

# Un cours sur les indices des prix

Steve Martin

été 2020

## Contents

<b>Aperçu</b>	<b>3</b>
<b>I Introduction aux indices de prix</b>	<b>3</b>
<b>1 Programme pour Introduction aux indices de prix</b>	<b>3</b>
<b>2 Qu'est-ce qu'un indice des prix et pourquoi en avons-nous besoin?</b>	<b>4</b>
2.1 Qu'est-ce qu'un indice des prix? . . . . .	4
2.2 Utilisations d'un indice des prix . . . . .	5
2.3 Pourquoi un indice des prix? . . . . .	6
<b>3 Formule de numéro d'index</b>	<b>7</b>
3.1 Indices arithmétiques des prix . . . . .	7
3.2 Indices arithmétiques communs . . . . .	8
3.3 Indices arithmétiques moins courants . . . . .	9
3.4 Indices géométriques des prix . . . . .	11
3.5 L'indice Fisher . . . . .	12
3.6 Exemple avec R: formules d'index . . . . .	13
<b>4 Calcul d'un indice des prix</b>	<b>14</b>
4.1 Sous-indices et agrégation . . . . .	14
4.2 Factorisation d'un index . . . . .	17
4.3 Rebasing . . . . .	18
4.4 Chaînage . . . . .	20
4.5 Exemple avec R: Cohérence dans l'agrégation et le rebasage . . . . .	22
<b>5 Affectation 1</b>	<b>24</b>
<b>II Théorie de l'indice des prix</b>	<b>26</b>
<b>6 Programme pour Théorie de l'indice des prix</b>	<b>26</b>
<b>7 Qu'est-ce que la théorie de l'indice des prix?</b>	<b>27</b>
<b>8 L'approche axiomatique</b>	<b>28</b>
8.1 Un indice de prix abstrait . . . . .	28
8.2 Les axiomes . . . . .	29
8.3 Exemples de compteur . . . . .	31
8.4 Les tests . . . . .	31

8.5	Un résultat impossible . . . . .	33
8.6	Motiver les indices Jevons et Laspeyres . . . . .	33
<b>9</b>	<b>L'approche économique</b>	<b>34</b>
9.1	Indice des prix des intrants . . . . .	34
9.2	Indice des prix à la production . . . . .	36
9.3	Indice du coût de la vie . . . . .	37
9.4	Approximation d'un indice économique . . . . .	37
<b>10</b>	<b>L'approche stochastique</b>	<b>39</b>
10.1	Le problème de prédiction optimal . . . . .	39
10.2	Les indices de prix géométriques comme meilleur prédicteur . . . . .	41
10.3	Des indices de prix plus généraux . . . . .	41
10.4	Inférence statistique . . . . .	42
<b>11</b>	<b>Affectation 2</b>	<b>43</b>
<b>III</b>	<b>Indices de prix de qualité constante</b>	<b>46</b>
<b>12</b>	<b>Programme pour Indices de prix de qualité constante</b>	<b>46</b>
<b>13</b>	<b>Introduction</b>	<b>46</b>
<b>14</b>	<b>Prix potentiels</b>	<b>47</b>
14.1	Indices de prix de qualité constante . . . . .	48
14.2	Indices de prix de transaction . . . . .	49
14.3	Identification . . . . .	50
14.4	Exemple . . . . .	50
<b>15</b>	<b>Indices de prix stratifiés</b>	<b>51</b>
15.1	Indépendance conditionnelle . . . . .	52
15.2	Chevauchement . . . . .	53
15.3	Problèmes de stratification . . . . .	54
15.4	Exemple avec R: Indice stratifié . . . . .	54
<b>16</b>	<b>Indices de prix hédoniques</b>	<b>55</b>
16.1	Indices hédoniques linéaires . . . . .	56
16.2	L'approche temporelle . . . . .	57
16.3	Exemple avec R: Hédonique . . . . .	58
<b>17</b>	<b>Affectation 3</b>	<b>60</b>
<b>IV</b>	<b>Construire un indice des prix avec R</b>	<b>63</b>
<b>18</b>	<b>Programme pour Construire un indice des prix avec R</b>	<b>64</b>
<b>19</b>	<b>Affectation 4</b>	<b>64</b>
19.1	Méthodologie . . . . .	65
19.2	Instructions . . . . .	66
	<b>Références</b>	<b>69</b>

## Aperçu

Ce cours donne une introduction moderne et simple aux indices des prix. L'accent mis sur les principales idées et concepts qui sont importants pour les utilisateurs et les compilateurs d'indices de prix; par conséquent, le cours n'est pas encyclopédique et plusieurs sujets importants sont soit traités rapidement, soit pas du tout traités. De nombreuses références sont fournies pour compléter le matériel de ce cours et étendre la portée des sujets traités. Dans l'ensemble, le cours repose fermement sur le domaine de la théorie et est conçu pour aider à favoriser la compréhension de ce qu'est un indice des prix. Plusieurs exemples pratiques utilisant le langage de programmation R sont donnés tout au long du cours, ainsi que plusieurs exercices appliqués. Mais la plupart des aspects les plus «réels» de l'élaboration d'un indice des prix sont ignorés. Ce n'est pas parce que ces choses ne sont pas importantes; il est plutôt difficile d'apprécier ces sujets plus pratiques sans une solide compréhension de la théorie.

Le cours est divisé en quatre modules qui s'appuient les uns sur les autres. Le premier module donne une introduction aux indices de prix, en se concentrant principalement sur les aspects informatiques d'un indice de prix. Le deuxième module donne une large introduction à la théorie des indices des prix, en mettant l'accent sur la motivation des indices des prix qui sont produits par les organismes statistiques nationaux. Le troisième module couvre les ajustements de qualité, l'un des principaux obstacles pratiques lors de la construction d'un indice des prix. Le quatrième module rassemble le tout avec un exercice étendu pour construire et analyser un indice des prix avec R.

Tous les modules ne sont pas nécessaires pour comprendre les indices de prix, et le matériel de ce cours peut être utilisé par des apprenants ayant une variété de parcours et d'objectifs. Dans tous les cas, le premier module devrait être le point de départ, car il fournit une introduction de base sans approfondir la théorie mathématique.

Ceux qui s'intéressent principalement aux aspects de calcul d'un indice de prix peuvent suivre le module 4 après le module 1, en sautant le matériel plus théorique des modules 2 et 3. Cette séquence est appropriée pour les personnes ayant peu ou pas de connaissances préalables en économie ou en statistique, et qui le font pas besoin d'utiliser un indice des prix pour le travail d'analyse.

Ceux qui ont déjà été familiarisés avec la théorie économique peuvent suivre le module 2 après le module 1 pour acquérir une compréhension plus approfondie du matériel du module 1, suivi du module 4 pour mettre ces idées en pratique. Cette séquence est appropriée pour ceux qui utilisent des indices de prix pour le travail analytique.

Les personnes ayant déjà été exposées à l'économétrie et aux statistiques peuvent suivre le module 3 après le module 2 pour en savoir plus sur certains défis pratiques lors de la construction d'un indice de prix, suivi à nouveau par le module 4. Cette séquence est appropriée pour ceux qui utilisent et construisent des indices de prix, car elle donne les outils fondamentaux nécessaires pour concevoir un indice des prix.

## Part I

# Introduction aux indices de prix

## 1 Programme pour Introduction aux indices de prix

Un indice des prix permet de condenser systématiquement les informations sur de nombreux prix pour une gamme de produits et services en une seule valeur. Les indices des prix sont largement utilisés pour analyser et comparer les conditions macroéconomiques dans le temps et ont des implications importantes pour les politiques gouvernementales.

L'objectif de ce module est de fournir une introduction aux indices de prix, avec une application aux indices de prix des entrées et sorties de production (indice des prix à la production), et des prix à la consommation

(indice des prix à la consommation). À la fin du module, une personne devrait:

1. Comprendre les raisons pour lesquelles les indices de prix sont compilés et pourquoi ils sont importants.
2. Familiarisez-vous avec les formules d'index communes et comment les manipuler.
3. Avoir une compréhension de base de la structure d'un indice de prix et de son calcul.

Ce module est utile à la fois pour les utilisateurs et les compilateurs d'indices de prix ayant peu de connaissances préalables sur les utilisations et la construction d'un indice de prix.

Ce module se compose de lectures autodirigées, accompagnées d'un devoir. Au total, environ 10 à 15 heures devraient être consacrées à ce module.

Il n'y a pas de prérequis pour ce module, bien que les individus devraient être à l'aise avec l'algèbre de base et les arguments mathématiques simples. Un cours d'introduction à la macroéconomie est utile.

L'évaluation de ce module est basée sur une tâche composée de 20 questions à choix multiples / vrai-faux qui s'appuient sur du matériel dans le contenu du cours et les lectures. La collaboration sur la mission est la bienvenue, mais chaque personne doit soumettre son propre travail unique. La réussite de ce module nécessite au moins 65% sur l'affectation.

Les lectures de ce module proviennent des chapitres 1, 2, 4 et 9 du Manuel de l'indice des prix à la production (ILO et al. 2004b) et / ou du Manuel de l'indice des prix à la consommation (ILO et al. 2004a), publiés par le FMI (disponibles gratuitement sur leur site Web). Comme il existe un chevauchement considérable entre ces sources, le manuel PPI est utilisé comme référence principale.

Veuillez envoyer un e-mail à l'instructeurs du cours (Steve Martin ou Shaoxiong Wang ou Anastasiia Nosach) si vous avez des questions ou si vous avez besoin d'aide concernant le matériel ou le devoir du cours.

## 2 Qu'est-ce qu'un indice des prix et pourquoi en avons-nous besoin?

Un indice des prix est un outil pour mesurer l'évolution du prix des biens et services au fil du temps. Les indices des prix ont une longue histoire et constituent une statistique macroéconomique importante et largement utilisée. Cette section du cours commence par présenter ce qu'est un indice des prix, à quoi ils servent et pourquoi un indice des prix est nécessaire.

Manuel PPI: Chapitre 1, sections A, K; Chapitre 2, sections A, B, D1, D2, E, F.

Manuel de l'IPC: chapitre 1, paragraphes 1.8 à 1.12; Chapitre 2.

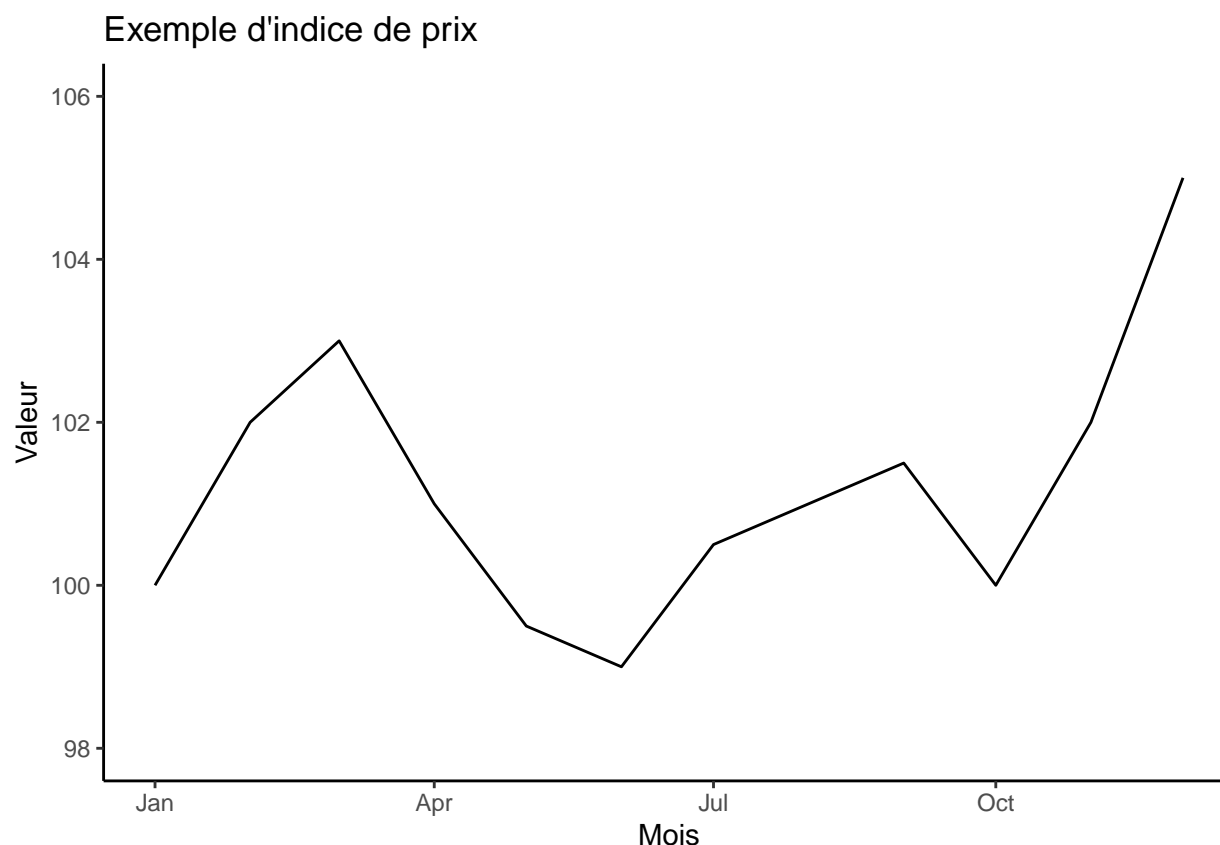
### 2.1 Qu'est-ce qu'un indice des prix?

Un indice de prix est une série chronologique économique qui donne la variation en pourcentage des prix d'une collection de biens et services entre un moment donné et un point de référence fixe (c'est-à-dire la période de base ou la période de référence).<sup>1</sup> Un indice de prix est généralement représenté de manière à soustraire 100 de la valeur de l'indice donne la variation en pourcentage des prix des biens et services capturés par l'indice. Par exemple, si un indice des prix des widgets a une valeur de 110 pendant une certaine période, alors les prix des widgets ont augmenté de 10% depuis la période de base. Cette représentation signifie également que l'indice a une valeur de 100 dans la période de base.

Le graphique ci-dessous donne un exemple d'indice de prix mensuel montrant la variation des prix de chaque mois de l'année par rapport à janvier (la période de base), de sorte que chaque point du graphique donne la variation en pourcentage des prix depuis janvier. Par exemple, en mars, l'indice avait une valeur de 103, ce qui signifie que les prix ont augmenté de 3% entre mars et janvier, tandis que l'indice avait une valeur de 99 en juin, ce qui signifie que les prix ont diminué de 1% entre janvier et juin.

---

<sup>1</sup>Un prix l'indice peut également être utilisé pour mesurer la différence de prix à différents endroits au même moment, bien qu'il s'agisse d'une utilisation moins courante d'un indice de prix.



Étant donné qu'un indice des prix ne mesure que la variation relative des prix depuis la période de base, le changement d'échelle de l'indice ne modifie pas son contenu économique. Par exemple, la division de 100 de toutes les valeurs du graphique ci-dessus ne modifie pas la forme du graphique ou la variation en pourcentage de la valeur d'index qui en résulte depuis la période de base. C'est pour cette raison que l'ampleur d'un indice des prix ne donne pas une mesure du niveau des prix à un moment donné - un indice des prix ne peut que mesurer la variation des prix au fil du temps.

L'utilité d'un indice des prix est qu'il permet de condenser systématiquement les informations sur de nombreux prix pour une gamme de biens et services en une seule valeur qui mesure l'évolution des prix de ces biens et services au fil du temps. Bien que les biens et services qui entrent dans la portée d'un indice des prix dépendent de son objet, tous les indices des prix donnent une mesure des mouvements de prix agrégés pour un ensemble bien défini de produits. Cela donne à son tour une mesure des pressions inflationnistes dans une économie et, par conséquent, les indices des prix sont omniprésents lors de l'examen des conditions et des politiques macroéconomiques.

## 2.2 Utilisations d'un indice des prix

Dans la pratique, il existe deux grandes catégories dans lesquelles les indices des prix publiés par les organismes statistiques nationaux entrent dans la catégorie: les indices des prix à la consommation (IPC) et les indices des prix à la production (IPP). Comme leur nom l'indique, les IPC sont axés sur la mesure des variations de prix des biens et services achetés par les consommateurs (consommation finale), tandis que les IPP se concentrent sur la mesure des prix facturés par les producteurs pour les biens et services produits et les prix payés pour les intrants dans la production. Malgré les nombreuses subtilités sur ce qui constitue un consommateur et ce qui constitue un producteur, les IPC et les IPP cherchent à mesurer la variation des prix au fil du temps pour les biens et services qui sont négociés dans une économie. Par conséquent, il y a souvent chevauchement considérable dans la façon dont les IPC et les IPP sont calculés et utilisés, et il n'est pas nécessaire de faire la distinction entre un indice comme un IPC ou un IPP — il suffit souvent de parler

d'un indice général des prix.

L'utilisation la plus importante d'un indice des prix, qu'il s'agisse d'un IPC ou d'un IPP, est probablement de mesurer l'inflation - le changement systématique des prix dans une économie au fil du temps. Au-delà d'être un indicateur macroéconomique important, la mesure de l'inflation affecte directement de nombreuses interactions économiques dans les économies modernes. Par exemple, de nombreuses banques centrales opèrent dans un régime de ciblage de l'inflation, selon lequel la politique monétaire est menée en partie pour maintenir le taux d'inflation dans une fourchette prédéterminée (par exemple, 1% à 3% par an). Ce régime politique nécessite une mesure fréquente et opportune de l'inflation, et l'instrument utilisé pour mesurer l'inflation peut avoir un impact direct sur la politique monétaire et, par conséquent, sur les taux d'intérêt. L'inflation affecte également les salaires et les contrats de services, les paiements de pension et les prestations de sécurité sociale, car ces contrats sont souvent indexés sur l'inflation — si les prix augmentent de 2% par an, les salaires augmentent également de 2% par an. L'indexation préserve la valeur des paiements au fil du temps sans avoir à constamment renégocier les contrats ou repenser les politiques. La mesure de l'inflation a donc des implications importantes sur le revenu des ménages et les revenus de l'industrie.

Au-delà de la mesure de l'inflation, l'une des principales utilisations des indices de prix pour les organismes statistiques nationaux est de déflater les valeurs agrégées dans un cadre de comptabilité nationale pour obtenir une mesure de l'évolution de la production de biens et services réels au fil du temps. Cela équivaut à trouver un indice de quantité pour condenser systématiquement les informations sur de nombreuses sorties pour une gamme de biens et services en une seule valeur qui mesure la variation de la quantité de biens et services produits au fil du temps. Bien qu'il existe de nombreux détails, la justification de la déflation des valeurs agrégées pour obtenir une mesure de la production réelle est assez simple. Si  $V_1/V_0$  est le rapport entre la valeur de la production de la période 1 et la valeur de la production de la période 0, et si cela peut être décomposé en un indice de quantité  $Q$  et un indice de prix  $I$  tels que  $V_1/V_0 = IQ$ , alors la variation de la production réelle entre la période 0 et la période 1 est simplement  $Q = V_1/(V_0I)$ . L'évolution de la valeur de la production au fil du temps - une quantité connue - peut être transformée en une mesure de l'évolution de la production réelle en la déflatant simplement avec un indice des prix.

### 2.3 Pourquoi un indice des prix?

Le point de départ pour comprendre l'utilité d'un indice des prix est de reconnaître qu'il n'y en a pas besoin si l'objectif est de comparer le prix d'un seul bien ou service au cours d'une période au prix du même bien ou service au cours d'une autre période. Si  $p_1$  est le prix d'un bien pendant la période 1 et  $p_0$  est le prix de ce même bien pendant la période 0, alors le prix relatif  $p_1/p_0$  donne sans ambiguïté la variation en pourcentage du prix entre la période 1 et la période 0. La nécessité d'un indice des prix découle d'un problème fondamental de comparaison d'une collecte de deux prix ou plus à deux moments pour arriver à une variation globale des prix.

Pour illustrer ce problème de comparaison, supposons qu'il existe deux biens, notés  $i$  et  $j$ , qui se vendent dans les périodes 0 et 1. Si le prix du bon  $i$  et le prix du bon  $j$  augmentent tous les deux entre les périodes 0 et la période 1, il n'est donc pas contesté de dire que les prix ont augmenté entre la période 0 et la période 1 (même si la quantité de prix a augmenté moins clairement). De même, si les deux prix baissent, il est naturel de dire que les prix ont baissé. Cependant, un problème se pose si le prix du bon  $i$  augmente entre la période 0 et période 1 tandis que le prix du bon  $j$  diminue. Dans ce cas, il n'est pas clair si les prix ont augmenté ou diminué — il n'y a pas de moyen évident de donner une direction unique à l'évolution des prix au fil du temps.

Une solution à ce problème de comparaison consiste à trouver un moyen d'agréger les prix relatifs pour de nombreux biens et services pour former un ratio unique, qui est ensuite interprété comme l'évolution globale des prix au fil du temps. C'est le travail d'un indice des prix. Bien qu'un indice des prix résout le problème de la comparaison, il introduit un nouveau défi: comment combiner au mieux les prix relatifs pour produire une mesure unique décrivant la variation globale des prix. Par conséquent, diverses formules d'indice-nombre ont été proposées pour construire un indice de prix.

### 3 Formule de numéro d'index

Il existe un grand nombre de formules d'indices qui peuvent être utilisées pour combiner des informations sur les prix à deux moments dans le temps pour produire un indice des prix, malgré l'objectif commun de mesurer les variations de prix au fil du temps. Cette section du cours couvre 10 formules d'indice-nombre importantes qui sont utilisées dans la pratique pour calculer un indice de prix. Le nom de la personne qui l'a développée est associé à chaque formule. Il est important d'associer les noms aux formules, car ils sont généralement désignés par leur nom dans l'application. Il s'agit malheureusement d'un exercice de mémorisation, car la plupart des formules d'index ont une forme similaire.

Pour aider à construire l'intuition et une compréhension plus approfondie de ce que font les formules de l'indice-nombre, ces différentes approches pour calculer un indice de prix peuvent être classées en gros comme des indices de prix arithmétiques ou des indices de prix géométriques. Toutes les formules de nombre d'index ne tombent pas dans ces deux groupes — par exemple, il existe également des indices de prix harmoniques, une catégorie qui peut chevaucher des indices arithmétiques — mais les indices arithmétiques et géométriques sont assez faciles à comprendre, avec la plupart des indices- formules de nombres entrant dans l'une de ces deux catégories.

En pratique, un indice de prix est multiplié par 100, de sorte que le changement en pourcentage de la valeur de l'indice sur deux périodes est simplement la valeur de l'indice moins 100. Il s'agit simplement d'une normalisation pratique qui n'a aucune incidence sur le contenu économique d'un indice de prix, et est ignoré dans cette section. Notez que toutes les formules d'index peuvent être multipliées par 100 pour les mettre sur l'échelle habituelle.

Manuel PPI: Chapitre 1, sections B1 - B3.

#### 3.1 Indices arithmétiques des prix

Un indice arithmétique des prix prend les prix relatifs pour une collection de biens et services sur deux périodes et les combine ensemble comme moyenne pondérée. Autrement dit, un indice arithmétique est simplement la variation moyenne du prix entre deux points dans le temps. Laissant les marchandises être énumérées par  $i = 1, \dots, n$ , un index arithmétique entre la période 0 et la période 1 a la forme générale<sup>2</sup>

$$I^A = \sum_{i=1}^n \omega_i \frac{p_{i1}}{p_{i0}},$$

où  $p_{it}$  est le prix du bon  $i$  dans la période  $t = 0, 1$ , et  $\omega_i \geq 0$  est le poids que le bon  $i$  reçoit dans le calcul de l'indice, tel que  $\sum_{i=1}^n \omega_i = 1$ .

Différents indices de prix arithmétiques correspondent à des cas particuliers de l'indice arithmétique général, selon le choix des pondérations. La plupart de ces choix utilisent des informations sur la quantité d'un bien vendu, de sorte que les poids donnent une mesure de l'importance économique d'un bien. Notons  $q_{it}$  la quantité de bon  $i$  consommée / produite dans la période  $t = 0, 1$ .

Avant d'examiner des indices arithmétiques spécifiques, il convient de noter qu'un indice arithmétique peut toujours être écrit comme le rapport des dépenses / recettes pour un "panier" de biens et services à deux points dans le temps, de sorte que pour tout ensemble de pondérations, il est implicite "quantités" telles que

$$I^A = \frac{\sum_{i=1}^n p_{i1} \tilde{q}_i}{\sum_{i=1}^n p_{i0} \tilde{q}_i},$$

---

<sup>2</sup>La lettre majuscule sigma  $\Sigma$  est l'opérateur de sommation. Pour une collection de nombres  $x_1, x_2, \dots, x_n$ ,  $\sum_{i=1}^n x_i$  signifie  $x_1 + x_2 + \dots + x_n$ .

où  $\tilde{q}_i = \alpha \omega_i / p_{i0}$  pour un certain facteur de proportionnalité  $\alpha$ .<sup>3</sup> Ainsi, un indice arithmétique peut toujours être interprété comme le rapport des dépenses nécessaires à l'achat d'un panier fixe de biens à deux moments (ou les revenus d'un panier fixe de biens à deux moments). Le choix du panier est lié un à un au choix des pondérations utilisées pour agréger les prix relatifs. Les deux représentations de l'indice arithmétique sont utilisées, car certains indices sont plus faciles à représenter sous une forme que l'autre.

### 3.2 Indices arithmétiques communs

Il existe six principaux indices de prix arithmétiques qui sont utilisés dans la pratique, chacun correspondant à une déclaration différente sur le poids qu'un prix relatif devrait recevoir dans le calcul de l'indice.

**Indice Carli.** La définition de  $\omega_i = 1/n$  entraîne l'index Carli

$$I_C^A = \frac{1}{n} \sum_{i=1}^n \frac{p_{i1}}{p_{i0}}.$$

L'indice Carli adopte une position neutre sur les pondérations et considère chaque prix relatif comme tout aussi important.

**Indice de Dutot.** La définition de  $\omega_i = p_{i0} / \sum_{j=1}^n p_{j0}$  entraîne l'index Dutot

$$I_D^A = \frac{\sum_{i=1}^n p_{i1}}{\sum_{i=1}^n p_{i0}}.$$

L'indice Dutot donne plus de poids aux prix apparentés qui ont un prix supérieur à la période 0, en comparant le prix moyen des biens de la période 1 au prix moyen de la période 0.

**Indice Lowe.** La définition de  $\omega_i = p_{i0} q_{ib} / \sum_{j=1}^n p_{j0} q_{jb}$  entraîne l'index Lowe

$$I_L^A = \frac{\sum_{i=1}^n p_{i1} q_{ib}}{\sum_{i=1}^n p_{i0} q_{ib}},$$

où  $q_{ib}$  est la quantité de biens  $i$  au cours d'une période de base  $b$ , généralement avant la période 0. Les pondérations de l'indice de Lowe sont des parts de dépenses / revenus "hybrides" pour le panier de biens et services dans la période  $b$  en utilisant les prix de la période 0.

**Indice de Laspeyres.** La définition de  $\omega_i = p_{i0} q_{i0} / \sum_{j=1}^n p_{j0} q_{j0}$  entraîne l'index de Laspeyres

$$I_L^A = \frac{\sum_{i=1}^n p_{i1} q_{i0}}{\sum_{i=1}^n p_{i0} q_{i0}}.$$

L'indice de Laspeyres pondère les prix relatifs en fonction de leur part des dépenses / recettes pour la période 0, et constitue un cas particulier de l'indice de Lowe.

**Indice de Paasche.** La définition de  $\omega_i = p_{i0} q_{i1} / \sum_{j=1}^n p_{j0} q_{j1}$  entraîne l'index de Paasche

---

<sup>3</sup>La valeur de  $\alpha$  n'a aucun impact sur l'indice, mais est nécessaire pour que la propriété interprète  $\tilde{q}_i$  comme une quantité. Par exemple, si  $\omega_i$  est la part des dépenses de la période 0 sur les biens  $i$ ,  $p_{i0} q_{i0} / \sum_{j=1}^n p_{j0} q_{j0}$ , puis  $\alpha = \sum_{j=1}^n p_{j0} q_{j0}$  pour que  $\tilde{q}_i = q_{i0}$ .



$$I_P^A = \frac{\sum_{i=1}^n p_{i1} q_{i1}}{\sum_{i=1}^n p_{i0} q_{i1}}.$$

Comme l'indice Laspeyres, l'indice Paasche est un cas particulier de l'indice Lowe et utilise des parts hybrides dépenses / recettes pour pondérer les prix relatifs.

