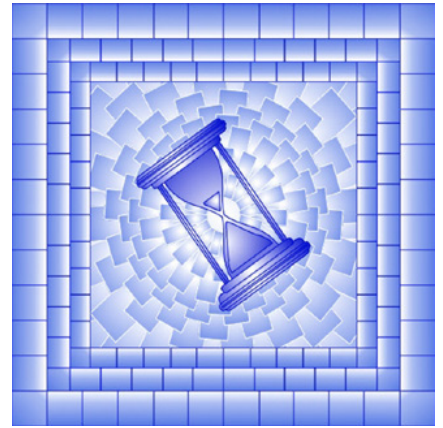


N° 62F0014M au catalogue  
ISBN 978-0-660-31807-3

# Méthodologie de l'Indice des prix des propriétés résidentielles (IPPR)

par Steve Martin

Date de diffusion : le 14 novembre 2019



Statistique  
Canada

Statistics  
Canada

Canada

---

## Comment obtenir d'autres renseignements

Pour toute demande de renseignements au sujet de ce produit ou sur l'ensemble des données et des services de Statistique Canada, visiter notre site Web à [www.statcan.gc.ca](http://www.statcan.gc.ca).

Vous pouvez également communiquer avec nous par :

**Courriel** à [STATCAN.infostats-infostats.STATCAN@canada.ca](mailto:STATCAN.infostats-infostats.STATCAN@canada.ca)

**Téléphone** entre 8 h 30 et 16 h 30 du lundi au vendredi aux numéros suivants :

- |   |                |
|---|----------------|
| • Service de renseignements statistiques                                    | 1-800-263-1136 |
| • Service national d'appareils de télécommunications pour les malentendants | 1-800-363-7629 |
| • Télécopieur   | 1-514-283-9350 |

### Programme des services de dépôt

- |                             |                |
|-----------------------------|----------------|
| • Service de renseignements | 1-800-635-7943 |
| • Télécopieur               | 1-800-565-7757 |

## Normes de service à la clientèle

Statistique Canada s'engage à fournir à ses clients des services rapides, fiables et courtois. À cet égard, notre organisme s'est doté de normes de service à la clientèle que les employés observent. Pour obtenir une copie de ces normes de service, veuillez communiquer avec Statistique Canada au numéro sans frais 1-800-263-1136. Les normes de service sont aussi publiées sur le site [www.statcan.gc.ca](http://www.statcan.gc.ca) sous « Contactez-nous » > « [Normes de service à la clientèle](#) ».

## Note de reconnaissance

Le succès du système statistique du Canada repose sur un partenariat bien établi entre Statistique Canada et la population du Canada, les entreprises, les administrations et les autres organismes. Sans cette collaboration et cette bonne volonté, il serait impossible de produire des statistiques exactes et actuelles.

Publication autorisée par le ministre responsable de Statistique Canada

© Sa Majesté la Reine du chef du Canada, représentée par le ministre de l'Industrie 2019

Tous droits réservés. L'utilisation de la présente publication est assujettie aux modalités de l'[entente de licence ouverte](#) de Statistique Canada.

Une [version HTML](#) est aussi disponible.

*This publication is also available in English.*

---

## Table des matières

<b>Introduction</b> .....	<b>4</b>
<b>1 Indice des prix des logements neufs (IPLN)</b> .....	<b>4</b>
1.1 Concepts et définitions.....	5
1.2 Données .....	5
1.3 Calcul de l'indice .....	5
1.3.1 Remplacement du modèle.....	6
<b>2 Indice des prix des appartements en copropriété neufs (IPACN)</b> .....	<b>7</b>
2.1 Concepts et définitions.....	7
2.2 Données .....	8
2.2.1 Échantillonnage.....	8
2.2.2 Nettoyage et filtrage.....	8
2.3 Calcul de l'indice .....	8
<b>3 Indice des prix de revente des propriétés résidentielles (IPRPR)</b> .....	<b>9</b>
3.1 Concepts et définitions.....	10
3.2 Données.....	10
3.2.1 Sources de données .....	10
3.2.2 Délai de collecte.....	11
3.2.3 Nettoyage et filtrage.....	11
3.3 Calcul de l'indice .....	12
3.3.1 La méthode des ventes répétées.....	12
3.3.2 Les indices GVR et AVR.....	13
3.3.3 Exemple élaboré de l'indice AVR.....	15
3.3.4 Représentativité de la population cible.....	17
3.3.5 Pondérations par l'inverse de la variance .....	17
3.4 Révision .....	18
3.4.1 Prise en compte de la révision dans le modèle de ventes répétées.....	18
3.4.2 Prise en compte de la révision en raison du délai de collecte.....	19
<b>4 Indice des prix des propriétés résidentielles (IPPR)</b> .....	<b>19</b>
<b>Références</b> .....	<b>20</b>

# Méthodologie de l'Indice des prix des propriétés résidentielles (IPPR)

par **Steve Martin**

## Introduction

Après la crise financière mondiale de 2008, le G-20 a désigné les indices des prix des biens immobiliers comme un indicateur important de la solidité financière. En lien avec ces efforts, les indices des prix des propriétés résidentielles constituent un ensemble de données de base nécessaires à l'analyse de la stabilité financière sous un nouveau niveau de la Norme spéciale de diffusion des données du FMI, appelée NSDD Plus. Afin de répondre à ces nouvelles exigences en matière de données et d'améliorer la pertinence des statistiques sur les prix des logements, le budget fédéral de 2016 a confié à Statistique Canada le mandat d'élaborer un indice officiel des prix des propriétés résidentielles (IPPR).

L'IPPR est un indice trimestriel, débutant au premier trimestre de 2017, qui couvre les prix des logements résidentiels neufs et en revente dans les régions métropolitaines de recensement (RMR) de Calgary, Montréal, Ottawa, Toronto, Vancouver et Victoria. Un indice agrégé des six RMR est aussi disponible. L'indice est composé de trois indices distincts produits à Statistique Canada. Les logements neufs sont couverts par l'Indice des prix des logements neufs (IPLN) et l'Indice des prix des appartements en copropriété neufs (IPACN), et les logements en revente sont couverts par l'Indice des prix de revente des propriétés résidentielles (IPRPR). Ces trois indices sont agrégés pour former l'IPPR.

Le présent document décrit les détails méthodologiques de l'IPLN, de l'IPACN et de l'IPRPR, ainsi que la façon dont ces trois indices sont agrégés pour former l'IPPR. Les sections 1 et 2 portent sur l'IPLN et l'IPACN, qui sont des indices de prix basés sur des enquêtes, alors que la section 3 décrit l'IPRPR, qui utilise une méthode de ventes répétées plus complexe. Comme le logement est un bien assez hétérogène, l'élaboration d'un indice de prix avec une interprétation de qualité constante est une importante considération méthodologique pour les trois indices. La section 4 décrit la façon dont l'IPLN, l'IPACN et l'IPRPR sont agrégés pour former l'IPPR.

## 1 Indice des prix des logements neufs (IPLN)

L'IPLN mesure la variation, au fil du temps, des prix de vente tels qu'établis par les constructeurs de nouveaux logements (maisons individuelles, jumelées et en rangée) dans 27 RMR. Aux fins de l'élaboration de l'IPPR, l'IPLN couvre les logements neufs à Calgary, Montréal, Ottawa, Toronto, Vancouver et Victoria. L'IPLN est un indice mensuel produit depuis janvier 1981. Pour produire un indice des prix de qualité constante, l'IPLN utilise une approche d'appariement de modèles, selon laquelle les prix des mêmes modèles de maisons sont comparés au fil du temps, et des ajustements explicites de la qualité sont effectués afin d'assurer la comparabilité des modèles. Les données sont recueillies mensuellement auprès des constructeurs dans le cadre d'une enquête au moyen d'un questionnaire électronique.

Contrairement à l'IPACN et à l'IPRPR, l'IPLN constitue sa propre série d'indices qui est distincte de l'IPPR<sup>1</sup>. L'IPPR utilise simplement les valeurs de l'IPLN à l'échelle municipale pour tenir compte des variations de prix des logements neufs. Par conséquent, la présente section porte sur les détails méthodologiques de l'IPLN en ce qui a trait à l'IPPR.

---

1. Voir la [BMDI](#) pour des plus amples renseignements sur l'IPLN.

## 1.1 Concepts et définitions

Le tableau 1.1 définit les concepts clés utilisés pour élaborer l'IPLN, du moins pour son utilisation dans l'IPPR.

**Tableau 1.1**  
**Concepts et définitions de l'IPLN**

Concept	Définition
Prix	Le prix de transaction ou le prix courant d'un modèle de maison déclaré par le constructeur au cours d'un mois donné, à l'exclusion de toute taxe de vente. Il s'agit du prix reçu par le constructeur, excluant les frais supplémentaires payés par l'acheteur.
Modèle	Plan d'étage et caractéristiques particulières d'une maison.
Échantillon	Voir la section 1.2.
Population cible	Tous les logements résidentiels neufs (maisons individuelles, jumelées et en rangée) mis en vente ou vendus à Calgary, Montréal, Ottawa, Toronto, Vancouver et Victoria au cours d'un mois donné.
Période de base de l'indice	Période pour laquelle l'indice est égal à 100. La période de base pour l'IPLN est décembre 2016=100.

