice. Ce terme est aussi employé pour désigner la plus récente de deux périodes. Le sens exact est en général clairement défini par le contexte.

Pondérations

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

Poststratification

Utilisé pour décrire les méthodes qui essayent d'annuler ou de réduire l'effet du changement de la composition de l'échantillon des propriétés vendues sur l'indice des prix des propriétés.

Prix d'offre

Le prix auquel une propriété est mise à la vente. Le prix d'offre peut être ajusté pendant le processus d'achat et de vente d'une maison jusqu'à l'obtention du prix final de la transaction.

Prix de demande (prix proposé par l'acheteur)

Le prix qu'un acheteur potentiel dit être prêt à payer pour une propriété.

Prix de vente (ou de transaction)

Le prix final auquel la propriété est vendue.

Prix imputé

Prix attribué à un article (par exemple une propriété) dont le prix est «manquant» pour une période donnée. Cela peut être réalisé en utilisant une méthode de régression hédonique. Voir aussi *imputation hédonique*.

L'expression «prix imputé» désigne aussi le prix attribué à un article qui n'est pas vendu sur le marché, tel qu'un bien ou un service produit par un agent économique et destiné à son autoconsommation, par exemple les services de logement dont bénéficie un propriétaire occupant et qui sont mesurés par la *méthode des loyers imputés*.

Produit

Terme générique désignant un bien ou un service. Les différents produits de l'échantillon choisis pour la collecte des prix sont souvent appelés «produits élémentaires».

Propriété représentative

Propriété (bien immobilier), ou catégorie de propriétés, qui représente une proportion élevée du total des dépenses d'un agrégat, et/ou dont la variation de prix moyenne sera vraisemblablement proche de la moyenne pour toutes les propriétés de l'agrégat.

Propriété résidentielle

Propriété sur laquelle sont autorisées des maisons unifamiliales, des maisons en rangée et des appartements selon divers régimes judiciaires.

Rapport de prix ou indice élémentaire

Rapport entre le prix d'un produit élémentaire déterminé pendant une période et le prix de ce même produit pendant une autre période.

Régression hédonique

Estimation d'un modèle hédonique qui met en œuvre les techniques de régression pour expliquer le prix d'une propriété en fonction de ses caractéristiques (relatives aussi bien aux bâtiments qu'à l'emplacement). Voir aussi *imputation hédonique* et *méthode hédonique des variables indicatrices temporelles*.

Stock de logements

Le nombre total de logements disponibles, en dehors des occupations passagères. Selon la définition adoptée, le stock de logements peut inclure ou non des mobile homes, etc.

Stratification

La stratification et la «repondération» d'un échantillon est une technique générale pour obtenir des résultats plus stables ou atténuer tout biais dû à la sélection de l'échantillon, y compris aux non-réponses.

Dans le contexte de l'indice des prix de l'immobilier résidentiel, l'échantillon de propriétés vendues est subdivisé en un nombre de strates (ou de cellules) relativement homogènes selon un nombre (limité) de caractéristiques qui déterminent le prix. Le prix moyen (valeur unitaire) ou la médiane des prix peut ensuite être utilisé pour calculer un indice de prix pour chaque strate. Dans une deuxième étape, ces indices de strate sont agrégés en utilisant les pondérations provenant des ventes ou des stocks. Cette méthode, dite *poststratification*, a souvent été utilisée pour ajuster les changements de composition de l'échantillon, ou les changements de composition de la qualité des maisons vendues.

La stratification peut aussi être utilisée avec d'autres méthodes d'ajustement des changements de qualité, comme les régressions hédoniques, la méthode des ventes répétées ou la méthode RPVE.

Système de comptabilité nationale (SCN)

Ensemble cohérent, homogène et intégré de comptes macroéconomiques, bilans et tableaux fondés sur des concepts, définitions, classifications et règles comptables adoptés au niveau international. Les revenus et les dépenses de consommation des ménages font partie des comptes répertoriés dans le SCN.

Test d'identité

Pour satisfaire à ce test, dans le cadre de l'approche axiomatique, l'indice des prix doit être égal à l'unité lorsque chaque prix demeure inchangé pendant les deux périodes.

Valeur

Prix multiplié par la quantité. La valeur des dépenses effectuées pour acquérir un ensemble de produits homogènes résulte de la multiplication unique de la composante «prix» par la composante «quantité». De même, la variation dans le temps de la valeur d'un ensemble de produits homogènes résulte d'une multiplication unique de la variation de la valeur unitaire par la variation des quantités totales. En revanche, un grand nombre de multiplications différentes de la composante «prix» par la composante «quantité» permettent de calculer la variation dans le temps de la valeur d'un ensemble de *produits hétérogènes*.

Dans le contexte du logement, la valeur peut aussi se rapporter à une seule propriété. Le «prix» d'une propriété est en fait sa valeur, qui est la somme du prix du bâtiment avec le prix du terrain sur lequel le bâtiment est construit.

Valeur aberrante

Qualifie généralement toute valeur extrême d'un ensemble de données recueillies par sondage. Dans le contexte de l'IPIR, cette expression désigne une valeur extrêmement forte ou faible pour un prix d'une propriété ou un rapport de prix, qui appelle des vérifications et qui devrait être supprimée quand elle est jugée inexacte.

Valeur de marché

La valeur d'une propriété à un certain moment, ou le prix auquel serait vendue la propriété sur le «marché libre».

Valeur estimée ou évaluation

Évaluation d'une propriété au prix du marché. Cette estimation peut être nécessaire pour obtenir un prêt hypothécaire. Dans certains pays, l'estimation est réalisée pour le compte du gouvernement pour des raisons de taxation (foncière). Les valeurs estimées des propriétés sont aussi appelées des évaluations. Voir aussi *méthode rapport prix de vente/évaluation (RPVE)*.

Valeur unitaire ou valeur moyenne

La valeur unitaire d'un ensemble de produits homogènes est la valeur totale des achats/ventes divisée par la somme des quantités. Il s'agit donc de la moyenne pondérée par les quantités des différents prix auxquels le produit est acheté/vendu. La valeur unitaire peut varier au fil du temps en cas de changement de la composition de l'ensemble de produits vendus à différents prix, même si les prix eux-mêmes ne changent pas.

Variation pure de prix

Variation du prix d'une propriété dont les caractéristiques n'ont pas changé, ou variation du prix de la propriété après ajustement au titre de la qualité (rénovations, agrandissements, dépréciation).

Vérification

Processus par lequel on vérifie et contrôle les prix relevés par les enquêteurs. Certaines vérifications peuvent être effectuées par ordinateur, au moyen de logiciels statistiques élaborés à cette fin. Voir aussi *apurement des données*.

Bibliographie

Abraham, J.M., and W.S. Schauman (1991), "New Evidence on Home Prices from Freddie Mac Repeat Sales," *American Real Estate and Urban Economics Association Journal* 19, 333–52.

Acadametrics (2009), House Price Indices—Fact or Fiction, www.acadametrics.co.uk.

Alterman, W., W.E. Diewert and R.C. Feenstra (1999)," *International Trade Price Indexes and Seasonal Commodities,*" Department of Labor, Bureau of Labor Statistics, Washington, D.C.: US Government Printing Office.

Australian Bureau of Statistics (ABS) (2005), "Renovating the Established House Price Index," Information Paper, Cat. No. 6417.0.

Australian Bureau of Statistics (ABS) (2006), "A Guide to House Price Indexes," Information Paper, Cat. No. 6464.

Bailey, M.J., R.F. Muth and H.O. Nourse (1963), "A Regression Method for Real Estate Price Construction," *Journal of the American Statistical Association* 58, 933–42.

Baldwin, A., A. Nakamura and M. Prud'homme (2010), "An Empirical Analysis of the Different Concepts for Owned Accommodation in the Canadian CPI: The Case of Ottawa, 1996–2005," paper presented at the joint UNECE and ILO Meeting on Consumer Price Indices, Geneva, May 10–12.