Il convient de noter que l'indice de Paasche est souvent calculé comme une moyenne harmonique pondérée, avec des parts de dépenses / recettes de la période 1 comme pondérations, de sorte que

$$I_P^A = \left( \sum_{i=1}^n \frac{\frac{p_{i1} q_{i1}}{\sum_{j=1}^n p_{j1} q_{j1}}}{\frac{p_{i1}}{p_{i0}}} \right)^{-1}.$$

Il s'agit d'un exemple d'indice de prix harmonique et d'un moyen pratique de calculer un indice de Paasche si seules les parts des dépenses / recettes de la période 1 sont connues, plutôt que les quantités de la période 1.

**Indice Young.** La définition de  $\omega_i = p_{ib} q_{ib} / \sum_{j=1}^n p_{jb} q_{jb}$  entraîne l'index Young

$$I_Y^A = \sum_{i=1}^n \frac{p_{ib} q_{ib}}{\sum_{j=1}^n p_{jb} q_{jb}} \frac{p_{i1}}{p_{i0}},$$

où  $q_{ib}$  est la quantité de bon  $i$  dans une période de base  $b$ , avec  $p_{ib}$  comme prix, généralement avant la période 0. L'indice Young utilise la période  $b$  dépenses / parts de revenus sous forme de pondérations.<sup>4</sup> L'indice Laspeyres est un cas particulier de l'indice Young.

Dans la plupart des applications, un indice de Laspeyres est la formule d'indice-nombre souhaitée pour un indice de prix arithmétique, en partie parce que les pondérations peuvent être observées séparément des informations sur les prix. Les pondérations pour un IPC, par exemple, peuvent provenir d'une enquête nationale représentative des dépenses des ménages, et donc seules les informations sur les prix doivent être collectées pour calculer un indice de Laspeyres. Bien que les pondérations de Young soient également directement observables, elles ne reflètent pas nécessairement les changements dans la composition des dépenses ou des sources de revenus entre la période  $b$  et la période 0. Alternativement, les indices Lowe et Paasche ont tous deux des pondérations qui doivent être explicitement calculées. Comme l'indice Young, l'indice Lowe utilise des informations de quantité obsolètes, tandis que l'indice Paasche utilise des informations de quantité de la période actuelle, ce qui signifie généralement qu'il ne peut pas être calculé en temps opportun.

Malgré l'objectif de calculer un indice de Laspeyres, dans la plupart des applications, un indice arithmétique est souvent calculé comme un indice de Young (ou parfois un indice de Lowe), car il est difficile d'obtenir des informations en temps opportun sur le partage des dépenses / recettes. L'attente des informations de pondération peut nécessiter de retarder la production d'un indice bien après la période 0. L'espoir est que les pondérations utilisées dans l'indice de Young offrent une approximation raisonnable des pondérations de Laspeyres.

### 3.3 Indices arithmétiques moins courants

Il existe une variété d'indices arithmétiques qui sont rarement utilisés dans la pratique, mais peuvent être utiles à connaître. Comme ces formules d'index sont rarement utilisées, ce matériau peut être sauté en toute sécurité.

<sup>4</sup>Dans certains cas, un indice de Young est utilisé pour faire référence à un indice arithmétique général, plutôt qu'à un indice arithmétique avec un ensemble particulier de pondérations.

**Indice de Palgrave.** La définition de  $\omega_i = p_{i1}q_{i1} / \sum_{j=1}^n p_{j1}q_{j1}$  entraîne l'index Palgrave

$$I_p^A = \sum_{i=1}^n \frac{p_{i1}q_{i1}}{\sum_{j=1}^n p_{j1}q_{j1}} \frac{p_{i1}}{p_{i0}}.$$

L'indice Palgrave utilise les parts de dépenses de la période 1 comme pondérations et constitue un cas particulier de l'indice Young.

**Index sans nom.** Réglage

$$\omega_i = \frac{1}{2} \frac{p_{i0}q_{i0}}{\sum_{j=1}^n p_{j0}q_{j0}} + \frac{1}{2} \frac{p_{i1}q_{i1}}{\sum_{j=1}^n p_{j1}q_{j1}}$$

donne une formule de numéro d'index sans nom,

$$I_U^A = \sum_{i=1}^n \left( \frac{1}{2} \frac{p_{i0}q_{i0}}{\sum_{j=1}^n p_{j0}q_{j0}} + \frac{1}{2} \frac{p_{i1}q_{i1}}{\sum_{j=1}^n p_{j1}q_{j1}} \right) \frac{p_{i1}}{p_{i0}}.$$

Cet indice est un mélange des indices Laspeyres et Palgrave, les pondérations étant données par la part moyenne des dépenses / recettes entre la période 0 et la période 1.

**Indice Drobisch.** Réglage

$$\omega_i = \frac{1}{2} \frac{p_{i0}q_{i0}}{\sum_{j=1}^n p_{j0}q_{j0}} + \frac{1}{2} \frac{p_{i0}q_{i1}}{\sum_{j=1}^n p_{j0}q_{j1}}$$

résultats dans l'indice Drobisch

$$I_d^A = \frac{1}{2} \frac{\sum_{i=1}^n p_{i1}q_{i0}}{\sum_{i=1}^n p_{i0}q_{i0}} + \frac{1}{2} \frac{\sum_{i=1}^n p_{i1}q_{i1}}{\sum_{i=1}^n p_{i0}q_{i1}}.$$

Cet indice est un mélange des indices Laspeyres et Paasche.

**Indice de Walsh.** Définition de  $\omega_i = p_{i0}\sqrt{q_{i0}q_{i1}} / \sum_{j=1}^n p_{j0}\sqrt{q_{j0}q_{j1}}$  entraîne l'index de Walsh

$$I_W^A = \frac{\sum_{i=1}^n p_{i1}\sqrt{q_{i0}q_{i1}}}{\sum_{i=1}^n p_{i0}\sqrt{q_{i0}q_{i1}}}.$$

Cet indice utilise un panier qui contient la moyenne géométrique des quantités de la période 0 et de la période 1.

**Indice Marshall-Edgeworth.** Définition de  $\omega_i = p_{i0}(q_{i0} + q_{i1}) / \sum_{j=1}^n p_{j0}(q_{j0} + q_{j1})$  résulte dans l'indice Marshall-Edgeworth

$$I_M^A = \frac{\sum_{i=1}^n p_{i1}(q_{i0} + q_{i1})/2}{\sum_{i=1}^n p_{i0}(q_{i0} + q_{i1})/2}.$$

Comme l'indice de Walsh, cet indice utilise un panier qui prend une moyenne des quantités de la période 0 et de la période 1.

**Indice Geary-Khamis.** Paramètre  $\omega_i = p_{i0}/(1/q_{i0} + 1/q_{i1}) / \sum_{j=1}^n p_{j0}/(1/q_{j0} + 1/q_{j1})$  se traduit par l'indice Geary-Khamis

$$I_G^A = \frac{\sum_{i=1}^n 2p_{i1}/(1/q_{i0} + 1/q_{i1})}{\sum_{i=1}^n 2p_{i0}/(1/q_{i0} + 1/q_{i1})}.$$

Comme les indices Walsh et Marshall-Edgeworth, cet indice utilise un panier qui prend en moyenne les quantités de la période 0 et de la période 1.

### 3.4 Indices géométriques des prix

Un indice de prix géométrique est entièrement analogue à un indice arithmétique, sauf que les prix relatifs sont agrégés avec une moyenne géométrique au lieu d'une moyenne arithmétique. Autrement dit, un indice géométrique général des prix est donné par<sup>5</sup>

$$I^G = \prod_{i=1}^n \left( \frac{p_{i1}}{p_{i0}} \right)^{\omega_i}.$$

Comme pour les indices arithmétiques, différents indices géométriques correspondent à différents choix pour les poids.<sup>6</sup>

**Indice de Jevons.** La définition de  $\omega_i = 1/n$  entraîne l'index Jevons

$$I_J^G = \prod_{i=1}^n \left( \frac{p_{i1}}{p_{i0}} \right)^{1/n},$$

ou équivalent,

$$I_J^G = \frac{\prod_{i=1}^n p_{i1}^{1/n}}{\prod_{i=1}^n p_{i0}^{1/n}}.$$

L'indice de Jevons est l'analogue géométrique de l'indice de Carli ou Dutot.

**Indice géométrique de Laspeyres.** La définition de  $\omega_i = p_{i0}q_{i0} / \sum_{j=1}^n p_{j0}q_{j0}$  entraîne l'index géométrique de Laspeyres

$$I_L^G = \prod_{i=1}^n \left( \frac{p_{i1}}{p_{i0}} \right)^{\frac{p_{i0}q_{i0}}{\sum_{j=1}^n p_{j0}q_{j0}}}.$$

<sup>5</sup>La lettre majuscule pi  $\Pi$  est l'opérateur du produit. Pour une collection de nombres  $x_1, x_2, \dots, x_n$ ,  $\prod_{i=1}^n x_i$  signifie  $x_1 \times x_2 \times \dots \times x_n$ .

<sup>6</sup>Il est intéressant de noter que les poids peuvent également être choisis de sorte que tout indice géométrique soit également un indice arithmétique (et vice versa). Autrement dit, pour tout  $\omega_i$ , il existe un  $\tilde{\omega}_i$  tel que  $\prod_{i=1}^n (p_{i1}/p_{i0})^{\omega_i} = \sum_{i=1}^n \tilde{\omega}_i p_{i1}/p_{i0}$ . Voir Balk (2008Section 4.2) pour voir à quoi ressemblent ces poids. Cadrer un index comme arithmétique ou géométrique est vraiment juste un cas de cadrage des poids.

Similaire à l'indice de Jevons, il s'agit de l'analogue géométrique de l'indice de Laspeyres. Il est également trivial de définir un indice géométrique de Young en utilisant la part des dépenses / recettes de la période- $b$  plutôt que celle de la période 0.<sup>7</sup>

**Indice de Törnqvist.** Réglage

$$\omega_i = \frac{1}{2} \frac{p_{i0}q_{i0}}{\sum_{j=1}^n p_{j0}q_{j0}} + \frac{1}{2} \frac{p_{i1}q_{i1}}{\sum_{j=1}^n p_{j1}q_{j1}}$$

résulte en l'indice de Törnqvist, qui est généralement exprimé en

$$\log(I_T^G) = \sum_{i=1}^n \left( \frac{1}{2} \frac{p_{i0}q_{i0}}{\sum_{j=1}^n p_{j0}q_{j0}} + \frac{1}{2} \frac{p_{i1}q_{i1}}{\sum_{j=1}^n p_{j1}q_{j1}} \right) \log \left( \frac{p_{i1}}{p_{i0}} \right).$$

L'indice de Törnqvist développe l'indice géométrique de Laspeyres en utilisant les parts de dépenses à la fois pour la période 0 et la période 1 pour former des pondérations (c'est-à-dire la part des dépenses moyennes entre la période 0 et la période 1).

L'indice de Jevons est généralement synonyme d'index géométrique, et il trouve application dans des situations où il n'y a pas d'informations de quantité pour former des poids (par opposition à l'utilisation d'un index Carli ou Dutot). Parfois, un indice de Jevons pondéré est utilisé comme raccourci pour l'indice géométrique général.

Un point à noter à propos des indices géométriques est qu'ils sont toujours plus petits que leurs homologues arithmétiques. Pour tout poids donné, on peut montrer que  $I^G \leq I^A$ , avec égalité uniquement lorsque tous les prix relatifs sont égaux ou tous les prix relatifs sauf un ont un poids nul. Par conséquent, un indice géométrique montre toujours une augmentation des prix plus faible dans le temps (ou une diminution plus importante) que l'indice arithmétique correspondant.<sup>8</sup> C'est un inconvénient important d'avoir un menu de nombres d'index parmi lesquels choisir, car le choix de la formule de nombre d'index a un impact sur la mesure de l'inflation qui en résulte.<sup>9</sup>

### 3.5 L'indice Fisher

Un indice de prix important qui n'est ni un indice géométrique ni un indice arithmétique est l'indice Fisher, qui est la moyenne géométrique de l'indice Laspeyres (arithmétique) et de l'indice Paasche (arithmétique),

$$\begin{aligned} I_F &= \sqrt{\frac{\sum_{i=1}^n p_{i1}q_{i0}}{\sum_{i=1}^n p_{i0}q_{i0}} \times \frac{\sum_{i=1}^n p_{i1}q_{i1}}{\sum_{i=1}^n p_{i0}q_{i1}}} \\ &= \sqrt{I_L^A \times I_P^A}. \end{aligned}$$

L'indice Fisher est souvent considéré comme un indice idéal car il traite symétriquement les informations de la période 0 et de la période 1.<sup>10</sup> En pratique cependant, l'indice de Fisher n'est pas fréquemment utilisé

<sup>7</sup>La définition d'un indice géométrique de Lowe est moins évidente. Par exemple, l'indice géométrique de Paasche utilise les parts des dépenses / recettes de la période 1 comme pondérations, au lieu des pondérations hybrides utilisées pour l'indice arithmétique de Paasche. Peut-être un meilleur nom pour cet index serait l'indice de Palgrave géométrique.

<sup>8</sup>La différence entre un indice arithmétique et un indice géométrique a tendance à être plus grande lorsque les prix relatifs sont plus dispersés, bien que cela n'est pas toujours le cas. Il est possible que la différence devienne plus grande lorsque la variance entre les prix relatifs devient plus petite — voir Lord (2002).

<sup>9</sup>Le même raisonnement peut être utilisé pour montrer qu'un indice de prix harmonique est toujours plus petit que l'indice géométrique correspondant — voir Bullen (2003II 2.1 Corollaire 2).

<sup>10</sup>L'indice Törnqvist est également considéré comme idéal pour une raison similaire, et ces types d'indices sont souvent appelés superlatifs.

par les organismes statistiques nationaux car il n'est pas opportun de le calculer. En effet, cela dépend de l'indice de Paasche, qui nécessite des informations sur les quantités de la période 1 en plus des prix de la période 1, ce qui a été historiquement impossible à collecter pour les organismes statistiques nationaux.

### 3.6 Exemple avec R: formules d'index

Toutes les formules de l'indice-nombre présentées dans cette section sont simplement des moyennes pondérées et sont assez faciles à calculer dans R étant donné les informations sur les prix et certains poids.

```
# Apportez dans la bibliothèque gpindex
library(gpindex) # install.packages("gpindex")

# Faites des prix apparentés
relatives <- c(1.1, 1.2, 0.9, 1.1)

# Faites des poids
weights <- c(0.25, 0.3, 0.3, 0.15)

# Calculer les indices
c(Carli = mean_arithmetic(relatives),
  Jevons = mean_geometric(relatives),
  Arithmétique = mean_arithmetic(relatives, weights),
  Géométrie = mean_geometric(relatives, weights))

##          Carli          Jevons Arithmétique  Géométrie
##    1.075000    1.069184    1.070000    1.063125
```

Le type d'indices arithmétiques et géométriques que cela calcule dépend entièrement de la façon dont les poids sont calculés. Habituellement, les pondérations proviennent des données sur les parts des dépenses / recettes, et la manière dont ces informations sont utilisées pour pondérer les prix relatifs affectera le type d'indice de prix calculé.

```
# Part des dépenses / recettes de la période de base
share0 <- c(0.25, 0.3, 0.3, 0.15)

# Part des dépenses / recettes de la période en cours
share1 <- c(0.2, 0.2, 0.4, 0.2)

# Calculer les indices
c(Laspeyres = mean_arithmetic(relatives, share0),
  Paasche = mean_harmonic(relatives, share1),
  Fisher = sqrt(mean_arithmetic(relatives, share0) *
    mean_harmonic(relatives, share1)),
  `Géométrie Laspeyres` = mean_geometric(relatives, share0),
  `Géométrie Paasche` = mean_geometric(relatives, share1),
  Tornqvist = mean_geometric(relatives, (share0 + share1) / 2))

##          Laspeyres          Paasche          Fisher
##    1.070000    1.025907    1.047721
## Géométrie Laspeyres  Géométrie Paasche  Tornqvist
##    1.063125    1.032976    1.047942
```

Il est intéressant de noter que tout indice géométrique peut être transformé en indice arithmétique par un changement approprié de poids. Cela permet de déterminer facilement la contribution de chaque prix relatif à la valeur de l'indice.

```

# Tornqvist index
mean_geometric(relatives, (share0 + share1) / 2)

## [1] 1.047942

# Changer les poids
new_weights <- weights_transmute(0, 1)(relatives, (share0 + share1) / 2)
mean_arithmetic(relatives, new_weights) # same as geometric mean

## [1] 1.047942

# Contribution de chaque relative
(relatives - 1) * new_weights

## [1] 0.02095439 0.04455317 -0.03600469 0.01629786

```

## 4 Calcul d'un indice des prix

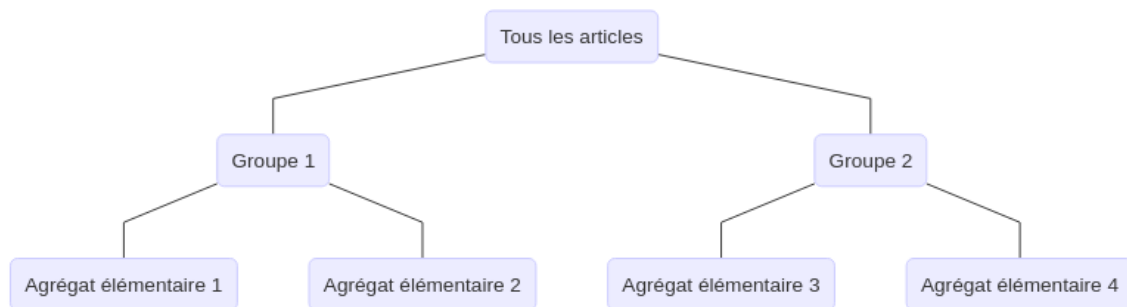
N'importe laquelle des formules de numéro d'index de la section précédente peut être utilisée pour créer un indice de prix qui mesure la variation de prix pour une seule collection de biens et services sur deux périodes. Dans l'application, cependant, les indices de prix sont calculés pour une variété de biens et services différents, et sont calculés sur de nombreuses périodes, pas seulement deux. Ces deux points ont des implications sur la façon dont les formules de l'indice-nombre dans la section précédente sont utilisées pour calculer un indice de prix dans la pratique. Cette section du cours explore comment les formules d'indice à deux périodes de la section précédente peuvent être utilisées pour calculer un indice de prix pour plusieurs groupes de biens et services sur plusieurs périodes.

Manuel PPI: Chapitre 4, section D; Chapitre 9, paragraphes 9.6 à 9.24 et sections B3, B4, C1 à C7.2.

Manuel de l'IPC: chapitre 1, paragraphes 1.147 à 1.164; Chapitre 4, paragraphes 4.16–4.33.

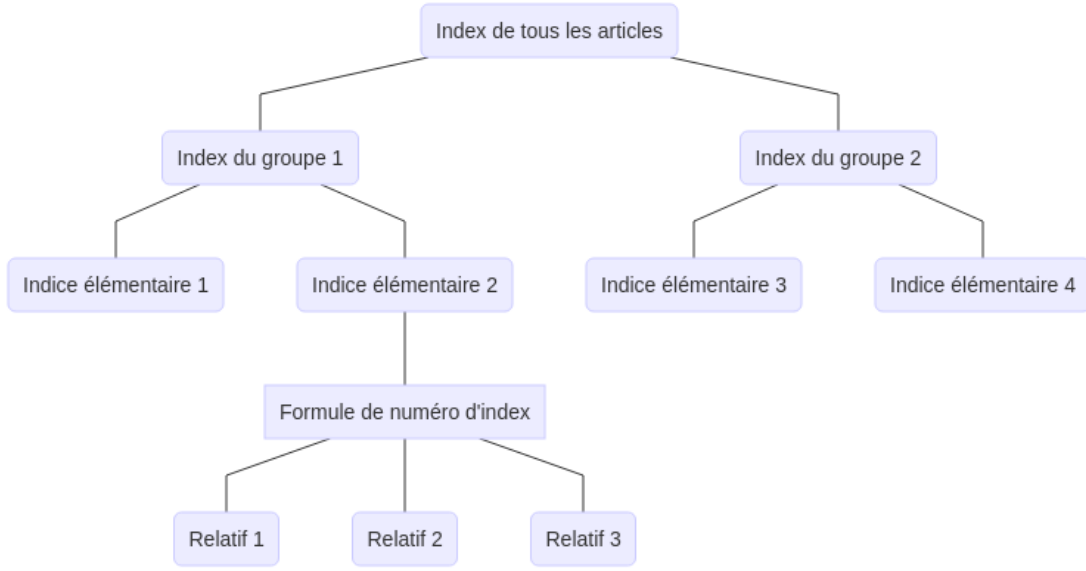
### 4.1 Sous-indices et agrégation

La plupart des indices des prix ont une structure hiérarchique (ou agrégée), de sorte que les biens et services sont divisés en groupes ou catégories de plus en plus larges qui s'étendent à tous les biens et services couverts par l'indice des prix. Par exemple, les marchandises d'un IPC peuvent être divisées en grandes catégories telles que la nourriture, le logement, le transport, etc. Ces catégories sont ensuite divisées en catégories plus petites; les aliments peuvent être divisés en viande, journal intime, volaille, etc. Ce type de structure présente l'avantage évident de fournir un indice de prix pour les sous-catégories de biens et services qui présentent souvent un intérêt à part entière (bien qu'il existe d'autres, plus raisons subtiles de vouloir structurer un indice de prix comme celui-ci). Le graphique ci-dessous donne un exemple d'une structure d'agrégation simple à trois niveaux.



Une structure hiérarchique ne pose pas de problème particulier pour créer un indice de prix — l’une des formules de numéro d’index de la section précédente peut toujours être utilisée pour agréger les prix relatifs des biens et services dans chaque groupe de la hiérarchie à produire une collection d’indices de prix pour chaque niveau de la hiérarchie. Le calcul de l’indice doit simplement être répété pour chaque groupe de biens et services dans la structure d’agrégation. Il s’agit de l’approche directe pour calculer un indice de prix avec une structure hiérarchique.

Cependant, avoir une structure hiérarchique suggère que le calcul peut être simplifié avec une procédure en deux étapes. Dans la première étape, les prix relatifs sont combinés au niveau d’agrégation le plus bas en utilisant l’une des formules de numéro d’index de la section précédente. Cela donne une collection d’indices de prix appelés indices élémentaires (ou élémentaires). Ces indices de prix donnent l’évolution des prix dans le temps pour de petits groupes de biens et services relativement homogènes appelés agrégats élémentaires qui constituent la base d’un indice de prix. Une fois les indices élémentaires calculés, les indices des niveaux d’agrégation supérieurs sont calculés en combinant les indices élémentaires à l’aide de pondérations. Ces pondérations fournissent une mesure de l’importance économique des biens et services dans chaque agrégat élémentaire afin de combiner les indices élémentaires pour former un indice de niveau supérieur. Si l’indice cible est un indice arithmétique, alors les agrégats élémentaires doivent être combinés avec une moyenne arithmétique pondérée; sinon, si un indice géométrique est l’indice cible, une moyenne géométrique pondérée doit être utilisée. Autrement dit, chaque agrégat élémentaire est traité comme un bien unique, les indices élémentaires étant les prix apparentés pour ces produits, et cette information est introduite dans une formule de nombre d’indice pour produire un indice de niveau supérieur.



Étant donné qu'il existe deux façons de calculer un indice de prix avec une structure hiérarchique, le calcul direct donne-t-il le même résultat que le calcul en deux étapes? La réponse est oui pour un indice arithmétique ou géométrique, à condition que les poids utilisés pour agréger les agrégats élémentaires soient formés en utilisant les poids sous-jacents pour l'indice cible. Par exemple, si l'objectif est de calculer un indice de Laspeyres, le poids d'un agrégat élémentaire devrait être la part des dépenses / revenus pour les biens de cet agrégat élémentaire au cours de la période 0. Comme les indices géométriques et arithmétiques sont calculés en tant que moyennes pondérées, il ne devrait pas être surprenant qu'un indice puisse être calculé en regroupant d'abord les parents de prix, en calculant les moyennes au niveau du groupe, puis en faisant la moyenne des moyennes au niveau du groupe pour former une moyenne globale. Par conséquent, ces indices seraient cohérents dans l'agrégation, bien que toutes les formules de nombre d'indices ne soient pas cohérentes dans l'agrégation. Par exemple, l'indice de Fisher n'est pas cohérent dans l'agrégation, et l'utilisation de la procédure en deux étapes pour le calculer donnera une réponse différente de celle du calcul direct.

Il est facile de montrer que tout indice arithmétique est cohérent en agrégation (le raisonnement pour le cas géométrique est presque identique). Supposons que les marchandises soient partitionnées en  $k = 1, \dots, m$  catégories distinctes et mutuellement exclusives — c'est-à-dire qu'il y a  $m$  agrégats élémentaires. Si la catégorie  $k$  contient  $n_k$  biens, alors l'index arithmétique de cette catégorie est

$$I_k^A = \sum_{i=1}^{n_k} \frac{\omega_{ik}}{\sum_{j=1}^{n_k} \omega_{jk}} \frac{p_{ik1}}{p_{ik0}},$$

où  $p_{ik1}/p_{ik0}$  est le prix relatif du bien  $i$  dans la catégorie  $k$  et  $\omega_{ik}$  est le poids du bien  $i$  dans la catégorie  $k$ . Ces indices  $m$  peuvent maintenant être agrégés ensemble pour former un indice global en prenant une moyenne pondérée de chaque indice, avec des pondérations données par  $\omega_k = \sum_{j=1}^{n_k} \omega_{jk}$ , de sorte que



$$\begin{aligned}
I^A &= \sum_{k=1}^m \omega_k I_k^A \\
&= \sum_{k=1}^m \omega_k \sum_{i=1}^{n_k} \frac{\omega_{ik}}{\sum_{j=1}^{n_k} \omega_{jk}} \frac{p_{ik1}}{p_{ik0}} \\
&= \sum_{k=1}^m \sum_{i=1}^{n_k} \omega_{ik} \frac{p_{ik1}}{p_{ik0}}.
\end{aligned}$$

Mais c'est exactement la même chose que si l'indice arithmétique avait été calculé en regroupant tous les parents de prix et en calculant directement l'indice, plutôt que de les diviser en groupes et de calculer l'indice en deux étapes.

Un avantage du calcul en deux étapes par rapport au calcul direct est que la formule d'indice utilisée pour calculer les indices élémentaires peut différer de la formule d'indice utilisée à des niveaux d'agrégation plus élevés. Cela peut sembler étrange de vouloir le faire, mais il existe un certain nombre de raisons théoriques et pratiques subtiles pour lesquelles une formule d'indice différente serait utilisée pour les indices élémentaires. Dans la plupart des applications, les indices élémentaires sont calculés à l'aide d'un indice de Jevons, en partie parce que des informations détaillées sur les quantités achetées ou vendues peuvent être difficiles à obtenir à des niveaux d'agrégation très faibles. Des niveaux plus élevés sont ensuite calculés comme un indice arithmétique, généralement un indice de Young ou Lowe.

## 4.2 Factorisation d'un index

Lorsqu'il est calculé sur plusieurs périodes, un indice des prix donne une mesure de la variation des prix dans le temps par rapport à une période de base fixe. Le calcul d'un indice de prix sur de nombreuses périodes ne pose pas de nouveaux défis — une fois qu'une période de base est sélectionnée, l'une des formules de nombre d'index dans la section précédente peut être directement appliquée pour produire un indice qui évolue au fil du temps en formant simplement des parents de prix qui calculez toujours la variation du prix par rapport à la période de base. Par exemple, la série d'indices géométriques calculés sur les périodes 0, 1 et 2, avec la période 0 comme période de base, est

$$1, \prod_{i=1}^n \left( \frac{p_{i1}}{p_{i0}} \right)^{\omega_i}, \prod_{i=1}^n \left( \frac{p_{i2}}{p_{i0}} \right)^{\omega_i}.$$

Ainsi, la valeur de l'indice de la période 1 donne la variation des prix entre la période 1 et la période 0, et l'indice de la période 2 donne la variation des prix entre la période 2 et la période 0. La valeur de la période 0 est 1 car le rapport des prix de la période 0 à la période 0 prix est toujours 1.

La plupart des indices de prix sont calculés fréquemment — généralement mensuellement ou trimestriellement — et il est utile de pouvoir calculer un indice de prix en utilisant uniquement la valeur d'indice de la période précédente et les prix relatifs les plus récents pour chaque bien couvert par l'indice. [L'une des raisons est que cela peut faciliter l'échantillonnage des informations sur les prix, car les mêmes unités n'ont pas besoin d'être échantillonnées à la fois pour la période actuelle et la période de base.] C'est-à-dire qu'il est utile de pouvoir factoriser un indice en deux termes : la valeur de l'indice de la période précédente et un indice qui dépend uniquement des prix de la période en cours et des prix de la période précédente. Une telle factorisation permet de calculer un indice période par période et évite d'avoir à conserver les données de prix pendant la période de base pendant la durée de vie d'un indice. Ceci est pratiquement significatif, car cela signifie qu'un indice peut toujours être produit même si les données de prix de la période de base sont perdues.