## 1.2 Données

Les données de l'IPLN sont recueillies à partir d'une enquête auprès des constructeurs d'habitations. La base de sondage pour l'IPLN est l'Enquête sur les permis de bâtir de Statistique Canada, et l'enquête utilise un plan d'échantillonnage à plusieurs étapes dans lequel des modèles représentatifs sont sélectionnés à l'intérieur de l'échantillon à chaque étape. La première étape d'échantillonnage consiste à communiquer avec le premier 15 % des promoteurs immobiliers au sein d'une RMR, en fonction de la valeur de leurs permis de bâtir, pour déterminer s'ils font partie du champ de l'enquête. Cela permet de s'assurer que les constructeurs qui aménagent des lotissements au complet sur de vastes parcelles de terrains sont inclus dans l'échantillon. Une fois qu'un constructeur est identifié comme faisant partie du champ de l'enquête, il sélectionne le projet domiciliaire avec le plus grand nombre de lots toujours disponibles pour la vente au sein d'une RMR, et identifie jusqu'à trois modèles de maisons les plus vendus dans ce projet. Cela permet de s'assurer que les mêmes modèles peuvent être suivis au fil du temps dans le même projet domiciliaire et que ces modèles sont largement représentatifs de l'activité sur le marché pour les logements neufs.

Un questionnaire électronique est utilisé pour recueillir des données sur les prix de ces modèles chaque mois. Si un modèle ne se vend pas au cours d'un mois donné, on demande au constructeur de fournir le prix annoncé. L'échantillon est mis à jour périodiquement à mesure que les projets domiciliaires se vendent et que les constructeurs entrent sur le marché et en sortent. Les données recueillies auprès des constructeurs immobiliers sont examinées manuellement pour en vérifier la cohérence et l'exhaustivité, et certaines données peuvent être modifiées ou supprimées en fonction du jugement.

## 1.3 Calcul de l'indice

L'IPLN est un indice à modèles appariés assez simple. Les prix sont stratifiés par RMR, constructeur et modèle pour produire un rapport de prix pour chaque modèle que chaque constructeur déclare dans l'enquête. La valeur de toute promotion ou de toutes améliorations ou ajout est soustraite du prix d'un modèle avant le calcul d'un rapport de prix. À condition que les modèles de maison ne changent pas au fil du temps, cette collecte de rapports de prix a une interprétation de qualité constante. Les rapports de prix pour chaque modèle sont ensuite agrégés à l'échelle des RMR à l'aide d'un indice de Jevons. Bien que l'IPLN soit calculé mensuellement, la moyenne des valeurs des trois indices au cours d'un trimestre est établie pour produire un indice trimestriel pour l'IPPR.

Pour rendre le calcul de l'indice explicite, supposons que  $p_{mbt}$  est le prix du modèle  $m$  par le constructeur  $b$  au moment  $t$ . Ces prix de modèle sont utilisés pour calculer un rapport de prix entre la période  $t-1$  et la période  $t$ ,

$p_{mbt} / p_{mbt-1}$ , pour chaque modèle que chaque constructeur déclare dans l'enquête. Pour produire un indice à l'échelle des RMR, les rapports de prix pour tous les modèles par tous les constructeurs sont agrégés avec un indice de Jevons

$$I_t^{t-1} = \prod_{b=1}^{B_t} \prod_{m=1}^{M_{bt}} \left( \frac{P_{mbt}}{P_{mbt-1}} \right)^{1/\sum_{b=1}^{B_t} M_{bt}},$$

où  $M_{bt}$  est le nombre de modèles produits par le constructeur  $b$ , et  $B_t$  est le nombre de constructeurs. Cet indice est ensuite enchaîné avec la valeur de l'indice de la période précédente  $I_{t-1}$  pour produire un indice

$I_t = I_t^{t-1} \cdot I_{t-1}$  allant de la période de base à la période  $t$ . Enfin, l'indice trimestriel à l'échelle des RMR est simplement la moyenne des valeurs des trois indices au cours du trimestre. Pour le trimestre débutant le mois de  $q$ , l'indice est

$$I_q = \frac{1}{3} \sum_{t=q}^{q+2} I_t.$$

L'ensemble d'indices trimestriels obtenu à l'échelle des RMR tient compte du côté des maisons neuves de l'IPPR.

### 1.3.1 Remplacement du modèle

Lorsqu'un modèle de maison n'est plus à vendre, ou n'est plus représentatif, et qu'il est remplacé par un autre modèle dans l'échantillon, un prix antérieur pour le modèle de remplacement est imputé à la première période où il apparaît dans l'échantillon. Cela permet d'utiliser immédiatement un nouveau modèle dans le calcul de l'indice à modèles appariés. L'imputation est effectuée à l'aide d'un modèle de régression linéaire (hédonique) qui établit un lien entre les prix des maisons et les caractéristiques observées (voir de Haan et Diewert (2013, chapitre 5) pour plus de détails). Un modèle distinct est calculé pour chacune des six villes. Aucune imputation n'est effectuée lorsqu'un nouveau modèle est ajouté à l'échantillon sans remplacer un ancien modèle ni lorsqu'un nouveau constructeur est ajouté à l'échantillon.

En supposant que  $P_{mbt}$  est le prix du modèle  $m$  par le constructeur  $b$  à la période  $t$ , le modèle de régression est fondé sur un modèle structurel pour les prix des maisons

$$\log(p_{mbt}) = \alpha + x_{mbt}\beta + z_{mbt}\gamma + d_b + d_t + \log(\epsilon_{mbt}),$$

où  $x_{mbt}$  est un vecteur (rangée) des caractéristiques du modèle,  $z_{mbt}$  est un vecteur des caractéristiques de l'emplacement,  $d_b$  et  $d_t$  sont des interceptions propres au constructeur et au temps, respectivement, et  $\epsilon_{mbt}$  est un terme d'erreur. Les caractéristiques du logement comprennent le logarithme de la taille du terrain et de la taille de la maison (en pieds ou mètres carrés), et les variables nominales pour le nombre de garages, le nombre de salles de bain et le nombre de chambres à coucher. Les caractéristiques de l'emplacement comprennent les variables nominales pour la région de tri d'acheminement de la propriété (les trois premiers chiffres du code postal). Ces données sur les caractéristiques sont recueillies auprès des constructeurs pendant le processus d'échantillonnage.

Le modèle de régression est estimé à l'aide d'une fenêtre dynamique de cinq ans de données recueillies pour l'IPLN. L'estimation est effectuée avec un estimateur M robuste, à l'aide de la fonction de perte bi-carrée (voir Amemiya (1985, section 2.3) ou Wooldridge (2010, chapitre 12) pour plus de détails sur l'estimation M). Selon les hypothèses du modèle de régression linéaire classique, cette méthode d'estimation est plus robuste pour les observations aberrantes des prix que l'estimateur MCO habituel.

Lorsqu'un nouveau modèle de maison est introduit dans l'échantillon, les caractéristiques du nouveau modèle et les caractéristiques de l'ancien modèle sont utilisées pour calculer une paire de prix ajustés à partir du modèle de régression. Le prix ajusté pour le nouveau modèle est ensuite soustrait du prix ajusté pour l'ancien modèle, et cette différence est ajoutée au prix pour l'ancien modèle afin d'imputer le prix antérieur pour le nouveau

modèle. Cela tient effectivement compte de la différence entre les caractéristiques de l'ancien modèle et du nouveau modèle, ce qui donne une imputation pour ce que le prix du nouveau modèle aurait été au cours de la période précédente. C'est-à-dire que l'intégration des caractéristiques d'un nouveau modèle  $n$  dans le modèle hédonique produit un prix ajusté  $\widehat{\log(p_n)}$ , et l'intégration des caractéristiques de l'ancien modèle  $o$  dans le modèle hédonique produit un prix ajusté  $\widehat{\log(p_o)}$ . La différence entre ces prix ajustés  $\widehat{\log(p_n)} - \widehat{\log(p_o)}$  est ensuite ajoutée au prix de l'ancien modèle  $\log(p_o)$  pour produire un prix antérieur pour le nouveau modèle  $\exp(\log(p_o) + \widehat{\log(p_n)} - \widehat{\log(p_o)})$ . Le rapport de prix imputé pour le nouveau modèle est alors simplement

$$\frac{p_n}{\exp(\log(p_o) + \widehat{\log(p_n)} - \widehat{\log(p_o)})},$$

et cela est utilisé directement dans le calcul de l'indice.