Balk, B.M. (1998), "On the Use of Unit Value Indices as Consumer Price Subindices," in Lane, W. (ed.), *Proceedings of the Fourth Meeting of the International Working Group on Price Indices*, Washington, D.C.: Bureau of Labor Statistics.

Balk, B.M. (2008), *Price and Quantity Index Numbers; Models for Measuring Aggregate Change and Difference*, New York: Cambridge University Press.

Belski, E., and J. Prakken (2004), "Housing Wealth Effects: Housing's Impact on Wealth Accumulation, Wealth Distribution and Consumer Spending," Working Paper W04–13, Harvard University, Joint Center for Housing Studies.

Berndt, E.R. (1991), *The Practice of Econometrics: Classic and Contemporary*, Reading, Massachusetts: Addison-Wesley Publishing Company.

Blades, D. (2009), "Housing in ICP 2011: Issues to be Resolved," paper presented at the Technical Advisory Group Meeting, October 1–2, 2009, World Bank, Washington, D.C.

Bostic, R.W., S.D. Longhofer and C.L. Readfearn (2007), "Land Leverage: Decomposing Home Price Dynamics," *Real Estate Economics* 35:2, 183–2008.

Bourassa, S.C., M. Hoesli and J. Sun (2006), "A Simple Alternative House Price Index Method," *Journal of Housing Economics* 15(1), 80–97.

Bover, O., and M. Izquierdo (2003), "Quality-adjusted Prices: Hedonic Methods and Implications for National Accounts," *Investigaciones Económicas* 27, 199–238.

Butler, J.S., Y. Chang and A. Crews Cutts (2005), "Revision Bias in Repeat-Sales Hone Price Indices," Freddie Mac Working Paper No. 05–03.

Calhoun, C.A. (1996), "OFHEO House Price Indexes: HPI Technical Description," Office of Federal Housing Enterprise Oversight, Washington, D.C.

Campbell, J.Y., and J.F. Cocco (2007), "How Do House Prices Affect Consumption? Evidence from Micro Data," *Journal of Monetary Economics* 54(3), 591–621.

Carare, A., and M.R. Stone (2003), "Inflation Targeting Regimes," IMF Working Paper WP 03/9.

Case, B., H.O. Pollakowski and W.M. Wachter (1991), "On Choosing Between House Price Index Methodologies," *American Real Estate and Urban Economics Association Journal* 19, 286–307.

Case, B., H. Pollakowski, and S. Wachter (1997), "Frequency of Transaction and House Price Modeling," *Journal of Real Estate Finance and Economics* 14(1–2), 173–87.

Case, B., and J.M. Quigley (1991), "The Dynamics of Real Estate Prices," *Review of Economics and Statistics* 22, 50–58.

Case, K.E., J.M. Quigley, and R.J. Shiller (2001), "Comparing Wealth Effects: The Stock Market Versus the Housing Market." National Bureau of Economic Research Working Paper Series, No. 8606.

Case, K.E., and R.J. Shiller (1987), "Prices of Single-Family Homes Since 1970: New Indexes for Four Cities," *New England Economic Review*, September–October, 45–56.

Case, K.E., and R.J. Shiller (1989), "The Efficiency of the Market for Single-family Homes," *The American Economic Review*, 79 (1), 125–37.

Case, K.E., and R.J. Shiller (1990), "Forecasting Price and Excess Returns in the Housing Market," *American Real Estate and Urban Economics Association Journal* 18, 253–73.

Case, B., and S. Wachter (2005), "Residential Real Estate Price Indices as Financial Soundness Indicators: Methodological Issues," p. 197–211 in *Real Estate Indicators and Financial Stability*, BIS Papers No. 21, Bank for International Settlements, Washington, D.C.

Chinloy, P., M. Cho and I.F. Megbolugbe (1997), "Appraisals, Transactions Incentives and Smoothing," *Journal of Real Estate Finance and Economics* 14(1), 45–55.

Chiodo, A.J., R. Hernandez-Murillo, and M.T. Owyang (2010), "Non-linear Effects of School Quality on House Prices," *Federal Reserve Bank of St. Louis Review*, May/June, 185–204.

Cho, M. (1996), "House Price Dynamics: A Survey of Theoretical and Empirical Issues," *Journal of Housing Research* 7(2), 145–172.

Clapham E., P. Englund, J.M. Quigley and C.L. Redfearn (2006), "Revisiting the Past and Settling the Score: Index Revision for House Price Derivatives," *Real Estate Economics* 34(2), 275–302.

Clapp, J.M. (1980), "The Elasticity of Substitution for Land: The Effects of Measurement Errors," *Journal of Urban Economics* 8, 255–63.

Clapp, J.M., and C. Giaccotto (1992), "Estimating Price Trends for Residential Property: A Comparison of Repeat Sales and Assessed Value Methods," *Journal of Real Estate Finance and Economics* 5(4), 357–74.

Clapp, J.M., and C. Giaccotto (1998), "Price Indices Based on the Hedonic Repeat-Sale Method: Application to the Housing Market," *Journal of Real Estate Finance and Economics* 16(1), 5–26.

Clapp, J.M., and C. Giaccotto (1999), "Revisions in Repeat Sales Price Indices: Here Today, Gone Tomorrow?" *Real Estate Economics* 27(1), 79–104.

Clapp, J.M., C. Giaccotto and D. Tirtiroglu (1991), "Housing Price Indices Based on All Transactions Compared to Repeat Subsamples," *AREUEA Journal* 19(3), 270–85.

Cleassens, S.M., A. Kose and M.E. Terrones (2008), "What Happens During Recessions, Crunches and Busts?" IMF Working Paper 08/274, International Monetary Fund, Washington, D.C.

Colwell, P.F. (1998), "A Primer on Piecewise Parabolic Multiple Regression Analysis via Estimations of Chicago CBD Land Prices," *Journal of Real Estate Finance and Economics* 17(1), 87–97.

Colwell, P.F., and G. Dilmore (1999), "Who was First? An Examination of an Early Hedonic Study," *Land Economics* 75(4), 620–26.

Congressional Budget Office (2007), "Housing Wealth and Consumer Spending," background paper, January. Disponible à http://www.cbo.gov.

Court, A.T. (1939), "Hedonic Price Indexes with Automotive Examples," in *The Dynamics of Automobile Demand*, General Motors Corporation, New York, 99–117.

Crone, T.M., and R.P. Voith (1992), "Estimating House Price Appreciation: A Comparison of Methods," *Journal of Housing Economics* 2(4), 324–38.

Crone, T.M., L.I. Nakamura and R.P. Voith (2009), "Hedonic Estimates of the Cost of Housing Services: Rental and Owner Occupied Units," p. 67–84 in *Price and Productivity Measurement, Volume 1: Housing*, W.E. Diewert, B.M. Balk, D. Fixler, K.J. Fox and A.O. Nakamura (eds.), Trafford Press and www.indexmeasures.com.

Dale-Johnson, D., (1982), "An Alternative Approach to Housing Market Segmentation Using Hedonic Price Data," *Journal of Urban Economics* 11(3), 311–32.

Davidson, R., and J.G. MacKinnon (1993), *Estimation and Inference in Econometrics*, Oxford: Oxford University Press.

Davis, M.A., and J. Heathcote (2007), "The Price and Quantity of Residential Land in the United States," *Journal of Monetary Economics* 54, 2595–2620.

Davis, M.A., and M.G. Palumbo (2008), "The Price of Residential Land in Large US Cities," *Journal of Urban Economics* 63, 352–84.

Destatis (2009), *Handbook on the Application of Quality Adjustment Methods in the Harmonised Index of Consumer Prices*, Statistics and Science, Vol. 13, Wiesbaden: Federal Statistical Office of Germany.

Diewert, W.E. (1974), "Intertemporal Consumer Theory and the Demand for Durables," *Econometrica* 42, 497–516.