La factorisation d'un indice géométrique est triviale; pour un ensemble de poids donné, l'index géométrique qui va de la période 0 à la période  $t$  peut toujours être écrit comme le produit de l'index géométrique qui va de la période 0 à la période  $k$  et de l'index géométrique qui va de la période  $k$  à la période  $t$ . C'est,

$$\begin{aligned} I^G(0, t) &= \prod_{i=1}^n \left( \frac{p_{it}}{p_{i0}} \right)^{\omega_i} \\ &= \prod_{i=1}^n \left( \frac{p_{ik}}{p_{i0}} \right)^{\omega_i} \times \prod_{i=1}^n \left( \frac{p_{it}}{p_{ik}} \right)^{\omega_i} \\ &= I^G(0, k) \times I^G(k, t). \end{aligned}$$

Cela devrait être profondément intuitif. Si les prix ont augmenté de 20% entre la période 0 et la période  $k$ , puis augmentent de 10% supplémentaires entre la période  $k$  et la période  $t$ , l'augmentation totale du prix de la période 0 à la période  $t$  est de 32%. Cela revient à multiplier une valeur d'index de 1,2 par une valeur d'index de 1,1, car le résultat est 1,32.

La factorisation d'un indice arithmétique est légèrement plus complexe, car elle nécessite de modifier les pondérations utilisées pour agréger les prix relatifs. L'index arithmétique qui s'étend de la période 0 à la période  $t$  peut être écrit comme

$$\begin{aligned} I^A(0, t) &= \sum_{i=1}^n \omega_i \frac{p_{it}}{p_{i0}} \\ &= I^A(0, k) \times \sum_{i=1}^n \tilde{\omega}_i \frac{p_{it}}{p_{ik}}, \end{aligned}$$

où

$$\tilde{\omega}_i = \frac{\omega_i \frac{p_{ik}}{p_{i0}}}{I^A(0, k)}.$$

Contrairement à l'index géométrique, la factorisation d'un index arithmétique nécessite de changer les pondérations de l'index allant de la période  $k$  à  $t$ . Parfois, cela s'appelle le prix mettant à jour les poids, mais cette terminologie est utilisée de manière incohérente. Malgré cette ride, cependant, l'idée de base est la même que dans le cas géométrique — un indice arithmétique peut toujours être décomposé en deux indices arithmétiques.

Intuitivement, la factorisation d'un indice bricole simplement la valeur de l'indice dans la période de base afin qu'il puisse continuer à partir d'une valeur précédente sans avoir besoin d'être recalculé dès le départ. Plutôt que de commencer l'indice à la valeur 1 de la période  $k$ , il part de la valeur de l'indice de la période  $k$ , s'appuyant essentiellement sur la variation cumulative des prix jusqu'à ce moment-là.

### 4.3 Rebasing

La période de base est un moment important pour un indice des prix, car elle donne le point de référence fixe par rapport auquel les variations de prix sont mesurées. Cependant, le fait de devoir choisir une période de base introduit un problème potentiel: ce qui fonctionne comme période de base pour une utilisation d'un indice de prix peut ne pas fonctionner pour une autre. Heureusement, le choix de la période de base n'a pas beaucoup d'impact sur le contenu économique d'un indice des prix — dans la plupart des cas, une nouvelle période de base peut être choisie pour un indice sans avoir à recalculer la série d'indices entière.

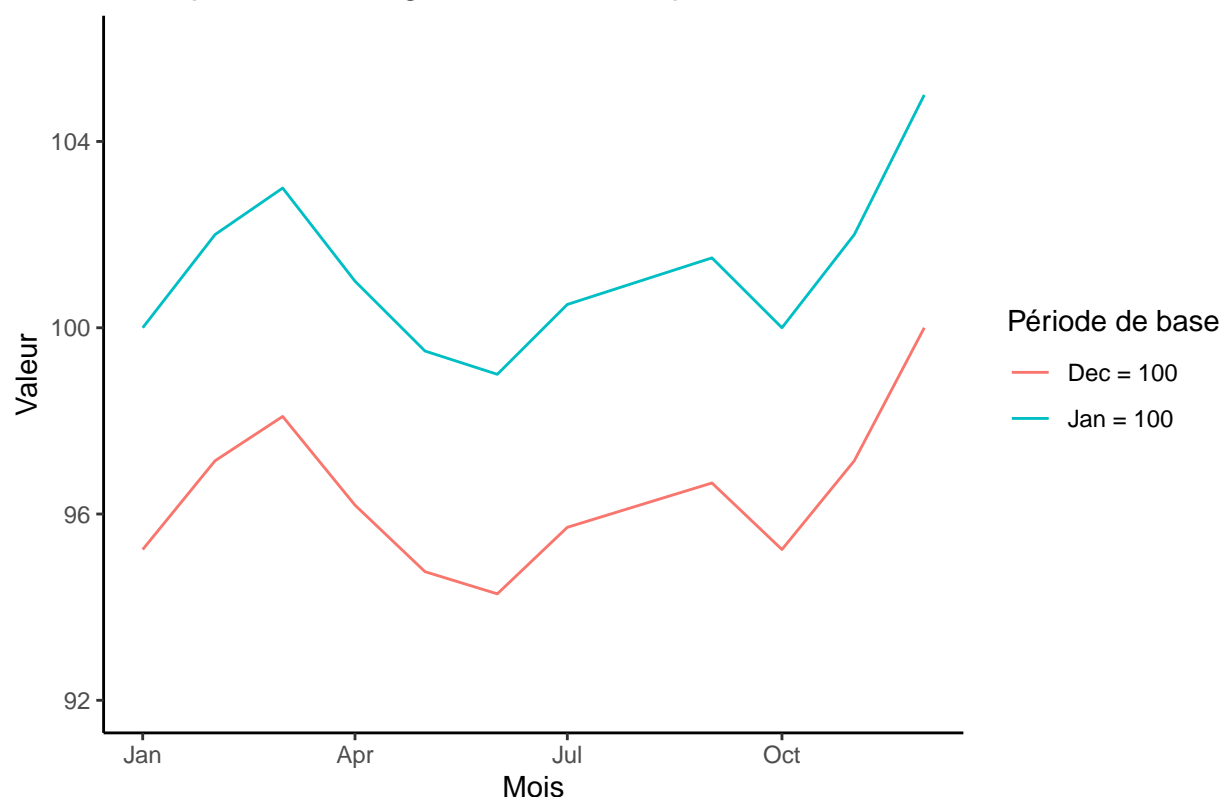
Pour un indice géométrique, le choix de la période de base n'a aucun impact sur le contenu économique d'un indice de prix — la variation de prix entre deux périodes quelconques est la même quelle que soit la période de base. La période de base standardise simplement l'indice à un niveau particulier. Cela peut être vu en rappelant qu'un indice géométrique peut toujours être factorisé en deux indices géométriques. Étant donné une série d'indices avec la période 0 comme période de base, la valeur de l'indice dans la période  $t$  avec la période  $k$  comme période de base peut toujours être trouvée en divisant simplement la valeur de l'indice dans la période  $t$  par la valeur de l'indice dans la période  $k$

$$\begin{aligned} I^G(k, t) &= \prod_{i=1}^n \left( \frac{p_{it}}{p_{ik}} \right)^{\omega_i} \\ &= \frac{\prod_{i=1}^n \left( \frac{p_{it}}{p_{i0}} \right)^{\omega_i}}{\prod_{i=1}^n \left( \frac{p_{ik}}{p_{i0}} \right)^{\omega_i}} \\ &= \frac{I^G(0, t)}{I^G(0, k)}. \end{aligned}$$

Par conséquent, la division de la série d'indices par la valeur d'index de la période  $k$  produit une série d'indices qui est la même que si la période  $k$  était initialement choisie comme période de base. C'est ce que l'on appelle le rebasage d'un indice et cela signifie qu'un changement de prix entre deux périodes peut être calculé en mettant simplement à l'échelle l'indice des prix. Cela explique également pourquoi la période de base obtient une valeur dans la série d'indices, sinon la série d'indices se rétrécirait chaque fois que l'indice est rebasé.

Le graphique ci-dessous donne un exemple de la façon dont le rebasage modifie la série chronologique. L'indice avec une période de base de décembre est simplement l'indice avec une période de base de janvier divisé par la valeur de l'indice en décembre (105). Cela déplace simplement l'indice avec janvier comme période de base vers le bas, et l'écrase un peu, mais ne change pas la forme relative de la série chronologique.

### Exemple de rebasage d'un indice de prix



La recadrage est plus complexe pour un index arithmétique. Bien qu'il fonctionne pour un indice de Lowe et Dutot de la même manière qu'un indice géométrique, en général, le rebasage d'un indice arithmétique modifie les pondérations utilisées pour la moyenne des prix relatifs en raison de la façon dont un indice arithmétique est factorisé. L'indice rebasé est toujours un indice arithmétique, mais les poids ne sont plus corrects. Par exemple, le rebasage d'un indice de Laspeyres modifie les pondérations afin qu'elles ne soient plus les parts de dépenses de la période de base — l'indice est toujours un indice arithmétique (un indice de Lowe), mais des produits qui ont connu une augmentation de prix plus importante depuis la période de base recevra plus de poids dans le calcul de l'indice. En pratique, cependant, cette ride théorique est généralement ignorée, et un indice de prix arithmétique est rebasé en divisant simplement la série d'indices entière par la valeur de l'indice dans la nouvelle période de base.

## 4.4 Chaînage

Les pondérations d'un indice donnent idéalement une mesure de l'importance économique des biens qui le composent, de sorte que les prix apparentés pour les biens et services de plus grande importance économique ont une plus grande influence sur l'évolution globale des prix entre les périodes. Dans l'application, les poids sont souvent réutilisés au fil du temps, car il est coûteux et long de mettre à jour en permanence les poids. Le maintien des mêmes poids période après période, cependant, signifie que les poids finiront par devenir obsolètes et pourraient ne pas saisir les changements dans l'importance économique de certains biens. Le chaînage d'un index permet de mettre à jour les pondérations d'un index et d'ajouter de nouvelles valeurs d'index aux séries d'index calculées avec les anciennes pondérations, sans avoir à réviser les anciennes séries d'index. Le chaînage peut également être utilisé pour modifier la composition des biens et services couverts par un indice des prix, et cela peut changer à mesure que de nouveaux biens sont introduits sur un marché et que des biens anciens disparaissent.

L'idée derrière le chaînage est la même que celle de la factorisation d'un index — divisez un index en deux morceaux qui peuvent être collés ensemble à une période de chevauchement commune. La différence est que,

bien que la factorisation d'un indice est simplement une astuce d'algèbre, le chaînage d'un index est une solution pratique pour produire un index avec une série chronologique plus longue, plutôt que de redémarrer l'index dans la période où les pondérations changent. Dans les deux cas, l'idée est de modifier la valeur de la période de base d'un indice pour refléter la variation cumulative des prix jusqu'à ce moment-là.

Les étapes de chaînage de tout type d'index sont les mêmes, alors considérez un index arithmétique allant de la période 0 à la période  $t$ , et supposez que les pondérations de cet indice changent au cours d'une période  $k$ . Cela pourrait être dû au fait que les pondérations précédentes sont devenues suffisamment obsolètes pour être modifiées, ou que les biens et services auxquels l'indice s'applique ont changé. Quelle que soit la raison, cela donne lieu à deux indices, l'un allant de la période 0 à la période  $k$ , utilisant des pondérations de la période 0

$$I^A(0, k) = \sum_{i=1}^n \omega_{i0} \frac{p_{ik}}{p_{i0}},$$

et un autre allant de la période  $k$  à la période  $t$  en utilisant les pondérations period- $k$

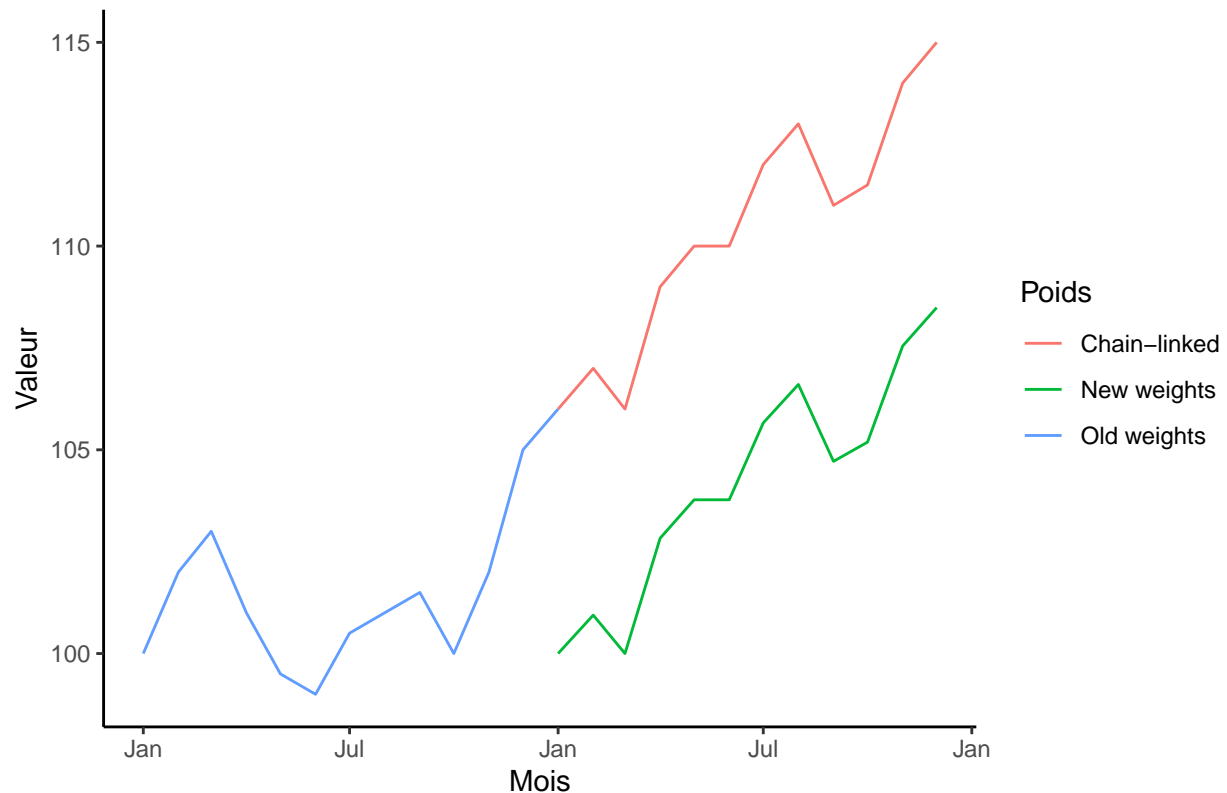
$$I^A(k, t) = \sum_{i=1}^m \omega_{ik} \frac{p_{it}}{p_{ik}}.$$

L'indice chaîné multiplie simplement ces deux indices ensemble, de sorte que

$$I^A(0, t) = \sum_{i=1}^n \omega_{i0} \frac{p_{ik}}{p_{i0}} \times \sum_{i=1}^m \omega_{ik} \frac{p_{it}}{p_{ik}}.$$

Le graphique ci-dessous donne un exemple de ce à quoi ressemble un indice chaîné.

### Exemple de chaînage d'un indice de prix



La partie importante d'un indice chaîné est la période de chevauchement  $k$  lorsque les pondérations changent et les prix de la période  $k$  sont utilisés avec deux ensembles de pondérations. S'il n'y a pas de période de chevauchement, alors l'index ne peut pas être enchaîné. Chaque indice peut être considéré comme un maillon de la chaîne, donnant le mouvement des prix pour cette partie du temps, et la période de chevauchement est l'endroit où les maillons se rencontrent.

#### 4.5 Exemple avec R: Cohérence dans l'agrégation et le rebasage

Il est utile de terminer cette section du cours avec quelques exemples dans R. Le premier exemple montre qu'un indice géométrique est cohérent dans l'agrégation. Le calcul de l'indice en deux étapes donne la même réponse que le calcul direct.

```
# Apportez dans la bibliothèque gpindex
library(gpindex)

# Faites des prix apparentés
dat <- data.frame(relative = c(1.1, 1.2, 0.9, 1.1),
                  weight = c(0.25, 0.3, 0.3, 0.15),
                  group = letters[c(1, 1, 2, 2)])

dat
```

```
##   relative weight group
## 1      1.1    0.25     a
## 2      1.2    0.30     a
## 3      0.9    0.30     b
## 4      1.1    0.15     b
```

```
# Calculer l'indice géométrique pour les parents des groupes a et b
index_ab <- sapply(split(dat, dat$group),
                   function(x) mean_geometric(x$relative, x$weight) * 100)
index_ab
```

```
##           a           b
## 115.34655  96.22603
```

```
# Indices agrégés de niveau inférieur
index_top <- mean_geometric(index_ab, as.numeric(tapply(dat$weight, dat$group, sum)))
index_top
```

```
## [1] 106.3125
```

```
# Identique au calcul direct
mean_geometric(dat$relative, dat$weight) * 100
```

```
## [1] 106.3125
```

Le deuxième exemple montre comment rebaser un indice de prix — divisez simplement la série d'indices par la valeur de l'indice dans la nouvelle période de base.

```
# Faire un index sur 12 périodes avec période 0 = 100
index <- data.frame(period = 0:11, value = c(100, sample(90:130, 11)))
index
```

```
##   period value
## 1      0   100
## 2      1   111
## 3      2    99
## 4      3   110
## 5      4   101
## 6      5   122
## 7      6   124
## 8      7   123
## 9      8    93
## 10     9   115
## 11    10   116
## 12    11   112
```

```
# Rebase l'indice pour que la période 6 = 100
transform(index, value = value / value[period == 6] * 100)
```

```
##   period   value
## 1      0 80.64516
## 2      1 89.51613
## 3      2 79.83871
## 4      3 88.70968
## 5      4 81.45161
## 6      5 98.38710
## 7      6 100.00000
## 8      7 99.19355
## 9      8 75.00000
## 10     9 92.74194
## 11    10 93.54839
## 12    11 90.32258
```

## 5 Affectation 1

Les réponses à ces questions proviennent à la fois du contenu du cours et des lectures. Chaque question vaut un point, pour un total de 20 points. La réussite de ce module nécessite au moins 65% (13 sur 20 correct). Envoyez vos réponses par e-mail à l'instructeurs du cours (Steve Martin ou Shaoxiong Wang ou Anastasiia Nosach) lorsque vous avez terminé.

**Question 1** Vrai ou faux: l'indice Jevons est toujours inférieur ou égal à l'indice Laspeyres.

**Question 2** Pourquoi l'indice Laspeyres est-il toujours plus grand que l'indice Paasche?

- a) Biais de substitution.
- b) Les quantités sont mesurées moins précisément dans l'indice de Paasche, de sorte que la variation d'échantillonnage est plus grand.
- c) Les tendances saisonnières de la consommation sont surestimées dans l'indice de Laspeyres.
- d) L'indice Laspeyres peut être plus petit que l'indice Paasche.
- e) Aucune de ces réponses.

**Question 3** Lesquels des éléments suivants sont des utilisations d'un indice des prix à la production?

- a) Indicateur des tendances inflationnistes des intrants des producteurs.
- b) Déflateur des comptes nationaux.
- c) Indicateur des tendances inflationnistes pour les biens et services de consommation.
- d) Marge d'ajustement cyclique de la balance des paiements.
- e) a et b.
- f) a, b et d.
- g) Tout ce qui précède.

**Question 4** Vrai ou faux: l'indice Jevons est toujours inférieur ou égal à l'indice Carli.

**Question 5** Vrai ou faux: un indice des prix à la consommation donne une mesure du niveau de prix global auquel sont confrontés les consommateurs.

**Question 6** Lesquels des éléments suivants sont des utilisations d'un indice des prix à la consommation?

- a) Indicateur des tendances inflationnistes pour les biens et services de consommation.
- b) Déflateur des comptes nationaux.
- c) Indicateur des tendances inflationnistes de la production des producteurs.
- d) Indexation des salaires.
- e) a et b.
- f) a, b et d.
- g) Tout ce qui précède.

**Question 7** Quels sont les trois types d'indices des prix à la production?

- a) Prix d'entrée, prix de sortie, prix de transfert.
- b) Prix d'entrée, valeur ajoutée, prix à la consommation.
- c) Prix de sortie, prix d'entrée, valeur ajoutée.
- d) Prix d'entrée, prix de sortie, valeur ajoutée.
- e) Aucune de ces réponses.



**Question 8** Qu'est-ce qu'un indice élémentaire?

- a) Le prix des intrants bruts pour une entreprise.
- b) Les indices de prix pour le niveau le plus bas d'une hiérarchie.
- c) Les prix individuels des répondants.
- d) La valeur des biens et services achetés au niveau le plus bas d'une hiérarchie.
- e) Aucune de ces réponses.

**Question 9** Quel est le classement habituel des indices Fisher, Laspeyres, Lowe et Paasche en fonction de leur ampleur (c'est-à-dire que la demande diminue lorsque les prix augmentent)?

- a) Laspeyres Lowe Paasche Fisher.
- b) Lowe Laspeyres Fisher Paasche.
- c) Paasche Fisher Laspeyres Lowe.
- d) Fisher Lowe Laspeyres Paasche.
- e) Aucune de ces réponses.

**Question 10** Vrai ou faux: les pondérations pour l'agrégation d'un indice doivent être basées sur l'échantillon de prix utilisé pour calculer les agrégats élémentaires.

**Question 11** Vrai ou faux: un indice de Fisher est considéré comme le meilleur type d'index car il est très opportun à produire.

**Question 12** Lesquels des éléments suivants sont des indices symétriques?

- a) Fisher, Laspeyres, Paasche.
- b) Jevons, Lowe.
- c) Fisher, Törnqvist.
- d) Jeune.
- e) Aucune de ces réponses.

**Question 13** Vrai ou faux: les indices Dutot et Jevons sont tous deux transitifs.

**Question 14** Les pondérations pour un indice des prix à la production devraient toujours être basées sur.

- a) Parts des revenus.
- b) Part des dépenses.
- c) Valeur ajoutée.
- d) Parts des revenus actuels et futurs.
- e) Aucune de ces réponses.

**Question 15** Vrai ou faux: tous les indices de prix peuvent être classés comme arithmétiques, géométriques ou harmoniques.

**Question 16** Les pondérations pour un indice des prix à la consommation peuvent provenir de quelles sources.

- a) Données du scanner.
- b) Enquête sur les dépenses de consommation.
- c) Comptes nationaux.
- d) Recensement de la population.

- e) a et b.
- f) Tout ce qui précède.
- g) Aucune de ces réponses.

**Question 17** Un indice a une valeur de 90 en janvier et une valeur de 130 en mai. Quelle est la valeur de l'indice en mai si janvier est défini comme période de base?

- a) 90,0.
- b) 100,0.
- c) 140,0.
- d) 144,4.
- e) 130,0.

**Question 18** Vrai ou faux: l'indice de Fisher est cohérent dans l'agrégation.

**Question 19** Un indice des prix a une valeur de 80 en mars et 90 en septembre. Quel est le pourcentage de variation des prix entre mars et septembre?

- a) 10%.
- b) 12,5%.
- c) -11%.
- d) Aucune de ces réponses.
- e) Il est impossible de savoir sans connaître la période de base.

**Question 20** Vous essayez de faire un index pour les widgets bleus et rouges. Le prix relatif des widgets rouges est de 1,2 et le prix relatif des widgets bleus est de 0,9. La part des dépenses consacrées aux widgets rouges est de 0,3 au cours de la période de base et de 0,25 au cours de la période en cours; la part des dépenses pour les widgets bleus est de 0,7 pour la période de base et de 0,75 pour la période en cours. Quelle est la valeur de l'indice géométrique de Laspeyres?

- a) 98,1.
- b) 96,7.
- c) 97,4.
- d) 99,0.
- e) 97,5.

## Part II

# Théorie de l'indice des prix

## 6 Programme pour Théorie de l'indice des prix

Pourquoi certaines formules de numéro d'indice sont-elles utilisées pour calculer un indice de prix? Répondre à cette question nécessite une théorie des indices de prix qui peut être utilisée pour évaluer les propriétés d'un indice de prix et déterminer si un numéro d'indice est meilleur qu'un autre.

L'objectif de ce module est de fournir les fondements théoriques d'un indice des prix, avec un accent particulier sur la motivation des formules d'indice-nombre qui sont utilisées dans la pratique. À la fin du module, une personne devrait:

1. Familiarisez-vous avec les trois approches théoriques pour développer un indice des prix.
2. Savoir comment les trois approches peuvent être utilisées pour motiver les numéros d'index couramment utilisés.

Ce module est utile à la fois pour les utilisateurs et les compilateurs d'indices de prix avec une compréhension de base de la façon de construire un indice de prix, et qui aimerait mieux comprendre pourquoi des formules d'indice-nombre particulières sont utilisées.

Ce module se compose de lectures autodirigées, accompagnées d'un devoir. Au total, environ 10 à 15 heures devraient être consacrées à ce module.

Prérequis: Introduction aux indices de prix, au moins un cours de micro-théorie intermédiaire et au moins un cours d'introduction aux probabilités et aux statistiques. Un cours avancé de micro-théorie est utile.

L'évaluation de ce module est basée sur une tâche composée de 20 questions à choix multiples / vrai-faux qui s'appuient sur du matériel dans le contenu du cours et les lectures. La collaboration sur la mission est la bienvenue, mais chaque personne doit soumettre son propre travail unique. La réussite de ce module nécessite au moins 65% sur l'affectation.

Les lectures de ce module proviennent des chapitres 15 à 18 du manuel PPI (ILO et al. 2004b) et / ou du manuel CPI (ILO et al. 2004a), publié par le FMI (disponible gratuitement sur leur site Internet), ainsi que de la revue des prix. théorie de l'indice par Balk (1995).

Veuillez envoyer un e-mail à l'instructeurs du cours (Steve Martin ou Shaoxiong Wang ou Anastasiia Nosach) si vous avez des questions ou si vous avez besoin d'aide concernant le matériel ou le devoir du cours.

## 7 Qu'est-ce que la théorie de l'indice des prix?

Il existe trois grandes approches pour motiver les formules d'indice des prix qui sont utilisées dans la pratique par les organismes statistiques: l'approche axiomatique, l'approche économique et l'approche stochastique. Chaque approche est utile en soi, et les trois donnent des perspectives et des compréhensions différentes pour les diverses formules de nombre d'indices qui sont utilisées pour construire des indices de prix. Ensemble, ces différentes approches constituent une théorie des indices de prix.

Chacune des trois approches peut être considérée comme répondant à une question différente sur la façon de construire un indice des prix. L'approche axiomatique tente de répondre à la question "Comment combiner les informations d'une collection de prix et de quantités pour satisfaire certaines propriétés?" Il s'agit essentiellement d'une approche théorique définie sur la meilleure façon de construire un numéro d'index. En revanche, l'approche économique tente de répondre à la question "Que doit mesurer un indice des prix?" Le but de l'approche économique est d'utiliser le mécanisme de la théorie microéconomique pour définir un indice des prix afin de mesurer l'expérience d'un changement de prix pour les consommateurs et les producteurs. L'approche stochastique est plus pragmatique dans sa portée et tente de répondre à la question "Comment les informations sur les prix peuvent-elles être utilisées pour prédire au mieux un changement de prix pour un bien ou un service?" Il s'agit essentiellement d'une question sur la meilleure façon de gonfler et de dégonfler les prix au fil du temps.

Cette partie du cours donne une introduction pratique à chacune des trois approches, en se concentrant sur la façon dont elles peuvent être utilisées pour motiver la manière dont les indices de prix sont construits dans la pratique. Bien qu'il ait une orientation plus pratique, cependant, ce matériel est toujours concerné par une théorie qui peut être assez technique. Plutôt que de se concentrer sur les détails techniques (dont ils sont nombreux), cette partie du cours donne un aperçu suffisant de la machinerie technique pour montrer ce qui se passe, tout en fournissant l'intuition économique et les idées générales de la théorie.