## 2 Indice des prix des appartements en copropriété neufs (IPACN)

L'IPACN mesure les variations, au fil du temps, des prix de vente tels qu'établis par les constructeurs de nouvelles unités résidentielles dans des immeubles d'appartements en copropriété à Calgary, Montréal, Ottawa, Toronto, Vancouver et Victoria. Il s'agit d'un indice trimestriel, commençant au premier trimestre de 2017, composé de six sous-indices (un pour chaque ville). Chaque sous-indice est calculé au moyen d'une méthode de valeur unitaire, selon laquelle le prix d'une unité est normalisé par sa superficie en pieds carrés pour donner un prix par pied carré. Des ajustements explicites de la qualité sont effectués avant le calcul de ces prix unitaires afin de produire un indice de qualité constante. Les données de l'IPACN sont recueillies mensuellement à partir d'une enquête auprès des constructeurs au moyen d'un questionnaire électronique.

### 2.1 Concepts et définitions

Le tableau 2.1 définit les concepts clés utilisés pour élaborer l'IPACN. Un aspect important du marché des copropriétés neuves est que les copropriétés se vendent souvent pendant la phase de prévente d'un immeuble, avant le début de la construction. Les prix au cours de la phase de prévente donnent un indicateur des prix des copropriétés neuves, mais peuvent ne pas refléter un transfert de l'acheteur au vendeur si, par exemple, le constructeur n'est pas en mesure de vendre suffisamment d'unités pour financer la construction de l'immeuble.

**Tableau 2.1**  
**Concepts et définitions de l'IPACN**

Concept	Définition
<b>Prix</b>	Le prix de transaction ou le prix annoncé d'une unité déclaré par le constructeur au cours d'un mois donné, à l'exclusion de toute taxe de vente. Il s'agit du prix reçu par le constructeur, excluant les frais supplémentaires payés par l'acheteur.
<b>Valeur unitaire</b>	Le prix d'une unité normalisée par sa superficie, ce qui donne un prix par pied ou mètre carré.
<b>Type d'unité</b>	Le nombre de chambres à coucher dans un appartement, avec ou sans coin-détente, dans l'une des catégories suivantes : une chambre à coucher, une chambre à coucher + coin-détente, deux chambres à coucher, deux chambres à coucher + coin-détente et trois chambres à coucher.
<b>Prévente</b>	Période au cours de laquelle les unités peuvent être achetées avant le début de la construction.
<b>Échantillon</b>	Voir la section 2.2.1.
<b>Population cible</b>	Tous les appartements en copropriété résidentiels neufs dans un immeuble de moins de 5 étages et de 5 étages ou plus mis en vente ou vendus à Calgary, Montréal, Ottawa, Toronto, Vancouver et Victoria au cours d'un mois donné.
<b>Période de base de l'indice</b>	Période pour laquelle l'indice est égal à 100. La période de base pour l'IPACN est 2017=100.

## 2.2 Données

### 2.2.1 Échantillonnage

Les données pour l'IPACN sont recueillies à partir d'une enquête auprès des constructeurs de copropriétés. La base de sondage est compilée à partir de multiples sources, y compris des demandes de zonage et de planification reçues des municipalités, des permis de bâtir, des associations de constructeurs, des sociétés d'assurance offrant des produits aux nouveaux acheteurs de maisons et des services gouvernementaux ou sans but lucratif de protection des acheteurs de maisons, des publicités et diverses sources Internet qui fournissent de l'information sur les immeubles à venir.

L'IPACN utilise un plan d'échantillonnage à plusieurs degrés dans lequel les unités sont sélectionnées dans l'échantillon à chaque étape. Le premier degré d'échantillonnage consiste à communiquer avec les promoteurs immobiliers de la base de sondage pour déterminer s'ils font partie du champ de l'enquête. Pour s'assurer que le même immeuble peut être suivi au fil du temps, si un promoteur immobilier est visé, on lui demande de déclarer jusqu'à quatre immeubles dans lesquels moins de 70 % d'au moins un des types de logements visés a été vendu. La deuxième étape de l'échantillonnage consiste à sélectionner l'un de ces immeubles dans l'échantillon. Un questionnaire électronique sert ensuite à recueillir des renseignements sur les prix auprès des promoteurs immobiliers pour un maximum de trois unités de chaque type dans un immeuble chaque mois. Les promoteurs immobiliers déclarent également toute prime appliquée à une unité (p. ex. la valeur d'une place de stationnement ou une meilleure orientation dans l'immeuble) et on leur demande un prix annoncé si aucune unité d'un type particulier n'est vendue ce mois-là. Les mêmes renseignements sur les primes sont également recueillis pour les prix annoncés. L'échantillon est périodiquement rafraîchi à mesure que les immeubles se vendent et que les promoteurs immobiliers entrent sur le marché et en sortent.

### 2.2.2 Nettoyage et filtrage

Les données recueillies auprès des promoteurs immobiliers sont examinées manuellement pour en vérifier la cohérence et l'exhaustivité, et certains dossiers peuvent être modifiés ou supprimés en fonction du jugement. En plus de ce nettoyage manuel, les rapports de prix (voir la section 2.3) supérieurs ou égaux à 3 écarts absolus par rapport à la médiane ne sont pas inclus dans le calcul de l'indice. Comme l'IPACN est fondé sur les prix de transaction/annoncés moyens, il s'agit d'un filtre standard pour éliminer les valeurs aberrantes qui peuvent avoir une grande influence sur les moyennes (p. ex. Rousseeuw et Hubert, 2011). Afin de nettoyer adéquatement les données, l'IPACN a une révision d'un trimestre. Cela est dû en partie à la petite taille de l'échantillon au cours de la plupart des mois.

## 2.3 Calcul de l'indice

Le calcul de l'indice pour l'IPACN est assez simple et ressemble à celui de l'IPLN. Premièrement, toute prime est soustraite du prix d'une unité pour en arriver à un prix rajusté en fonction de la qualité pour une unité de référence « sans superflu ». Le prix rajusté en fonction de la qualité est ensuite normalisé par la superficie en pieds carrés d'une unité pour en arriver à un prix unitaire rajusté en fonction de la qualité. Les unités sont stratifiées par RMR, immeuble et type d'unité, et un indice géométrique non pondéré est calculé pour chaque strate, ce qui donne un rapport de prix pour chaque strate. La combinaison de la stratification et de l'ajustement explicite de la qualité signifie que le même type d'unité dans chaque immeuble est comparé au fil du temps, ce qui donne à ces rapports de prix une interprétation de qualité constante<sup>2</sup>. Ces rapports de prix par strate sont ensuite agrégés à l'échelle des RMR à l'aide d'un indice de Jevons. L'IPACN est calculé mensuellement, et les trois valeurs de l'indice dans un trimestre sont calculées en moyenne pour produire un indice trimestriel.

Pour rendre le calcul de l'indice explicite, supposons que  $p_{usbt}$  est le prix de l'unité  $u$  de type  $s$  dans l'immeuble  $b$  au moment  $t$ , supposons que  $\Delta_{usbt}$  est la valeur de la prime pour cette unité, et supposons que  $a_{usbt}$  est sa superficie en pieds carrés. Le prix unitaire ajusté en fonction de la qualité est calculé comme suit :

2. Deux caractéristiques dignes de mention pour lesquelles aucun ajustement n'est effectué sont le nombre de salles de bain dans une unité et une superficie habitable extérieure (p. ex. balcon), car les constructeurs ne sont pas en mesure d'établir le prix de ces caractéristiques.



$$\rho_{usbt} = \frac{p_{usbt} - \Delta_{usbt}}{a_{usbt}}.$$

Ces prix unitaires sont utilisés dans un indice géométrique pour produire une collection d'indices au niveau des strates entre la période  $t-1$  et la période  $t$ ,

$$I_{sbt}^{t-1} = \frac{\prod_{u=1}^{U_{sbt}} (\rho_{usbt})^{1/U_{sbt}}}{\prod_{u=1}^{U_{sbt-1}} (\rho_{usbt-1})^{1/U_{sbt-1}}},$$

où  $U_{sbt}$  est le nombre d'unités vendues de type  $s$  dans l'immeuble  $b$  au moment  $t$ . Pour produire un indice à l'échelle des RMR, les rapports de prix au sein d'une RMR pour chaque type d'unité dans chaque immeuble sont agrégés avec un indice de Jevons

$$I_t^{t-1} = \prod_{b=1}^{B_t} \prod_{s=1}^{S_{bt}} (I_{sbt}^{t-1})^{1/\sum_{b=1}^{B_t} S_{bt}},$$

où  $S_{bt}$  est le nombre de types d'unités dans l'immeuble  $b$  et  $B_t$  est le nombre d'immeubles. Ces indices d'une période à une autre sont enchaînés avec la valeur de l'indice de la période précédente pour donner la valeur de l'indice de la période en cours

$$I_t = I_t^{t-1} \cdot I_{t-1},$$

où  $I_{t-1}$  est l'indice qui va de la période de base à la période  $t-1$ . Si un nouvel immeuble est introduit dans l'échantillon au cours d'une période, il n'y a aucune tentative d'imputer des prix antérieurs pour les unités de cet immeuble. Cela signifie qu'un immeuble n'est pas inclus dans le calcul de l'indice au cours de la première période où il est introduit dans l'échantillon.