Diewert, W.E. (1980), "Aggregation Problems in the Measurement of Capital," p. 433–528 in *The Measurement of Capital*, D. Usher (ed.), Chicago: The University of Chicago Press.

Diewert, W.E. (1983), "The Treatment of Seasonality in a Cost of Living Index," p. 1019–45 in *Price Level Measurement*, W.E. Diewert and C. Montmarquette (eds.), Ottawa: Statistics Canada.

Diewert, W.E. (1998), "High Inflation, Seasonal Commodities and Annual Index Numbers," *Macroeconomic Dynamics* 2, 456–71.

Diewert, W.E. (1999), "Index Number Approaches to Seasonal Adjustment," *Macroeconomic Dynamics* 3, 1–21.

Diewert, W.E. (2002), "Harmonized Indexes of Consumer Prices: Their Conceptual Foundations," *Swiss Journal of Economics and Statistics* 138, 547–637.

Diewert, W.E. (2003a), "Hedonic Regressions: A Consumer Theory Approach," p. 317–48 in *Scanner Data and Price Indexes*, Studies in Income and Wealth, Volume 64, R.C. Feenstra and M.D. Shapiro (eds.), NBER and University of Chicago Press.

Diewert, W.E. (2003b), "Hedonic Regressions: A Review of Some Unresolved Issues," paper presented at the 7th Meeting of the Ottawa Group, Paris, May 27–29.

Diewert, W.E. (2005), "Issues in the Measurement of Capital Services, Depreciation, Asset Price Changes and Interest Rates," p. 479–542 in *Measuring Capital in the New Economy*, C. Corrado, J. Haltiwanger and D. Sichel (eds.), Chicago: University of Chicago Press.

Diewert, W.E. (2009a), "Durables and Owner Occupied Housing in a Consumer Price Index," p. 445–500 in *Price Index Concepts and Measurement*, W.E. Diewert, J. Greenlees and C. Hulten (eds.), Studies in Income and Wealth, Volume 70, Chicago: The University of Chicago Press.

Diewert, W.E. (2009b), "The Paris OECD–IMF Workshop on Real Estate Price Indexes: Conclusions and Future Directions," p. 87–116 in *Price and Productivity Measurement: Volume 1—Housing*, Trafford Press and www.indexmeasures.com.

Diewert, W.E. (2010), "Alternative Approaches to Measuring House Price Inflation," Discussion Paper 10–10, Department of Economics, University of British Columbia, Vancouver, Canada, V6T 1Z1, December.

Diewert, W.E., Y. Finkel and Y. Artsev (2009), "Empirical Evidence on the Treatment of Seasonal Products: The Israeli Experience," p. 53–78 in *Price and Productivity Measurement: Volume 2; Seasonality*, W.E. Diewert, B.M. Balk, D. Fixler, K.J. Fox and A.O. Nakamura (eds.), Trafford Press.

Diewert, W.E., J. de Haan and R. Hendriks (2010), "The Decomposition of a House Price index into Land and Structures Components: A Hedonic Regression Approach," Discussion Paper 10–01, Department of Economics, University of British Columbia, Vancouver, Canada, V6T 1Z1.

Diewert, W.E., J. de Haan and R. Hendriks (2011), "Hedonic Regressions and the Decomposition of a House Price Index into Land and Structure Components," Discussion Paper 11–01, Department of Economics, The University of British Columbia, Vancouver, Canada, V6T 1Z1.

Diewert, W.E., S. Heravi and M. Silver (2009), "Hedonic Imputation Versus Time Dummy Hedonic Indexes," p. 161–96 in *Price Index Concepts and Measurement*, W.E. Diewert, J. Greenlees and C. Hulten (eds.), NBER Studies in Income and Wealth, Chicago: University of Chicago Press.

Diewert, W.E., and D.A. Lawrence (2000), "Progress in Measuring the Price and Quantity of Capital," p. 273–326 in *Econometrics Volume 2: Econometrics and the Cost of Capital: Essays in Honor of Dale W. Jorgenson*, L.J. Lau (ed.), Cambridge, MA: The MIT Press.

Diewert, W.E., and P. von der Lippe (2010), "Notes on Unit Value Index Bias," *Journal of Economics and Statistics* 230(6), 690–708.

Diewert, W.E., and A.O. Nakamura (2009), "Accounting for Housing in a CPI," p. 7–32 in *Price and Productivity Measurement, Volume 1: Housing*, W.E. Diewert, B.M. Balk, D. Fixler, K.J. Fox and A.O. Nakamura (eds.), Trafford Press and www.indexmeasures.com.

Diewert, W.E., A.O. Nakamura and L.I. Nakamura (2009), "The Housing Bubble and a New Approach to Accounting for Housing in A CPI," *Journal of Housing Economics* 18, 156–71.

Dodge, D. (2003), "It All Starts with the Data," Remarks by David Dodge, Governor of the Bank of Canada to the Conference of European Statisticians, Geneva. Disponible à www.bankofcanada.ca/en/speeches/2003.

Dreiman, M.H., and A. Pennington-Cross (2004), "Alternative Methods of Increasing the Precision of Weighted Repeat Sales House Prices Indices," *Journal of Real Estate Finance and Economics* 28(4), 299–317.

Duffy, D. (2009), "Measuring House Price Change," Economic and Social Research Institute, Working Paper 291.

Edelstein, R.H., and D.C. Quan (2006), "How Does Appraisal Smoothing Bias Real Estate Returns Measurement?," *Journal of Real Estate Finance and Economics* 32(1), 41–60.

Eiglsperger, M. (2010), "Residential Property Price Statistics for the Euro Area and the European Union," in *Housing Market Challenges in Europe and the United States*, P. Arestis, P. Mooslechner and K. Wagner (eds.), Basingstoke: Palgrave Macmillan.

Englund, P., J.M. Quigley and C.L. Redfearn (1998), "Improved Price Indexes for Real Estate: Measuring the Course of Swedish Housing Prices," *Journal of Urban Economics* 44(2), 171–96.

Eurostat (2001), Handbook on Price and Volume Measures in the National Accounts. Luxembourg: Office for Official Publications of the European Communities.

Eurostat (2005), "Final Report of the Joint NA-PPP Task Force on Dwelling Services," paper presented at the Meeting of the Eurostat Working Parties on National Accounts and PPP, November 23, Luxembourg.

Eurostat (2012), *Technical Manual on Owner-Occupied Housing for Harmonised Index of Consumer Prices*, Draft 2.0, March 2012. See also the Statistics Explained article: Housing price statistics—house price index at http://epp.eurostat.ec.europa.eu/statistics_explained.

Eurostat (2012), *Experimental House price indices in the Euro Area and the European Union* at http://epp.eurostat.ec.europa.eu/portal/page/portal/hicp/methodology/hps/house_price_index.

Eurostat, Fonds monétaire international, OCDE, Nations Unies, Banque mondiale (1993), *Système de comptabilité nationale 1993*, Luxembourg : Eurostat. Cité comme *SCN (1993)*.

Eurostat, Fonds monétaire international, OCDE, Nations Unies, Banque mondiale (2009), *Système de comptabilité nationale 2008*, Luxembourg : Eurostat. Cité comme *SCN (2008)*.

Fenwick, D. (2005), "Statistics on Real Estate Prices: The Need for a Strategic Approach," p. 368–72 in *Real Estate Indicators and Financial Stability*, BIS Papers No. 21, Bank for International Settlements, Washington, D.C.: The International Monetary Fund.

Fenwick, D. (2006), "Real Estate Prices: The Need for a Strategic Approach to the Development of Statistics to Meet User Needs," paper presented at the OECD–IMF Workshop on Real Estate Price Indices held in Paris, November 6–7, 2006.

Finocchiaro, D., and V. Queijo von Heideken (2007), "Do Central Banks React to House Prices?," Working Paper Series 217, Sveriges Riksbank (Central Bank of Sweden), (revised 01 Aug 2009).

