Méthodologie statistique

N°M0503

INTRODUCTION A LA PRATIQUE DES INDICES STATISTIQUES

Note de cours

Jean-Pierre BERTHIER

Document de travail

Institut National de la Statistique et des Etudes Economiques

INSTITUT NATIONAL DE LA STATISTIQUE ET DES ÉTUDES ÉCONOMIQUES
Série des Documents de Travail
de la
DIRECTION DES STATISTIQUES DEMOGRAPHIQUES ET SOCIALES

Unité « Méthodes Statistiques »

Série des Documents de Travail
Méthodologie Statistique

N° M0503

INTRODUCTION A LA PRATIQUE DES INDICES STATISTIQUES

Note de cours

Jean-Pierre BERTHIER
INSEE, Division Agriculture

Ce document est issu d'une formation effectuée au CEPE, auquel vont mes remerciements.
Je remercie également le referee dont les remarques ont contribué à enrichir ce document.

Introduction à la pratique des indices statistiques
Notes de cours

Jean-Pierre Berthier
INSEE Direction des statistiques d'entreprises

Résumé

Ce document de travail vise à fournir une vision d'ensemble des principales questions méthodologiques liées à la construction pratique des indices statistiques. Il s'appuie sur des exemples concrets pour montrer le type de questions qui se posent, dans un contexte plus général. A vocation pédagogique, il est constitué de notes de cours.

Mots-clés : indices ; indices de prix et de volume ; indices chaînés

Introduction to the practice of statistical indexes
Notes of course

Abstract

This working paper aims at giving a general view on the main methodological questions raised by the practical construction of statistical indexes. It presents concrete examples in order to look at what kinds of questions are raised, in a more general context. It is a teaching material based on notes of course.

Keywords : index numbers ; price and volume indexes ; chained index

SOMMAIRE

Introduction

Les indices constituent un domaine peu enseigné. A cet état de fait, on peut avancer plusieurs éléments d'explication :

- D'une part, il se situe au carrefour de la statistique et de l'économie.
- D'autre part, ce domaine soulève assez peu d'intérêt parmi les professeurs, de statistique comme d'économie, car il n'est pas au centre de leur problématique et passe souvent pour être un simple outil sans difficulté.

D'où la tentation, pour chacun des deux enseignements, de considérer qu'il relève de l'autre discipline.

Cependant les indices sont extrêmement utilisés et il est essentiel à la fois de savoir les interpréter correctement et d'être capable de concevoir des indices pertinents. C'est en fait un domaine très intéressant et complexe du fait même qu'il recoupe des problématiques diverses qui interfèrent :

- Il fait intervenir des considérations proprement statistique (ou mathématique) et d'autres directement liées au domaine d'application (souvent économique, mais qui peut être très varié).

- Sur un autre plan, il fait intervenir à la fois des aspects pratiques et des aspects méthodologiques. La difficulté, et l'intérêt, résident ici dans le fait que ces deux aspects sont intimement liés et ne peuvent pas être traités séquentiellement. Il peut être utile d'illustrer ici ce point, quitte à anticiper sur ce qui sera précisé plus avant dans ce document.

Pour le praticien, la première des questions qui se posent est de savoir ce qu'il veut mesurer. La réponse à cette question apparemment très banale s'avère en fait souvent redoutable. Un indice (agrégé) étant par essence un résumé de l'information, tout dépend du point de vue auquel on se place. C'est le très classique problème des effets de structure avec la question sous-jacente de savoir ce que l'on considère comme « produits » différents. Prenons un exemple.

Pour le calcul de l'indice des prix à la consommation (IPC), deux produits identiques mais vendus dans des types de magasins différents (grande ou petite surface) doivent-ils être considérés comme un seul et même produit ? La réponse apportée est non, mais il s'agit bien sûr d'un choix délicat qui peut être (et se trouve souvent) discuté.

Il est intéressant de s'interroger sur la nature de ce choix. Il signifie qu'un achat représente à la fois le produit acheté mais aussi la fourniture d'un service, en matière de proximité par exemple. Sur longue période, il est certain que l'incidence sur l'évolution des prix est importante. Le choix adopté, pour logique qu'il soit, signifie aussi que lorsque l'on calcule l'évolution du volume de la consommation en divisant classiquement l'évolution de sa valeur par l'indice de prix, on intègre de fait l'évolution du volume de service qui va avec. Les consommateurs pourraient donc garder rigoureusement la même alimentation mais voir le volume de leur consommation alimentaire baisser en achetant de plus en plus en grande surface.

Mais il faut aussi insister sur le fait que le choix à effectuer peut être très contraint par des considérations pratiques. Dispose-t-on, ou peut-on disposer à un coût raisonnable, des informations nécessaires aux différents calculs envisageables ? Dans le cas présent, il serait coûteux de disposer de l'information permettant de changer d'optique.