Enfin, l'indice trimestriel à l'échelle des RMR est simplement la moyenne des valeurs des trois indices au cours du trimestre. Pour le trimestre débutant le mois de  $q$ , l'indice est

$$I_q = \frac{1}{3} \sum_{t=q}^{q+2} I_t.$$

L'ensemble d'indices trimestriels obtenu à l'échelle des RMR tient compte du côté des appartements en copropriété neufs de l'IPPR.

### 3 Indice des prix de revente des propriétés résidentielles (IPRPR)

L'IPRPR mesure la variation des prix de transaction au fil du temps pour la revente de maisons et d'appartements en copropriété à Calgary, Montréal, Ottawa, Toronto, Vancouver et Victoria. Il s'agit d'un indice trimestriel, commençant au premier trimestre de 2017, composé de 12 sous-indices, soit un pour chaque type de propriété (maison et appartement en copropriété) dans chacune des six villes. Chaque sous-indice est calculé à l'aide de la méthode de ventes répétées, une méthode internationalement acceptée pour l'élaboration d'un indice de prix de qualité constante, comme le décrit le *Manuel sur les indices des prix des propriétés résidentielles* d'Eurostat (FMI,

2015). La collecte, l'intégration, la vérification et le calcul des données se font en partenariat avec Teranet et la Banque Nationale.<sup>3</sup>

### 3.1 Concepts et définitions

Le tableau 3.1 définit les concepts clés utilisés pour élaborer l'IPPR. Veuillez noter que le concept de la date de vente d'une propriété est la date de clôture, date à laquelle la propriété est transférée du vendeur à l'acheteur et par la suite inscrite au registre foncier. La date de clôture est postérieure à la date à laquelle l'acheteur et le vendeur s'entendent sur le prix de transaction de la propriété.

**Tableau 3.1**  
**Concepts et définitions de l'IPPR**

Concept	Définition
<b>Prix</b>	Prix de transaction final à la date de clôture de la vente d'une propriété et inscrit au registre foncier provincial.
<b>Date de vente</b>	Date de clôture de la vente d'une propriété.
<b>Paire de ventes</b>	Prix et dates de vente consécutifs pour la même propriété physique.
<b>Échantillon</b>	Toutes les maisons individuelles et jumelées résidentielles, toutes les maisons en rangée et tous les appartements en copropriété à Calgary, Montréal, Ottawa, Toronto, Vancouver et Victoria qui ont été vendus au moins deux fois depuis le 1 <sup>er</sup> janvier 1998 et qui figurent dans les bases de données du registre foncier.
<b>Population cible</b>	Toutes les maisons individuelles et jumelées, toutes les maisons en rangée et tous les appartements en copropriété à Calgary, Montréal, Ottawa, Toronto, Vancouver et Victoria, admissibles à la revente, qui ont été vendus entre le 1 <sup>er</sup> janvier 1998 et la période en cours.
<b>Période de base de l'indice</b>	Période pour laquelle l'indice est égal à 100. La période de base pour l'IPPR est 2017=100.

### 3.2 Données

#### 3.2.1 Sources de données

Les données sur les transactions immobilières pour l'IPPR proviennent des bureaux provinciaux d'enregistrement immobilier de l'Alberta, de la Colombie-Britannique, de l'Ontario et du Québec, de 1998 à la période en cours. Étant donné que chaque vente de propriété au Canada est enregistrée dans son bureau provincial d'enregistrement immobilier respectif, ces données tiennent compte de toutes les transactions immobilières au cours de cette période. L'IPPR comprend seulement les transactions pour les maisons individuelles et jumelées résidentielles, les maisons en rangée et les appartements en copropriété dans les RMR de Calgary, de Montréal, d'Ottawa, de Toronto, de Vancouver et de Victoria. Ces données sont recueillies et traitées par Teranet et la Banque Nationale.

Les données sur les transactions de chaque registre foncier provincial sont fournies mensuellement. Ces transactions sont ensuite appariées à la base de données sur les propriétés de Teranet afin de créer un historique de vente pour chaque propriété. Des paires de ventes sont créées pour chaque propriété qui a été vendue deux fois, tenant compte des prix de transaction et des dates de clôture pour les deux ventes de cette propriété; des paires de ventes sont créées pour des ventes consécutives pour des propriétés qui ont été vendues trois fois ou plus. Les propriétés qui n'ont été vendues qu'une seule fois (p. ex. les propriétés nouvellement construites) sont exclues. Le tableau 3.2 donne un exemple fictif des données sur les paires de ventes obtenues.

**Tableau 3.2**  
**Exemple de données sur les paires de ventes**

Adresse	Type de propriété	Date de vente	Prix de vente	Date de vente précédente	Prix de vente précédent
123, rue Fausse	Copropriété	08/01/2018	250 000	01/02/2014	200 000
321, promenade Erronée	Maison	18/01/2018	500 000	04/06/2005	400 000
321, promenade Erronée	Maison	04/06/2005	400 000	15/12/1999	350 000

3. Les indices des prix de revente des propriétés résidentielles (IPPR) sont produits par Teranet Inc. (Teranet) et la Banque Nationale du Canada (BNC) selon une méthodologie personnalisée élaborée en collaboration avec Statistique Canada. Teranet et la BNC ont octroyé à Statistique Canada une licence d'utilisation de l'IPPR à des fins de publication et de distribution.

### 3.2.2 Délai de collecte

Bien que les données du registre foncier soient reçues des registres fonciers provinciaux tous les mois, il y a un délai entre le moment où les ventes sont enregistrées dans les registres fonciers et celui où ces données sont reçues par Teranet et la Banque Nationale. Ce délai est particulièrement important en Colombie-Britannique. Le tableau 3.3 donne un exemple de la proportion cumulative des ventes reçues par province à la fin de chaque mois, pour un mois fixe  $M$ . En raison de ce délai de collecte, l'IPPR a une révision d'un trimestre pour veiller à ce que les données recueillies soient suffisantes pour produire des valeurs d'indice fiables.

**Tableau 3.3**  
**Part moyenne des ventes réalisées par province à chaque mois**

Province	Période M	Période M+1	Période M+2	Période M+3	Période M+4	Période M+5
	pour cent					
Alberta	92	100	100	100	100	100
Colombie-Britannique	43	94	97	99	99	100
Ontario	90	95	97	100	100	100
Québec	83	83	83	83	83	88

### 3.2.3 Nettoyage et filtrage

Les données de l'IPPR proviennent des sources administratives — et sont par conséquent assez épurées —, bien qu'il faille effectuer un certain filtrage pour supprimer les transactions immobilières qui ne sont pas appropriées pour l'élaboration de l'IPPR, ainsi que les valeurs aberrantes qui peuvent avoir une grande influence sur l'indice. Cela comprend la suppression de paires de ventes pour lesquelles une des transactions peut ne pas être sans lien de dépendance (p. ex. un legs) ou peut être une vente à tout prix, ou pour lesquelles la fluctuation des prix entre les ventes est si extrême que cela laisse entendre que la qualité de la propriété peut avoir changé (p. ex. en raison de rénovations). Ces filtres sont appliqués à chaque RMR et à chaque type de propriété séparément et sont résumés dans l'ordre dans lequel ils sont appliqués au tableau 3.4.

Avant l'application de ces filtres, on utilise une série de filtres pour supprimer les transactions qui peuvent faire partie d'une transaction fractionnée de constructeur ou d'une transaction en bloc de promoteur immobilier (c.-à-d. une vente groupée de propriétés multiples), car ces types de transactions ne sont pas compris dans le champ de l'IPPR. Les groupes de cinq propriétés ou plus dans la même région de tri d'acheminement (trois premiers chiffres du code postal d'une propriété), vendues à la même date et pour le même prix, sont traitées comme une transaction en bloc/fractionnée. La transaction pour chaque propriété du groupe est supprimée lorsqu'il s'agit de la transaction la plus récente pour chaque propriété.

Les transactions dans un groupe peuvent revenir si le prix de vente subséquent pour au moins 75 % des propriétés dans le groupe est au moins 75 % du prix de la transaction en bloc/fractionnée, et, pour chaque vente subséquente pour chaque propriété dans le groupe, il existe au plus une autre propriété dans la même région de tri d'acheminement qui s'est vendue au même prix à la même date. Cela permet d'utiliser des transactions en bloc/fractionnées si le prix de ces transactions est proche du prix de vente subséquent pour la plupart des propriétés dans la transaction en bloc/fractionnée.