FMI (2006), *Compilation Guide on Financial Soundness Indicators*, Washington, D.C.: IMF.

FMI (2007), "Review of Exchange Arrangements, Restrictions, and Controls," Disponible à http://www.imf.org/external/np/pp/2007/eng.

FMI, OIT, OCDE, CEE-NU et Banque mondiale (2009), *Export and Import Price Index Manual,* Washington, D.C. : FMI. Cité comme Manuel de l'XMPI (2009).

Francke, M.K. (2008), "The Hierarchical Trend Model," p. 164–80 in *Mass Appraisal Methods: An International Perspective for Property Valuers*, T. Kauko and M. Damato (eds.), Oxford: Wiley-Blackwell.

Francke, M.K. (2010), "Repeat Sales Index for Thin Markets: A Structural Time Series Approach," *Journal of Real Estate Finance and Economics* 41(1), 24–52.

Francke, M.K., and G.A. Vos (2004), "The Hierarchical Trend Model for Property Valuation and Local Price Indices," *Journal of Real Estate Finance and Economics* 28(2/3), 179–208.

Garner, T.I., and K. Short (2009), "Accounting for Owner-Occupied Dwelling Services: Aggregates and Distributions," *Journal of Housing Economics* 18, 273–79.

Garner, T.I., and R. Verbrugge (2009a), "The Puzzling Divergence of Rents and User Costs, 1980–2004: Summary and Extensions," p. 125–46 in *Price and Productivity Measurement, Volume 1: Housing*, W.E. Diewert, B.M. Balk, D. Fixler, K.J. Fox and A.O. Nakamura (eds.), Trafford Press and www.indexmeasures.com.

Garner, T.I., and R. Verbrugge (2009b), "Reconciling User Costs and Rental Equivalence: Evidence from the US Consumer Expenditure Survey," *Journal of Housing Economics* 18, 172–92.

Gatzlaff, D.H., and D. Ling (1994), "Measuring Changes in Local House Prices: An Empirical Investigation of Alternative Methodologies," *Journal of Urban Economics* 35(2), 221–44.

Gatzlaff, D.H., and D.R. Haurin (1997), "Sample Selection Bias and Repeat-Sales Index Estimates," *Journal of Real Estate Finance and Economics* 14(1), 33–50.

Gelfand, A.E., M.D. Ecker, J.R. Knight and C.F. Sirmans (2004), "The Dynamics of Location in Home Price," *Journal of Real Estate Finance and Economics* 29(2), 149–66.

Geltner, D. (1996), "The Repeated-Measures Regression-Based Index: A Better Way to Construct Appraisal-Based Indexes of Commercial Property Value," *Real Estate Finance* 12(4), 29–35.

Geltner, D., and H. Pollakowski (2006), "A Set of Indices for Trading Commercial Real Estate Based on the Real Capital Analytics Transaction Prices Database," Working Paper, MIT Center for Real Estate: CREDL.

Goetzmann, W.N. (1992), "The Accuracy of Real Estate Indices: Repeat Sales Estimators," *Journal of Real Estate Finance and Economics* 5(1), 5–53.

Goetzmann, W., and L. Peng (2002), "The Bias of RSR Estimator and the Accuracy of Some Alternatives," *Real Estate Economics* 30(1), 13–39.

Goetzmann, W.N., and M. Spiegel (1997), "A Spatial Model of Housing Returns and Neighborhood Substitutability," *Journal of Real Estate Finance and Economics* 14(1–2), 11–31.

Goodhart, C. (2001), "What Weight Should Be Given to Asset Prices in Measurement of Inflation?" *Economic Journal* 111, 335–56.

Goodhart, C., and B. Hofmann (2007), "Financial Conditions Indices," in *House Prices and the Macroeconomy: Implications for Banking and Price Stability*, Charles Goodhart (ed.), Oxford: Oxford University Press.

Goodman, A.C., and T.G. Thibodeau (2003), "Housing Market Segmentation and Hedonic Prediction Accuracy," *Journal of Housing Economics* 12(3), 181–201.

Gouriéroux, C., and A. Laferrère (2009), "Managing Hedonic House Price Indices: The French Experience," *Journal of Housing Economics* 18, 206–13.

Griliches, Z. (1961, 1971a), "Hedonic Price Indexes for Automobiles: An Econometric Analysis for Quality Change," in Z. Griliches (eds.), *Price Indexes and Quality Change: Studies in New Methods of Measurement*, Cambridge MA: Harvard University Press.

Griliches, Z. (1971b), "Introduction: Hedonic Price Indexes Revisited," p. 3–15 in *Price Indexes and Quality Change*, Z. Griliches (ed.), Cambridge MA: Harvard University Press.

Gudnason, R. (2004), "Simple User Costs and Rentals," paper presented at the 8th Ottawa Group Meeting, Helsinki, August 23–25.

Gudnason, R., and G. Jónsdóttir (2006), "House Price Index, Market Prices and Flow of Services Methods," paper presented at the OECD–IMF Workshop on Real Estate Price Indices held in Paris, November 6–7, 2006.

Gudnason, R., and G.R. Jónsdóttir (2009), "Owner Occupied Housing in the Icelandic CPI," p. 85–88 in *Price and Productivity Measurement, Volume 1: Housing*, W.E. Diewert, B.M. Balk, D. Fixler, K.J. Fox and A.O. Nakamura (eds.), Trafford Press and www.indexmeasures.com.

Gyourko, J., and A. Saiz (2004), "Reinvestment in the Housing Stock: The Role of Construction Costs and the Supply Side," *Journal of Urban Economics* 55, 238–56.

de Haan, J. (2004), "Direct and Indirect Time Dummy Approaches to Hedonic Price Measurement," *Journal of Social and Economic Measurement* 29, 427–43.

de Haan, J. (2009), "Comment on Hedonic Imputation versus Time Dummy Hedonic Indexes," p. 196–200 in *Price Index Concepts and Measurement*, W.E. Diewert, J.S. Greenlees and C.R. Hulten (eds.), Studies in Income and Wealth 70, Chicago: University of Chicago Press.

de Haan, J. (2010a), "Hedonic Price Indexes: A Comparison of Imputation, Time Dummy and Other Approaches," *Journal of Economics and Statistics* 230(6), 772–91.

de Haan, J. (2010b), "An Appraisal–Based Generalized Regression Estimator of House Price Change," paper presented at the Economic Measurement Group Workshop, 1–3 December 2010, Sydney, Australia.

de Haan, J., and H. van der Grient (2011), "Eliminating Chain Drift in Price Indexes Based on Scanner Data," *Journal of Econometrics* 161, 36–46.

de Haan, J., E. van der Wal and P. de Vries (2009), "The Measurement of House Prices: A Review of the Sales Price Appraisal Method," *Journal of Economic and Social Measurement* 34(2–3), 51–86.

Halvorsen, R., and H.O. Pollakowski (1981), "Choice of Functional Form for Hedonic Price Equations," *Journal of Urban Economics* 10, 37–49.

Haurin, D.R., and P.H. Hendershott (1991), "House Price Indexes: Issues and Results," *American Real Estate and Urban Economics Association Journal* 19, 259–69.

Heston, A., and A.O. Nakamura (2009a), "Reported Prices and Rents of Housing: Reflections of Costs, Amenities or Both?" p. 117–24 in *Price and Productivity Measurement, Volume 1: Housing*, W.E. Diewert, B.M. Balk, D. Fixler, K.J. Fox and A.O. Nakamura (eds.), Trafford Press and www.indexmeasures.com.

Heston, A., and A.O. Nakamura (2009b), "Questions about the Equivalence of Market Rents and User Costs for Owner Occupied Housing," *Journal of Housing Economics* 18, 273–79.

Hicks, J.R. (1946), *Value and Capital*, Second Edition, Oxford: Clarendon Press.