## 8 L'approche axiomatique

Un indice des prix combine des informations sur les prix et les quantités à deux moments afin de produire un rapport de prix entre ces moments. À moins qu'il n'y ait qu'un seul prix à chaque moment, cependant, il n'y a pas de moyen unique de combiner les prix au fil du temps — il existe une infinité de formules de numéros d'index possibles qui peuvent être utilisées pour calculer un indice de prix. Cet embarras des richesses n'est pas souhaitable, car il introduit un choix considérable sur la manière de mesurer un changement de prix. L'approche axiomatique tente de résoudre ce problème en considérant une classe générale d'indices de prix et en trouvant quels membres de cette classe satisfont certaines propriétés intuitives ou raisonnables qu'un indice de prix devrait satisfaire. Cela équivaut à faire une déclaration normative sur la façon dont un indice des prix devrait se comporter — quels sont les axiomes fondamentaux qui définissent un indice des prix? L'idée est de commencer par une classe très large de formules d'indices possibles et de définir des axiomes pour réduire les formules d'indices qui ne se comportent pas comme un indice de prix.

Le résultat idéal de l'approche axiomatique est une formule d'indice unique et unique qui satisfait un ensemble minimal de conditions non controversées qu'un indice de prix devrait satisfaire. Cela signifierait qu'il n'y a qu'une seule formule d'indice qui devrait être utilisée pour créer un indice de prix. La réalité, cependant, est que des compromis doivent être faits, et le maintien de certaines conditions nécessite d'ignorer les autres. Par conséquent, une distinction est faite entre les axiomes — déclarations fondamentales que tout indice de prix raisonnable devrait satisfaire — et les tests, qui sont des propriétés souhaitables mais pas aussi fondamentales que les axiomes. Un petit ensemble d'axiomes clés est toujours conservé, tandis que les tests peuvent être mutuellement incohérents, et des choix doivent être faits quant aux tests qu'un index doit satisfaire en fonction de son objectif. Les axiomes peuvent être considérés comme le premier tamis à éliminer toutes les formules de nombre d'index déraisonnables, avec des tests agissant comme un second tamis pour affiner davantage ce qui reste.

Cette section du cours commence le voyage dans la théorie de l'indice des prix en donnant un bref aperçu de l'approche axiomatique de la théorie de l'indice des prix. Cela a été un domaine d'étude considérable, et par conséquent il y a beaucoup de matériel intéressant qui n'est pas couvert dans ce cours. Le lecteur intéressé peut consulter ILO et al. (2004aChapitre 16) ou ILO et al. (2004bChapter 16) pour plus de détails, ou Balk (2008).

Balk (1995), ignorez la section 4 et l'annexe.

Manuel PPI: Chapitre 1, section C.

### 8.1 Un indice de prix abstrait

Dans sa forme abstraite, un indice de prix est une fonction  $I$  qui prend quatre arguments — prix période-0  $p_0$ , période- $t$  prix  $p_t$ , quantités période-0  $q_0$  et période- $t$  quantités  $q_t$ , pour  $n$  biens et services distincts — et renvoie une valeur unique. Il s'agit d'une règle qui transforme les informations sur les prix et les quantités à deux moments différents en un seul chiffre qui résume la variation des prix. Pour simplifier la notation, les prix et les quantités sont ici des vecteurs de prix et de quantités à un moment donné; par exemple, avec  $n$  biens et services distincts, les prix de la période 0 sont donnés par  $p_0 = (p_{10}, p_{20}, \dots, p_{n0})$ . Cette notation est utile, car un indice de prix est un moyen de distiller des informations pour de nombreux prix et quantités en une seule valeur; par exemple, avec 25 biens et services, l'indice des prix prend 100 pièces pour information (50 prix et 50 quantités) et la transforme en une seule information. Pour une collection particulière de prix et de quantités, la valeur de l'indice est donnée par  $I(p_t, p_0, q_t, q_0)$ .

La plupart des indices de prix peuvent être exprimés sous cette forme abstraite. Par exemple, un indice de Laspeyres est simplement

$$I(p_t, p_0, q_t, q_0) = \frac{\sum_{i=1}^n p_{it} q_{i0}}{\sum_{i=1}^n p_{i0} q_{i0}}.$$

Les formules de numéros d'index qui n'utilisent pas d'informations sur les quantités ne nécessitent aucun traitement spécial — la valeur de l'indice est juste indépendante des quantités vendues. Par exemple, l'indice Jevons est

$$I(p_t, p_0, q_t, q_0) = \prod_{i=1}^n \left( \frac{p_{it}}{p_{i0}} \right)^{1/n}.$$

Bien que la plupart des formules d'indice correspondent à cette représentation abstraite, les indices de prix qui utilisent des informations dans une période autre que la période 0 ou la période  $t$ , comme un indice Lowe ou Young, ne le font pas. C'est un point important, car l'approche axiomatique ignore implicitement ces types d'indices de prix dès le départ, plutôt que d'utiliser des axiomes pour évaluer leur caractère raisonnable en tant qu'indice de prix.

## 8.2 Les axiomes

Il existe cinq axiomes clés que tout indice de prix devrait satisfaire pour toutes les collections de prix et de quantités. Ces axiomes peuvent être considérés comme généralisant les propriétés du rapport des prix entre la période  $t$  et la période 0 pour un seul bien ou service à des environnements avec de nombreux biens et services.

1. **Monotonie**  $I(p'_t, p_0, q_t, q_0) > I(p_t, p_0, q_t, q_0)$  if  $p'_t \geq p_t$  and  $p'_t \neq p_t$  et  $I(p_t, p'_0, q_t, q_0) < I(p_t, p_0, q_t, q_0)$  if  $p'_0 \geq p_0$  et  $p'_0 \neq p_0$ .
2. **Homogénéité linéaire**  $I(cp_t, p_0, q_t, q_0) = cI(p_t, p_0, q_t, q_0)$  pour tout  $c > 0$ .
3. **Identité**  $I(p_0, p_0, q_t, q_0) = 1$ .
4. **Homogénéité du degré zéro**  $I(cp_t, cp_0, q_t, q_0) = I(p_t, p_0, q_t, q_0)$  pour tout  $c > 0$ .
5. **Invariance dimensionnelle** Si  $A$  est une matrice diagonale de nombres positifs, alors  $I(Ap_t, Ap_0, A^{-1}q_t, A^{-1}q_0) = I(p_t, p_0, q_t, q_0)$ .

Tous ces axiomes sont assez simples à comprendre d'après leur représentation mathématique, peut-être à l'exception de l'invariance dimensionnelle, et sont assez intuitifs. La monotonie signifie simplement qu'un indice de prix augmente dans les prix de la période  $t$  et diminue dans les prix de la période 0 — les prix  $t$  de la période plus grande produisent des valeurs d'indice plus élevées et les prix de la période 0 plus grands produisent des valeurs d'indice plus petites. Cela semble être une condition nécessaire pour qu'un indice des prix puisse mesurer de manière significative l'inflation, car un indice qui ne satisfait pas à la monotonie peut diminuer lorsque les prix augmentent.

L'homogénéité linéaire est probablement la moins intuitive des axiomes, et dit que multiplier tous les prix de la période  $t$  par une constante équivaut à multiplier l'indice entier par une constante, de sorte qu'une augmentation proportionnelle des prix se traduit par une augmentation proportionnelle du indice. Pour voir pourquoi cela a du sens, considérons deux villes (A et B) qui ont toutes les mêmes prix dans la période 0. Si tous les prix dans la ville A augmentent de deux fois plus que les prix dans la ville B entre la période 0 et la période  $t$ , les quantités achetées restent les mêmes dans les deux villes (car les prix relatifs sont les mêmes dans les deux villes), alors il est raisonnable de dire que les prix dans la ville A ont augmenté deux fois plus que dans la ville B.

L'axiome identitaire est simple et indique qu'un indice des prix ne montre pas de changement de prix si les prix ne changent pas. Comme l'axiome de monotonie, cela semble être une exigence nécessaire pour un indice des prix pour mesurer l'inflation; sinon, un indice des prix pourrait montrer une variation des prix lorsque les prix ne changent pas avec le temps.

L'homogénéité du degré zéro signifie que la multiplication de tous les prix par une constante n'a aucun impact sur un indice — la devise dans laquelle les prix sont mesurés n'affecte pas la valeur d'un indice de prix. Une formule d'indice qui ne satisfait pas à l'homogénéité du degré zéro peut donner une mesure différente de

l'inflation en fonction du prix des devises et ne peut pas être utilisée pour comparer l'inflation dans différents pays.

L'invariance dimensionnelle semble complexe, mais elle dit simplement que le changement des unités de mesure ne change pas la valeur de l'indice. Un indice de prix ne devrait pas changer si tous les prix sont multipliés par une constante et toutes les quantités sont divisées par la même constante — un indice de prix ne devrait pas dépendre des unités de mesure.

Par exemple, il est simple de voir que l'indice de Laspeyres satisfait les cinq axiomes.

1. Pour la monotonie,

$$I(p'_t, p_0, q_t, q_0) = \frac{\sum_{i=1}^n p'_{it} q_{i0}}{\sum_{i=1}^n p_{i0} q_{i0}} > \frac{\sum_{i=1}^n p_{it} q_{i0}}{\sum_{i=1}^n p_{i0} q_{i0}} = I(p_t, p_0, q_t, q_0)$$

chaque fois que chaque  $p'_{it} \geq p_{it}$ , avec au moins un strictement supérieur. L'inverse est vrai si chaque  $p'_{i0} \geq p_{i0}$ , avec au moins un strictement supérieur.

2. Pour une homogénéité linéaire,

$$I(cp_t, p_0, q_t, q_0) = \frac{\sum_{i=1}^n cp_{it} q_{i0}}{\sum_{i=1}^n p_{i0} q_{i0}} = \frac{c \sum_{i=1}^n p_{it} q_{i0}}{\sum_{i=1}^n p_{i0} q_{i0}} = cI(p_t, p_0, q_t, q_0)$$

pour tout  $c > 0$ .

3. Pour l'identité,

$$I(p_0, p_0, q_t, q_0) = \frac{\sum_{i=1}^n p_{i0} q_{i0}}{\sum_{i=1}^n p_{i0} q_{i0}} = 1.$$

4. Pour l'homogénéité du degré zéro,

$$I(cp_t, cp_0, q_t, q_0) = \frac{\sum_{i=1}^n cp_{it} q_{i0}}{\sum_{i=1}^n cp_{i0} q_{i0}} = \frac{c \sum_{i=1}^n p_{it} q_{i0}}{c \sum_{i=1}^n p_{i0} q_{i0}} = I(p_t, p_0, q_t, q_0)$$

pour tout  $c > 0$ .

5. Pour l'invariance dimensionnelle, si

$$A = \begin{bmatrix} a_1 & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & a_n \end{bmatrix}$$

puis

$$I(Ap_t, Ap_0, A^{-1}q_t, A^{-1}q_0) = \frac{\sum_{i=1}^n a_i p_{it} a_i^{-1} q_{i0}}{\sum_{i=1}^n a_i p_{i0} a_i^{-1} q_{i0}} = \frac{\sum_{i=1}^n p_{it} q_{i0}}{\sum_{i=1}^n p_{i0} q_{i0}} = I(p_t, p_0, q_t, q_0).$$

Il devrait être assez sûr de dire que tout indice de prix raisonnable devrait satisfaire ces axiomes. Bien qu'il existe de nombreux indices de prix qui le font, les axiomes fonctionnent pour exclure des formules de nombre d'indices déraisonnables qui n'ont aucun sens pour mesurer les variations de prix.

### 8.3 Exemples de compteur

Pour la plupart, il est difficile de noter un prix décent à mi-chemin qui ne satisfait pas aux cinq axiomes de la section précédente. Les exemples pathologiques abondent, mais il est plus intéressant de regarder des cas raisonnables où une formule d'indice ne satisfait pas tous les axiomes.<sup>11</sup> Deux exemples intéressants d'indices de prix raisonnables qui ne satisfont pas aux cinq axiomes sont l'indice de Dutot et l'indice de valeur médiane (c'est-à-dire l'indice qui prend le prix médian relatif).

L'indice de Dutot satisfait tous les axiomes sauf l'invariance dimensionnelle, tandis que l'indice de valeur moyenne satisfait tous les axiomes sauf la monotonie. Pour voir pourquoi l'indice Dutot ne satisfait pas à l'invariance dimensionnelle, supposons qu'il n'y a que deux biens (c'est-à-dire  $n = 2$ ) avec  $p_{1t} = 1$ ,  $p_{2t} = 2$  et  $p_{10} = p_{20} = 1$ , et avec

$$A = \begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix}.$$

Dans ce cas

$$I(Ap_t, Ap_0, A^{-1}q_t, A^{-1}q_0) = \frac{1 \times 1 + 2 \times 2}{1 \times 1 + 2 \times 1} = 5/3$$

et

$$I(p_t, p_0, q_t, q_0) = \frac{1+2}{1+1} = 3/2 \neq I(Ap_t, Ap_0, A^{-1}q_t, A^{-1}q_0),$$

échouant ainsi l'axiome d'invariance dimensionnelle. Cela officialise simplement le problème bien connu de l'indice Dutot selon lequel tous les biens et services doivent être mesurés dans les mêmes unités pour être utiles.

Pour voir que l'indice de prix médian ne satisfait pas à la monotonie, supposons qu'il existe 3 biens tels que  $p_{1t} = 1$ ,  $p_{2t} = 2$ ,  $p_{3t} = 3$  et  $p_{10} = p_{20} = p_{30} = 1$ . L'indice médian renvoie la valeur 2 (le prix médian relatif). Maintenant, si  $p'_{1t} = 1$ ,  $p'_{2t} = 2$  et  $p'_{3t} = 4$ , l'index médian renvoie toujours 2; les prix ont augmenté, mais la valeur de l'indice reste la même.

### 8.4 Les tests

En plus des cinq axiomes, il existe trois tests importants qui peuvent être souhaitables pour un indice de prix. Ces tests sont des propriétés agréables à avoir d'un index, mais ne sont pas aussi fondamentaux que les axiomes. Ils agissent comme un moyen de réduire davantage l'ensemble des indices de prix qui satisfont les cinq axiomes.

La formulation de ces tests nécessite la notion d'indice de quantité, une fonction  $Q$  identique à un indice de prix sauf que le rôle des prix et des quantités est inversé. Un indice de quantité mappe les quantités et les prix pour produire un nombre, bien qu'au lieu de donner la variation du prix au fil du temps, un indice de quantité donne la variation des quantités physiques au fil du temps. Un indice de quantité devrait également satisfaire les cinq axiomes (changer les prix avec les quantités).

Avec un indice de quantité en main, les trois tests clés sont les suivants.

<sup>11</sup>Deux cas pathologiques intéressants ont à voir avec l'échec de la monotonie pour les indices géométriques Laspeyres et Paasche. Bien que le Laspeyres géométrique augmente toujours dans les prix de la période- $t$ , il peut également augmenter dans les prix de la période-0. Pour cela, il faut que  $p_{i1}/p_{i0} \geq \exp(1)I(p_1, p_0, q_t, q_0) \approx 2,72I(p_1, p_0, q_t, q_0)$  pour au moins un bon  $i$ , de sorte qu'au moins certaines marchandises subissent un changement de prix qui est considérablement plus grand que le changement de prix pour les autres marchandises. De même, si le Paasche géométrique diminue toujours dans les prix de la période 0, il peut également diminuer dans les prix de la période  $t$ . Pour que cela soit nécessaire, il faut que  $p_{i1}/p_{i0} \leq I(p_1, p_0, q_t, q_0)/\exp(1) \approx I(p_1, p_0, q_t, q_0)/2,72$  pour au moins un bon  $i$ . Voir Balk (2008Section 3.12).

1. **Test de circularité**  $I(p_t, p_0, q_t, q_0) = I(p_k, p_0, q_k, q_0)I(p_t, p_k, q_t, q_k)$ .
2. **Test de produit**  $I(p_t, p_0, q_t, q_0)Q(q_t, q_0, p_t, p_0) = \frac{\sum_{i=1}^n p_{it}q_{it}}{\sum_{i=1}^n p_{i0}q_{i0}}$ .
3. **Cohérence dans l'agrégation** Pour toute partition de  $n$  marchandises en  $k = 1, \dots, m$  groupes distincts, il existe une fonction  $\psi$  telle que  $\psi(I(p_t, p_0, q_t, q_0), V_0, V_t) = \sum_{k=1}^m \psi(I(p_{kt}, p_{k0}, q_{kt}, q_{k0}), V_{k0}, V_{kt})$ , où  $V_{kt} = \sum_{i=1}^{n_k} p_{ikt}q_{ikt}$  est la valeur des biens  $n_k$  pour le groupe  $k$  dans la période  $t$ , avec  $V_t = \sum_{k=1}^m V_{kt}$  étant la valeur totale de la période  $t$ , et si  $I(p_{kt}, p_{k0}, q_{kt}, q_{k0}) = c$  pour tous les groupes  $m$ , alors  $I(p_t, p_0, q_t, q_0) = c$ .

Le test de circularité est simple et exprime l'idée qu'un indice entre deux périodes peut être calculé en enchaînant l'indice pour des périodes consécutives ensemble. Le test de circularité légitime l'utilisation du chaînage lors du calcul d'un indice et l'utilisation des variations de prix d'une période à l'autre pour calculer un indice avec une période de base fixe. Il légitime également le rebasage d'un indice.

Le test de produit est également simple et indique que la variation de la valeur globale entre deux périodes peut être décomposée en un indice des prix et un indice des quantités. Les tests de produits légitiment l'utilisation d'indices de prix pour déflater les valeurs agrégées dans un cadre comptable national.

Le test de cohérence dans l'agrégation semble compliqué, mais exprime une idée simple: un indice de prix devrait être un agrégat de sous-indices, où l'agrégation ne dépend que de l'indice lui-même et de la valeur totale des biens pour chaque sous-indice. La cohérence dans l'agrégation légitime le calcul hiérarchique d'un indice de prix, où les indices de prix sont calculés pour des catégories de biens de plus en plus larges en utilisant la même formule d'indice tout au long, avec des parts de valeur comme poids.

Par exemple, l'indice Laspeyres satisfait le test du produit et le test de cohérence dans l'agrégation.<sup>12</sup> Pour le test du produit, laissez l'indice de quantité être

$$Q(q_t, q_0, p_t, p_0) = \frac{\sum_{i=1}^n p_{it}q_{it}}{\sum_{i=1}^n p_{it}q_{i0}},$$

(c'est l'indice de quantité de Paasche) pour que

$$\begin{aligned} I(p_t, p_0, q_t, q_0)Q(q_t, q_0, p_t, p_0) &= \frac{\sum_{i=1}^n p_{it}q_{i0}}{\sum_{i=1}^n p_{i0}q_{i0}} \frac{\sum_{i=1}^n p_{it}q_{it}}{\sum_{i=1}^n p_{it}q_{i0}} \\ &= \frac{\sum_{i=1}^n p_{it}q_{it}}{\sum_{i=1}^n p_{it}q_{i0}}. \end{aligned}$$

Pour le test de cohérence dans l'agrégation, laissez  $\psi(I(p_t, p_0, q_t, q_0), V_0, V_t) = V_0 I(p_t, p_0, q_t, q_0)$ , pour que le test devienne

$$I(p_t, p_0, q_t, q_0) = \sum_{k=1}^m \frac{V_{k0}}{V_0} I(p_{kt}, p_{k0}, q_{kt}, q_{k0}).$$

L'indice de Laspeyres satisfait évidemment à cette condition puisque  $V_{k0}/V_0$  est la part des dépenses / revenus de la période 0 pour les produits du groupe  $k$ .

<sup>12</sup>L'indice Laspeyres ne satisfait pas le test de circularité, car un indice Laspeyres devrait être multiplié par un indice Lowe pour renvoyer un indice Laspeyres.



## 8.5 Un résultat impossible

Idéalement, il y aurait un seul indice des prix qui satisfait les cinq axiomes et les trois tests des sections précédentes. Cela signifierait qu'il n'y a qu'une seule formule de numéro d'index qui devrait être utilisée, au moins aux fins de créer des statistiques nationales. Cependant, il peut être démontré qu'il n'y a pas d'indice de prix qui satisfait l'axiome d'identité, le test de circularité et le test de produit (Balk 1995, Théorème 1), et donc *a fortiori* il n'y a pas d'indice qui satisfasse les cinq axiomes et les trois tests. En effet, il existe un compromis entre un indice qui peut être enchaîné sur des périodes successives (test de circularité) et un indice qui s'inscrit dans un cadre comptable national (test de produit). Ainsi, l'utilité d'un test particulier dépend de l'utilisation de l'indice, et cela est important pour sélectionner un meilleur indice dans le cadre axiomatique.

## 8.6 Motiver les indices Jevons et Laspeyres

Bien qu'aucun indice ne satisfasse les cinq axiomes et les trois tests, les cinq axiomes, ainsi que le test de circularité, définissent de manière unique un indice de prix comme étant un indice de prix géométrique avec des poids qui ne changent pas dans le temps (Balk 1995, Theorem 11 et Remark 11.1). Autrement dit, un indice de prix satisfait les cinq axiomes et le test de circularité si et seulement si

$$I(p_t, p_0, q_t, q_0) = \prod_{i=1}^n \left( \frac{p_{it}}{p_{i0}} \right)^{\omega_i},$$

où  $\sum_{i=1}^n \omega_i = 1$ . Notez que la constance des poids dans le temps exclut l'indice de Törnqvist.

En désactivant le test de circularité pour le test de produit, les indices arithmétiques Laspeyres et Paasche sont les seuls indices qui satisfont aux cinq axiomes, le test de produit et la cohérence dans le test d'agrégation (Balk 1995, Corollary 16). Autrement dit, l'indice des prix doit être soit

$$I(p_t, p_0, q_t, q_0) = \frac{\sum_{i=1}^n p_{it} q_{i0}}{\sum_{i=1}^n p_{i0} q_{i0}}$$

ou

$$I(p_t, p_0, q_t, q_0) = \frac{\sum_{i=1}^n p_{it} q_{it}}{\sum_{i=1}^n p_{i0} q_{it}}.$$

Ces résultats frappants aident à motiver l'utilisation de ces indices simples et donnent des indications pratiques sur le moment de les utiliser. Un indice géométrique est utile lorsque le chaînage d'un indice est relativement plus important que la déflation des valeurs agrégées, tandis qu'un indice arithmétique de Laspeyres (ou Paasche) est utile lorsque la déflation des valeurs agrégées dans un cadre de comptabilité nationale est relativement plus importante que l'utilisation d'un calcul chaîné. Cela permet d'expliquer pourquoi les indices de prix produits par les organismes statistiques ont tendance à utiliser un indice géométrique pour calculer les indices élémentaires et un indice arithmétique pour calculer les indices de niveau supérieur. Au niveau le plus bas de la hiérarchie d'un indice, il est important de pouvoir calculer les indices élémentaires en utilisant les variations de prix d'une période à l'autre, mais ces indices n'ont pas tendance à être utilisés pour déflater les valeurs agrégées. À des niveaux d'agrégation plus élevés, il est nécessaire de pouvoir déflater les valeurs agrégées, mais pas besoin de chaîner les variations de prix d'une période à l'autre pour calculer l'indice, car il est agrégé à partir des indices élémentaires.

Il convient de noter que l'indice Fisher se présente généralement comme l'idéal théorique dans un cadre axiomatique.<sup>13</sup> Cela se produit parce que le test du produit est maintenu sans exiger la cohérence de l'agrégation, ainsi que d'imposer d'autres tests. Par conséquent, l'indice Fisher peut être plus approprié si un indice de prix n'est pas calculé à l'aide d'une structure hiérarchique. Mais il est important de noter que ce n'est pas acquis d'avance — l'un des principaux résultats de l'approche axiomatique est que l'indice de prix idéal dépend de son objectif, et qu'il n'existe aucun indice de prix universellement applicable.

## 9 L'approche économique

L'approche axiomatique de la construction d'un indice des prix consiste à trouver un moyen de combiner les informations sur les prix et les quantités afin de remplir certaines conditions. L'accent est essentiellement fixé sur la théorie et l'objectif est de définir une relation entre les prix dans le temps qui possède les propriétés qu'un indice des prix devrait avoir. À l'exception peut-être du test du produit, l'approche axiomatique n'a pas de contenu économique - il s'agit d'un exercice purement mathématique. L'approche économique, en revanche, cherche à examiner le phénomène économique sous-jacent qu'un indice des prix devrait mesurer, en utilisant la théorie microéconomique d'une entreprise représentative et compétitive (pour un indice des prix à la production) ou d'un consommateur représentatif (pour un indice des prix à la consommation). La manière exacte dont les informations sur les prix doivent être combinées pour mesurer une variation des prix dans le temps est entièrement déterminée par le concept économique à mesurer.

Bien que l'approche économique offre un moyen parcimonieux de construire un indice des prix, la validité de cette approche repose sur des hypothèses solides concernant le comportement des entreprises / des consommateurs et les conditions du marché. L'hypothèse la plus importante - et peut-être la plus forte - est qu'il existe une seule entreprise ou un seul consommateur représentatif sur le marché sans pouvoir de marché (c'est-à-dire qu'il prend les prix du marché comme indiqué). Avoir une entreprise représentative n'est pas une hypothèse restrictive si le marché est concurrentiel, mais un comportement concurrentiel est indispensable (ILO et al. 2004b, chapitre 18).<sup>14</sup> Cela est problématique car la plupart des marchés sont caractérisés par un certain degré de concurrence imparfaite, entraînant la désintégration de la théorie. Inversement, avoir un consommateur de prendre les prix comme indiqué est une hypothèse standard, mais avoir un consommateur représentatif est difficile à justifier et ne peut pas être facilement assoupli (Pollak 1980; Kirman 1992).<sup>15</sup> Par conséquent, l'approche économique est mieux perçue comme un moyen cohérent de conceptualiser un indice des prix plutôt que comme une recette pour trouver un indice de prix supérieur, et est la moins applicable des trois approches.

Cette section du cours donne une brève introduction à l'approche économique. Comme pour l'approche axiomatique, l'accent est mis sur la motivation des formules d'index utilisées dans la pratique, et il y a beaucoup de matériel intéressant qui n'est pas couvert. Contrairement à l'approche axiomatique, différents types d'indices nécessitent une analyse distincte car ils mesurent différents phénomènes économiques.

Manuel PPI: Chapitre 1, section E; Chapitre 17, paragraphes 17.1–17.26 et 17.61–17.67.

Manuel de l'IPC: Chapitre 1, paragraphes 1.85–1.107.

### 9.1 Indice des prix des intrants

L'objectif d'un indice des prix des intrants est de mesurer la variation du coût des intrants intermédiaires d'une entreprise représentative, en maintenant les extrants à un certain niveau de référence. Ceci est analogue à un indice à panier fixe, sauf que maintenant les conditions économiques auxquelles l'entreprise est

---

<sup>13</sup>Ou parfois les indices Törnqvist ou Walsh, selon les axiomes particuliers — voir Balk and Diewert (2001) et Balk (2008Sections 3.6.4 et 3.6.5.).

<sup>14</sup>L'entreprise est également supposée faire des choix pour maximiser la rentabilité. Bien qu'il ne s'agisse pas d'une hypothèse controversée, cela ne doit pas être la seule motivation de l'entreprise — voir Martin (2019).

<sup>15</sup>Le consommateur représentatif est supposé choisir la consommation pour maximiser l'utilité. Bien qu'il s'agisse d'une hypothèse standard, elle est de moins en moins acceptée au fil du temps, car les considérations comportementales font désormais partie de la théorie microéconomique.

confrontée sont maintenues fixes, plutôt que la quantité d'intrants utilisée. Cette distinction est pratiquement importante et signifie qu'il n'y a pas de limite à la capacité de l'entreprise de substituer des intrants plus chers à des intrants moins chers lorsque les prix changent. Un indice des prix des intrants peut également être considéré comme un indice du coût de production, analogue à un indice du coût de la vie pour un indice des prix à la consommation — un indice des prix des intrants mesure l'expérience de l'entreprise en matière de changement d'intrants des prix.

Pour corriger la notation, soit  $y$  la quantité de production produite par une entreprise,  $q$  la quantité d'intrants disponibles et  $p$  le prix des intrants primaires. Comme pour l'approche axiomatique, ce sont des vecteurs de valeurs (c'est-à-dire que  $y$  est un vecteur avec la quantité de chaque production produite par l'entreprise). En prenant  $y$  comme fixe et en supposant que l'entreprise n'a aucun pouvoir de marché sur le marché des intrants (de sorte que l'entreprise traite  $p$  comme fixe), une entreprise qui maximise ses profits choisira ses intrants,  $q$ , pour minimiser le coût d'atteindre les résultats  $y$ . Le résultat de cet exercice est la fonction de dépense de l'entreprise,  $e(p, y)$ , qui donne le coût minimum d'achat des intrants nécessaires pour produire  $y$  unités de production.<sup>16</sup> Parce que l'entreprise est un maximiseur de profit, ce minimum le coût du produit sera en accord avec leur coût de production réel sous différentes configurations de prix pour les intrants et les extrants produits.