**Tableau 3.4**  
**Filtres de données pour les paires de ventes dans l'IPPR**

Filter	Justification
Prix de transaction inférieur ou égal à 10 000 dollars.	Ces transactions peuvent ne pas être des transactions sans lien de dépendance (p. ex. legs).  Ces transactions peuvent être des ventes à tout prix ou des transactions spéculatives (de Haan et Diewert, 2013, section 6.11), ou des propriétés revendues pour lesquelles il y a un changement important dans la qualité de la propriété (p. ex. Jansen et coll., 2008; S&P Dow Jones, 2018).
Période de rétention inférieure à 6 mois.	Il peut y avoir un changement dans la qualité d'une propriété qui entraîne un changement de prix exceptionnellement important entre les transactions, ou une erreur dans la saisie de données pour l'un des prix de transaction. Comme l'IPPR est fondé sur les prix de transaction moyens, cela élimine également les valeurs aberrantes qui peuvent avoir une grande influence sur les moyennes (p. ex. Rousseeuw et Hubert, 2011).
Rendement annualisé supérieur ou égal à 3 écarts absolus médians par rapport à la médiane.	

### 3.3 Calcul de l'indice

La méthode des ventes répétées offre un moyen d'élaborer un indice de prix de qualité constante, exploitant de multiples ventes pour la même propriété au fil du temps pour contrôler les différences de qualité invariantes entre les propriétés. D'autres méthodes d'élaboration d'un indice de qualité constante (p. ex. hédonique ou stratification) exigent des caractéristiques de propriété, comme l'âge de la propriété, qui ne sont pas accessibles dans les données du registre foncier. Voir Hansen (2009) pour une comparaison des différentes approches d'élaboration d'un indice des prix des propriétés.

En pratique, il y a un certain nombre de choix méthodologiques à faire lors de la mise en œuvre d'un indice de ventes répétées. La présente section décrit la méthode de ventes répétées et souligne la saveur particulière de l'indice de ventes répétées utilisé pour élaborer l'IPPR. Voir Wang et Zorn (1997) et de Haan et Diewert (2013, chapitre 6) pour une vue d'ensemble de la méthode des ventes répétées, et Jansen et coll. (2008) pour une mise en application.

En raison du petit nombre de transactions pour les appartements en copropriété, le sous-indice des appartements en copropriété est calculé pour chaque trimestre. Pour les maisons, l'indice est calculé mensuellement, et les valeurs de l'indice résultantes sont calculées en moyenne sur chaque trimestre pour produire un indice trimestriel.

#### 3.3.1 La méthode des ventes répétées

Il existe deux grandes catégories d'indices de prix de ventes répétées : l'indice géométrique de ventes répétées (indice GVR) de type Jevons proposé par Bailey et al. (1963) et l'indice arithmétique de ventes répétées de type Laspeyres (indice AVR) proposé par Shiller (1991)<sup>4</sup>. Les indices GVR et AVR montrent souvent des fluctuations de prix semblables au fil du temps (p. ex. Shiller, 1991). L'IPPR utilise l'indice arithmétique de ventes répétées décrit dans Shiller (1991, section II), semblable à celui utilisé par S&P Dow Jones (2018).

En plus des versions géométrique et arithmétique de l'indice des ventes répétées, il existe divers schémas de pondération qui peuvent être utilisés pour pondérer les rapports de prix dans le calcul de l'indice (p. ex. Case et Shiller, 1987; Abraham et Schauman, 1991; Calhoun, 1996). Il s'agit de pondérations par l'inverse de la variance conçues pour corriger les différences dans la variance des prix des transactions pour les propriétés ayant des périodes de rétention différentes qui peuvent compliquer l'élaboration d'intervalles de confiance pour l'indice. Bien que les pondérations aient une incidence directe sur les valeurs de l'indice, dans la pratique, ces pondérations peuvent avoir une incidence tout au plus marginale sur l'indice (p. ex. Goetzmann, 1992; Hansen, 2009), surtout pour les grands échantillons. Les indices pondérés, cependant, reposent sur plus d'hypothèses que leurs homologues non pondérés et ne peuvent être calculés si leurs pondérations ne peuvent pas être calculées. Des études antérieures ont également révélé que les indices non pondérés ne sont pas inférieurs aux indices pondérés (de Haan et Diewert, 2013, section 6.14). Par conséquent, comme les intervalles de confiance ne sont pas déclarés pour l'IPPR, les pondérations de l'inverse de la variance ne sont pas utilisées pour calculer l'IPPR.

4. Il y a aussi des indices hybrides de ventes répétées qui peuvent faire des ajustements hédoniques de la qualité à l'aide d'un modèle à effets aléatoires (voir de Haan et Diewert, 2013, sections 6.25 à 6.27), bien que ces méthodes ne puissent pas être utilisées avec les données du registre foncier, car ces données n'ont pas de caractéristiques de propriété.

### 3.3.2 Les indices GVR et AVR

Par le passé, l'indice GVR a précédé l'indice AVR, commençant avec l'article précurseur de Bailey et coll. (1963), et il est plus facile de comprendre l'indice AVR en élaborant d'abord l'indice GVR. En supposant que les périodes sont indexées par  $t \in \{0, 1, \dots, T\}$  et les propriétés sont indexées par  $i \in \{1, 2, \dots, N\}$ , le point de départ de l'indice GVR est un modèle structurel (hédonique) des prix des propriétés

$$\log(p_{it}) = \log(P_t) + x_{it}\theta + \log(\epsilon_{it}),$$

où  $p_{it}$  est le prix de transaction de la propriété  $i$  au moment  $t$ ,  $P_t$  est un prix à l'échelle municipale commun illustrant les fluctuations de prix agrégés,  $x_{it}$  est un vecteur (rangée) des caractéristiques de la propriété (p. ex. nombre de chambres à coucher de la propriété  $i$  au moment  $t$ ),  $\theta$  est un vecteur de prix implicites (hédoniques), et  $\epsilon_{it}$  est un terme d'erreur<sup>5</sup>. Il s'agit simplement d'un modèle hédonique avec des variables nominales de temps dans lequel les propriétés peuvent se vendre plus d'une fois (p. ex. de Haan et Diewert, 2013, chapitre 5). Dans le contexte de ce modèle, l'indice de prix (géométrique) de qualité constante dans la période  $\tau$  avec la période de base 0, indiqué par  $I_\tau^G$ , est  $I_\tau^G \equiv P_\tau / P_0$ . Il est important de noter que  $P_t$  n'est pas aléatoire – c'est un paramètre qui régit la répartition conjointe des prix des propriétés.

En supposant que les caractéristiques des propriétés ne changent pas au fil du temps (c.-à-d.  $x_{it} = x_i$ , pour tous les  $t$ ) et que chaque propriété se vend deux fois, la transformation de différences premières peut être utilisée pour livrer

$$\begin{aligned} \log\left(\frac{p_{is(i)}}{p_{if(i)}}\right) &= \log\left(\frac{P_{s(i)}}{P_{f(i)}}\right) + \log\left(\frac{\epsilon_{is(i)}}{\epsilon_{if(i)}}\right) \\ &= \sum_{t=1}^T D_{it} \log\left(\frac{P_t}{P_0}\right) + \log\left(\frac{\epsilon_{is(i)}}{\epsilon_{if(i)}}\right), \end{aligned}$$

où  $s(i)$  donne le moment de la deuxième vente de la propriété  $i$ ,  $f(i)$  donne le moment de la première vente de la propriété  $i$ , et  $D_{it}$  est une variable nominale qui prend la valeur 1 si une propriété se vend pour la deuxième fois lors de la période  $t$  (c.-à-d.  $s(i) = t$ ), -1 si la propriété se vend pour la première fois lors de la période  $t$  (c.-à-d.  $f(i) = t$ ), et 0 dans les autres cas. L'hypothèse selon laquelle les caractéristiques des propriétés ne changent pas au fil du temps signifie que la variation en pourcentage du prix d'une propriété suit la variation agrégée en pourcentage des prix des propriétés, jusqu'à une erreur additive. Les propriétés qui se vendent trois fois ou plus peuvent être intégrées à la transformation des différences premières en traitant des paires consécutives de ventes comme des propriétés distinctes.

En supposant que les termes d'erreur sont strictement exogènes, de sorte que  $E[\log(\epsilon_{it}) | D_{i1}, D_{i2}, \dots, D_{iT}] = 0$  – une hypothèse généralisée dans les applications de données recueillies au moyen d'un panel (p. ex. Wooldridge, 2010, chapitre 10) – l'hypothèse selon laquelle les caractéristiques des propriétés ne changent pas au fil du temps permet d'identifier l'indice de prix à partir de la régression linéaire

5. Les modèles structurels de Case et Shiller (1987) et de Shiller (1991, section V) sont des cas spéciaux de ce modèle.

$$\log \left( \frac{P_{is(i)}}{P_{if(i)}} \right) = \sum_{t=1}^{\tau} D_{it} \gamma_t + \log \left( \frac{\epsilon_{is(i)}}{\epsilon_{if(i)}} \right),$$

afin que  $I_{\tau}^G \equiv P_{\tau} / P_0 = \exp(\gamma_{\tau})$ . La transformation des différences premières transforme un modèle structurel qui dépend des caractéristiques des propriétés en une équation d'estimation qui dépend seulement du moment de la vente d'une propriété<sup>6</sup>.