De même, parmi la boite à outil des instruments mis sa disposition par la théorie, et dont aucun n'est complètement satisfaisant, le praticien choisira largement en fonction de considérations pratiques. Dans bien des cas, si un indice de Laspeyres est utilisé plutôt qu'un Paasche ou, mieux, qu'un Fisher, c'est avant tout parce que l'on ne dispose pas de pondérations propres à chaque période (pour l'IPC les pondérations sont déduites des comptes nationaux annuels, alors que l'indice est mensuel). Ceci ne signifie cependant pas que tous les indices se valent pour le praticien. Ainsi, face au vieillissement de ces pondérations, le calcul de l'IPC s'effectue à l'aide d'un indice de Laspeyres, mais chaîné annuellement. Au total, il convient de trouver la meilleure adéquation, voire le meilleur compromis, entre ce que l'on veut mesurer, le type d'indice à retenir et l'ensemble des problèmes pratiques liés à son établissement.

Une autre caractéristique de l'étude des indices est que, si l'on peut se satisfaire d'un appareillage mathématique très modeste, l'usage des indices est assez subtil et plein de pièges. Il donne lieu par ailleurs à des développements théoriques très sophistiqués, qui ne seront pas véritablement abordés ici, même si l'existence de certains aspects pourra affleurer dans la suite de l'exposé. Pour mémoire, signalons simplement que de grands économistes, comme Solow ou Samuelson, ont contribué à développer cette théorie, sous son aspect économique principalement lié à la théorie du consommateur et du producteur (indices à utilité constante et mesure de la productivité globale des facteurs). Par ailleurs, la recherche théorique sur les indices est un domaine toujours actif, malgré l'ancienneté des problèmes posés.

De ces considérations introductives, il ressort que si l'étude des indices est quelque peu sacrifiée dans les cursus de formation initiale, elle présente un grand intérêt en temps que formation continue : celle-ci permettra non seulement de combler une éventuelle lacune, mais aussi de mieux enrichir la présentation par des considérations sur les liens entre des questions méthodologiques et des contraintes opérationnelles.

Dans le présent document, on se concentrera sur les aspects méthodologiques du choix d'un indice et de son calcul, mais en les insérant autant que possible dans un contexte plus large.

1 Qu'est-ce qu'un indice?

1.1 Les indices élémentaires

Définition : soit X_t une variable fonction du temps, l'indice d'évolution de X entre une date 0 et une date t est défini par

$$\mathbf{I_{t/0} = X_t / X_0}$$

Remarques :

- il est habituel d'écrire les indices sous la forme **102,1 au lieu de 1,021**. L'écriture des propriétés des indices est cependant alourdie inutilement par ce facteur 100, et l'équation ci-dessus reste privilégiée dans les considérations théoriques. On raisonne au facteur 100 près.

- Un indice est un nombre sans dimension, indépendant du choix des unités.

Exemple pour le fret SNCF (cf. tableau « indices élémentaires » en page 44) :

indice de quantité entre 1994 et 2000 : $I_{2000/94} = 46{,}5/44{,}8 = 1{,}038$ (soit 103,8)

Propriétés des indices élémentaires

- **Réversibilité** (entre les dates notées 1 et 2)

$$\mathbf{I_{2/1} = 1 / I_{1/2}}$$

- **Circularité (ou transitivité)**

$$\mathbf{I_{t/0} = I_{t/t-1} \times I_{t-1/t-2} \times \ldots \times I_{1/0}}$$

On dit que l'on peut chaîner les évolutions.

- **Conséquence**

$$\mathbf{I_{t2/t1} = I_{t2/0} / I_{t1/0}}$$

- **Partage volume-prix**

Si valeur = quantité x prix, alors **Ival = Iq x Ip**

1.2 Introduction aux indices synthétiques

Exemple du trafic SNCF

Il y a le fret, mais aussi le trafic voyageur (cf. tableau « Indices synthétiques » en page 45). On a deux séries qui ne sont pas dans les mêmes unités : tonnes-km et voyageurs-km. Les quantités ne peuvent donc pas être ajoutées. Seules les valeurs peuvent l'être : un indice élémentaire peut être calculé à partir de la valeur totale fret+voyageur.

Problème d'agrégation :

- Il s'agit de créer un indice unique (série) qui résume l'évolution du « volume » de trafic, et un autre qui résume l'évolution des prix.

- Il n'est pas envisageable pour définir l'indice de prix de ne faire intervenir que les séries de prix ; et de même pour les volumes/quantités : l'impact d'une hausse de prix du fret doit dépendre du poids du fret vis à vis du trafic voyageur.

- L'indice doit être neutre par rapport à un changement d'unité des quantités.