Après avoir défini la fonction de dépense de l'entreprise, il est alors simple de définir un indice de coût des intrants entre la période 0 et la période  $t$  comme

$$I^C = \frac{e(p_t, y_0)}{e(p_0, y_0)},$$

où  $y_0$  est la quantité de production produite par l'entreprise au cours de la période 0. L'indice des prix des intrants mesure une variation des prix des intrants en comparant les dépenses nécessaires pour satisfaire un ensemble fixe de conditions économiques (de production) dans le temps.

Il existe en fait toute une famille d'indices de prix des intrants selon le niveau auquel les quantités sont fixées dans la fonction de dépense. L'indice des prix des intrants défini ci-dessus est un indice des prix des intrants de type Laspeyres car les conditions économiques auxquelles l'entreprise est confrontée sont fixées à leurs valeurs de période 0; un index de type Paasche fixe la sortie à son niveau période- $t$ . Cependant, l'intuition est la même dans tous les cas, donc l'attention est limitée à l'indice des prix des intrants de type Laspeyres ci-dessus pour garder la présentation simple.

Il est utile de comprendre en quoi l'indice des prix des intrants diffère d'un indice de Laspeyres afin d'apprécier l'approche économique de la construction d'un indice des prix. Il est facile de voir que

$$e(p_0, y_0) = \sum_{i=1}^n p_{i0} q_{i0};$$

la valeur de la fonction de dépenses de la période 0 est simplement la dépense observable sur les intrants de la période 0, sinon l'entreprise ne pourrait pas fonctionner pour maximiser le profit de la période 0. Maintenant, dans la période  $t$ ,

$$e(p_t, y_0) \leq \sum_{i=1}^n p_{it} q_{i0}.$$

---

<sup>16</sup>Formellement,  $e(p, y) = \min_{q \in V(y)} p \cdot q$ , où  $V(y)$  est l'ensemble des entrées qui produisent au moins  $y$  unités de sortie. Cette fonction est bien définie dans des conditions de régularité assez douces sur l'ensemble  $V(y)$  (c'est-à-dire qu'elle est non vide et fermée) — voir McFadden (1978).

Évidemment, les intrants utilisés dans la période 0 sont suffisamment bons pour produire la quantité de production dans la période 0, mais ce choix d'intrants n'a pas besoin de minimiser le coût de production aux prix de la période  $t$ . Si le prix d'un intrant augmente entre la période 0 et la période  $t$ , l'entreprise se substituera généralement à cet intrant et utilisera davantage un intrant relativement moins cher afin de minimiser les coûts tout en produisant  $y_0$  unités de production. Cette substitution ne se reflète pas dans le choix des intrants  $q_0$ , et donc cet ensemble d'intrants coûtera plus cher que l'ensemble d'intrants minimisant les coûts qui peut produire  $y_0$  unités de production à la période  $t$  prix.

Ces deux expressions peuvent maintenant être utilisées pour montrer que

$$\frac{e(p_t, y_0)}{e(p_0, y_0)} = \frac{e(p_t, y_0)}{\sum_{i=1}^n p_{i0} q_{i0}} \leq \frac{\sum_{i=1}^n p_{it} q_{i0}}{\sum_{i=1}^n p_{i0} q_{i0}}.$$

L'indice de Laspeyres surestime donc le mouvement des prix entre la période 0 et la période  $t$  car il ne prend pas en compte la substitution des intrants d'un changement de prix qui fait partie de l'indice des prix des intrants — il y a une substitution positive biais. Ceci est la clé de l'approche économique — il existe une divergence entre les formules de référence standard utilisées pour mesurer un changement de prix et la variation économique de prix — et est l'élan pour dériver des formules de référence qui peuvent être utilisés pour calculer l'indice économique des prix des intrants.

## 9.2 Indice des prix à la production

Un indice des prix à la production est analogue à un indice des prix des intrants; la seule différence est que l'indice compare la valeur de la production au fil du temps, plutôt que le coût des intrants, en maintenant les conditions économiques fixes. À cette fin, un indice des prix à la production compare le revenu maximal qu'une entreprise représentative pourrait recevoir à deux moments dans le temps, étant donné un ensemble fixe d'intrants pour la production.

En modifiant la notation de la section précédente, que  $p$  soit le prix des sorties  $y$ , plutôt que le prix des entrées  $q$ . En prenant des intrants fixés à  $q$ , une entreprise qui maximise ses profits choisit ses sorties  $y$  pour maximiser ses revenus, étant donné les prix  $p$ . Cela se traduit par la fonction de revenu de l'entreprise,  $r(p, q)$ , qui donne le revenu maximal que l'entreprise pourrait gagner étant donné les prix du marché et une quantité fixe d'intrants.<sup>17</sup> Comme pour un indice de prix d'entrée, l'indice de prix de sortie est défini comme

$$I^R = \frac{r(p_t, q_0)}{r(p_0, q_0)}.$$

Un indice des prix à la production mesure la valeur de la production d'une entreprise au fil du temps, compte tenu d'une quantité fixe d'intrants.

Contrairement à un indice des prix des intrants, un indice des prix des extrants est toujours plus grand qu'un indice de Laspeyres. Dans la période 0, les revenus d'une entreprise sont tout simplement

$$r(p_0, q_0) = \sum_{i=1}^n p_{i0} y_{i0};$$

la valeur de la fonction de revenu de la période 0 est simplement la période de revenu observable 0, sinon l'entreprise ne pourrait pas fonctionner pour maximiser le profit de la période 0. Maintenant, dans la période  $t$ ,

---

<sup>17</sup>Formellement,  $r(p, q) = \max_{y \in Y(q)} p \cdot y$ , où  $Y(q)$  est l'ensemble des sorties qui produisent avec au plus  $q$  unités d'entrée. Cette fonction est bien définie dans des conditions de régularité assez douces sur l'ensemble  $Y(q)$  (c'est-à-dire qu'elle est compacte) — voir McFadden (1978).

$$r(p_t, q_0) \geq \sum_{i=1}^n p_{it} y_{i0}$$

car l'entreprise produira relativement plus de produits dont le prix a augmenté, tout en utilisant la même quantité d'intrants. L'entreprise se substitue à des productions moins rentables, ce qui ne se reflète pas dans le choix de l'entreprise pour la production de la période 0. Par conséquent,

$$\frac{r(p_t, q_0)}{r(p_0, q_0)} = \frac{r(p_t, q_0)}{\sum_{i=1}^n p_{i0} y_{i0}} \geq \frac{\sum_{i=1}^n p_{it} y_{i0}}{\sum_{i=1}^n p_{i0} y_{i0}};$$

un indice de Laspeyres sous-estime une augmentation des prix dans le temps car il ne prend pas en compte la substitution des sorties dans le temps qui fait partie de l'indice des prix des sorties. Il s'agit de la direction opposée du biais de substitution pour un indice des prix des intrants, et signifie que la direction du biais de substitution dépend de ce que l'indice des prix tente de mesurer.

### 9.3 Indice du coût de la vie

Un indice du coût de la vie est entièrement analogue à un indice des prix des intrants, sauf qu'il mesure le coût pour un consommateur représentatif associé à l'atteinte d'un niveau d'utilité fixe à partir de la consommation — comment les dépenses d'un consommateur changent-elles pour les garder aussi bien qu'à la période 0? La théorie est exactement la même qu'un indice des prix des intrants, sauf que les extrants  $y_0$  sont remplacés par une mesure de l'utilité de la consommation pendant la période 0 dans la fonction de dépense. Autrement dit, un indice du coût de la vie est

$$I^C = \frac{e(p_t, u_0)}{e(p_0, u_0)},$$

où  $u_0$  est l'utilité du consommateur par rapport à la consommation de la période 0. À l'exception d'un changement d'interprétation, la mécanique d'un indice du coût de la vie est identique à un indice des prix des intrants et ne sera pas répétée.

### 9.4 Approximation d'un indice économique

Un indice des prix économiques ne peut pas être calculé directement dans la pratique. Il n'y a aucun espoir de calculer un indice des prix des intrants ou des prix des extrants sans connaître la technologie de production de l'entreprise, car la forme de la fonction de dépenses ou de revenus des entreprises est inconnue. Le problème est encore plus grave pour un indice du coût de la vie, car il nécessite la connaissance des préférences des consommateurs pour construire la fonction de dépense du consommateur. Le mieux que l'on puisse faire est d'approximer un indice économique en utilisant des données sur les prix et les quantités observables. Les indices arithmétiques de Laspeyres et géométriques de Laspeyres offrent une bonne approximation soit du prix des intrants, soit de l'indice des prix des extrants, soit de l'indice du coût de la vie, au moins selon une approximation de premier ordre.

L'argument pour approximer un indice économique avec un Laspeyres arithmétique ou un Laspeyres géométrique est le même pour un indice des prix des intrants, un indice des prix des extrants et un indice du coût de la vie, et donc l'attention se limite à l'approximation d'un prix des intrants indice. Pour montrer que l'indice de Laspeyres se rapproche de l'indice des prix des intrants, rappelons que

$$e(p_0, y_0) = \sum_{i=1}^n p_{i0} q_{i0}.$$

On peut maintenant montrer que la meilleure approximation linéaire de la fonction de dépense avec les prix de la période  $t$  est

$$e(p_t, y_0) \approx \sum_{i=1}^n p_{it} q_{i0}$$

chaque fois que les prix de la période  $t$  ne sont pas trop différents des prix de la période 0.<sup>18</sup> Par conséquent

$$\frac{e(p_t, y_0)}{e(p_0, y_0)} \approx \frac{\sum_{i=1}^n p_{it} q_{i0}}{\sum_{i=1}^n p_{i0} q_{i0}},$$

au moins pour les petits changements de prix.

De même, pour les Laspeyres géométriques, la meilleure approximation linéaire est basée sur

$$\log(e(p_t, y_0)) - \log(e(p_0, y_0)) \approx \sum_{i=1}^n \frac{p_{i0} q_{i0}}{\sum_{j=1}^n p_{j0} q_{j0}} \log\left(\frac{p_{it}}{p_{i0}}\right).$$

Donc

$$\begin{aligned} I^C &= \exp(\log(e(p_t, y_0)) - \log(e(p_0, y_0))) \\ &\approx \exp\left(\sum_{i=1}^n \frac{p_{i0} q_{i0}}{\sum_{j=1}^n p_{j0} q_{j0}} \log\left(\frac{p_{it}}{p_{i0}}\right)\right) \\ &= \prod_{i=1}^n \left(\frac{p_{it}}{p_{i0}}\right)^{\omega_{i0}} \end{aligned}$$

où  $\omega_{i0} = p_{i0} q_{i0} / \sum_{j=1}^n p_{j0} q_{j0}$  est le poids des dépenses de la période 0 pour un bon  $i$ . Comme un indice géométrique est toujours plus petit que son indice arithmétique correspondant et que l'indice de Laspeyres est trop grand par rapport à l'indice des prix des intrants (c.-à-d. Que l'erreur d'approximation est toujours positive en raison du biais de substitution), l'indice géométrique de Laspeyres donnera une meilleure approximation de l'indice des prix des intrants que de l'indice de Laspeyres chaque fois que l'erreur d'approximation de l'indice géométrique est positive.

Dans de nombreux cas, l'indice de Fisher ou de Törnqvist offre une meilleure approximation à un indice des prix des intrants que l'indice de Laspeyres ou l'indice géométrique de Laspeyres. L'argument pour lequel ces indices peuvent être une meilleure approximation est plus complexe — voir ILO et al. (2004aChapitre 17) et ILO et al. (2004bChapitre 17) pour plus de détails. Le coût de cette meilleure approximation, cependant, est le besoin de plus d'informations sur les quantités de la période  $t$ , qui peuvent ou non être disponibles au moment où l'indice est réellement calculé.<sup>19</sup>

<sup>18</sup>Cela découle du lemme de Shephard (McFadden 1978), ce qui implique que  $e(p_t, y_0) = \sum_{i=1}^n p_{it} q_{i0} + \eta(p_t - p_0) \|p_t - p_0\|$ , où  $\eta(p_t - p_0) \rightarrow 0$  lorsque  $p_t \rightarrow p_0$ .

<sup>19</sup>Dans certains cas, ces approximations sont exactes. Un indice de Laspeyres sera un véritable indice des prix des intrants si l'entreprise a une fonction de production Leontief; un indice de Fisher sera également exact dans ce cas. De même, un Laspeyres géométrique sera un véritable indice des prix des intrants si l'entreprise a des fonctions de production Cobb-Douglas; un indice de Törnqvist sera également exact dans ce cas. L'indice de prix de Lloyd-Moulton est un cas intéressant car il est exact chaque fois que l'entreprise a une élasticité constante de la fonction de production de substitution, mais il ne nécessite que des informations sur la quantité pour la période 0.

## 10 L'approche stochastique

Les approches axiomatiques et économiques pour motiver un indice des prix prennent les prix comme fixes. Comme son nom l'indique, l'approche stochastique ne le fait pas; au lieu de cela, les prix sont traités comme aléatoires, du moins du point de vue de celui qui établit l'indice des prix. Cela ne veut pas dire que les prix sont choisis ou déterminés au hasard, mais simplement qu'il y a une distribution des prix des biens et services échangés, et l'observation d'un prix donné s'apparente à un tirage aléatoire de cette distribution. L'avantage évident de cette approche est qu'elle est conforme à la pratique habituelle de calcul d'un indice des prix à l'aide d'un échantillon de données, permettant un traitement explicite des propriétés statistiques d'un indice des prix.<sup>20</sup>

Au niveau conceptuel, cependant, l'approche stochastique motive un indice des prix par l'un de ses principaux usages: dégonfler et gonfler les prix dans le temps. L'acte de dégonfler / gonfler les prix est essentiellement un acte de prédiction — quel serait le prix d'un bien ou d'un service dans une autre période? — et un indice de prix peut être motivé en examinant comment choisir une valeur unique qui prédit le mieux un changement de prix pour une population de biens vendus entre deux périodes.

Cette section du cours conclut le voyage dans la théorie de l'indice des prix en donnant une introduction un peu plus approfondie à l'approche stochastique que celle vue ailleurs. En effet, l'approche stochastique est probablement la plus pratique des trois approches. Bien qu'il ne soit pas aussi élevé que les approches axiomatiques ou économiques, il offre le modèle le plus utile pour produire réellement un indice des prix en tant que statistique macroéconomique.

Manuel PPI: Chapitre 1, section D; Chapitre 16, section D.

### 10.1 Le problème de prédiction optimal

Le point de départ de l'approche stochastique est une distribution de  $n$  prix relatifs entre la période 0 et la période  $t$ , le prix relatif du bon  $i$  étant donné par  $p_{it}/p_{i0}$ , pour la population (univers) de tous les biens et services  $n$  à mesurer par l'indice des prix. À chaque bien  $i$  est associée la probabilité que  $P_i$  observe ce bien transigé sur les deux périodes dans la population de biens mesurée par l'indice, ce qui donne à son tour une mesure de la fréquence d'observation de différentes valeurs pour les changements de prix au fil du temps. Autrement dit, un prix relatif  $p_t/p_0$  est traité comme une variable aléatoire dans l'approche stochastique, dont la distribution dépend de la distribution sous-jacente des biens et services qui constituent la portée de l'indice des prix.<sup>21</sup>

Étant donné une distribution des prix relatifs, le meilleur prédicteur d'un changement de prix entre la période 0 et la période  $t$  est la valeur  $I^A$  qui minimise l'erreur de prédiction attendue. Formellement, le meilleur prédicteur résout le problème suivant<sup>22</sup>

$$\min_I E \left[ \left( \frac{p_t}{p_0} - I \right)^2 \right],$$

où  $E$  est l'opérateur de valeur attendue, repris par la distribution du prix relatif. Comme cela est bien connu, la solution à ce problème consiste à définir  $I^A = E(p_t/p_0)$  (Lehmann and Casella 1998, chapitre 1, exemple

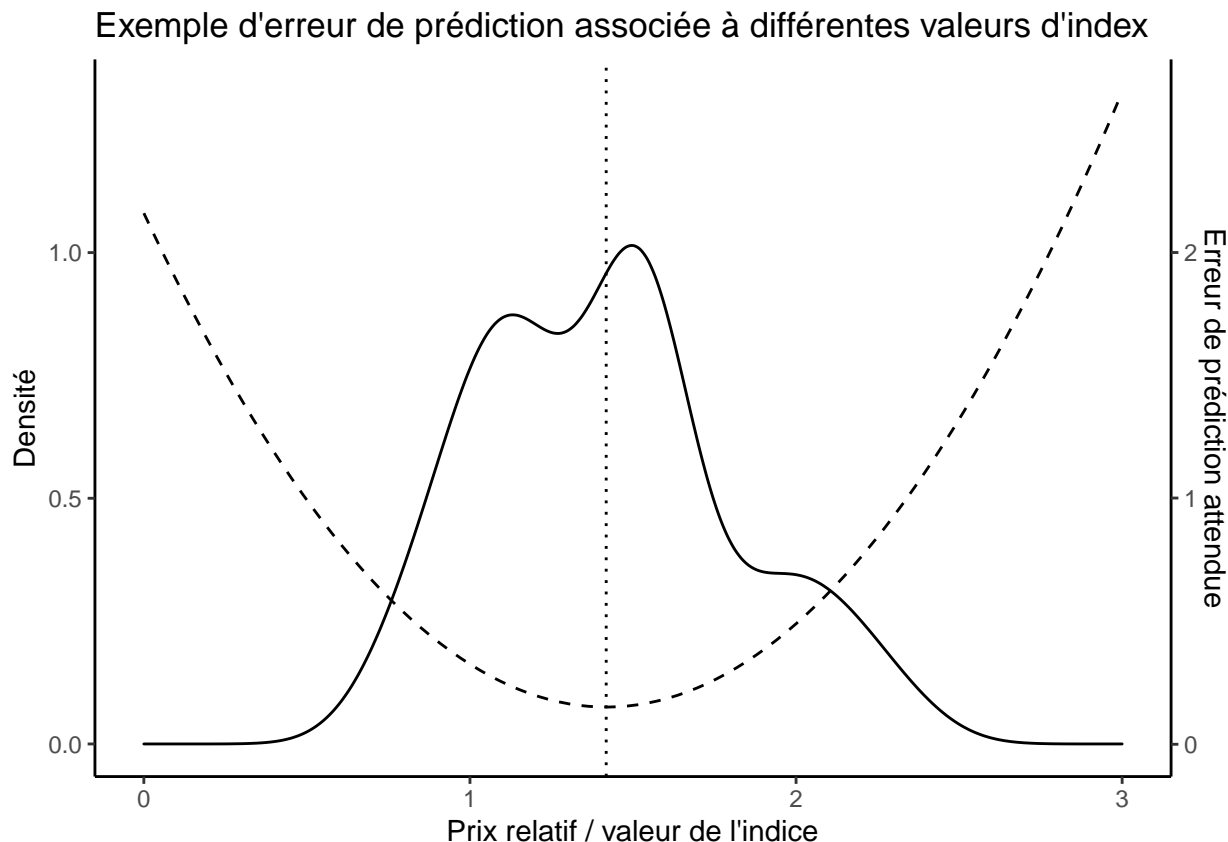
<sup>20</sup>Formellement, dans l'approche stochastique, les biens et les services sont modélisés comme appartenant à un espace de probabilité  $(\Omega, \mathcal{F}, P)$ , où  $\Omega$  est l'ensemble (fini) de biens et services qui sont traités,  $\mathcal{F}$  est l'ensemble de puissance de  $\Omega$ , et  $P$  est une mesure de probabilité qui donne la probabilité d'observer un bien ou un service dans  $\Omega$  traité. Il y a ensuite un mappage de  $\Omega$  à  $\mathbb{R}_+^2$  qui donne le prix d'un bien pendant la période 0 et la période  $t$ .

<sup>21</sup>On a tendance à penser que cette distribution est gaussienne, mais cela signifie que certains biens et services ont des prix négatifs. Dans la plupart des cas, les prix relatifs ne peuvent pas être distribués gaussiens.

<sup>22</sup>Il s'agit d'un problème de prédiction utilisant une fonction de perte quadratique, pour garantir que l'erreur de prédiction est toujours positive et pour formaliser l'idée qu'une grande erreur de prévision est plus coûteuse que de nombreuses petites erreurs de prévision. Différentes fonctions de perte donnent généralement différents prédicteurs optimaux; par exemple, la fonction de perte absolue renvoie la médiane comme meilleur prédicteur, qui trouve des applications pour les indices des prix des logements. Si les parents de prix sont distribués symétriquement, alors toutes les fonctions de perte convexes symétriques donnent le même indice de prix (Lehmann and Casella 1998, Chapitre 1 Corollaire 7.19).

7.17); c'est-à-dire que le meilleur prédicteur d'un changement de prix entre deux périodes est le changement attendu de prix. S'il peut sembler inhabituel d'exprimer un indice des prix en termes de valeur attendue, cela souligne simplement qu'un indice des prix est à juste titre une caractéristique de la distribution des prix dans la population. C'est une caractéristique clé de l'approche stochastique, et c'est ce qui permet une analyse significative des propriétés statistiques d'un indice des prix lorsqu'il est calculé avec un échantillon de données sur les prix de la population.

La figure donne un exemple de ce à quoi ressemble le problème de prédiction. Étant donné une distribution des prix relatifs (ligne continue), il y a une erreur de prédiction attendue associée à chaque valeur qu'un indice de prix peut prendre (ligne pointillée). Le meilleur choix pour un indice des prix comme outil de prévision est la valeur qui minimise l'erreur de prédiction attendue (ligne pointillée).



Comme la valeur attendue n'est qu'une moyenne pondérée, l'indice  $I^A = E(p_t/p_0)$  est simplement un indice arithmétique qui fait la moyenne des prix relatifs. Ce qui est intéressant, c'est que les poids ont maintenant une nouvelle interprétation — ils sont la probabilité d'observer un bien ou un service transigé dans la population. C'est,

$$I^A = \sum_{i=1}^n \frac{p_{it}}{p_{i0}} P_i.$$

Différentes formules arithmétiques d'indice de prix — par exemple, Laspeyres, Paasche, Lowe — correspondent donc à différentes déclarations sur ce qui détermine la probabilité d'observer une transaction. Par exemple, avec un indice de Laspeyres, les probabilités sont données par la part des dépenses ou des revenus de la période 0. Cela revient à dire que la probabilité d'observer une transaction pour un bien est la probabilité de dépenser ou de recevoir un dollar pour ce bien au cours de la période 0.



## 10.2 Les indices de prix géométriques comme meilleur prédicteur

Un indice géométrique peut être motivé de la même manière qu'un indice arithmétique. Formellement, un index géométrique est la valeur  $I^G$  qui résout

$$\min_I E \left[ \left( \log \left( \frac{p_t}{p_0} \right) - \log(I) \right)^2 \right],$$

la solution à laquelle est  $I^G = \exp(E[\log(p_t/p_0)])$ . Mais ce n'est qu'un index géométrique, comme

$$\begin{aligned} \exp \left( E \left[ \log \left( \frac{p_t}{p_0} \right) \right] \right) &= \exp \left( \sum_{i=1}^n P_i \log \left( \frac{p_{it}}{p_{i0}} \right) \right) \\ &= \prod_{i=1}^n \left( \frac{p_{it}}{p_{i0}} \right)^{P_i}, \end{aligned}$$

avec des poids égaux à la probabilité d'observer un prix relatif. Il convient de noter qu'en raison des logarithmes du problème de prédiction, un indice géométrique peut également être motivé par la recherche d'une valeur  $I^G$  telle que  $I^G \times p_{i0}$  prédit le mieux  $p_{it}$  (ou  $p_{it}/I^G$  prédit le mieux  $p_{i0}$ ). Il s'agit d'une propriété non partagée avec les indices arithmétiques. Alors qu'un indice arithmétique prédit les variations de prix au fil du temps, un indice géométrique prend le prix d'un bien pendant la période 0, le gonfle / le dégonfle avec l'indice des prix et utilise le résultat pour prédire le prix de ce bien pendant la période  $t$ . Par conséquent, un indice géométrique s'intègre mieux dans l'approche stochastique car il correspond directement à la façon dont un indice des prix est utilisé pour gonfler et dégonfler les prix dans le temps dans la pratique.<sup>23</sup>

L'indice de Törnqvist apparaît généralement comme le meilleur indice de l'approche stochastique car il sélectionne une valeur sensible pour représenter la probabilité d'observer un prix relatif. L'idée derrière les probabilités de Törnqvist est que la probabilité d'observer un prix particulier relatif à un bien dépend de la part des dépenses / revenus de ce bien sur les deux périodes. Ainsi, l'indice de Törnqvist définit les probabilités sous forme de parts de dépenses moyennes:

$$P_i = \frac{1}{2} \frac{p_{i0}q_{i0}}{\sum_{j=1}^n p_{j0}q_{j0}} + \frac{1}{2} \frac{p_{it}q_{it}}{\sum_{j=1}^n p_{jt}q_{jt}}.$$

L'inconvénient de ces probabilités est qu'elles nécessitent des informations pour les parts des dépenses / recettes de la période 0 et de la période  $t$ , ce qui peut ne pas être connu au moment du calcul de l'indice.

## 10.3 Des indices de prix plus généraux

Les indices arithmétiques et géométriques sont des cas particuliers d'une classe plus large d'indices de prix qui résolvent le problème de prédiction suivant:

$$\min_I E \left[ \left( \left( \frac{p_t}{p_0} \right)^r - I^r \right)^2 \right],$$

---

<sup>23</sup>Motiver un indice des prix pour résoudre le problème  $\min_I E[(p_t - p_0 I)^2]$  donne  $I = E(p_t p_0) / E(p_0^2)$ , qui est un indice des prix assez inhabituel.

pour certains  $r \neq 0$ . La solution à ce problème est  $I = E((p_t/p_0)^r)^{1/r}$ . Fixer  $r = 1$  donne un indice arithmétique, alors que prendre  $r \rightarrow 0$  donne un indice géométrique (Bullen 2003, III 1 Theorem 2).

Ce qui est utile dans cette approche, c'est que de nombreux autres types d'indices de prix correspondent à différents choix de  $r$ , et ceux-ci peuvent à leur tour être motivés pour résoudre un type de problème de prédiction. La définition de  $r = -1$ , par exemple, produit un indice de prix harmonique

$$I^H = \left( \sum_{i=1}^n \frac{P_i}{p_{it}/p_{i0}} \right)^{-1}.$$

Lorsque la probabilité d'observer un bien est donnée par sa part de dépenses période- $t$ , il s'agit de l'indice de Paasche. Fixer  $r = 1 - \sigma$ , où  $\sigma$  est l'élasticité de substitution, donne l'indice de prix de Lloyd-Moulton

$$I^{LM} = \left( \sum_{i=1}^n P_i \left( \frac{p_{it}}{p_{i0}} \right)^{1-\sigma} \right)^{1/(1-\sigma)}$$

lorsque les probabilités sont des dépenses / recettes pour la période 0.

Un point intéressant à propos de ces types d'indices de prix plus généraux est que, pour un ensemble de pondérations donné, la valeur de l'indice est plus grande lorsque le paramètre  $r$  augmente (Bullen 2003, III 3.1 Theorem 1). Cela donne le résultat familier qu'un indice arithmétique est plus grand qu'un indice géométrique, qui à son tour est plus grand qu'un indice harmonique, mais il peut également être utilisé pour classer des types d'indices plus exotiques comme l'indice Lloyd-Moulton.

## 10.4 Inférence statistique

Dans la pratique, la distribution complète des prix relatifs n'est généralement pas connue et un indice est calculé à partir d'un échantillon de prix collectés auprès des producteurs ou des détaillants. Cela signifie que les valeurs de l'indice sont calculées avec un estimateur pour  $I^A$  ou  $I^G$ , selon qu'un indice arithmétique ou géométrique est l'indice cible. Bien que les sujets de l'échantillonnage et de l'inférence statistique se compliquent rapidement et dépassent le cadre de ce cours, il convient d'examiner comment les indices arithmétiques et géométriques se comportent lorsqu'ils sont calculés avec un échantillon aléatoire de données de prix.<sup>24</sup> L'un des avantages de l'approche stochastique est qu'elle donne un aperçu du problème d'estimation d'un indice de prix avec un échantillon.

Avec l'échantillonnage aléatoire, une approche naturelle pour estimer l'indice arithmétique consiste à remplacer la valeur attendue — une moyenne de population — par la moyenne de l'échantillon. Cela donne un estimateur de la méthode des moments  $\hat{I}^A = 1/n_s \sum_{i=1}^{n_s} p_{it}/p_{i0}$ , où  $n_s$  est la taille de l'échantillon, qui n'est qu'un index Carli. Si les parents de prix sont échantillonnés au hasard, alors il est facile de voir que  $E(\hat{I}^A) = I^A$ , donc l'indice de Carli est un estimateur sans biais pour l'indice arithmétique.