Il est instructif de dériver la forme de l'indice GVR comme indice pour faire le lien avec l'indice AVR. En

supposant que  $N_f(\tau)$  est l'ensemble des propriétés qui se vendent pour la première fois lors de la période

$\tau$ ,  $N_s(\tau)$  est l'ensemble des propriétés qui se vendent pour la deuxième fois lors de la période  $\tau$ , et

$N(\tau) = |N_f(\tau)| + |N_s(\tau)|$  (le nombre de propriétés qui se vendent lors de la période  $\tau$ ), il peut être démontré que

$$I_{\tau}^G = \prod_{i \in N_f(\tau)} \left( \frac{P_{i\tau}}{P_{is(i)}} \frac{I_{s(i)}^G}{I_{s(i)}^G} \right)^{\frac{1}{N(\tau)}} \prod_{i \in N_s(\tau)} \left( \frac{P_{i\tau}}{P_{if(i)}} \frac{I_{f(i)}^G}{I_{f(i)}^G} \right)^{\frac{1}{N(\tau)}}.$$

L'indice GVR est simplement un indice de Jevons à modèles appariés avec une variante. Plutôt que d'utiliser uniquement les transactions immobilières qui se produisent au cours de la période 0 et de la période  $\tau$ , l'indice lui-même est utilisé pour extrapoler les prix au fil du temps pour toutes les propriétés qui se vendent au cours de la période  $\tau$  en baissant les prix des ventes qui ne se produisent pas au cours de la période de base à l'aide de l'indice de cette période. Cela permet à toutes les propriétés qui se vendent au cours de la période  $\tau$  d'être utilisées dans le calcul de l'indice, que la propriété soit vendue ou non au cours de la période 0.

Comme solution de remplacement à un indice géométrique, Shiller (1991) propose l'indice AVR, indiqué par  $I_{\tau}^A$ , qui remplace simplement les moyennes géométriques de l'indice GVR par des moyennes arithmétiques :

$$I_{\tau}^A = \frac{\sum_{i \in N_f(\tau)} P_{i\tau} + \sum_{i \in N_s(\tau)} P_{i\tau}}{\sum_{i \in N_f(\tau)} \frac{P_{is(i)}}{I_{s(i)}^A} + \sum_{i \in N_s(\tau)} \frac{P_{if(i)}}{I_{f(i)}^A}}.$$

Les rapports de prix sont formés de la même façon que l'indice GVR, sauf que maintenant, un indice de Laspeyres est utilisé pour combiner les rapports de prix, plutôt qu'un indice de Jevons. Il s'agit de l'indice utilisé pour calculer l'IPPR.

Pour calculer l'indice AVR, il faut résoudre un système d'équations pour calculer l'indice pour chaque période. Comme dans le cas de l'indice GVR, l'indice AVR peut être calculé comme une régression linéaire, bien que maintenant avec un ensemble de variables instrumentales – cela constitue une façon pratique de calculer l'indice et de déterminer ses propriétés statistiques. En supposant que

6. Si les caractéristiques des propriétés changent systématiquement au fil du temps, alors  $D_{it}$  est corrélée avec l'erreur  $\log(\epsilon_{is(i)} / \epsilon_{if(i)})$ , bien que le sens de cette corrélation dépend de l'existence d'une dépréciation nette ou d'une valorisation des propriétés au fil du temps.

$$Y_i = \begin{cases} p_{if(i)} & \text{si } f(i) = 0 \\ 0 & \text{si } f(i) > 0 \end{cases},$$

et

$$X_{it} = \begin{cases} -p_{if(i)} & \text{si } f(i) = t \\ p_{is(i)} & \text{si } s(i) = t \\ 0 & \text{autrement} \end{cases},$$

l'indice AVR est le réciproque de l'estimateur des variables instrumentales (VI) pour la régression

$$Y_i = \sum_{t=1}^T X_{it} \beta_t + v_i,$$

avec  $D_{it}$  comme instrument pour  $X_{it}$ . En supposant que  $X_i = (X_{i1}, X_{i2}, \dots, X_{iT})$  et  $D_i = (D_{i1}, D_{i2}, \dots, D_{iT})$ , toute la série d'indices AVR de la période 1 à la période  $T$  est calculée comme suit :

$$(I_1^A, I_2^A, \dots, I_T^A)' = \text{diag} \left( \left[ \begin{array}{c} \sum_{i=1}^N D_i' X_i \\ \sum_{i=1}^N D_i' Y_i \end{array} \right]^{-1} \right)^{-1}.$$

La validité de l'estimateur VI repose sur un bilan selon lequel l'indice à calculer est un indice arithmétique (Shiller, 1991, p. 115). Par exemple, pour un échantillon de transactions de ventes répétées de propriétés au cours de  $T$  périodes, l'estimateur VI est convergent dans des conditions assez faibles dans le cadre du processus d'échantillonnage (p. ex. White, 2001, théorème 3.15; Wooldridge, 2010, théorèmes 5.1 et 8.1), de sorte que l'estimateur pour l'indice AVR converge en probabilité vers l'indice AVR de population (c.-à-d. il est sans biais dans les grands échantillons).

### 3.3.3 Exemple élaboré de l'indice AVR

L'exemple non négligeable le plus simple d'un indice de ventes répétées comporte trois périodes — une première période 0 qui sert de période de base, suivie des périodes 1 et 2 — et trois maisons, étiquetées comme  $a$ ,  $b$  et  $c$ . La maison  $a$  se vend pour la première fois au cours de la période 1 et pour la deuxième fois au cours de la période 2; la maison  $b$  se vend pour la première fois au cours de la période 0 et pour la deuxième fois au cours de la période 2; et la maison  $c$  se vend pour la première fois au cours de la période 0 et pour la deuxième fois au cours de la période 1. Le tableau 3.5 résume ces données.

**Tableau 3.5**  
**Données sur les paires de ventes**

Maison	Date de vente	Prix de vente	Date de vente précédente	Prix de vente précédent
$a$	2	$P_{a2}$	1	$P_{a1}$
$b$	2	$P_{b2}$	0	$P_{b0}$
$c$	1	$P_{c1}$	0	$P_{c0}$

Avec ces données, l'indice AVR est

$$I_1^A = \frac{P_{a1} + P_{c1}}{\frac{P_{a2}}{I_2^A} + P_{c0}} = \frac{P_{a1}}{I_2^A} \cdot \frac{P_{a2}}{I_2^A} + \frac{P_{c1}}{P_{c0}} \cdot \frac{P_{c0}}{\frac{P_{a2}}{I_2^A} + P_{c0}}$$

et

$$I_2^A = \frac{P_{a2} + P_{b2}}{\frac{P_{a1}}{I_1^A} + P_{b0}} = \frac{P_{a2}}{I_1^A} \cdot \frac{P_{a1}}{I_1^A} + \frac{P_{b2}}{P_{b0}} \cdot \frac{P_{b0}}{\frac{P_{a1}}{I_1^A} + P_{b0}}$$

Cela ressemble à un indice de Laspeyres pur à modèles appariés, sauf que la maison  $a$  peut être incluse dans le calcul de l'indice en baissant son prix pour obtenir un prix d'une pseudo-période  $0^7$ . Toutefois, cela signifie que l'indice est défini par un système d'équations — une pour chaque période — qui doivent être résolues pour obtenir l'indice pour une période donnée. L'indice AVR est défini simultanément pour chaque période.