Hill, R.C., J.R. Knight and C.F. Sirmans (1997), "Estimating Capital Asset Price Indexes," *Review of Economics and Statistics* 79(2), 226–33.

Hill, R.J. (2004), "Superlative Index Numbers: Not All of Them Are Super," *Journal of Econometrics* 130(1), 25–43.

Hill, R.J (2011), "Hedonic Price Indexes for Housing," OECD Statistics Working Paper 2011/1, OECD Publishing, Paris.

Hill, R.J., and D. Melser (2008), "Hedonic Imputation the Price Index Problem: An Application to Housing," *Economic Inquiry* 46(4), 593–609.

Hill, R.J., D. Melser and I. Syed (2009), "Measuring a Boom and Bust: The Sydney Housing Market 2001–2006," *Journal of Housing Economics* 18, 193–205.

Hill, R.J., D. Melser and B. Reid (2010), "Hedonic Imputation with Geospatial Data: An Application of Splines to the Housing Market," Mimeo.

Hill, T.P. (1988), "Recent Developments in Index Number Theory and Practice," *OECD Economic Studies* 10, 123–48.

Hill, T.P. (1993), "Price and Volume Measures," p. 379–406 in *SNA (1993).*

Hoffmann, J., and C. Kurz (2002), "Rent Indices for Housing in West Germany: 1985 to 1998," Discussion Paper 01/02, Economic Research Centre of the Deutsche Bundesbank, Frankfurt.

Hoffmann, J., and J.R. Kurz-Kim (2006), "Consumer Price Adjustment Under the Microscope: Germany in a Period of Low Inflation," Discussion Paper Series 1, Economic Studies, No. 16/2006, Deutsche Bundesbank, Frankfurt am Main.

Hoffmann, J., and A. Lorenz (2006), "Real Estate Price Indices for Germany: Past, Present and Future," paper presented at the OECD–IMF Workshop on Real Estate Price Indexes, Paris, November 6–7.

Hwang, M., and J.M. Quigley (2004), "Selectivity, Quality Adjustment and Mean Reversion in the Measurement of House Values," *Journal of Real Estate and Finance and Economics* 28(2/3), 161–78.

Ivancic, L., W.E. Diewert and K.J. Fox (2011), "Scanner Data, Time Aggregation and the Construction of Price Indexes," *Journal of Econometrics* 161, 24–35.

Jansen, S.J.T., P. de Vries, H.C.C.H. Coolen, C.J.M. Lamain and P. Boelhouwer (2008), "Developing a House Price Index for the Netherlands," *Journal of Real Estate and Finance and Economics* 37(2), 163–86.

Jarociński, M., and F.R. Smets (2008), "House Prices and the Stance of Monetary Policy," *Federal Reserve Bank of St. Louis Review* 90(4), 339–65.

Katz, A.J. (2009), "Estimating Dwelling Services in the Candidate Countries: Theoretical and Practical Considerations in Developing Methodologies Based on a User Cost of Capital Measure," p. 33–50 in *Price and Productivity Measurement, Volume 1: Housing*, W.E. Diewert, B.M. Balk, D. Fixler, K.J. Fox and A.O. Nakamura (eds.), Trafford Press and www.indexmeasures.com.

Knight, J.R., J. Dombrow and C.F. Sirmans (1995), "A Varying Parameters Approach to Constructing House Price Indexes," *Real Estate Economics* 23 (2), 187–205.

Koev, E., and J.M.C. Santos Silva (2008), "Hedonic Methods for Decomposing House Price Indices into Land and Structure Components," unpublished paper, Department of Economics, University of Essex, England, October.

Lancaster, K.J. (1996), "A New Approach to Consumer Theory," *Journal of Political Economy* 74, 132–57.

Lebow, D.E., and J.B. Rudd (2003), "Measurement Error in the Consumer Price Index: Where do we Stand?," *Journal of Economic Literature* 41, 159–201.

Leishman, C. (2000), "Estimating Local Housing Market Price Indices Using Land Registry Data," *Housing Finance* 47 (August), 55–60.

Leishman, C., C. Watkins and W.D. Fraser (2002), "The Estimation of House Price Indices Based on Repeat Sales Regression Using Land Registry Data," Report to RICS Education Trust, London.

Leventis, A. (2006), "Removing Appraisal Bias from a Repeat Transactions House Price Index: A Basic Approach," paper presented at the OECD–IMF Workshop on Real Estate Price Indices held in Paris, November 6–7, 2006.

Li, W., M. Prud'homme and K. Yu (2006), "Studies in Hedonic Resale Housing Price Indices," paper presented at the OECD–IMF Workshop on Real Estate Price Indices held in Paris, November 6–7, 2006.

Lim, S., and M. Pavlou (2007), "An Improved National House Price Index Using Land Registry Data," RICS Research Paper Series, Vol. 7, No. 11.

Luüs, C. (2005), "The Absa Residential Property Market Database for South Africa: Key Data Trends and Implications," p. 149–170 in *Real Estate Indicators and Financial Stability*, BIS Papers No. 21, Bank for International Settlements, Washington, D.C.: IMF.

Makaronidis, A., and K. Hayes (2006), "Owner Occupied Housing for the HICP," paper presented at the OECD–IMF Workshop on Real Estate Price Indices, Paris, November 6–7, 2006.

Malpezzi, S. (2003), "Hedonic Pricing Models: A Selective and Applied Review," in *Housing Economics and Public Policy*, T.O'Sullivan and K. Gibb (eds.), Malder, MA: Blackwell.

Malpezzi, S., L. Ozanne and T. Thibodeau (1987), "Microeconomic Estimates of Housing Depreciation," *Land Economics* 63, 372–85.

Mark, J.H., and M.A. Goldberg (1984), "Alternative Housing Price Indices: An Evaluation," *American Real Estate and Urban Economics Association Journal* 12(1), 30–49.

Marola, B., D. Santos, R. Evangelista and V. Guerreiro (2012), "Working Towards a Comparable House Price Inflation Measure in Europe," paper presented at the UNECE Meeting of the Group of Experts on Consumer Price Indices held in Geneva, 30 May–1 June 2012 (www.unece.org/stats/).

McDonald, J.F. (1981), "Capital-Land Substitution in Urban Housing: A Survey of Empirical Estimates," *Journal of Urban Economics* 9, 190–211.

McDonald, C., and M. Smith, (2009), "Developing stratified housing price measures for New Zealand," Discussion Paper No. 2009/07, Reserve Bank of New Zealand, Wellington.

McMillen, D.P. (2003), "The Return of Centralization to Chicago: Using Repeat Sales to Identify Changes in House Price Distance Gradients," *Regional Science and Urban Economics* 33, 287–304.

Meese, R., and N. Wallace (1991), "Nonparametric Estimation of Dynamic Hedonic Price Models and the Construction of Residential Housing Price Indices," *American Real Estate and Urban Economics Association Journal* 19, 308–32.

Mudgett, B.D. (1955), "The Measurement of Seasonal Movements in Price and Quantity Indexes," *Journal of the American Statistical Association* 50, 93–98.

Muellbauer, J. (1974), "Household Production Theory, Quality and the 'Hedonic Technique'," *American Economic Review* 64, 977–94.

Muth, R.F. (1971), "The Derived Demand for Urban Residential Land," *Urban Studies* 8, 243–54.

Nabarro, R., and T. Key (2003), "Performance Measurement and Real Estate Lending Risk," paper presented at the IMF/BIS Conference on Real Estate Indicators and Financial Stability, October 2003.

Nations Unies (2009), *Guide pratique pour l'établissement d'indices des prix à la consommation*, New York et Genève: Nations Unies.

Ohnishi, T., T. Mizuno, C. Shimizu and T. Watanabe (2010), "On the Evolution of the House Price Distribution," Discussion Paper No. 61, Research Center for Price Dynamics.