Un estimateur naturel pour l'indice géométrique est  $\hat{I}^G = \prod_{i=1}^{n_s} (p_{it}/p_{i0})^{1/n_s}$ , l'index Jevons. Contrairement à l'indice de Carli, cependant, l'indice de Jevons est un estimateur biaisé de l'indice géométrique. Il est à nouveau simple de montrer que  $E(\hat{I}^G) \geq I^G$ , avec  $E(\hat{I}^G) = I^G$  ne tenant que dans des circonstances très particulières — l'indice de Jevons surestime systématiquement l'indice géométrique (c'est-à-dire qu'il est biaisé à la hausse).<sup>25</sup> Bien que le biais soit un inconvénient de l'indice de Jevons, le biais n'est pas la seule propriété statistique importante d'un estimateur, et l'indice de Jevons a d'autres propriétés statistiques souhaitables (par exemple, c'est une constante estimateur de l'indice géométrique et, dans certaines circonstances, atteint

<sup>24</sup>Voir Balk (2008Chapitre 5) pour une introduction à l'échantillonnage et à l'inférence statistique pour les indices de prix.

<sup>25</sup>L'indice de Jevons n'est pas biaisé dans les cas où la distribution d'échantillonnage est dégénérée, ou la variance des prix relatifs est zéro. Voir Lehmann and Casella (1998Chapitre 1 Théorème 7.5).

le maximum d'efficacité de vraisemblance lié).<sup>26</sup> Il est possible d'ajuster le biais à la hausse des Jevons en divisant les Jevons index par  $1 + \sigma^2/(2n)$ , où  $\sigma^2$  est la variance des logarithmes du prix (Kennedy 2003, 41), mais cela n'est généralement pas fait dans la pratique.

## 11 Affectation 2

Les réponses à ces questions proviennent à la fois du contenu du cours et des lectures. Chaque question vaut un point, pour un total de 20 points. La réussite de ce module nécessite au moins 65% (13 sur 20 correct). Envoyez vos réponses par e-mail à l'instructeurs du cours (Steve Martin ou Shaoxiong Wang ou Anastasiia Nosach) lorsque vous avez terminé.

**Question 1** Vrai ou faux: l'indice de Laspeyres est un indice du coût de la vie si les consommateurs ne se substituent pas à des produits plus chers lorsque les prix changent.

**Question 2** Vrai ou faux: le test d'inversion du temps indique qu'un indice des prix entre la période 0 et la période 1 est identique à l'inverse de l'indice des prix entre la période 1 et la période 0.

**Question 3** Quelle est la différence entre les approches économique et axiomatique des indices?

- a) Les prix et les quantités sont indépendants dans l'approche axiomatique et dépendants dans l'approche économique.
- b) L'approche économique traite les prix comme fixes, tandis que l'approche axiomatique traite les prix comme aléatoires.
- c) L'approche économique consiste en une série de tests basés sur la théorie économique pour évaluer la pertinence d'un numéro d'index, tandis que l'approche axiomatique utilise les axiomes fondamentaux de la théorie des ensembles pour évaluer dans quelle mesure un numéro d'index compare les listes de prix.
- d) L'approche axiomatique et l'approche économique sont équivalentes.
- e) Aucune de ces réponses.

**Question 4** Quels poids l'indice Palgrave utilise-t-il pour agréger les prix relatifs entre la période 0 et la période 1?

- a) Part des dépenses / recettes de la période 0.
- b) Part des dépenses / recettes de la période 1.
- c) La moyenne géométrique des parts des dépenses / recettes de la période 0 et de la période 1.
- d) La moyenne harmonique des parts des dépenses / recettes de la période 0 et de la période 1.
- e) Aucune de ces réponses.

**Question 5** Vrai ou faux: l'indice de Fisher est hautement considéré comme un numéro d'index car il satisfait aux axiomes les plus importants et correspond toujours à un indice du coût de la vie.

**Question 6** Vrai ou faux: le test d'inversion du temps indique qu'un indice de prix et un indice de quantité ont la même forme fonctionnelle.

**Question 7** Quel axiome l'indice Dutot ne satisfait-il pas?

- a) Monotonie.
- b) Continuité.

---

<sup>26</sup>L'indice de Carli est également un estimateur cohérent de l'indice arithmétique sous échantillonnage aléatoire, mais il ne peut pas être un estimateur du maximum de vraisemblance car cela nécessiterait des prix apparentés d'avoir une distribution gaussienne, ce qui signifie que les prix relatifs peuvent être négatifs. En revanche, l'indice de Jevons est un estimateur du maximum de vraisemblance pour l'indice géométrique si les parents de prix ont une distribution log-normale, ce qui implique que les prix sont toujours positifs.

- c) Invariance dimensionnelle.
- d) Circularité.
- e) Aucune de ces réponses.

**Question 8** Vrai ou faux: il est impossible qu'un indice géométrique de Laspeyres soit un indice du coût de la vie.

**Question 9** Quel type d'indice faut-il utiliser pour calculer un indice de quantité de Laspeyres en dégonflant la variation de la valeur globale entre deux périodes?

- a) Indice de Palgrave.
- b) Indice de Laspeyres.
- c) Indice de Paasche.
- d) Indice de Fisher.
- e) Indice géométrique de Laspeyres.

**Question 10** Vrai ou faux: l'indice de Jevons est un estimateur biaisé de l'indice géométrique dans la population sous échantillonnage aléatoire.

Les 4 questions suivantes utilisent les données du tableau suivant qui contiennent une population de trois parents de prix, chacun avec une probabilité égale, et tous les échantillons possibles de cette population.

Population des parents de prix	Échantillon 1	Échantillon 2	Échantillon 3
1.2	1.2	1.2	1.1
1.1	1.1	1,35	1,35
1,35	...	...	...

**Question 11** Quelle est la valeur de l'indice géométrique dans la population?

- a) 121,2
- b) 100
- c) 120,5
- d) 129,8
- e) 98,7

**Question 12** Quelle est la valeur moyenne (arithmétique) des indices de Jevons dans chaque échantillon?

- a) 121,2
- b) 119,4
- c) 121,3
- d) 100
- e) 123,1

**Question 13** Quelle est la valeur du biais dans l'indice de Jevons?

- a) 0
- b) -0,2
- c) 0,1
- d) 0,8

e) -1,3

**Question 14** Quelle est la valeur du biais dans l'indice Carli?

a) 1,1

b) -0,4

c) -0,2

d) 0,7

e) 0

**Question 15** Vrai ou faux: si un consommateur représentatif a une fonction d'utilité d'élasticité de substitution constante, l'ampleur du biais de substitution augmente à mesure que l'élasticité de substitution augmente.

Les quatre questions suivantes utilisent les données du tableau suivant qui contiennent des informations sur les prix et les quantités pour deux entrées sur deux périodes, ainsi qu'un véritable indice des prix des entrées.

Période	Prix 1	Prix 2	Quantité 1	Quantité 2	Indice des prix des intrants
0	100	60	141	391	100,0
1	120	80	160	360	128,0

**Question 16** Quelle est la valeur de l'indice de Laspeyres?

a) 101,0

b) 133,3

c) 128,0

d) 128,3

e) 99,3

**Question 17** Quelle est la valeur de l'indice de Paasche?

a) 127,7

b) 131,2

c) 101,0

d) 99,1

e) 130,0

**Question 18** Quelle est la valeur de l'indice Jevons?

a) 120,1

b) 128,2

c) 99,0

d) 126,5

e) 98,4

**Question 19** Vrai ou faux: l'indice Jevons calculé à la question 18 est plus proche de l'indice des prix des intrants que l'indice Fisher.

**Question 20** Vrai ou faux: il est possible de construire un indice des prix des intrants pour une entreprise qui détient un pouvoir de marché pour les produits qu'elle vend.

## Part III

# Indices de prix de qualité constante

## 12 Programme pour Indices de prix de qualité constante

Un indice des prix devrait mesurer un pur mouvement des prix dans le temps. Dans la pratique, cependant, les biens et services changent avec le temps, ce qui peut contaminer un indice des prix — sans informations supplémentaires, il n'y a aucun moyen de dissocier un mouvement des prix d'un changement dans la qualité d'un bien ou d'un service.

L'objectif de ce module est de toucher à certains sujets avancés dans la construction d'indices de prix, avec un accent particulier sur les ajustements de qualité et les données manquantes. À la fin du module, une personne devrait:

1. Familiarisez-vous avec la nécessité de procéder à des ajustements de qualité dans un indice des prix et le lien avec les données manquantes.
2. Comprendre l'importance de la stratification pour construire un indice des prix.
3. Avoir une compréhension de base des indices de prix qui utilisent des méthodes économétriques pour faire face aux changements de qualité.

Ce module est utile pour les compilateurs d'indices de prix ayant une compréhension de la théorie de base et de la construction d'un indice de prix, et qui souhaitent obtenir une compréhension plus approfondie de la façon de traiter les ajustements de qualité et les données manquantes.

Ce module se compose de lectures autodirigées, accompagnées d'un devoir. Au total, environ 10 à 15 heures devraient être consacrées à ce module

Prérequis: Théorie de l'indice des prix, au moins un cours d'introduction à l'économétrie. Un cours avancé d'économétrie ou de probabilité et de statistiques est utile.

L'évaluation de ce module est basée sur une tâche composée de 20 questions à choix multiples / vrai-faux qui s'appuient sur du matériel dans le contenu du cours et les lectures. La collaboration sur la mission est la bienvenue, mais chaque personne doit soumettre son propre travail unique. La réussite de ce module nécessite au moins 65% sur l'affectation.

Les lectures de ce module proviennent des chapitres 17 et 21 du manuel PPI (ILO et al. 2004b) et / ou du manuel CPI (ILO et al. 2004a), publiés par le FMI (disponibles gratuitement sur leur site Web), ainsi que des chapitres 4 à 6 de le Manuel sur les indices des prix des propriétés résidentielles (ILO et al. 2013), publié par Eurostat (disponible gratuitement sur leur site Web).<sup>27</sup>

Veuillez envoyer un e-mail à l'instructeurs du cours (Steve Martin ou Shaoxiong Wang ou Anastasiia Nosach) si vous avez des questions ou si vous avez besoin d'aide concernant le matériel ou le devoir du cours.

## 13 Introduction

La prise en compte des différences de qualité entre les biens est une préoccupation permanente lors de la construction d'un indice des prix. L'objectif d'un indice de prix est de capturer un mouvement de prix pur

---

<sup>27</sup>Bien que le manuel se concentre sur les prix des propriétés résidentielles, les idées s'appliquent à tous les biens ou services, et au matériel est relativement bien présenté.

sur deux périodes, mais s'il y a des différences systématiques dans les marchandises comparées au fil du temps qui affectent également le prix — soi-disant différences de qualité — alors un prix pur le mouvement ne peut pas être capturé en utilisant uniquement les prix de transaction. Sans information préalable, il n'y a aucun moyen de démêler un mouvement de prix pur d'un mouvement de prix en raison de l'évolution de la qualité au fil du temps. La mesure idéale d'un mouvement de prix pur est un indice des prix de qualité constante qui maintient les déterminants des prix fixés d'une période à l'autre. Cela permet une comparaison de pommes à pommes en utilisant les prix de transaction, et toute différence de prix entre les périodes doit refléter un mouvement de prix pur.

Il existe un certain nombre de techniques disponibles pour lutter contre les différences de qualité se glissant dans un indice des prix. L'approche la plus simple et la plus intuitive est probablement l'indice de modèle apparié pur, dans lequel les prix de paires de produits similaires sont comparés dans le temps. En se concentrant sur des paires de produits similaires, il devient moins probable que les différences de qualité entre différents produits puissent contaminer la mesure d'un mouvement de prix pur. D'autres techniques plus exotiques, à savoir la stratification et l'hédonique, offrent la même promesse d'un indice des prix de qualité constante, généralement au détriment d'appareils économétriques supplémentaires.

Une caractéristique intéressante de ces différentes méthodes - modèle, stratification et hédonie appariés - est qu'elles ne sont vraiment pas si différentes. Chacun s'appuie sur la même hypothèse clé pour produire un indice de qualité constante; ce qui diffère entre les méthodes, c'est simplement la manière dont elles mettent en œuvre cette hypothèse. Ce point est fondamental pour comprendre comment ces différentes approches peuvent fournir un indice de qualité constante et comment elles sont liées les unes aux autres.

Dans un souci de simplicité, ce cours se concentre sur les indices géométriques des prix. Les concepts d'un indice géométrique sont directement applicables aux indices arithmétiques, bien qu'il y ait quelques détails supplémentaires à s'inquiéter dans le cas arithmétique.<sup>28</sup> Dans la plupart des applications, les ajustements de qualité sont effectués à l'aide d'une formule d'indice géométrique. Tout au long de ce cours, l'attention est concentrée sur la construction d'un indice de qualité constante lorsque l'ensemble de la population des transactions de biens et services est connue. Ce n'est presque jamais le cas dans la pratique, mais cela facilite l'exposition en ignorant les problèmes associés à l'échantillonnage et à l'inférence statistique. Il souligne également que la construction d'un indice de qualité constante est fondamentalement un problème au niveau de la population, et non un problème d'échantillonnage. Une fois qu'un indice de qualité constante est défini dans la population, il est simple de l'estimer avec un échantillon de transactions en remplaçant les quantités au niveau de la population par leurs homologues de l'échantillon — voir Manski (1988).

## 14 Prix potentiels

La formalisation d'un indice de qualité constante nécessite d'étendre l'approche stochastique aux indices en généralisant le concept de prix. À cette fin, définissez un prix potentiel comme le prix pour lequel un bien ou un service se vendrait à un moment donné, qu'il se vende ou non — un prix potentiel est un prix contrefactuel pour un bien ou un service à un point dans le temps. Cela contraste avec un prix de transaction, qui est le prix observé pour lequel un bien se vend au moment où il se vend réellement. La comparaison des prix potentiels à deux moments donne un indice de qualité constante, car les produits comparés sont nécessairement maintenus fixes dans le temps. Il s'agit simplement d'un pur mouvement de prix sur deux périodes, faisant abstraction des variations de prix dues aux différences dans les caractéristiques des marchandises vendues au cours de ces périodes.

Le concept de prix potentiel est extrêmement utile. Dans le modèle d'indice de prix standard, seuls les prix et les quantités peuvent changer au fil du temps. Cependant, avec les prix potentiels, la composition de ce qui se vend au fil du temps peut également changer. Cette généralisation permet de définir un indice de prix lorsque non seulement les prix et les quantités évoluent dans le temps, mais aussi ce qui vend évolue dans le temps.

Un prix potentiel est un type particulier de résultat potentiel qui constitue la base du Rubin Causal Model,

---

<sup>28</sup>Voir Lee (2016Chapitre 1) et Manski (2007Chapitre 7) pour certains des détails.

un modèle performant pour faire l'inférence causale dans la littérature sur l'évaluation de programme. Le modèle causal de Rubin s'articule autour du récit d'une expérience, où un groupe d'individus reçoit un traitement, l'autre groupe servant de référence contrôlée. Le but de l'expérience est d'obtenir une mesure de l'impact causal du traitement. Dans le contexte d'un indice des prix, un indice de qualité constante est l'effet causal du temps sur les prix, et il est utile de garder à l'esprit le récit d'une expérience. Manski (2007Chapitre 7), Angrist and Pischke (2009Chapitres 2 à 4) et Wooldridge (2010Chapitre 21) fournissent chacun une excellente présentation du cadre de résultats potentiels. Lee (2016) donne un traitement d'une longueur de livre.

Pour opérationnaliser le concept d'un prix potentiel dans le cadre stochastique, supposons qu'il existe  $n$  biens uniques qui peuvent se vendre dans la période  $t = 0$  ou la période  $t = 1$ . Soit  $p_i(1)$  le prix du bien  $i$  s'il devait être vendu au cours de la période 1, et que  $p_i(0)$  soit le prix du bien  $i$  s'il devait vendre pendant la période 0, quel que soit le moment où le bien se vend réellement. Si le bien  $i$  se vend effectivement au cours de la période 1, le prix  $p_i(1)$  est le prix de transaction observé — il peut être observé lorsque le bien se vend au cours de la période 1 — alors que  $p_i(0)$  est contre-factuel, et donc non observable. De même, si un bon  $i$  se vend pendant la période 0, alors  $p_i(0)$  est le prix de transaction, avec  $p_i(1)$  contre-factuel. Ainsi, pour chaque prix de transaction, il existe également un prix contrefactuel, ce qui permet de construire un prix relatif  $p_i(1)/p_i(0)$  malgré un bien qui ne se vend pas dans les deux périodes.

Traiter chaque bien comme unique et le vendre une seule fois peut sembler étrange, mais c'est la façon appropriée de modéliser les marchandises au fil du temps. Cela signifie que le même produit vendu à deux moments différents est en réalité deux produits différents, ce qui permet à la qualité d'un même produit de changer avec le temps. Par exemple, les aliments pour chats peuvent se vendre à la fois pendant la période 0 et la période 1, mais ceux-ci sont traités comme des produits distincts car la qualité des aliments pour chats peut varier entre la période 0 et la période 1, si, par exemple, la taille de la boîte est plus petite dans période 1. Le fait que chaque bien soit unique permet également aux produits de disparaître avec le temps, ou de ne pas se vendre à chaque période, et toutes ces questions peuvent être traitées dans un seul cadre en utilisant la notion de prix potentiel. Il est clair que cela s'applique plus à certains types de produits (par exemple, le logement, les ordinateurs, les voitures) que d'autres (par exemple, la nourriture), mais le cadre est suffisamment général pour couvrir tous les types de biens et services.

## 14.1 Indices de prix de qualité constante

La motivation d'un indice des prix de qualité constante se déroule exactement comme dans le cadre stochastique, sauf que maintenant les prix potentiels sont utilisés à la place des prix de transaction. L'idée est qu'il existe une distribution des prix potentiels apparentés, un pour chaque bien, et qu'un indice des prix agit comme le meilleur prédicteur de l'évolution des prix au fil du temps. Cependant, en utilisant des prix potentiels, les mêmes produits sont comparés dans le temps, même si un produit tel qu'il apparaît sur le marché change entre la période 0 et la période 1. Formellement, un indice de prix (géométrique) de qualité constante est donné par

$$I^Q = \exp \left( E \left[ \log \left( \frac{p(1)}{p(0)} \right) \right] \right) \\ = \prod_{i=1}^n \left( \frac{p_i(1)}{p_i(0)} \right)^{P_i},$$

où  $P_i$  est la probabilité d'observer une bonne  $i$  dans la population de biens. Un indice de qualité constante est une généralisation de l'indice géométrique standard qui permet explicitement aux biens, ainsi qu'aux prix, de changer dans le temps.

Afin de simplifier la notation, laissez  $\rho(t) = \log(p(t))$ . Avec cette nouvelle notation, un indice de prix de qualité constante est donné par



$$\log(I^Q) = E(\rho(1)) - E(\rho(0)).$$

Cette notation est pratique car elle permet d'écrire un indice des prix de qualité constante comme une différence dans les prix potentiels moyens, bien que l'indice soit toujours une moyenne géométrique des prix potentiels apparentés.<sup>29</sup>

## 14.2 Indices de prix de transaction

Le défi de la construction d'un indice de prix de qualité constante est que les prix potentiels ne sont pas observables, et seules les informations sur les prix de transaction et le moment où un bien se vend peuvent être utilisées pour calculer un indice. Il y a un problème de données manquantes. Cela peut être vu en liant les prix de transaction aux prix potentiels

$$\rho = \rho(0) + t(\rho(1) - \rho(0)),$$

où  $\rho$  est le prix de transaction (log) et  $t$  donne la période de vente (soit  $t = 1$  ou  $t = 0$ ). Les seules informations qui peuvent être observées à partir des transactions sur le marché sont  $\rho$  et  $t$  — le prix pour lequel un bien a été vendu et le moment où il a effectivement vendu.

Compte tenu des informations sur les prix de transaction, un indice géométrique des prix de transaction tente de reproduire l'indice de qualité constante en comparant le prix de transaction moyen pour les marchandises qui se vendent au cours de la période 1 avec le prix de transaction moyen des marchandises qui se vendent au cours de la période 0, et est ainsi donné par

$$\begin{aligned} \log(I^T) &= E(\rho|t=1) - E(\rho|t=0) \\ &= E(\rho(1)|t=1) - E(\rho(0)|t=0). \end{aligned}$$

L'indice des prix de transaction diffère généralement de l'indice de qualité constante, car il compare les prix potentiels pour les produits qui se vendent au cours de la période 1 aux prix potentiels pour les produits qui se vendent au cours de la période 0, plutôt que pour tous les produits. Tout changement dans la composition des marchandises vendues entre la période 0 et la période 1 contamine la mesure d'un mouvement de prix pur, car les mêmes marchandises ne sont pas comparées dans le temps. Ce pour quoi les marchandises vendues au cours de la période 0 se sont vendues peut différer de ce que les marchandises vendues au cours de la période 1 auraient vendu au cours de la période 0, de sorte que la comparaison du prix de transaction moyen confond un changement de prix avec un changement dans la composition des marchandises vendues à différents points dans le temps.<sup>30</sup>

<sup>29</sup>Il s'agit d'un indice de type Törnqvist, car il donne la variation du prix de tous les biens et services échangés entre la période 0 et la période 1. Un indice de type Laspeyres utilise uniquement la distribution des biens qui effectuent des transactions pendant la période 0, de sorte que  $\log(I^Q) = E(\rho(1)|t=0) - E(\rho(0)|t=0)$ ; un indice de type Paasche utilise la distribution des biens qui transigent pendant la période 1, de sorte que  $\log(I^Q) = E(\rho(1)|t=1) - E(\rho(0)|t=1)$ . Dans le langage d'une expérience, un indice de qualité constante de type Törnqvist est l'effet moyen du traitement, tandis qu'un indice de type Laspeyres est l'effet moyen du traitement sur les non traités, et un indice de type Paasche est l'effet moyen du traitement sur les traités.

<sup>30</sup>Revenant au récit d'une expérience, les marchandises qui se vendent au cours de la période 1 sont le groupe de traitement et les marchandises qui se vendent au cours de la période 0 sont le groupe témoin, l'indice des prix de transaction donnant la différence moyenne dans le résultat de l'expérience. Cela peut correspondre ou non à l'effet causal du traitement, selon la façon dont les groupes de traitement et de contrôle sont formés. Par exemple, supposons que le traitement visite un hôpital et que le résultat soit une mesure de la santé. Le simple fait de comparer les résultats pour la santé de ceux qui ont récemment visité un hôpital à ceux qui ne l'ont pas fait, tout en ignorant la façon dont ces groupes sont formés, conduirait à conclure à tort que la visite d'un hôpital rend les gens en mauvaise santé, comme en moyenne ceux qui ont récemment visité l'hôpital sont en moins bonne santé que ceux qui n'en ont pas. Cela est dû au fait que ceux qui visitent un hôpital ont de toute façon de moins bons résultats en matière de santé, d'où une comparaison pommes-orange.

Notez que cet indice n'est pas nécessairement basé sur des prix relatifs, car le même nombre de biens peut ne pas se vendre à chaque période. Si  $n(t)$  désigne l'ensemble des biens qui se vendent au cours de la période  $t$ , l'indice géométrique des prix de transaction est un rapport de moyennes géométriques,

$$\log(I^T) = \frac{\prod_{i \in n(1)} p_i(1)^{P_i|t=1}}{\prod_{i \in n(0)} p_i(0)^{P_i|t=0}},$$

où  $P_i|t$  est la probabilité conditionnelle d'observer un bon  $i$  pendant la période  $t$ . Il s'agit simplement d'une généralisation de l'indice géométrique des prix habituel lorsque différents biens se vendent au fil du temps.

### 14.3 Identification

Il n'est pas possible de construire un indice des prix de qualité constante avec seulement la connaissance des prix de transaction et quand les marchandises se vendent. Tout ce que l'on peut savoir, c'est l'indice des prix de transaction, qui n'est pas nécessairement en accord avec l'indice de qualité constante. Progresser vers l'identification d'un indice des prix de qualité constante avec des informations sur les prix de transaction nécessite de faire des hypothèses.

L'hypothèse la plus simple qui identifie l'indice des prix de qualité constante avec l'indice des prix de transaction est l'hypothèse qu'il n'y a pas de différence systématique entre les marchandises qui transigent à différents moments, de sorte que le mouvement des prix potentiels est le même que le évolution des prix de transaction observables. Cela revient à supposer que les prix potentiels sont (statistiquement) indépendants du moment où un bien se vend réellement,  $\{p(1), p(0)\} \perp t$ , donc la distribution des prix potentiels des biens qui vendre au cours de la période 0 est la même que la distribution des prix potentiels pour les biens qui se vendent au cours de la période 1.

Si cette hypothèse d'indépendance est vérifiée, alors  $E(\rho(t)|t) = E(\rho(t))$  pour  $t = 0, 1$  et ainsi, tant que  $0 < P(t=1) < 1$ , afin qu'il y ait des biens qui se vendent dans les deux périodes,

$$\begin{aligned} \log(I^Q) &= E(\rho(1)) - E(\rho(0)) \\ &= E(\rho(1)|t=1) - E(\rho(0)|t=0) \\ &= \log(I^T). \end{aligned}$$

Autrement dit, l'indice de qualité constante est identique à l'indice des prix de transaction observable.<sup>31</sup>

Dans la pratique, supposer que les prix potentiels sont indépendants du moment où un bien se vend est une hypothèse assez forte. À moins qu'il n'y ait des raisons de croire que les marchandises se vendent au hasard, il s'agit probablement d'une hypothèse inappropriée. Néanmoins, il sert d'hypothèse fondamentale qui justifie l'utilisation de méthodes plus sophistiquées pour construire un indice des prix de qualité constante.

### 14.4 Exemple

Il vaut la peine de passer par un exemple pour corriger les idées discutées jusqu'à présent. Supposons que l'objectif soit de construire un indice de Jevons de qualité constante pour les aliments pour chats. Les

<sup>31</sup>L'un des principaux éléments de la littérature sur l'évaluation des programmes est qu'il est souvent plus facile d'obtenir un effet de traitement pour une sous-population que pour l'ensemble de la population. Par exemple, l'effet moyen du traitement sur les personnes traitées est identifié dans des conditions plus faibles que l'effet moyen du traitement sur l'ensemble de la population. Étant donné qu'un indice des prix de qualité constante n'est qu'un effet de traitement moyen, il convient de se demander si l'indépendance peut être assouplie tout en offrant un indice de qualité constante. Dans certains cas, la réponse est oui, si un indice des prix s'applique à une sous-population de biens (par exemple, la population de biens qui se vendent réellement au cours de la période 1, qui est l'effet moyen du traitement sur les produits traités). Cela signifie que les outils de la littérature sur l'évaluation des programmes — tels que les variables instrumentales pour les effets moyens du traitement local ou la discontinuité de régression — peuvent être utilisés pour calculer un indice des prix de qualité constante, bien qu'il s'agisse encore d'un domaine relativement nouveau.

aliments pour chats peuvent se présenter en trois variétés: poulet ( $c$ ), foie ( $l$ ) ou saumon ( $s$ ). Les aliments pour chats au poulet et au saumon se vendent au cours de la période 1, et seuls les aliments pour chats au foie se vendent au cours de la période 0. Dans ce contexte, l'indice des prix de qualité constante est

$$IQ = \frac{(p_c(1)p_l(1)p_s(1))^{1/3}}{(p_c(0)p_l(0)p_s(0))^{1/3}}.$$

Cet indice compare les prix potentiels au fil du temps pour les trois variétés d'aliments pour chats, et donc la qualité des aliments pour chats est maintenue fixe dans le temps.

En pratique, tout ce qui peut être calculé avec des informations sur les prix de transaction est

$$IT = \frac{(p_c(1)p_s(1))^{1/2}}{(p_l(0)p_l(0))^{1/2}}.$$

Cet indice compare le prix moyen des aliments pour chats de la période 1 au prix moyen des aliments pour chats de la période 0.