Pour obtenir une solution sous forme fermée pour l'indice AVR, veuillez noter que

$$D \equiv \begin{bmatrix} D_a \\ D_b \\ D_c \end{bmatrix} = \begin{bmatrix} -1 & 1 \\ 0 & 1 \\ 1 & 0 \end{bmatrix},$$

$$X \equiv \begin{bmatrix} X_a \\ X_b \\ X_c \end{bmatrix} = \begin{bmatrix} -P_{a1} & P_{a2} \\ 0 & P_{b2} \\ P_{c1} & 0 \end{bmatrix},$$

et

$$Y \equiv \begin{bmatrix} Y_a \\ Y_b \\ Y_c \end{bmatrix} = \begin{bmatrix} 0 \\ P_{b0} \\ P_{c0} \end{bmatrix}.$$

L'indice AVR provient de l'estimateur VI pour la régression linéaire

$$Y = X\beta + v$$

avec  $D$  comme variable instrumentale. La condition du moment (orthogonalité) de l'estimateur VI,  $\hat{\beta}$ , est

$$D'X \cdot \hat{\beta} = D'Y$$

$$\begin{bmatrix} P_{a1} + P_{c1} & -P_{a2} \\ -P_{a1} & P_{a2} + P_{b2} \end{bmatrix} \begin{bmatrix} \hat{\beta}_1 \\ \hat{\beta}_2 \end{bmatrix} = \begin{bmatrix} P_{c0} \\ P_{b0} \end{bmatrix},$$

la solution à laquelle est

7. Avec un indice pur à modèles appariés, l'indice de la période 1 est simplement  $P_{c1}/P_{c0}$  et l'indice de la période 2 est  $P_{b2}/P_{b0}$ .



$$\begin{bmatrix} \hat{\beta}_1 \\ \hat{\beta}_2 \end{bmatrix} = \frac{1}{(p_{a1} + p_{c1})(p_{a2} + p_{b2}) - p_{a1}p_{a2}} \begin{bmatrix} p_{a2} + p_{b2} & p_{a2} \\ p_{a1} & p_{a1} + p_{c1} \end{bmatrix} \begin{bmatrix} p_{c0} \\ p_{b0} \end{bmatrix}.$$

L'indice AVR pour la période  $t$  est simplement  $1 / \hat{\beta}_t$ , et donc

$$I_1^A = \frac{(p_{a1} + p_{c1})(p_{a2} + p_{b2}) - p_{a1}p_{a2}}{p_{c0}(p_{a2} + p_{b2}) + p_{b0}p_{a2}}$$

et

$$I_2^A = \frac{(p_{a1} + p_{c1})(p_{a2} + p_{b2}) - p_{a1}p_{a2}}{p_{b0}(p_{a1} + p_{c1}) + p_{c0}p_{a1}}.$$

Malgré la simplicité conceptuelle de l'AVR en tant qu'indice à modèles appariés, il a néanmoins une structure non linéaire assez complexe.

### 3.3.4 Représentativité de la population cible

La population cible de l'IPPR est composée de toutes les propriétés qui sont admissibles à la revente et qui ont été vendues depuis janvier 1998. Dans la pratique, les données sur les paires de ventes ne sont accessibles que pour les propriétés qui se vendent deux fois ou plus au cours de cette période; les propriétés qui ne se vendent qu'une fois sont omises de l'échantillon. Il s'agit d'un problème de sélection de l'échantillon – les propriétés à ventes répétées ne sont peut-être pas représentatives de toutes les propriétés faisant l'objet de transactions – et l'indice de ventes répétées qui en résulte pourrait ne pas tenir compte de la fluctuation des prix pour la population cible de toutes les propriétés faisant l'objet de transactions. La production d'un indice représentatif repose sur l'hypothèse qu'il n'y ait pas de différences systématiques dans les prix de vente latents et les périodes de rétention entre les propriétés qui font l'objet d'une seule transaction et celles qui font l'objet de deux ou plus. (Voir Wooldridge (2010, théorème 19.1) pour connaître les conditions précises dans lesquelles la sélection de l'échantillon peut être ignorée avec un estimateur VI.) Des études antérieures ont trouvé des preuves à l'appui de cette hypothèse (voir de Haan et Diewert, 2013, section 6.17).

Étant donné que l'IPPR met l'accent sur les propriétés en revente, les propriétés qui ne se vendent qu'une seule fois parce qu'elles sont nouvellement construites ne contribuent pas à un échantillon sélectionné. Les seules différences entre la population cible et l'échantillon de transactions accessibles sont les propriétés vendues avant janvier 1998 et une seule fois depuis. Ces propriétés ne servent pas au calcul de l'IPPR, mais entrent dans le champ de la population cible puisque ces propriétés sont à la fois admissibles à la revente et ont été vendues après janvier 1998. Cet écart entre la population cible et l'échantillon disparaîtra avec le temps.

### 3.3.5 Pondérations par l'inverse de la variance

Case et Shiller (1987) soutiennent que la variance des prix de transaction des paires de ventes augmente avec la période de rétention d'une propriété, auquel cas le terme d'erreur dans la régression de l'indice GVR peut être hétéroscédastique<sup>8</sup>. Cela signifie que les erreurs-types MCO habituelles pour l'indice GVR sont incohérentes et que l'estimateur MCO n'est plus la variance minimale; la même chose s'applique à l'estimateur VI pour l'indice AVR. Si la relation entre la période de rétention et la variance des prix de transaction est connue, les estimateurs de moindres carrés généralisés (MCG) et de variables instrumentales généralisées (VIG), utilisant des pondérations par l'inverse de la variance, sont des solutions plus efficaces que leurs homologues non pondérés, et fournissent un estimateur convergent pour leurs erreurs-types (White, 2001, théorème 4.62; Wooldridge, 2010, théorème 8.5).

8. Selon l'hypothèse de l'exogénéité stricte,  $E(\log(p_{m(t)}/p_{f(t)}) | D_{t1}, D_{t2}, \dots, D_{tT})$  est linéaire, de sorte que  $\text{var}(\log(p_{m(t)}/p_{f(t)}) | D_{t1}, D_{t2}, \dots, D_{tT}) = \text{var}(\log(\epsilon_{m(t)}/\epsilon_{f(t)}) | D_{t1}, D_{t2}, \dots, D_{tT})$ .

L'hétéroscédasticité n'est pas particulièrement problématique pour l'IPPR; comme c'est le cas pour la plupart des indices de prix nationaux, les erreurs-types ne sont pas déclarées pour l'IPPR, et il y a un échantillon suffisamment grand pour que l'efficacité asymptotique ne soit pas préoccupante (Wang et Zorn, 1997, section 4.4)<sup>9</sup>. Toutefois, l'utilisation des pondérations par l'inverse de la variance modifie les valeurs de l'indice. Cela est problématique puisque les estimateurs de MCG et de VIG exigent des hypothèses plus puissantes que les estimateurs MCO et VI habituels (p. ex. la relation entre la variance et la période de rétention doit être connue), et le défaut de ces hypothèses peut porter atteinte à l'utilité de ces estimateurs (p. ex. Angrist et Pischke, 2009, section 3.4.1; Wooldridge, 2010, section 4.2.3). Il n'y a pas non plus de garantie que les pondérations par l'inverse de la variance peuvent être calculées à n'importe quel moment dans le temps (p. ex. Calhoun, 1996), et comme les pondérations ont une incidence sur les valeurs de l'indice, l'indice ne peut pas être calculé si les pondérations échouent. Par conséquent, l'IPPR n'utilise pas de pondérations par l'inverse de la variance<sup>10</sup>.

### 3.4 Révision

#### 3.4.1 Prise en compte de la révision dans le modèle de ventes répétées

Un désavantage de tout indice de ventes répétées est qu'il fait l'objet d'une révision perpétuelle. Pour calculer l'indice pour une période, il faut le calculer pour toutes les périodes et, à mesure que de nouvelles données deviennent accessibles, cela modifiera les valeurs de l'indice pour les périodes précédentes.

L'IPPR évite la révision en utilisant un raccordement des fluctuations pour mettre à jour l'indice lorsque de nouvelles périodes de données deviennent accessibles. Avec cette approche, la fluctuation des prix de la série calculée à partir des données les plus récentes est enchaînée avec la dernière valeur de l'indice de la série originale, évitant ainsi la révision de la série originale. Cette méthode d'enchaînement successif des indices est utilisée avec les indices de prix hédoniques pour éviter ce même type de révision (p. ex. de Haan et Diewert, 2013, section 5.18)<sup>11</sup>.

Pour fixer la notation, supposons que  $I_0^S, \dots, I_T^S$  est une série d'indices de prix de ventes répétées allant de la période 0 à la période  $T$ , calculés à partir des premières périodes  $S \leq T$  de données. Cette série peut être mise à jour avec un raccordement des fluctuations de la façon suivante : Premièrement, avec les périodes  $T + 1$

de données accessibles, calculer la série d'indices  $I_0^T, \dots, I_{T+1}^T$ ; c'est-à-dire recalculer toute la série en utilisant toutes les données accessibles. Pour ensuite mettre à jour la série originale d'indices qui s'étend jusqu'à la

période  $T$ , il suffit de calculer la valeur de l'indice dans la période  $T + 1$  comme  $I_T^S \cdot I_{T+1}^T / I_T^{T+1}$ , et d'ajouter cette valeur à la série originale. Ainsi, la série originale d'indices devient

$$I_0^S, I_1^S, \dots, I_T^S, I_T^S \cdot \frac{I_{T+1}^T}{I_T^{T+1}}.$$

L'incidence de toute dérive de ce type de raccordement dans l'indice peut facilement être évaluée au fil du temps en comparant l'indice calculé à l'aide de toutes les données à l'indice raccordé, et cela fait partie du travail d'assurance de la qualité effectué lors de la production de l'IPPR. À condition que la série d'indices historiques soit relativement stable au fil du temps, il devrait y avoir une dérive minimale du raccordement.