OIT, FMI, OCDE, Eurostat, Nations Unies, Banque mondiale (2004), *Manuel de l'indice des prix à la consommation — Théorie et pratique*, Genève : OIT. Cité comme Manuel de l'IPC (2004).

OIT, FMI, OCDE, Eurostat, Nations Unies, Banque mondiale (2004), *Manuel de l'indice des prix à la production — Théorie et pratique*, Washington, D.C. : FMI. Cité comme Manuel de l'IPP (2004).

Pakes, A. (2003), "A Reconsideration of Hedonic Price Indexes with an Application to PCs," *American Economic Review* 93(5), 1576–93.

Palmquist, R.B. (1980), "Alternative Techniques for Developing Real Estate Price Indexes," *Review of Economics and Statistics* 62, 442–48.

Palmquist, R.B. (1982), "Measuring Environmental Effects on Property Values without Hedonic Regressions," *Journal of Urban Economics* 11, 333–47.

Pennington-Cross, A. (2005), "Aggregation Bias and the Repeat Sales Price Index," p. 323–35 in *Real Estate Indicators and Financial Stability*, BIS Papers No. 21, Bank for International Settlements, Washington, D.C.: IMF.

Phang, S.Y. (2004), "House Prices and Aggregate Consumption: Do They Move Together? Evidence from Singapore," *Journal of Housing Economics* 13, 101–19.

Plosser, C.I. (2007), "House Prices and Monetary Policy," Speech for the European Economics and Financial Centre Distinguished Speakers Series, July 11, 2007, Monograph, Federal Reserve Bank of Philadelphia.

Prasad, N., and A. Richards (2006), "Measuring House Price Growth—Using Stratification to Improve Median-based Measures," Research Discussion Paper No. 2006–04, Reserve Bank of Australia.

Prasad, N.L., and A. Richards (2008), "Improving Median Housing Price Indexes Through Stratification," *Journal of Real Estate Research* 30(1), 45–71.

Quigley, J.M. (1995), "A Simple Hybrid Model for Estimating Real Estate Price Indexes," *Journal of Housing Economics* 4(1), 1–12.

Reinhart, C.M., and K.S. Rogoff (2009), *This Time is Different; Eight Centuries of Financial Folly*, Princeton and Oxford: Princeton University Press.

Rosen, S. (1974), "Hedonic Prices and Implicit Markets: Product Differentiation in Pure Competition," *Journal of Political Economy* 82, 34–55.

Rossini, P., and P. Kershaw (2006), "Developing a Weekly Residential Price Index Using the Sales Price Appraisal Ratio, paper presented at the Twelfth Annual Pacific Rim Real Estate Society Conference, Auckland, 22–25 January.

Saarnio, M. (2006), "Housing Price Statistics at Statistics Finland," paper presented at the OECD–IMF Workshop on Real Estate Price Indices, Paris, 6–7 November 2006.

Samuelson, Paul A., and S. Swamy (1974), "Invariant Economic Index Numbers and Canonical Duality: Survey and Synthesis," *American Economic Review* 64, 566–93.

Scheffé, H. (1959), *The Analysis of Variance*, New York: John Wiley and Sons.

Shi, S., M. Young and B. Hargreaves (2009), "Issues in Measuring a Monthly House Price Index in New Zealand," *Journal of Housing Economics* 18(4), 336–50.

Shiller. R.J. (1991), "Arithmetic Repeat Sales Price Estimators," *Journal of Housing Economics* 1(1), 110–26.

Shiller, R.J. (1993a), "Measuring Asset Values for Cash Settlement in Derivative Markets: Hedonic Repeated Measures Indices and Perpetual Futures," *Journal of Finance* 48(3), 911–31.

Shiller, R.J. (1993b), *Macro Markets*, Oxford: Oxford University Press.

Shiller, R.J. (2005), "Comments on Session 'Aggregation Issues'," p. 336–39 in *Real Estate Indicators and Financial Stability*, BIS Papers No. 21, Bank for International Settlements, Washington, D.C.: The International Monetary Fund.

Shimizu, C., and K.G. Nishimura (2006), "Biases in Appraisal Land Price Information: The Case of Japan," *Journal of Property Investment & Finance* 24 (2), 150–75.

Shimizu, C., and K.G. Nishimura (2007), "Pricing Structure in Tokyo Metropolitan Land Markets and Its Structural Changes: Pre-bubble, Bubble, and Post-bubble Periods," *Journal of Real Estate Finance and Economics* 35(4), 495–96.

Shimizu, C., K.G. Nishimura and Y. Asami (2004), "Search and Vacancy Costs in the Tokyo Housing Market: Attempt to Measure Social Costs of Imperfect Information," *Regional and Urban Development Studies* 16(3), 210–30.

Shimizu, C., K.G. Nishimura and T. Watanabe (2010), "Housing Prices in Tokyo: A Comparison of Hedonic and Repeat Sales Measures," *Journal of Economics and Statistics* 230/6, 792–813.

Shimizu, C., H. Takatsuji, H. Ono and K.G. Nishimura (2010), "Structural and Temporal Changes in the Housing Market and Hedonic Housing Price Indices," *International Journal of Housing Markets and Analysis* 3(4), 351–68.

Shimizu, C., K.G. Nishimura and T. Watanabe (2011), "House Prices from Realtors, Magazines, and Government: A Statistical Comparison," Mimeo.

Silver, M. (2009a), "Do Unit Value Export, Import, and Terms of Trade Indices Represent or Misrepresent Price Indices?" IMF Staff Papers 56, 297–322, IMF, Washington, D.C.

Silver, M. (2009b), "Unit Value Indices," Chapter 2 in XMPI Manual (2009).

Silver, M. (2010), "The Wrongs and Rights of Unit Value Indices," *Review of Income and Wealth*, Series 56, Special Issue 1, S206–S223.

Standard and Poor's (2009), *S&P/Case–Shiller Home Price Indices; Index Methodology*, New York: Standard and Poor's.

Statistics Portugal (Instituto Nacional de Estatistica) (2009), "Owner-Occupied Housing: Econometric Study and Model to Estimate Land Prices, Final Report," paper presented to the Eurostat Working Group on the Harmonization of Consumer Price Indices," March 26–27, Luxembourg: Eurostat.

Steele, M., and R. Goy (1997), "Short Holds, the Distributions of First and Second Sales, and Bias in the Repeat-Sales Price Index," *Journal of Real Estate Finance and Economics* 14(1–2), 133–54.

Stone, R. (1956), *Quantity and Price Indices in National Accounts*, Paris: OECD.

Thibodeau, T.G. (1992), "Residential Real Estate Prices: 1974–1983," The Blackstone Company, Mount Pleasant (MI).

Thibodeau, T.G. (2003), "Marking Single-Family Property Values to Market," *Real Estate Economics* 31(1), 1–22.

Thorsnes, P. (1997), "Consistent Estimates of the Elasticity of Substitution between Land and Non-Land Inputs in the Production of Housing," *Journal of Urban Economics* 42, 98–108.

Triplett, J.E. (2006), *Handbook on Hedonic Indexes and Quality Adjustments in Price Indexes; Special Application to Information and Technology Products*, Directorate for Science, Technology and Industry, Paris: OECD.

Triplett, J.E., and R.J. McDonald (1977), "Assessing the Quality Error in Output Measures: The Case of Refrigerators," *The Review of Income and Wealth* 23(2), 137–56.

Tse, R.Y.C. (2002), "Estimating Neighbourhood Effects in House Prices: Towards a New Hedonic Model Approach," Urban Studies 39(7), 1165–80.

Turvey, R., (1989), *Consumer Price Indices: An ILO Manual*, Geneva: International Labour Office.

UK Department of the Environment (1982), "A New Index of Average House Prices," *Economic Trends* 348, 134–38.

van der Wal, E., D. ter Steege and B. Kroese (2006), "Two Ways to Construct a House Price Index for the Netherlands: The Repeat Sale and the Sale Price Appraisal Ratio," paper presented at the OECD–IMF Workshop on Real Estate Price Indices, Paris, November 6–7, 2006.