Si les prix potentiels sont indépendants du temps, de sorte qu'il n'y a pas de différences systématiques entre le foie, le poulet et la nourriture pour chat de saumon qui affectent le prix, alors, au moins dans cet exemple,  $p_c(1) = p_l(1) = p_s(1)$  et  $p_c(0) = p_l(0) = p_s(0)$ , de sorte que

$$IQ = \frac{(p_c(1)p_l(1)p_s(1))^{1/3}}{(p_c(0)p_l(0)p_s(0))^{1/3}} = \frac{(p_c(1)p_s(1))^{1/2}}{(p_l(0)p_l(0))^{1/2}} = IT;$$

l'indice des prix de transaction est égal à l'indice de qualité constante et donne donc le mouvement pur des prix des aliments pour chats. Même si les aliments pour chats au poulet et au saumon ne se vendent pas pendant la période 0, l'indépendance garantit que le prix des aliments pour chats au foie sert de bonne comparaison pour ce que les aliments pour chats au poulet et au saumon auraient vendus pendant la période 0.

Si, au contraire, les aliments pour foie et poulet pour chats sont systématiquement moins chers que les aliments pour chats à saumon, par exemple parce qu'ils sont de qualité inférieure, alors  $p_c(1) = p_l(1) = p_s(1)$  et  $p_c(0) = p_l(0) < p_s(0)$ , de sorte que  $p_s(1)/p_s(0) < p_s(1)/p_l(0)$ . Donc

$$IQ = \frac{(p_c(1)p_l(1)p_s(1))^{1/3}}{(p_c(0)p_l(0)p_s(0))^{1/3}} < \frac{(p_c(1)p_s(1))^{1/2}}{(p_l(0)p_l(0))^{1/2}} = IT;$$

l'indice des prix de transaction montre une augmentation plus importante des prix au fil du temps (ou une diminution plus faible) parce que la nourriture pour chat de saumon se serait vendue plus que la nourriture pour chat de foie au cours de la période 0. Comparaison du prix de la nourriture pour chat de saumon au prix de la nourriture pour chat de foie confond un changement de prix au fil du temps avec un changement de la qualité de ce qui se vend au fil du temps. L'indépendance entre les prix potentiels et le temps exclut ce type de différences systématiques entre les produits qui se vendent réellement à différentes périodes.

## 15 Indices de prix stratifiés

Le point de départ d'un indice de prix stratifié est une répartition des biens selon un ensemble de caractéristiques observables qui déterminent le prix, de sorte que chaque bien est placé dans une strate basée

sur ces caractéristiques. Par exemple, un schéma de stratification simple consiste à séparer les biens en fonction de la géographie, de sorte que les biens soient regroupés par lieu de vente. Mais les biens peuvent être partitionnés selon des règles plus complexes - toute combinaison de caractéristiques peut, en principe, être utilisée pour stratifier les biens en groupes distincts. Un indice de prix stratifié est alors simplement un ensemble de sous-indices, un pour chaque strate, ainsi qu'un ensemble de pondérations pour agréger ces sous-indices en un indice de prix global.

L'utilité de la stratification est qu'elle fournit un moyen de justifier une hypothèse d'indépendance qui donne à chaque sous-indice une interprétation de qualité constante. Plutôt que d'exiger une indépendance totale entre les prix potentiels et le moment où les marchandises se vendent, l'indépendance ne doit être subordonnée qu'aux caractéristiques utilisées pour la stratification. C'est-à-dire que l'indépendance ne doit se maintenir que pour chaque strate individuellement, plutôt que pour toutes les strates simultanément. Avec l'indépendance, la stratification contrôle de manière non paramétrique les caractéristiques qui peuvent confondre les variations de prix dans le temps avec les modifications de la composition des marchandises vendues à différents moments. Ces sous-indices stratifiés peuvent ensuite être agrégés pour obtenir un indice global de qualité constante (c'est le travail du poids) car l'indice de chaque strate a une interprétation de qualité constante.

Une approche stratifiée pour construire un indice des prix de qualité constante n'est qu'une application directe des résultats de la section précédente. Il est utile de commencer par des indices stratifiés car ces indices nécessitent le moins d'hypothèses et, dans un certain sens, les autres approches pour construire un indice de qualité constante tentent d'imiter un indice stratifié.

Manuel RPPI: Chapitre 4.

Manuel PPI: Chapitre 7, sections A, C.

## 15.1 Indépendance conditionnelle

La configuration de l'indice stratifié suit de près la configuration de l'indice général de qualité constante dans la section précédente, sauf que les caractéristiques d'un bien doivent être explicitement modélisées. Pour ce faire, supposons que  $X$  soit un vecteur (aléatoire) de caractéristiques observables sur lesquelles les biens peuvent être stratifiés. Dans le contexte du logement, par exemple,  $X$  peut inclure la superficie en pieds carrés, le nombre de chambres et l'âge de la maison. Chaque strate correspond à une réalisation différente de  $x$  pour  $X$  (par exemple, maisons de 1500 à 2500 pieds carrés, 3 chambres à coucher et âgées de 10 à 25 ans).

L'indice géométrique des prix de transaction entre la période 0 et la période 1 pour la strate  $x$  compare les prix de transaction moyens dans le temps pour les biens de la strate  $x$ , et est donné par

$$\log(I_x^T) = E(\rho|X = x, t = 1) - E(\rho|X = x, t = 0).$$

Si le prix potentiel est indépendant du temps, en fonction des caractéristiques utilisées pour stratifier les biens, les sous-indices de chaque strate, calculés avec les prix de transaction, ont une interprétation de qualité constante. Formellement, ceci est l'hypothèse d'indépendance conditionnelle  $\{p(1), p(0)\} \perp t|X$ , l'analogue intra-strate de l'hypothèse d'indépendance de la section précédente. Pour chaque strate, si un bien se vend pendant la période 0 ou la période 1 est essentiellement dû au hasard, il n'y a donc pas de différences systématiques entre les biens qui se vendent pendant la période 0 et la période 1 qui influencent les prix au sein d'une strate. Autrement dit, la seule raison pour laquelle les prix potentiels pourraient changer avec le temps est due à un changement dans la composition des caractéristiques  $X$  — une fois que ces caractéristiques sont maintenues fixes, tout changement des prix observés dans le temps doit être un pur changement de prix.

Formellement, avec une indépendance conditionnelle, il faut que  $E(\rho(t)|X, t) = E(\rho(t)|X)$  pour  $t = 0, 1$ , et donc

$$\begin{aligned}\log(I_x^T) &= E(\rho(1)|X = x, t = 1) - E(\rho(0)|X = x, t = 0) \\ &= E(\rho(1)|X = x) - E(\rho(0)|X = x).\end{aligned}$$

L'indépendance conditionnelle donne un indice de qualité constante pour chaque strate qui coïncide avec l'indice des prix de transaction pour cette strate.

Dans le cas extrême où les biens sont regroupés en paires dans le temps, de sorte que chaque strate contient deux biens, l'indice stratifié n'est qu'un indice de modèle apparié pur. Pour voir cela, énumérez la population de paires par  $i = 1, \dots, n_p$  de sorte que

$$I^Q = \prod_{i=1}^{n_p} \left( \frac{p_{i1}}{p_{i0}} \right)^{\omega_i},$$

où  $P(X = x) = \omega_i$  et  $p_{it}$  est le prix du bien en paire  $i$  qui se vend pendant la période  $t = 0, 1$ . L'indépendance conditionnelle est valable pour un indice de modèle apparié si les biens de chaque paire sont suffisamment similaires dans le temps pour que le prix de transaction du bien qui se vend au cours de la période 0 donne une bonne référence pour ce que le bien qui vend au cours de la période 1 aurait vendu au cours de la période 0. Cela montre également que les formules d'index standard sont des cas particuliers de l'indice stratifié et supposent implicitement une indépendance conditionnelle.

Il convient de souligner que l'indépendance conditionnelle n'est pas une hypothèse vérifiable. C'est intrinsèquement une hypothèse économique sur la façon dont les marchandises se vendent au fil du temps. Mais c'est ce qui permet d'identifier un indice de qualité constante à partir d'un indice de prix de transaction.

## 15.2 Chevauchement

Pour agréger les sous-indices de chaque strate en un indice de prix global, il faut que  $0 < P(t = 1|X = x) < 1$  — cela signifie que, pour chaque strate, certains biens doivent vendre à chaque période. Il s'agit de la condition de chevauchement (ou de soutien commun) qui garantit qu'un indice des prix peut être construit pour chaque strate. Contrairement à l'indépendance conditionnelle, le chevauchement est vérifiable.

Lorsque la condition de chevauchement tient, avec l'indépendance conditionnelle,

$$\begin{aligned}\log(I^Q) &= E(\rho(1)) - E(\rho(0)) \\ &= E[E(\rho(1)|X) - E(\rho(0)|X)] \\ &= E[E(\rho(1)|X, t = 1) - E(\rho(0)|X, t = 0)] \\ &= E[E(\rho|X, t = 1) - E(\rho|X, t = 0)],\end{aligned}$$

pour que

$$I^Q = \prod_x (I_x^T)^{P(X=x)}.$$

L'indice de qualité constante est simplement un produit pondéré de chaque sous-indice, calculé avec les prix de transaction, où les pondérations correspondent à la probabilité qu'un bien appartienne à une strate particulière. Toutes ces informations sont observables, et donc l'indice de qualité constante est identifié.

### 15.3 Problèmes de stratification

Un avantage évident de l'approche stratifiée pour la construction d'un indice de qualité constante est qu'elle ne s'appuie que sur deux hypothèses pour fournir un tel indice - indépendance conditionnelle et chevauchement. Ces hypothèses ont également l'avantage d'être relativement transparentes. L'indépendance conditionnelle et le chevauchement sont des conditions intuitivement simples, bien que l'indépendance conditionnelle ne soit pas une hypothèse vérifiable.

Le défi avec une approche stratifiée est de trouver le bon équilibre entre ces deux hypothèses. Si la stratification est trop granulaire, alors la condition de chevauchement peut échouer — un bien avec une combinaison particulière de caractéristiques peut ne pas se vendre au cours d'une période donnée, de sorte que le sous-indice de cette strate n'est pas défini. Quand vient le temps d'agréger chaque sous-indice, il ne sera pas possible de l'agréger sur la population des strates, et donc l'indice global de qualité constante n'est pas identifié à partir des prix de transaction.<sup>32</sup> Ceci est problématique, car une partition plus granulaire des caractéristiques peut donner plus de crédibilité à l'hypothèse d'indépendance conditionnelle. C'est plus facile à voir lorsque les marchandises sont regroupées en paires (la partition la plus granulaire), de sorte que l'indice des prix de transaction est un indice de modèle apparié. Dans ce cas, le chevauchement échoue lorsqu'un bien ne se vend ni dans la période 0 ni dans la période 1, et donc le prix relatif d'une paire de produits ne peut pas être construit.<sup>33</sup>

Un problème connexe avec la stratification est que, si la stratification se produit pour de nombreuses caractéristiques d'un bien,  $X$  peut être de grande dimension. Dans ce cas, la taille des cellules pour chaque strate peut être très petite, et les indices spécifiques à la strate peuvent être très sensibles à l'ajout ou à la suppression de transactions pour un bien. Le chevauchement et la dimensionnalité limitent tous deux la capacité de créer des groupes de biens homogènes pour lesquels il est plus facile de motiver l'indépendance conditionnelle.

### 15.4 Exemple avec R: Indice stratifié

Le calcul d'un indice stratifié est simple en pratique. Une bonne façon de le faire est d'utiliser une simple régression linéaire.

Considérez les données suivantes avec des transactions sur deux périodes pour deux groupes de marchandises.

```
# Faire quelques données
df <- data.frame(period = c(0, 0, 0, 1, 1, 1, 1, 1),
                  group = letters[1:2],
                  price = 1:8)
df
```

##	period	group	price
## 1	0	a	1
## 2	0	b	2
## 3	0	a	3
## 4	1	b	4
## 5	1	a	5
## 6	1	b	6
## 7	1	a	7
## 8	1	b	8

<sup>32</sup>Une façon d'atténuer ce problème est de construire un indice des prix pour une sous-population de biens. Par exemple, un indice de type Paasche requiert uniquement des transactions dans la période 0 et la période 1 pour les strates dans lesquelles un bien se vend dans la période 1,  $P(t = 1|X = x) < 1$ . Cela a également l'avantage d'affaiblir l'hypothèse d'indépendance conditionnelle de sorte que seuls les prix potentiels de la période 0 doivent être conditionnellement indépendants du temps,  $p(0) \perp t|X$ .

<sup>33</sup>Une façon intéressante de réduire le chevauchement dans un indice de modèle apparié consiste à calculer l'indice sur plus de deux périodes. Voir Kirby-McGregor and Martin (2019) pour un exemple.

Les indices pour chaque groupe peuvent être calculés avec une seule régression linéaire, puis agrégés avec une moyenne.

```
# Apportez dans la bibliothèque gpindex
library(gpindex)

# Calculer les indices au niveau des strates avec une régression linéaire
mdl <- lm(log(price) ~ group + group:period - 1, df)

# Transformer les coefficients de régression en indices
index <- exp(coef(mdl)[-seq_len(length(unique((mdl$model$group))))]))

# Agrégat, en supposant que les deux strates ont un poids égal
mean_geometric(index) * 100

## [1] 313.886
```

## 16 Indices de prix hédoniques

L'approche stratifiée pour construire un indice de qualité constante répartit les biens selon des caractéristiques observables pour aider à justifier une hypothèse d'indépendance conditionnelle entre les prix potentiels et le moment où un bien se vend. Il en résulte une collection d'indices de qualité constante pour chaque strate qui peuvent être identifiés par les prix de transaction intra-strate. L'astuce avec l'approche stratifiée consiste à trouver une stratification suffisamment grossière pour justifier l'indépendance conditionnelle tout en garantissant la condition de chevauchement. Sinon, si le chevauchement échoue, l'indice global de qualité constante ne peut pas être construit à partir des indices de prix de transaction spécifiques à la strate car il n'est pas possible de construire un index de prix de transaction pour chaque strate.

Une motivation pour l'hédonique est de spécifier un modèle pour les prix à extrapoler à travers les strates lorsque le chevauchement échoue. Il peut y avoir une liste de caractéristiques d'un bien qui sont convenues de saisir de manière appropriée ce qui confond les changements de prix au fil du temps avec les changements de la qualité sous-jacente d'un bien, mais il n'est pas possible de stratifier les prix le long de cette liste de caractéristiques en raison d'un échec de chevauchement. L'idée derrière l'hédonique est de modéliser  $E(\rho|X, t)$  avec un modèle paramétrique,  $h(X, t)$ , qui ne nécessite pas de chevauchement (ou nécessite une forme de chevauchement plus faible). Avec l'indépendance conditionnelle, l'indice des prix de qualité constante est alors simplement

$$\log(I^Q) = E[E(\rho|X, t = 1) - E(\rho|X, t = 0)] = E[h(X, 1) - h(X, 0)].$$

Chaque indice géométrique hédonique a cette forme, et peut être considéré comme utilisant le modèle  $h$  pour construire des prix relatifs modélisés qui comparent les marchandises avec les mêmes caractéristiques au fil du temps, puis les agréger avec un indice géométrique.

Il convient de noter que le modèle hédonique reçoit généralement une interprétation structurelle (ILO et al. 2013, chapitre 5), de sorte qu'il capture la contribution structurelle de chaque caractéristique au prix. C'est en effet de là que vient le nom d'hédonique. Le problème de donner au modèle hédonique  $h$  une interprétation structurelle est qu'il requiert une quantité considérable de connaissances préalables, essentiellement pour comprendre comment les caractéristiques individuelles d'un produit déterminent son prix. Cependant, cela est totalement inutile une fois qu'une hypothèse d'indépendance conditionnelle est établie — le modèle hédonique est simplement un modèle pour la relation entre les prix de transaction moyens et les caractéristiques et le temps. Par conséquent, le label "hédonique" n'est pas très utile, et est potentiellement trompeur. Cependant, dans l'intérêt de conserver une documentation plus large, l'étiquette sera maintenue. Ce qui est important à garder à l'esprit, c'est que le modèle hédonique est simplement un modèle pour savoir comment le prix de transaction moyen se rapporte aux caractéristiques d'un bien et quand il se vend.

Manuel RPPI: chapitres 5 à 6.

Manuel PPI: Chapitre 21, sections A, C, D.

## 16.1 Indices hédoniques linéaires

La forme canonique du modèle hédonique  $h$  est une fonction linéaire des caractéristiques d'un bien, de sorte que  $h(X, 0) = \alpha_0 + X\beta_0$  et  $h(X, 1) = \alpha_1 + X\beta_1$ , et donc

$$E(\rho|X, t) = \alpha_0 + t(\alpha_1 - \alpha_0) + X\beta_0 + t \cdot X(\beta_1 - \beta_0).$$

Avec une forme linéaire pour le modèle hédonique, les prix de transaction sont exprimés comme une régression linéaire sur les caractéristiques d'un bien — les paramètres dans un modèle hédonique ne sont que des coefficients de régression de population. Le vecteur de paramètre  $\beta_t$  est généralement interprété comme un vecteur de prix implicites pour les caractéristiques non marchandes d'un bien. Cette interprétation découle d'une vue structurelle d'un modèle hédonique — le but ici est simplement de spécifier un modèle paramétrique pour la fonction de prix conditionnelle,  $E(\rho|X, t)$ , donc  $\beta_t$  aucune interprétation particulière n'est nécessaire.<sup>34</sup>

Avec un modèle hédonique linéaire, et une indépendance conditionnelle entre les prix potentiels et le temps,

$$\log(I^Q) = E(h(X, 1) - h(X, 0)) = \alpha_1 - \alpha_0 + E(X)(\beta_1 - \beta_0).$$

La spécification d'un modèle pour  $E(\rho|X, t)$  a pour effet qu'un indice de prix de qualité constante a maintenant une forme paramétrique simple qui peut être utilisée pour assouplir la condition de chevauchement.<sup>35</sup> Rien empêche d'utiliser un modèle non linéaire plus complexe, mais en application, le modèle hédonique est généralement linéaire.<sup>36</sup>

Ce type d'indice hédonique est parfois appelé modèle d'imputation hédonique, car il s'agit de l'indice géométrique formé en prenant le rapport des prix moyens prédits à partir des régressions linéaires du prix (log) sur les caractéristiques à deux moments. Comme mentionné dans la section précédente, cependant, tous les indices de prix hédoniques ont cette forme, que le prix moyen soit modélisé en fonction linéaire des caractéristiques ou non. Par conséquent, le label “imputation hédonique” n'est pas très utile et conduit à une variété d'indices de prix hédoniques “différents” qui sont vraiment tous la même chose.

Un avantage de spécifier un modèle linéaire pour  $E(\rho|X, t)$  est que l'index résultant a une forme très intuitive:<sup>37</sup>

$$\begin{aligned} \log(I^Q) = & \underbrace{E(\rho|t=1) - E(\rho|t=0)}_{\text{indice des prix de transaction}} \\ & + \underbrace{[E(X|t=0) - E(X|t=1)][\beta_0 P(t=1) + \beta_1 P(t=0)]}_{\text{terme de correction}}. \end{aligned}$$

<sup>34</sup>Il y a quelques problèmes théoriques avec la spécification d'un modèle hédonique linéaire — voir Hausman (2003) pour un exemple quand un IPC est censé mesurer les changements du coût de la vie, et Rosen (1974), généralement considéré comme le fondement conceptuel de l'approche hédonique.

<sup>35</sup>Le chevauchement n'est pas complètement disparu — les marchandises doivent encore être vendues au cours des deux périodes et, à chaque instant, aucune des caractéristiques de  $X$  ne peut être une combinaison linéaire des autres — mais n'est pas aussi onéreuse qu'avec un indice stratifié.

<sup>36</sup>Il est facile de montrer qu'un ensemble de coefficients de régression linéaire pour la régression de  $X$  sur  $y$  résout le problème suivant:  $\min_{b \in \mathbb{R}^k} E[(E(y|X) - Xb)^2]$ . Autrement dit, une régression linéaire donne l'approximation linéaire de l'erreur quadratique moyenne minimale à la fonction d'attente conditionnelle. Par conséquent, même si le prix moyen est une fonction non linéaire des caractéristiques, le modèle hédonique linéaire peut toujours donner une bonne approximation.

<sup>37</sup>Pour le montrer directement, notez que  $E(\rho|t) = E(E(\rho|X, t)) = E(\alpha_t + X\beta_t|t) = \alpha_t + E(X|t)\beta_t$ .



À condition que l'indépendance conditionnelle soit respectée et que le prix de transaction moyen soit une fonction linéaire des caractéristiques des marchandises vendues, l'indice de qualité constante peut être décomposé en un indice géométrique des prix de transaction et un terme de correction qui capture le changement dans la composition de caractéristiques du produit au fil du temps. Le terme de correction prend la variation moyenne des caractéristiques au fil du temps et l'utilise pour ajuster l'indice des prix de transaction — le terme  $\beta_0 P(t=1) + \beta_1 P(t=0)$  détermine si l'augmentation de la présence d'une caractéristique augmente en moyenne les prix ou non. Par exemple, si les produits de la période 1 ont, en moyenne, plus de caractéristiques positivement associées au prix, le terme de correction exercera une influence négative sur l'indice des prix de transaction pour arriver à un indice de qualité constante. En effet, l'indice des prix de transaction affichera une augmentation des prix au fil du temps en partie parce que les marchandises vendues au cours de la période 1 sont de meilleure qualité que celles vendues au cours de la période 0, et se seraient vendues plus pendant la période 0 que les marchandises qui vendus au cours de la période 0. Par conséquent, l'indice des prix de transaction surestime la variation pure des prix, d'où le terme de correction négatif.

Le calcul d'un indice de prix hédonique linéaire est facile, car

$$\begin{aligned} E(\rho|X, t) &= \alpha_0 + t(\alpha_1 - \alpha_0) + X\beta_0 + t \cdot X(\beta_1 - \beta_0) \\ &= \alpha_0 + t(\alpha_1 - \alpha_0 + E(X)(\beta_1 - \beta_0)) + X\beta_0 + t(X - E(X))(\beta_1 - \beta_0) \\ &= \alpha_0 + t \log(I^Q) + X\beta_0 + t(X - E(X))(\beta_1 - \beta_0). \end{aligned}$$

L'indice hédonique n'est alors que le coefficient d'une variable temps fictive dans une régression linéaire, et est donc extrêmement facile à calculer.

Il convient de conclure cette section en notant que, dans presque tous les cas, il est supposé que le prix de transaction moyen est une fonction linéaire des caractéristiques. Il existe cependant un cas intéressant où cette hypothèse est toujours correcte. Si les marchandises sont partitionnées par leurs caractéristiques, de sorte que chaque combinaison de  $x$  pour  $X$  appartient à son propre groupe et obtient son propre paramètre dans le modèle hédonique à chaque instant, alors  $E(\rho|X, t)$  est nécessairement linéaire. Dans ce cas, l'indice de prix hédonique est simplement un indice de prix stratifié. De cette façon, l'indice de prix hédonique linéaire est un type d'indice de prix plus général que l'indice stratifié et trouve sa valeur lorsque l'indice stratifié ne peut pas être calculé en raison d'un échec de chevauchement.

## 16.2 L'approche temporelle

L'indice hédonique de la section précédente peut être simplifié pour rendre le calcul de l'indice plus facile avec une approche appelée approche temporelle.<sup>38</sup> L'approche factice temporelle prend simplement le modèle hédonique linéaire de la section précédente et le restreint de sorte que  $\beta_0 = \beta_1 = \beta$ . Cela ne change rien de fondamental à propos de l'index, sauf qu'il nécessite désormais moins de paramètres à calculer, comme

$$E(\rho|X, t) = \alpha_0 + t \log(I^Q) + X\beta.$$

Une caractéristique intéressante de l'indice temporel est que, comme l'indice hédonique linéaire général, il s'agit d'un type d'index stratifié lorsque les biens sont répartis selon leurs caractéristiques observables. Pour voir cela, notez que, avec l'hypothèse que  $\beta_0 = \beta_1 = \beta$ , le modèle hédonique devient  $h(X, t) = \alpha_0 + t(\alpha_1 - \alpha_0) + X\beta$ . Lorsque les biens sont stratifiés en fonction de  $X$ , alors  $h$  est saturé en  $X$  de telle sorte qu'il existe un  $\beta_x$  unique pour chaque combinaison de  $x$  pour  $X$  — chaque strate obtient son propre paramètre dans la régression. Dans ce cas, le résultat dans Angrist and Pischke (2009Section 3.3.1) s'applique, de sorte que

---

<sup>38</sup>Encore une fois, c'est un mauvais nom, car l'approche «d'imputation hédonique» récupère l'indice comme le coefficient sur une variable factice temporelle.

$$\log(I^Q) = \sum_x [E(\rho|X = x, t = 1) - E(\rho|X = x, t = 0)]\omega_x,$$

où

$$\omega_x = \frac{\text{var}(t|X = x)P(X = x)}{\sum_x \text{var}(t|X = x)P(X = x)}.$$

Cet indice est de la même forme que l'indice géométrique stratifié dans la section précédente, sauf que les poids dépendent à la fois de la probabilité qu'un bien appartienne à une strate particulière et de la variance des dates de vente au sein d'une strate. Dans le cas spécial où chaque strate contient une paire de biens,  $\omega_x = P(X = x)$ , et l'indice factice temporel se réduit à l'indice de modèle apparié pur.

Le point à retenir ici est que lorsque les biens sont stratifiés en fonction de leurs caractéristiques, l'approche factice temporelle pour la construction d'un indice de qualité constante et l'approche stratifiée ne sont pas des approches fondamentalement différentes pour opérationnaliser une hypothèse d'indépendance conditionnelle. Les deux approches nécessitent la condition de chevauchement et des indices de prix de transaction intra-strates agrégés pour produire un indice de prix global. L'utilité de l'approche temporelle vient lorsque le chevauchement échoue — de sorte que les biens ne peuvent pas être stratifiés selon un ensemble particulier de caractéristiques — car l'approche temporelle fournit une correction paramétrique qui donne à l'indice des prix de transaction une constante - interprétation de qualité au détriment d'une hypothèse supplémentaire.

Dans la pratique, le modèle hédonique factice est plus populaire que l'approche d'imputation hédonique, en partie parce qu'il est plus simple, plus facile à calculer et moins gourmand en données. L'inconvénient de l'approche temporelle est qu'elle est plus restrictive et nécessite une hypothèse supplémentaire. Dans la plupart des cas, cependant, les deux approches donnent des résultats similaires.

### 16.3 Exemple avec R: Hédonique

Le calcul d'un indice de prix hédonique linéaire est très facile dans R.

```
# Apportez quelques données
df <- read.csv("csv/data.csv")
df
```

```
##      period price chicken liver salmon
## 1         0      2        1      0      0
## 2         0      4        1      0      0
## 3         0      3        1      0      0
## 4         0      4        1      0      0
## 5         0      5        1      0      0
## 6         0      1         0      1      0
## 7         0      3         0      1      0
## 8         0      2         0      1      0
## 9         0      1         0      1      0
## 10        0      1         0      1      0
## 11        0      5         0      0      1
## 12        0      7         0      0      1
## 13        0      8         0      0      1
## 14        0      4         0      0      1
## 15        1      5         1      0      0
## 16        1      3         1      0      0
## 17        1      2         0      1      0
```

```
## 18      1      2      0      1      0
## 19      1      1      0      1      0
## 20      1      3      0      1      0
## 21      1      1      0      1      0
## 22      1      1      0      1      0
## 23      1      9      0      0      1
## 24      1      8      0      0      1
## 25      1      5      0      0      1
```

Il s'agit d'un ensemble de données simple pour la nourriture pour chat, avec trois caractéristiques — est-ce de la nourriture pour chat de poulet, de la nourriture pour chat de foie ou de la nourriture pour chat de saumon? Pour évaluer l'impact d'un indice de prix hédonique, il est utile de calculer un indice de prix de transaction simple en regroupant les transactions pour les trois types d'aliments pour chats ensemble.

```
# Calculer l'indice de prix commun
pooled <- lm(log(price) ~ period, df)
exp(coef(pooled)[2]) * 100
```

```
##      period
## 93.86748
```

L'indice hédonique linéaire général est facile à calculer avec une simple régression linéaire. La seule astuce consiste à se souvenir de supprimer la moyenne de chaque caractéristique dans le terme d'interaction.

```
# Calculer l'indice d'imputation hédonique
hedonic_imputation <- lm(log(price) ~ period + liver + salmon +
  period:(I(liver - mean(liver)) + I(salmon - mean(salmon))),
  df)
exp(coef(hedonic_imputation)[2]) * 100
```

```
##      period
## 112.2817
```

C'est le même que le calcul en deux étapes qui est normalement effectué pour un indice d'imputation hédonique.

```
# Régression période 0
hi0 <- lm(log(price) ~ liver + salmon, df, subset = period == 0)

# Régression de la période 1
hi1 <- lm(log(price) ~ liver + salmon, df, subset = period == 1)

# Calculer l'indice à l'aide des prix prévus
exp(mean(predict(hi1, df) - predict(hi0, df))) * 100
```

```
## [1] 112.2817
```

C'est également la même chose que le calcul manuel de l'indice stratifié.

```
# Apportez dans la bibliothèque gpindex
library(gpindex)

# Poids de la nourriture pour chats par sa fréquence
weights <- colSums(df[-(1:2)])

# Calculer un indice pour chaque type de nourriture pour chat
strata_indices <-
  sapply(list(subset(df, chicken == 1),
    subset(df, liver == 1),
```

```
subset(df, salmon == 1)),
function(x) with(x, mean_geometric(price[period == 1]) / mean_geometric(price[period == 0]))
)

# Agrégat
mean_geometric(strata_indices, weights) * 100
```

```
## [1] 112.2817
```

Les résultats sont très similaires si un modèle temporel est utilisé à la place.

```
# Indice temporel factice
time_dummy <- lm(log(price) ~ period + liver + salmon, df)
exp(coef(time_dummy)[2]) * 100
```

```
## period
## 112.2246
```

L'ajout de pondérations pour les transactions dans chaque strate peut facilement être effectué avec une régression pondérée.

```
# Faites des poids
df$weights <- 1:25 / sum(1:25)

# Calculer l'indice d'imputation hédonique pondéré
hedonic_imputation <- lm(log(price) ~ period + liver + salmon +
                          period:(I(liver - weighted.mean(liver, weights)) +
                                     I(salmon - weighted.mean(salmon, weights))),
                          df, weights = weights)
exp(coef(hedonic_imputation)[2]) * 100
```

```
## period
## 111.1855
```

Cela donne la même réponse que le calcul en deux étapes.

```
# Régression période 0
hi0 <- lm(log(price) ~ liver + salmon, df, subset = period == 0, weights = weights)

# Régression de la période 1
hi1 <- lm(log(price) ~ liver + salmon, df, subset = period == 1, weights = weights)

# Calculer l'indice à l'aide de prix prédits pondérés
exp(weighted.mean(predict(hi1, df) - predict(hi0, df), df$weights)) * 100
```

```
## [1] 111.1855
```

## 17 Affectation 3

Les réponses à ces questions proviennent à la fois du contenu du cours et des lectures. Chaque question vaut un point, pour un total de 20 points. La réussite de ce module nécessite au moins 65% (13 sur 20 correct). Envoyez vos réponses par e-mail à l'instructeurs du cours (Steve Martin ou Shaoxiong Wang ou Anastasiia Nosach) lorsque vous avez terminé.