La résolution de ce problème passe par le fait que les indices de prix vont aussi faire intervenir les quantités (et les indices de volumes feront intervenir les prix), à travers l'utilisation des valeurs, éventuellement de façon implicite (cf.[1] en bibliographie)

Conclusion :

- Il y a plusieurs résumés (indices) possibles.

- Aucun résumé n'est (en général) parfait et les différents indices n'ont pas toutes les « bonnes » propriétés des indices élémentaires (cf. [2] et [3]).

- Les indices les plus classiques sont les indices de Laspeyres et de Paasche qui sont présentés dès la partie suivante. Mais le choix des indices est plus vaste.

- Deux questions essentielles se trouvent posées : comment construire ces indices ? Lequel choisir (et en fonction de quoi)?

2. Les indices classiques

2.1 Laspeyres et Paasche

Idée de départ :

Pour définir un indice prix, on va calculer l'indice élémentaire d'une variable ayant la dimension d'une valeur, en prenant en compte l'évolution des prix mais en neutralisant l'évolution des quantités (et inversement pour les indices de volume).

Définitions :

Si la variable dont on neutralise l'évolution est fixée à sa valeur à la date initiale, on a défini un **Laspeyres** (dans les formules qui suivent pour plus de lisibilité on omettra les indices de sommation des différents produits pour ne garder que les périodes 1 ou 2) :

$$\mathbf{Lp_{2/1} = \Sigma\, q_1p_2/ \Sigma\, q_1p_1 \qquad et \qquad Lq_{2/1} = \Sigma\, p_1q_2/ \Sigma\, p_1q_1}$$

Si la variable dont on neutralise l'évolution est fixée à sa valeur à la date finale, on a défini un **Paasche** :

$$\mathbf{Pp_{2/1} = \Sigma\, q_2p_2/ \Sigma\, q_2p_1} \qquad \text{et} \qquad \mathbf{Pq_{2/1} = \Sigma\, p_2q_2/ \Sigma\, p_2q_1}$$

Pour bien comprendre, explicitons le cas à deux produits (q, p) et (q', p'), entre les dates 1 et 2. S'agissant d'un indice de prix, on aura au numérateur les prix de la période 2 et au dénominateur ceux de la période 1. L'indice sera donc du type :

$$(q\, p_2 + q'\, p'_2)/(q\, p_1 + q'\, p'_1)$$

l'ensemble des termes de quantité relevant de la même période. S'agissant d'un indice de Laspeyres, ce sera la période 1. On aura donc :

$$Lp_{2/1} = (q_1\, p_2 + q'_1\, p'_2)/(q_1\, p_1 + q'_1\, p'_1)$$

De la même façon, l'indice de Paasche des volumes fera intervenir les quantités de la période 2 et les quantités de la période 1 (puisque c'est un indice de volume), mais tous les prix seront ceux de la période 2 (puisque c'est un Paasche) :

$$Pq_{2/1} = (p_2\, q_2 + p'_2\, q'_2)/(p_2\, q_1 + p'_2\, q'_1)$$

Exemple avec le trafic SNCF (tableau Indices synthétiques, page 41)

$$L_{vol2000/1994} = 114{,}6 \text{ alors que } P_{vol2000/1994} = 115{,}2$$

A ce stade, rien ne permet de dire lequel est le meilleur, les années 1994 et 2000 devant jouer a priori des rôles symétriques.

Une autre façon de voir : les indices de Laspeyres et de Paasche comme moyenne des indices élémentaires

- Le Laspeyres est égal à la moyenne arithmétique des indices élémentaires correspondants (prix ou quantités), pondérée par les valeurs de la période initiale.
- Le Paasche est égal à la moyenne harmonique des indices élémentaires correspondants, pondérée par les valeurs de la période finale.

Exemple SNCF (cf. tableau Indices synthétiques, page 45) :

$L_{vol2000/1994}$= [1661,71.(46,5/44,8)+3012,40.(52,3/43,4)]/(1661,71+3012,40)=1,146

(1535,18+3850,48)/ $P_{vol2000/1994}$ =1535,18/(46,5/44,8) + 3850,48/(52,3/43,4)
d'où $P_{vol2000/1994}$ =1,152

Suivant les cas, l'une ou l'autre approche peut être préférée (mais naturellement, le résultat est le même).

Propriétés élémentaires (d'autres propriétés ou défauts de propriété seront examinés plus loin : chaînage, agrégation)

- Ni L, ni P **ne vérifient la transitivité** : c'est un inconvénient majeur qui heurte le bon sens : si le PIB en volume, dont l'évolution mesure la croissance économique, est multiplié par 1,02 entre les années 0 et 1, et s'il est multiplié par 1,03 entre les années 1 et 2, on ne peut pas dire qu'il est exactement multiplié par 1,02 x 1,03 entre les années 0 et 2. On verra plus loin les réponses que l'on peut apporter (chaînage, base fixe).