Verbrugge, R. (2008), "The Puzzling Divergence of Aggregate Rents and User Costs, 1980–2004," *The Review of Income and Wealth* 54, 671–99.

de Vries, P., J. de Haan, E. van der Wal and G. Mariën (2009), "A House Price Index Based on the SPAR Method," *Journal of Housing Economics* 18(3), 214–23.

Wallace, N.E., and R.A. Meese (1997), "The Construction of Residential Housing Price Indices: A Comparison of Repeat-Sales, Hedonic-Regression and Hybrid Approaches," *Journal of Real Estate Finance and Economics* 14(1–2), 51–73.

Wang, T., and P.M. Zorn (1997), "Estimating House Price Growth with Repeat Sales Data: What's the Aim of the Game?" *Journal of Housing Economics* 6(2), 93–118.

Wenzlick, R. (1952), "As I See the Fluctuations in the Selling Prices of Single Family Residences," *The Real Estate Analyst* 21 (December 24), 541–48.

Wolverton, M.L., and J. Senteza (2000), "Hedonic Estimates of Regional Constant Quality House Prices," *Journal of Real Estate Research* 19(3), 235–53.

Wood, R. (2005), "A Comparison of UK Residential House Price Indices," p. 212–27 in *Real Estate Indicators and Financial Stability*, BIS Papers No. 21, Bank for International Settlements, Washington, D.C.: The International Monetary Fund.

Wyngarden, H. (1927), *An Index of Local Real Estate Prices*, Michigan Business Studies, Volume 1, Number 2, Ann Arbor: University of Michigan.

Zhu, H. (2005), "The Importance of Property Markets for Monetary Policy and Financial Stability," p. 9–29 in *Real Estate Indicators and Financial Stability*, Volume 21, Bank for International Settlements (ed.).

Index

achat et vente d'un logement
- délais d'acquisition et de vente (Japon) 10.56–57
- processus d'achat et de vente 9.3–9.6, 10.59

actualisation (g)
actualité 9.8
agents immobiliers 9.10–12
agrégat (g)
agrégat élémentaire (g)
agrégation (g)
- exemples d'agrégation 11.17–11.25
- première phase d'agrégation 4.13–15
- seconde phase d'agrégation 4.16–18
- ventes et stocks de logement 8.13–17

ajustement hédonique de la qualité 5.32–33
ajustement de la qualité des bâtiments 5.49–52
- dans différents pays 10.18

approche axiomatique (g)
approche économique (g)
approche par les dépenses engagées ou par les paiements (g), 9.30–31
approche par les paiements (g) 3.15
apurement des données (g) 5.9, 6.11–12, 7.22, 7.33
bâtiments
- prix des bâtiments 5.3

biais (g)
- biais de sélection de l'échantillon 4.3, 6.15–18
- biais de la valeur unitaire 4.15
- biais dû à l'omission de variables 5.6

biais de la valeur unitaire *voir* biais
biais de sélection de l'échantillon (g) *voir* biais
bien durable de consommation (g) 3.40
bien immobilier représentatif (g)
bien immobilier résidentiel (g)
bien normalisé 5.20
biens (g)
biens saisonniers (g)
bruit de transaction 6.18
cadre de Mudgett–Stone 4.42–45
caractéristiques (g) 3.18, 4–29, 5.1
- adjonction de caractéristiques liées aux bâtiments 8.10–12
- taille du bâtiment 4.34
- taille du terrain (parcelle) 4.34

chaînage (g)
changement de qualité (g)
- ajustement de la qualité du bâtiment 8.10–11
- indice des prix avec ajustement de la qualité des bâtiments 8.23

champ de l'indice (g), 12.7–9
comparabilité (entre les pays) 9.43–44, 10.13
composante (g)
compte de patrimoine 3.8
concept (ou approche) des acquisitions (nettes) (g) 3.15, 9.28
concept d'utilisation (g)
concordance de l'agrégation (g)
concordance des estimations mensuelles et trimestrielles 3.32–33
correction des variations saisonnières 3.37–38, 12.38–40
coûts d'usage
- biens durables en général 3.39–52
- des logements occupés par leur propriétaire 3.53–68

coût d'usage *ex ante* 3.48
coûts de transaction 3.14
couverture (g), 9.6, 9.8, 9.34–35
décomposition en composantes «terrain» et «bâtiments» 8.1–57
dépendance spatiale 5.7
dépréciation (g)
dépréciation dite du fiacre centenaire (ou de l'ampoule électrique) 3.51
dépréciation linéaire (ou géométrique) 3.51, 5.49–52
dépréciation nette 4.9, 8.5–8.9
- taux de dépréciation 3.48, 8.5

domaine (g) *voir* champ de l'indice
données administratives 9.1
données sur les prix à différents stades 9.5–7
échantillon (g)
échantillon et probabilité 4.34–35
erreur systématique *voir* biais
études de cas
- Afrique du Sud 10.78–93
- Allemagne 10.41–49
- Canada 10.27–40
- Colombie 10.72–77
- Inde 10.65–71
- Japon 10.50–58
- Royaume-Uni 10.59–64

exactitude 9.35
- compromis entre fréquence et exactitude 9.38

exigences des usagers 9.25
famille d'indices des prix de l'immobilier résidentiel 3.30
formation brute de capital 3.13–14
habitations existantes (g)
imputation hédonique (g) 5.25–34, 5.65–69, 11.40–44
- double imputation 5.27–29
  - indice de Fisher à double imputation 5.29, 5.68
  - indice de Laspeyres à double imputation 5.27, 5.66, 11.43
  - indice de Paasche à double imputation 5.28, 5.67
  - indices avec imputation reposant sur la moyenne géométrique 5.30–34
- indices avec imputation reposant sur la moyenne arithmétique 5.26–29
- modélisation hédonique 5.2–5