- 
9. Les pondérations par l'inverse de la variance n'offrent pas de correction pour les variations de la qualité dans une propriété au fil du temps. Les pondérations initiales de Case et Shiller (1987) accordaient moins de poids aux propriétés ayant des périodes de rétention plus longues, ce qui donnait aux propriétés des périodes plus longues entre les ventes, et donc plus d'occasions de changements de la qualité, moins de poids dans le calcul de l'indice parce que la variance des prix augmentait avec le temps de rétention. Jansen et coll. (2008), toutefois, trouvent une relation non monotone entre la période de rétention et la variance, de sorte que ces pondérations ne peuvent être ajustées selon les différences de qualité au sein d'une propriété. Si les hypothèses requises pour la pondération par l'inverse de la variance sont maintenues, alors les indices pondérés et non pondérés convergent en probabilité vers la même valeur – si la pondération offre une correction pour les variations de la qualité au fil du temps, alors l'indice pondéré devrait converger vers une valeur différente que l'indice non pondéré.
10. Les erreurs-types robustes face à l'hétéroscédasticité inconnue peuvent facilement être calculées pour l'indice AVR sans changer les valeurs de l'indice, et elles sont proposées par Shiller (1991, section II). Il est également facile de généraliser ces erreurs-types pour tenir compte des erreurs corrélées en série au fil du temps pour les propriétés qui se vendent trois fois ou plus (p. ex. White, 2001, chapitre 6; Wooldridge, 2010, chapitre 8).
11. Dans le cas des indices hédoniques, une fenêtre dynamique est parfois utilisée pour éviter la dérive dans l'indice en raison de l'enchaînement à haute fréquence. Ce n'est toutefois pas logique dans le cas d'un indice de ventes répétées, car l'utilisation d'une fenêtre dynamique éliminerait des propriétés qui ont été vendues pour la première fois au début de la série.

### 3.4.2 Prise en compte de la révision en raison du délai de collecte

L'IPPR fait l'objet d'une révision trimestrielle pour tenir compte du délai dans l'entrée des données des registres fonciers. Cette révision signifie que l'indice est calculé deux fois pour chaque période. Par exemple, lors du calcul de l'indice au premier trimestre de 2018, l'indice est d'abord calculé au deuxième trimestre de 2018 au moyen de toutes les données reçues au premier trimestre de 2018, et est ensuite calculé à nouveau au troisième trimestre de 2018 une fois que la majorité des données du premier trimestre de 2018 ont été reçues des registres fonciers au deuxième trimestre de 2018.

Cette révision signifie que l'indice doit être raccordé à deux séries d'indices différentes. En utilisant la notation ci-dessus, l'indice préliminaire est calculé comme suit :

$$I_0^S, I_1^S, \dots, I_T^S, I_T^S \cdot \frac{I_{T+1}^{T+1}}{I_T^{T+1}}$$

et l'indice révisé est calculé comme suit :

$$I_0^S, I_1^S, \dots, I_T^S, I_T^S \cdot \frac{I_{T+1}^{T+2}}{I_T^{T+2}}$$

Cette méthode de raccordement permet une révision trimestrielle de l'indice, de sorte que des données supplémentaires puissent être recueillies à partir des registres fonciers, tout en évitant la révision perpétuelle de l'indice de ventes répétées.

## 4 Indice des prix des propriétés résidentielles (IPPR)

L'IPPR regroupe les indices à l'échelle des RMR provenant de l'IPLN, de l'IPACN et de l'IPRPR pour produire un indice des prix des propriétés résidentielles à Calgary, Montréal, Ottawa, Toronto, Vancouver, Victoria et un indice agrégé des six RMR. La population cible de l'IPPR est l'ensemble des populations cibles pour chacun des trois indices des composantes. Chacun des quatre indices (maison neuve, appartement en copropriété neuf, maison en revente, appartement en copropriété en revente) est agrégé avec un indice de Young, les pondérations des ventes tenant compte de la part de valeur des propriétés neuves par rapport aux propriétés en revente, et les maisons par rapport aux appartements en copropriété, vendus dans chaque RMR. L'IPPR est un indice trimestriel, puisque l'IPACN et l'IPRPR sont trimestriels, commençant au premier trimestre de 2017. Pour rester conforme à l'IPACN et à l'IPRPR, l'IPPR fait l'objet d'une révision trimestrielle.

Les pondérations de l'IPPR sont tirées du Relevé des logements écoulés sur le marché de la Société canadienne d'hypothèques et de logement et de l'inventaire des transactions de ventes répétées de Teranet et de la Banque Nationale<sup>12</sup>. Ces deux sources tiennent compte de la valeur de toutes les transactions de ventes nouvelles et de ventes répétées respectivement pour les maisons individuelles et jumelées résidentielles, les maisons en rangée, et les appartements en copropriété dans un immeuble de moins de 5 étages ou de 5 étages ou plus. Par conséquent, les valeurs agrégées sont comparables afin de produire une part de valeur pour les propriétés neuves par rapport aux propriétés en revente, ainsi que pour les maisons par rapport aux appartements en copropriété. La période de base des pondérations correspond aux trois années civiles précédant l'année courante de l'indice, et ces pondérations sont mises à jour annuellement. Pour éviter tout chevauchement avec la période de révision, les pondérations sont mises à jour au deuxième trimestre de l'année.

12. Pour aider à réduire l'écart entre la population cible et l'échantillon, on utilise les ventes répétées de 1990 à la période en cours pour calculer les pondérations. Bien que les données de 1990 à 1998 ne conviennent pas au calcul d'un indice dans toutes les RMR, cela n'a aucune incidence sur les pondérations puisqu'elles n'utilisent que les transactions à partir de 2014. L'avantage de cette période plus longue est que les propriétés qui ont été vendues entre 1990 et 1998 et de nouveau pendant la période de base des pondérations font partie de la population cible et sont maintenant utilisées pour calculer les pondérations, alors que seules les données sur les ventes répétées à partir de 1998 omettent ces propriétés.

## Références

- ABRAHAM, J. M. et W. S. SCHAUMAN. 1991. « New evidence on home prices for Freddie Mac repeat sales », *Real Estate Economics*, vol. 19, n° 3, p. 333 à 352.
- AMEMIYA, T. 1985. *Advanced Econometrics*, Harvard University Press.
- ANGRIST, J. et J.-S. PISCHKE. 2009. *Mostly Harmless Econometrics*. Princeton University Press.
- BAILEY, M., R. MUTH et H. NOURSE. 1963. « A regression method for real estate price index construction », *Journal of the American Statistical Association*, vol. 58, n° 304, p. 933 à 942.
- CALHOUN, C. 1996. *OFHEO House Price Indexes: HPI Technical Description*. Office of Federal Housing Enterprise Oversight. Extrait de [http://www.ofheo.gov/Media/Archive/house/hpi\\_tech.pdf](http://www.ofheo.gov/Media/Archive/house/hpi_tech.pdf).
- CASE, K. et R. SHILLER. 1987. « Prices of single-family homes since 1970: New indexes for four cities », *New England Economic Review*, p. 45 à 56.
- DE HAAN, J. et W. E. DIEWERT (dir.). 2013. *Handbook on Residential Property Prices Indices (RPPIs)*. Eurostat.
- GOETZMANN, W. 1992. « The accuracy of real estate indices: Repeat sale estimators », *Journal of Real Estate Finance and Economics*, vol. 5, n° 1, p. 5 à 53.
- HANSEN, J. 2009. « Australian house prices: A comparison of hedonic and repeat-sales measures », *Economic Record*, vol. 85, n° 269, p. 132 à 145.
- FMI. 2015. *The Special Data Dissemination Standard Plus: Guide for Adherents and Users*. Extrait de <https://www.imf.org/external/pubs/ft/sdds/guide/plus/2015/sddsplus15.pdf>.
- JANSEN, S., P. DE VRIES, H. COOLEN, C. LAMAIN et P. BOELHOUWER. 2008. « Developing a house price index for The Netherlands: A practical application of weighted repeat sales », *Journal of Real Estate Finance and Economics*, vol. 37, n° 2, p. 163 à 186.
- ROUSSEEUW, P. J. et M. HUBERT. 2011. « Robust statistics for outlier detection », *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, vol. 1, n° 1, p. 73 à 79.
- SHILLER, R. 1991. « Arithmetic repeat sales price estimators », *Journal of Housing Economics*, vol. 1, n° 1, p. 110 à 126.
- S&P DOW JONES. Avril 2018. *S&P CoreLogic Case-Shiller Home Price Indices Methodology*. Extrait de <https://us.spindices.com/index-family/real-estate/sp-corelogic-case-shiller>.
- WANG, F. et P. ZORN. 1997. « Estimating house price growth with repeat sales data: What's the aim of the game? », *Journal of Housing Economics*, vol. 6, p. 93 à 118.
- WHITE, H. 2001. *Asymptotic Theory for econometricians* (version révisée), Emerald Group Publishing.
- WOOLDRIDGE, J. 2010. *Econometric Analysis of Cross Section and Panel Data* (2<sup>e</sup> édition), MIT University Press.