**Question 1** Vrai ou faux: l'indice hédonique factice temporel est un estimateur biaisé de l'indice géométrique de qualité constante.

**Question 2** Vrai ou faux: un indice de prix hédonique ne nécessite aucune hypothèse pour être un indice de qualité constante une fois que les caractéristiques des marchandises vendues sont connues.

**Question 3** Un indice hédonique linéaire est égal à l'indice de prix Jevons dans quelles circonstances.

- a) Lorsque les caractéristiques moyennes des biens vendus au cours de la période 1 sont égales aux caractéristiques moyennes des biens vendus au cours de la période 0.
- b) Lorsque les conditions de chevauchement sont remplies.
- c) Lorsque les paramètres du modèle hédonique ne changent pas avec le temps.
- d) a et c.
- e) Aucune de ces réponses.

**Question 4** Quel ajustement de qualité nécessite le plus grand nombre d'hypothèses?

- a) Stratification.
- b) Hédonique utilisant l'approche des caractéristiques.
- c) Hédonique utilisant l'approche temps-factice.
- d) Hédonique utilisant l'approche de la double imputation.

**Question 5** Vrai ou faux: un indice de vente répétée peut être motivé / dérivé comme un type d'indice hédonique.

**Question 6** La taille du contenant pour la nourriture pour chat de saumon a diminué entre la période 0 et la période 1 dans les exemples avec R pour les indices de prix hédoniques. Si les autres types d'aliments pour chats sont restés les mêmes, que dire des indices hédoniques?

- a) Ils surestiment la baisse des prix entre la période 0 et la période 1.
- b) Ils sous-estiment l'augmentation des prix entre la période 0 et la période 1.
- c) Ils surestiment la hausse des prix entre la période 0 et la période 1.
- d) Le modèle hédonique sous-estime l'augmentation et le modèle stratifié surévalue l'augmentation.
- e) Aucune de ces réponses.

Les quatre questions suivantes concernent le type d'ajustement de qualité suivant. Une stratégie courante pour construire un indice de qualité constante lorsque l'emballage d'un produit change consiste à diviser le prix par la taille de l'emballage (par exemple, le prix par litre de savon à vaisselle), de sorte que le prix relatif du produit  $i$  devienne  $p_{i1}/p_{i0} \times s_{i0}/s_{i1}$ , où  $s_{it}$  est la taille du paquet dans la période  $t$ . Un indice basé sur cette approche peut être calculé avec une régression linéaire, comme

$$E(\rho|i, s, t) = \alpha_i + t \log(I) + \log(s).$$

**Question 7** Vrai ou faux: l'ajustement du prix en fonction de la taille de l'emballage est un type d'index hédonique factice.

**Question 8** Quelle est l'hypothèse d'indépendance conditionnelle qui fait fonctionner cet ajustement de la qualité?

- a) La seule différence entre le même produit pendant la période 0 et la période 1 est la taille de l'emballage.
- b) Les produits de la période 1 sont un bon contrefactuel pour les produits de la période 0.
- c) Chaque produit se vend au cours des deux périodes.
- d) La seule différence entre les produits au fil du temps est due à la taille de l'emballage.

- e) L'indépendance conditionnelle n'est pas nécessaire.

**Question 9** Quelle hypothèse supplémentaire est faite pour que le simple fait de diviser le prix par la taille de l'emballage s'adapte aux changements de qualité, plutôt que d'avoir à estimer le modèle hédonique?

- a) Les paramètres du modèle hédonique ne changent pas avec le temps.
- b) L'erreur est que le modèle hédonique est homoscédastique.
- c) L'élasticité entre le prix et la taille de l'emballage est de un.
- d) Aucune hypothèse supplémentaire n'est faite.

**Question 10** Quelle est l'hypothèse de chevauchement requise pour calculer cet indice?

- a) Une transaction doit avoir lieu pour chaque combinaison de taille de produit-emballage dans la période 0 et la période 1.
- b) Chaque produit doit être vendu dans les deux périodes.
- c) Chaque produit doit être vendu dans les deux périodes et la taille de l'emballage doit changer pour au moins un produit.
- d) Aucune hypothèse de chevauchement n'est nécessaire.

**Question 11** Vrai ou faux: les indices de prix d'imputation temporelle et factice hédonique donnent toujours des valeurs différentes.

Les deux questions suivantes portent sur une variation de l'indice hédonique factice en fonction des médianes conditionnelles, de sorte que  $\text{med}(\rho|X, t) = \alpha_0 + t(\alpha_1 - \alpha_0) + X\beta$ . Avec l'indépendance conditionnelle,  $\alpha_1 - \alpha_0 = \log(\text{med}(p(1))/\text{med}(p(0)))$ .

**Question 12** Pourquoi l'indice de l'indice temporel médian est-il une mauvaise mesure pour un indice de qualité constante?

- a) Il compare ce que le produit médian de la période 1 se vendrait au cours de la période 1 à ce que le produit médian de la période 0 se vendrait au cours de la période 0, et ne compare donc pas les mêmes biens au fil du temps.
- b) Il s'agit d'un estimateur biaisé d'un indice de qualité constante.
- c) Il n'est pas possible de construire un estimateur du maximum de vraisemblance pour un indice basé sur les médianes.
- d) Il ne satisfait pas l'axiome de monotonie de l'approche axiomatique des indices de prix.
- e) Il n'y a rien de mal à l'index temporel médian.

**Question 13** Vrai ou faux: étant donné que la médiane est moins sensible aux valeurs aberrantes que la moyenne, l'indice médian temps factice trouve application dans les paramètres où il peut y avoir des observations de prix aberrantes.

Les quatre questions suivantes utilisent les données suivantes sur les transactions pour deux produits sur deux périodes.

Période	Produit	Prix 1	Prix 2	Prix 3
0	A	2	2	...
1	A	1	2	...
0	B	4	...	...
1	B	2	4	8

**Question 14** Quelle est la valeur de l'indice Jevons lorsque toutes les informations sur les prix sont regroupées pour les deux périodes?

- a) 100
- b) 98,3
- c) 104,7
- d) 106,4
- e) 89,8

**Question 15** Quelle est la valeur de l'indice stratifié, où chaque produit appartient à une strate distincte et toutes les strates ont un poids égal, calculé à l'aide d'un indice de Jevons?

- a) 104,7
- b) 98,3
- c) 84,1
- d) 110,6
- e) 101,3

**Question 16** Pourquoi l'indice stratifié montre-t-il une baisse des prix alors que l'indice groupé montre une augmentation des prix?

- a) L'indice groupé est un indice de qualité constante, tandis que l'indice stratifié est biaisé vers le bas.
- b) L'indice groupé utilise toutes les données sur les prix, contrairement à l'indice stratifié.
- c) Le produit B se vend à un prix plus élevé que le produit A, et relativement plus de produit B vendu au cours de la période 2, mais uniquement le produit Un prix modifié au cours de la période 2
- d) Aucune de ces réponses.

**Question 17** Vrai ou faux: si chaque produit ne change pas dans le temps, l'indice stratifié est un indice de qualité constante.

**Question 18** Vrai ou faux: l'indice de répétition des ventes est un type d'indice de modèle apparié.

**Question 19** Vrai ou faux: il est possible de mélanger la stratification avec l'hédonique pour détendre le chevauchement.

**Question 20** On peut montrer que

$$\underbrace{E(\rho|t=1) - E(\rho|t=0)}_{\text{indice des prix de transaction}} = \underbrace{E(\rho(1)|t=0) - E(\rho(0)|t=0)}_{\text{Indice de qualité constante semblable à Laspeyres}} + \underbrace{E(\rho(1)|t=1) - E(\rho(1)|t=0)}_{\text{biais de sélection}}.$$

Vrai ou faux: les prix potentiels de la période 0 étant-ils corrélés avec le temps, de sorte que les produits vendus au cours de la période 1 se seraient vendus moins cher au cours de la période 0 que les produits vendus à la période 0, affecter l'utilisation des prix de transaction pour identifier une constante de type Laspeyres -indice de qualité?

## Part IV

# Construire un indice des prix avec R

## 18 Programme pour Construire un indice des prix avec R

La mise en pratique de la théorie de l'indice des prix nécessite des outils de calcul. Un outil particulièrement utile pour construire et analyser des indices de prix est le langage de programmation R.

L'objectif de ce module est de fournir une expérience pratique de la construction et de l'analyse d'un indice des prix avec R. À la fin de ce module, une personne devrait:

1. Comprendre comment appliquer la théorie de l'indice des prix pour construire un indice des prix.
2. Savoir comparer les indices de prix de différentes sources.
3. Connaître la manière de construire et d'analyser un indice de prix avec R.

Ce module est utile pour les compilateurs d'indices de prix ayant une compréhension de base de la théorie et de la construction d'un indice de prix, et qui souhaitent mettre la théorie en pratique et acquérir une compréhension plus approfondie de R.

Ce module consiste en une tâche de grande envergure qui permet aux apprenants de construire et d'analyser un indice de prix en R sur une semaine. Le rythme du module est autogéré, mais une semaine entière devrait être consacrée à ce travail. Toutes les données pour cette affectation ont été générées au hasard.

Prérequis: une introduction aux indices de prix et une compréhension de base de la théorie de l'indice des prix R. et des indices de prix de qualité constante, ainsi qu'une compréhension intermédiaire de R, sont utiles.

L'évaluation de ce module est basée sur une tâche unique qui s'appuie sur les connaissances antérieures des apprenants sur la théorie de l'indice des prix. Le devoir se compose de 4 questions (plus une question bonus). La collaboration sur la mission est encouragée, mais chaque personne doit soumettre son propre travail unique. Les réponses à la mission, ainsi que le code de travail pour chaque question (à l'exclusion de la question bonus) valent 15%. L'instructeur doit être capable de comprendre et d'exécuter votre code; voir `exemple.R` pour un exemple. Chaque question correctement répondue vaut 10% supplémentaires. Répondre correctement à la question bonus vaut 20%. La réussite de ce module nécessite une note de 60% ou plus.

Veuillez envoyer un e-mail à l'instructeurs du cours (Steve Martin ou Shaoxiong Wang ou Anastasiia Nosach) si vous avez des questions ou si vous avez besoin d'aide concernant le matériel ou le devoir du cours.

## 19 Affectation 4

Les widgets sont un produit fortement réglementé dans chacune des 10 provinces, et le gouvernement fédéral limite les mouvements de widgets entre les provinces. Récemment, en réponse au tollé général suscité par le caractère inaccessible des widgets, le groupe canadien d'analyse économique réglementaire a proposé que le gouvernement fédéral autorise le commerce interprovincial de widgets pour améliorer l'accès aux marchés et réduire les prix. Avant d'envisager de concevoir une nouvelle politique, le gouvernement fédéral aimerait toutefois évaluer comment les prix des widgets ont augmenté ces dernières années. Un indice des prix est nécessaire pour les marchés de widgets provinciaux et pour le pays dans son ensemble.

Le groupe Analyse réglementaire de la réglementation canadienne produit un indice mensuel des prix à l'échelle du Canada pour les widgets (c.-à-d. Aucune ventilation provinciale), bien qu'il y ait peu d'information sur leurs sources de données ou leur méthodologie. La société privée Monsterweb, un courtier en widgets, établit également un indice des prix à l'échelle du Canada pour les widgets. Étant donné que les marchés provinciaux des widgets sont isolés les uns des autres, un indice des prix provincial est nécessaire pour évaluer comment les prix des widgets ont évolué au fil du temps et s'il est nécessaire que le gouvernement prenne des mesures politiques.



Les widgets sont de différents types, et une enquête a été utilisée pour collecter des devis mensuels pour une sélection de widgets représentatifs des types “A” à “J” de janvier 2018 à décembre 2019. Ces données sont stockées dans le système de prix globaux. Ces données d’enquête sont complétées par des données accessibles au public pour les transactions de widgets de type «K» qui n’ont pas pu être échantillonnées au cours de cette période, ainsi que la valeur des widgets traités dans chaque province en 2018 et 2019. Notez que les widgets de type «I» étaient un nouveau produit en 2019, et ces types de widgets ont remplacé les widgets de type “J” dans l’échantillon en 2019. En raison de la variation régionale des ventes de widgets, le même nombre de devis n’a pas pu être collecté pour tous les types de widgets dans toutes les provinces.

Toutes ces données sont résumées comme suit.

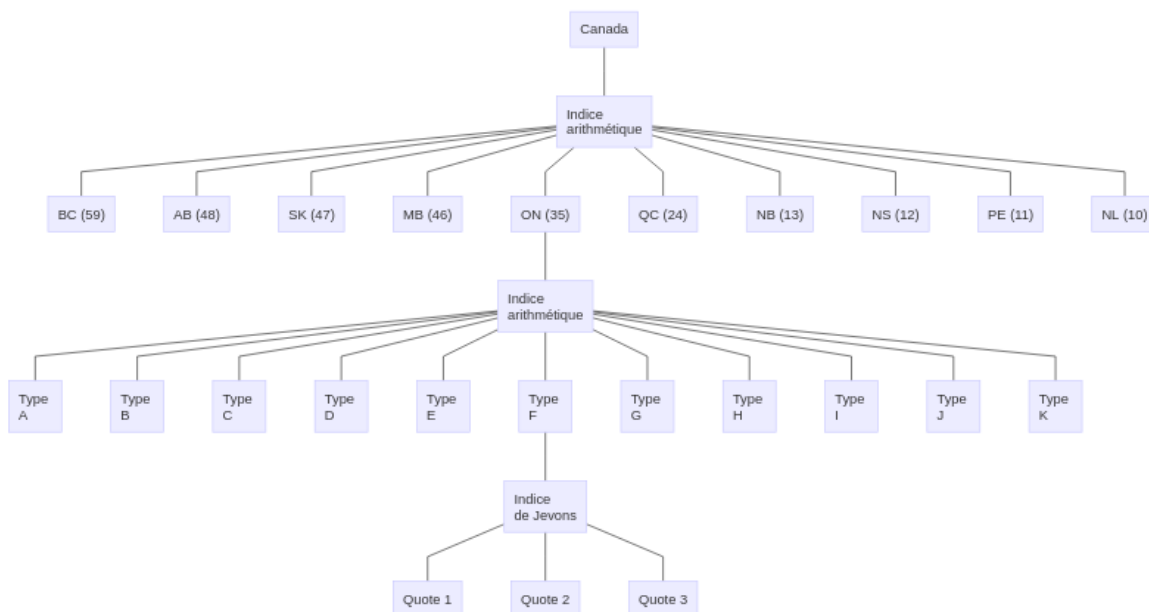
Nom de fichier	Données
dat_gps	Offres de prix mensuelles pour les types de widget “A” à “J” de janvier 2018 à décembre 2019 à partir du système de prix globaux.
dat_micro	Microdonnées accessibles au public pour les transactions quotidiennes des widgets de type “K” de janvier 2018 à décembre 2019.
poids	Valeur de la transaction de widget (en milliers de dollars) dans chaque province pour 2018 et 2019.
index_crea	L’indice des prix du groupe canadien d’analyse économique réglementaire.
index_mw	Indice des prix de Monsterweb.

Votre objectif est de construire un indice mensuel des prix des widgets, de janvier 2018 à décembre 2019, pour chacune des 10 provinces et le pays dans son ensemble, et de valider cet indice avec ceux produits par le Groupe canadien d’analyse économique réglementaire et Monsterweb. Vous devrez utiliser cet index pour répondre aux questions suivantes.

1. Quelle province a connu la plus forte augmentation des prix depuis le début de 2018? De combien les prix ont-ils changé?
2. Quelle province a connu la plus faible augmentation des prix depuis le début de 2018? De combien les prix ont-ils changé?
3. Quelle a été la plus importante variation des prix d’un mois à l’autre? Dans quelle province et quel mois cela s’est-il produit?
4. Tracez les trois indices au niveau du Canada sur un graphique. Quel indice montre la plus forte augmentation en pourcentage des prix des widgets nationaux depuis le premier trimestre 2019?
5. (Bonus) Calculez les coefficients de variation (CV) pour les indices provinciaux. En utilisant les règles de qualité habituelles pour la diffusion des données à Statistique Canada, quels indices provinciaux devraient être diffusés avec un avertissement (c.-à-d. Un CV supérieur à 33,3)?

## 19.1 Méthodologie

Les widgets vont être stratifiés par province et par type de widget pour produire des indices élémentaires au niveau de la province par type de widget. Ces indices élémentaires doivent être calculés avec un indice de Jevons. Les indices provinciaux et canadiens vont être des indices arithmétiques qui agrègent les indices élémentaires en utilisant des pondérations de part de valeur. Les pondérations des indices arithmétiques vont être mises à jour en 2019, avec janvier 2019 comme mois de liaison. La structure d’agrégation est résumée dans le graphique ci-dessous.



## 19.2 Instructions

Ces instructions décrivent un moyen simple de construire l'indice des prix des widgets, mais n'hésitez pas à utiliser une approche différente si c'est plus facile. Le but est de comprendre le processus, plutôt que de concevoir le système le plus efficace en termes de calcul.

Exécutez le code suivant pour récupérer tous les fichiers de données et les mettre dans votre environnement de travail. Notez que ces données ne sont pas parfaites. N'oubliez pas de vérifier vos calculs intermédiaires pour vous assurer que tout est comme prévu (par exemple, avec la fonction `summary`).

```
source('https://raw.githubusercontent.com/ppd-dpp/cours-indices-des-prix/master/scripts/get_data.R')
```

Le `dplyr`, `tidyr`, `gpindex` et les packages `calpr` seront utilisés dans ce module. Installez-les comme suit s'ils ne sont pas déjà installés. (Astuce: il vaut la peine de parcourir la documentation dans les packages `gpindex` et `calpr`.)

```
install.packages(c("dplyr", "tidyr", "gpindex", "devtools"))
devtools::install_github("ppd-dpp/calpr")
```

### 1. Faites les poids

À l'aide du fichier «poids», créez un ensemble de poids qui donne la part de valeur de chaque produit vendu dans chaque province chaque année, et un autre qui donne la part de valeur de tous les produits vendus dans chaque province chaque année. Ce premier ensemble de poids sera utilisé pour agréger les indices élémentaires pour obtenir 10 indices au niveau de la province, et le deuxième ensemble de poids sera utilisé pour agréger les indices au niveau de la province en un indice national. Il devrait y avoir 200 poids dans le premier fichier et 20 poids dans le second fichier.

```
library(dplyr)
library(tidyr)
library(gpindex)
```

```
library(calpr)

# Make provincial weights
weights_prov <- weights %>%
  group_by(year, province) %>% # group all sales values by reference year and province
  summarize(weight = sum(weight)) %>% # value of all sales in each province in each year
  mutate(share_prov = weights_scale(weight)) %>% # scale to sum to 1
  select(-weight) # drop the weight column as it won't be needed again

# Make product weights
weights_prod <- weights %>%
  # your code here %>%
  mutate(share_prod = weights_scale(weight)) %>%
  select(-weight)
```

## 2. Calculez la géométrie du produit K

En utilisant les microdonnées pour le produit K dans le fichier `dat_micro`, calculez la moyenne géométrique des prix dans chaque province pour chaque mois de chaque année de référence. Ce fichier devrait contenir 250 géométries.

```
dat_micro <- dat_micro %>%
  mutate(period = year_month(period)) %>% # turn transaction dates into month-by-year dates
  # your code here %>%
  summarize(geomean = mean_geometric(price)) # calculate the geometric average
```

## 3. Calculez la géométrie des produits A à J

En utilisant les données du système de prix globaux dans `dat_gps`, calculez la moyenne géométrique des prix pour chaque produit dans chaque province dans chaque mois de chaque année de référence, pour les produits A à J. Combinez ces données avec les données préparées à l'étape 2. Le fichier résultant devrait avoir 2 500 géométries.

```
dat_index <- dat_gps %>%
  # your code here %>%
  gather(key = "quote", value = "price", price1:price3) %>% # melt into a long dataset
  group_by(province, product, year, period) %>%
  # your code here %>%
  bind_rows(dat_micro) # append to dat_micro
```

## 4. Calculer les indices élémentaires période sur période

Pour chaque produit dans chaque province dans chaque année de référence, calculez le rapport des moyennes géométriques calculées à l'étape 3, en vous assurant d'avoir un prix relatif de 1 en janvier 2018 et janvier 2019 (janvier 2019 sera le mois de liaison lorsqu'il est le temps de chaîner l'index.)

```
dat_index <- dat_index %>%
  # your code here
```

## 5. Prix mettre à jour les poids des produits

Fusionnez les pondérations au niveau du produit par province de l'étape 1 dans l'ensemble de données créé à l'étape 4 et utilisez les indices élémentaires période sur période pour mettre à jour les pondérations.

```
dat_index <- dat_index %>%
  # your code here %>%
  mutate(share_prod_u = weights_update(cumprod(lag(rel, default = 1)), share_prod, scale = FALSE))
```

## 6. Calculez l'indice provincial

Calculez un indice arithmétique pour chaque province au cours de chaque année de référence, à l'aide des indices élémentaires d'une période à l'autre et des pondérations mises à jour des prix des étapes 4 et 5.

```
index_prov <- dat_index %>%
  group_by(period, year, province) %>%
  # your code here
```

## 7. Calculez l'indice au niveau du Canada

Fusionnez les pondérations provinciales à l'étape 1 avec les indices provinciaux de l'étape 6 pour calculer un indice canadien pour chaque année de référence.

```
index <- index_prov %>%
  # your code here
  group_by(period, year) %>%
  # your code here
  mutate(province = 0) %>% # province 0 is the Canada-level
  bind_rows(index_prov)
```

## 8. Chaîne des indices 2018 et 2019

Enchaînez les indices au niveau de la province de 2018 et 2019 et du niveau du Canada ensemble, en utilisant janvier 2019 comme mois de liaison, pour obtenir un indice de janvier 2018 à décembre 2019 avec janvier 2018 comme période de base (généralement, la période de base serait une année civile, mais il n'y a que 24 mois de données).

```
index_chain <- index %>%
  group_by(province) %>%
  # your code here
  group_by(period, province) %>%
  summarize(index = unique(index)) %>% # remove the duplicate rows for the link month
  ungroup()
```

Le résultat final devrait être un fichier avec 11 indices de prix, un pour chaque province et un indice national, de janvier 2018 à décembre 2019, avec janvier 2018 comme période de base (= 100).

La validation de l'indice au niveau du Canada nécessite de mettre l'indice du groupe canadien d'analyse économique réglementaire et l'indice Monsterweb sous une forme commune.

## 9. Trimestre par rapport aux indices mensuels

Transformez l'indice que vous venez de créer et l'indice du groupe Analyse réglementaire de la réglementation canadienne dans `index_crea` en un indice trimestriel, avec le 1er trimestre 2019 comme période de base, en prenant la moyenne des trois valeurs d'indice de chaque trimestre.

```
my_index <- index_chain %>%
  filter(province == '0') %>% # keep only the Canada-level index
  mutate(period = year_quarter(period)) %>%
  group_by(period) %>%
  summarize(index = mean(index)) %>%
  mutate(index = index/index[period == '2019-01-01'] * 100) # rebase to Jan 2019

index_crea <- index_crea %>%
  # your code here %>%
  # your code here %>%
  # your code here %>%
```

```
# your code here %>%
# your code here %>%

index_mw <- index_mw %>%
# your code here
```

#### 10. Ajoutez les trois indices ensemble

Rassemblez les indices de l'étape 9 dans un ensemble de données avec l'index de Monsterweb dans `index_mw`. Ils devraient tous être trimestriels avec le premier trimestre 2019 comme période de base (= 100).

## Références

- Angrist, J., and J-S. Pischke. 2009. *Mostly Harmless Econometrics*. Princeton University Press.
- Balk, B. M. 1995. "Axiomatic Price Index Theory: A Survey." *International Statistical Review/Revue Internationale de Statistique* 63 (1): 69–93.
- . 2008. *Price and Quantity Index Numbers*. Cambridge University Press.
- Balk, B. M., and W. E. Diewert. 2001. "A Characterization of the törnqvist Price Index." *Economics Letters* 72 (3): 279–81.
- Bullen, P. S. 2003. *Handbook of Means and Their Inequalities*. Springer.
- Hausman, J. 2003. "Sources of Bias and Solutions to Bias in the Consumer Price Index." *Journal of Economic Perspectives* 17 (1): 23–44.
- ILO, IMF, OECD, Eurostat, UN, and World Bank. 2004a. *Consumer Price Index Manual: Theory and Practice*. International Monetary Fund.
- . 2004b. *Producer Price Index Manual: Theory and Practice*. International Monetary Fund.
- . 2013. *Handbook on Residential Property Prices Indices (RPPIs)*. Eurostat.
- Kennedy, P. 2003. *A Guide to Econometrics*. 5th ed. MIT University Press.
- Kirby-McGregor, M., and S. Martin. 2019. "An r Package for Calculating Repeat-Sale Price Indices." *Romanian Statistical Review*, no. 3: 17–33.
- Kirman, A. P. 1992. "Whom or What Does the Representative Individual Represent?" *Journal of Economic Perspectives* 6 (2): 117–36.
- Lee, M-J. 2016. *Matching, Regression Discontinuity, Difference in Differences, and Beyond*. Oxford University Press.
- Lehmann, E. L., and G. Casella. 1998. *Theory of Point Estimation*. 2nd ed. Springer.
- Lord, N. 2002. "Does Smaller Spread Always Mean Larger Product?" *The Mathematical Gazette* 86 (506): 273–74.
- Manski, C. F. 1988. *Analog Estimation Methods in Econometrics*. Chapman Hall.
- . 2007. *Identification for Prediction and Decision*. Harvard University Press.
- Martin, S. 2019. "Moral Management in Competitive Markets." *Journal of Economics & Management Strategy* 28 (3): 541–60.
- McFadden, D. 1978. "Cost, Revenue, and Profit Functions." In *Production Economics: A Dual Approach to Theory and Applications*, edited by M. Fuss and D. McFadden. North-Holland.
- Pollak, R. A. 1980. "Group Cost-of-Living Indexes." *The American Economic Review* 70 (2): 273–78.

- Rosen, S. 1974. "Hedonic Prices and Implicit Markets: Product Differentiation in Pure Competition." *Journal of Political Economy* 82 (1): 34–55.
- Wooldridge, J. M. 2010. *Econometric Analysis of Cross Section and Panel Data*. MIT University Press.