indice-chaîne (g) 4.19–20
indice de la médiane (g) 11.17
indice de la moyenne (g)
indice de Laspeyres géométrique (g)
indice de niveau inférieur (g) *voir* (première phase d') agrégation
indice des prix à la consommation (IPC) (g)
indice des prix de Jevons (g)
indice des prix de Laspeyres (g) 4.16–18
indice des prix de Lowe (g), 8.51–52
indice des prix de Paasche (g) 4.16–18
indice des stocks 4.21–24, 5.23–24, 8.15–17
- à imputation hédonique 8.50–52
indice du coût de la construction 8.37
indice (idéal) des prix de Fisher (g) 4.16–18, 11.19
indice superlatif (g) 11.19
indice symétrique (g)
indice Törnqvist-Theil 11.19
indices à base fixe 4.18
indices à pondération fixe (g)
informations tirées d'expertises 9.22
ligne brisée continue 8.25–30
logement occupé par son propriétaire (g)
logements traditionnels (dans les pays en développement), 9.48–54
loyer approximatif 3.50
manque d'efficacité (de la méthode des ventes répétées) 6.19–20
marché de l'investissement locatif 6.16
marché du logement informel (g), 9.48–54, 10.85–94
ménages (g)
mesures de tendance centrale 4.1, 11.3–10
métadonnées 10.8, 10.22–26
méthode de coût de production 8.2–4
méthode des fenêtres mobiles (g) 4.42–45, 8.43–48
méthode des loyers imputés (g) 3.15, 3.61
méthode de stratification (g) 3.20, 4.4–4.12, 12.16–21
- dans différents pays 10.17, 11.11–16
- indices hédoniques stratifiés 5.35–40
- segmentation du marché 4.11–12, 5.37
méthode des ventes répétées (g) 3.21, 6.1–32, 11.45–56, 12.29–32
- équation d'estimation des ventes répétées 6.4
- méthode des ventes répétées reposant sur la moyenne arithmétique 6.9
- modèles hybrides 6.26–27
- période de possession 6.7
- processus aléatoire gaussien 6.4
- révisions 6.21
- technique des moindres carrés pondérés 11.52–56
- utilisation de l'information sur les évaluations 6.20
- utilisation des informations sur les frais d'entretien et de rénovation 6.23
méthode du coût d'opportunité 3.70–72
méthode du coût d'opportunité financier 3.46
méthode du coût d'usage (g) 3.15, 3.40–52
- méthode simplifiée du coût d'usage 3.65
méthode hédonique des prix des caractéristiques (g), 5.20–24
méthode (hédonique) des variables indicatrices temporelles (g) 5.5, 5.11–18
- indice des variables indicatrices temporelles 5.11
méthode des périodes consécutives 11.38–39
- modèle de régression hédonique linéaire à variables indicatrices temporelles 5.57–60
- modèle hédonique linéaire avec ajustement de la qualité des bâtiments 5.61–63
- modèle log-linéaire à variables indicatrices temporelles 5.45–48, 11.32–39
- modèle log-log à variables indicatrices temporelles 5.53–55
méthode RPVE (g) 7.4–35, 12.33–37
- biais du changement qualitatif 7.15–16
- cadre de régression généralisée 7.28
- estimateur d'échantillon d'un IPIR de stock 7.13
- hypothèses de modèle 7.26–27
- indice RPVE arithmétique 7.10–12
- indice RPVE de type Paasche 7.8
- modèle descriptif de régression 7.25
- ratio de l'estimation des prix de vente 7.7
- méthodes basées sur l'estimation 3.24, 7.1–6
- modèle des constructeurs 8.3–4
modèles hybrides (g) *voir* méthode des ventes répétées
modèle paramétrique (linéaire ou log-linéaire) 5.2, 11.28–31
- imputation unique 5.27
moyenne mobile 4.45
multicolinéarité 5.8, 8.3, 8.24
non unicité (g)
pays en développement 9.48–54
période de référence de l'indice (g)
période de référence des prix (g)
période de référence ou de base (g)
période en cours, ou période de comparaison (g)
périodicité de l'indice des prix de l'immobilier résidentiel 3.29–31, 9.38–39
point de rupture 4.2
politique de révision 3.35–36
pondérations (g)
- logements non formels (informels ou traditionnels) 10.89–90
- période de référence des pondérations (g)
- pondérations de stocks 4.6
- pondérations des ventes 4.6
- sources de données pour calculer les pondérations 9.45–47
poststratification (g) *voir* stratification
prix de demande (g), 9.14–15
prix demandé (g), 9.10–13
prix de transaction (définitif)
- des cadastres et administrations fiscales 9.19–20
- des institutions de crédit hypothécaire 9.18

prix de vente (ou de transaction) (g)
prix de valorisation
- des administrations fiscales 9.21
- des institutions de crédit hypothécaire 9.16–17
prix imputé (g)
problème général de démarrage (d'une série) 7.35
production réelle du secteur des services immobiliers résidentiels 3.9–12
produit (g)
produits appariés (g)
rapport de prix (g)
ratio du loyer à la valeur (prix) *voir* taux de capitalisation
régression hédonique (g) 3.22, 12.22–28
- exemples de méthodes de régression hédonique 11.26–44
- utilisation de restrictions de monotonie 8.31–35
- utilisation des informations exogènes sur les bâtiments 8.36–40, 8.53–56
regroupement des données (transversales) 5.12–13
repondération (g)
reproductibilité 6.12
restrictions de monotonie 8.31–35
révisions 5.17, 9.40–42
saisonnalité (traitement dans un indice des prix de l'immobilier résidentiel) 4.42–46
sociétés de crédit hypothécaire 9.14–18
sources des données (dans différents pays) 10.12–13
stabilité des coefficients 11.37
statistiques conçues pour une finalité précise 9.7–8
- conformité des sources de données aux besoins 9.23–26
stock de logements (g)
- approximation de la valeur du stock de logements 8.51–52
Système de comptabilité nationale (SCN) (g)
- cadre conceptuel à l'élaboration des indices des prix de l'immobilier résidentiel 3.6–14, 12.4
taux d'inflation anticipés (pour bâtiments et terrains) 3.55, 3.59
taux de capitalisation 3.70–71
techniques de régression
- régression non linéaire 8.6
- régression pour la version pondérée des ventes répétées 6.13, 11.52–56
- régression selon la méthode des moindres carrés 5.5
- régression selon la méthode des moindres carrés pondérés 5.15
terrain
- décomposition en composantes «terrain» et «bâtiments» 8.1–57
- indice des prix des terrains à bâtir (Allemagne) 10.41–42
- prix des terrains 5.3
test d'identité (g)
théories du cycle de vie 6.16
utilisations des indices des prix de l'immobilier résidentiel
- comparaisons interrégionales et internationales 2.28–30
- composante du patrimoine 2.14–15
- déflateur dans la comptabilité nationale 2.21–23
- indicateur de la stabilité ou de la solidité financière 2.16–20
- indicateur macroéconomique de la croissance économique 2.8–11
- logements occupés par leur propriétaire dans les indices des prix à la consommation (IPC) 2.27
- politique monétaire et de ciblage de l'inflation 2.12–13
valeur (g)
valeur aberrante (g) *voir* apurement des données
valeur de marché (g)
valeur estimée ou évaluation (g) 7.2, 7.17–20
valeur unitaire ou valeur moyenne (g) 3.32, 4.15
variation pure de prix (g)
vérification (g) *voir* apurement des données

European Commission

Handbook on Residential Property Prices Indices (RPPIs)

Luxembourg: Publications Office of the European Union

2013 — 179 pp. — 21 x 29.7 cm

Theme: Economy and finance
Collection: Methodologies & Working papers

ISBN 978-92-79-25984-5
ISSN 1977-0375
doi:10.2785/34007
Cat. No KS-RA-12-022-EN-C

Manuel des indices des prix de l'immobilier résidentiel

Pour la plupart d'entre nous, l'achat d'un bien immobilier (logement) est la transaction la plus importante de notre vie. La résidence d'un ménage absorbe en effet la plus forte proportion de ses dépenses et constitue en même temps son actif le plus précieux. Les *indices des prix de l'immobilier résidentiel* (IPIR) mesurent le rythme d'évolution des prix de l'immobilier résidentiel au fil du temps.

Les IPIR sont des statistiques cruciales non seulement pour les particuliers et les ménages du monde entier, mais aussi pour les responsables des politiques économiques et monétaires. Pour les professionnels, ils servent, par exemple, à surveiller les déséquilibres macroéconomiques et les risques auxquels est exposé le secteur financier.

Pour la première fois, ce *Manuel* offre des orientations complètes pour l'établissement des IPIR et explique en détail les méthodes et les meilleures pratiques utilisées pour les calculer. Il examine aussi les concepts économiques et statistiques qui sous-tendent cette démarche et définit les principes qui régissent les choix méthodologiques et pratiques en la matière. Ce *Manuel* s'adresse avant tout aux statisticiens officiels chargés de produire les indices des prix de l'immobilier résidentiel, mais il sera utile aussi à toutes les parties qui s'intéressent au calcul de ces indices en offrant un cadre méthodologique et pratique harmonisé et en répondant ainsi aux critères généraux que doivent remplir les IPIR.

Le *Manuel des indices des prix de l'immobilier résidentiel* a été rédigé par d'éminents universitaires spécialistes de la théorie des indices et par des experts réputés du calcul des IPIR. La production de l'ouvrage a été coordonnée par Eurostat, le Bureau statistique de l'Union européenne, avec la collaboration de l'Organisation internationale du travail (OIT), le Fonds monétaire international (FMI), l'Organisation de coopération et de développement économiques (OCDE), la Commission économique des Nations Unies pour l'Europe et la Banque mondiale.

Office des publications

ISBN: 978-1-49830-156-5