- Ni L, ni P ne vérifient la réversibilité, mais on a les propriétés suivantes :

$$\mathbf{L_{2/1} = 1/P_{1/2} \quad et \quad P_{2/1} = 1/L_{1/2}}$$

- Partage volume-prix : ni L ni P ne le vérifie (le produit Lp x Lq n'est pas égal à Ival) mais on a :

$$\mathbf{Lp \times Pq = Ival \ et \ Pp \times Lq = Ival}$$

Intuitivement (par exemple pour la première égalité) : Lp fait passer de (q_0, p_0) à (q_0, p_1) et Pq de (q_0, p_1) à (q_1, p_1) qui correspond à la valeur. Ceci montre qu'à un Laspeyres des prix il faut associer un Paasche des volumes (et inversement).

Ces deux derniers couples de propriétés renforcent l'idée d'une certaine symétrie entre L et P, symétrie cachée dans l'interprétation comme moyennes d'indices élémentaires.

Positions relatives de L et P

a) Un certain nombre d'arguments font penser qu'**il est assez naturel que l'indice de Laspeyres soit supérieur à l'indice de Paasche** : moyenne arithmétique pour l'un, harmonique pour l'autre ; propriété assurée lorsque la part en valeur d'un produit évolue en sens inverse de son prix relatif ; L tend à sur-estimer l'impact des prix sur le pouvoir d'achat du consommateur car il ne prend pas en compte les substitutions possibles entre produits consommés, alors que P tend à le sous-estimer (cf. encadré « Théorie du consommateur et indices de prix »).

b) **Mais ce n'est aucunement une règle générale** : cf. le contre exemple avec les indices de volume du trafic SNCF calculés précédemment : L = 114,6 et P = 115,2

Une telle situation peut mériter cependant un examen. Le trafic voyageur augmente fortement alors même que son prix tend aussi à augmenter. En l'occurrence, il s'agit d'un effet TGV qui doit nous faire réfléchir sur le sens des indices calculés (cf. Homogénéité-hétérogénéité).

2.2 Les autres indices classiques

a) Deux indices faisant intervenir de façon symétrique les deux dates

Indice de Fisher : $F^2 = L.P$ (le Fisher est la moyenne géométrique des indices de Laspeyres et de Paasche). F est réversible ($F_{2/1} = 1/F_{1/2}$) mais non transitif.

Indice de Tornqvist : $T = I_1^{s1}.\ I_2^{s2}$ C'est la moyenne géométrique des indices élémentaires, les pondérations $s_i = ½\ (w_i(t_1)+w_i(t_2))$ étant égales à la moyenne des parts en valeur aux deux dates : $w_i(t_1)=val_i(t_1)/val_{totale}(t_1)$.

Ces deux indices, très voisins numériquement, sont moins utilisés que L ou P mais peuvent être considérés comme meilleurs, étant des indices « superlatifs » (cf. [4]).

b) Deux indices à vocation théorique (pour mémoire)

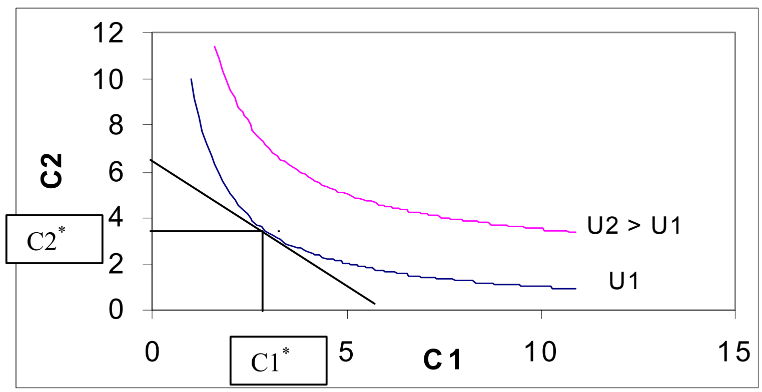
Indices à utilité constante (cf.[5] et encadré « Théorie du consommateur et indices de prix ») :
- Il se place dans le cadre de la théorie microéconomique du consommateur
- Il est égal à l'évolution du revenu nécessaire pour maintenir l'utilité du panier consommé.
- Suivant la forme de la fonction d'utilité, cet indice peut prendre la forme de divers indices classiques.

Indice de Divisia
- Il correspond à l'indice obtenu en chaînant en continu les indices classiques.
- Par construction, il est transitif, mais peut réserver de mauvaise surprises. En particulier, lors d'un retour à la situation initiale, l'indice peut être différent de 1 (cf. [6]).

Théorie du consommateur et indices de prix

On considère un consommateur et un certain nombre de biens (ou services) pouvant être consommés. Pour la clareté de l'exposé, on ne prendra ici que deux biens. Le consommateur peut, potentiellement, consommer différents paniers de consommation, définis par les quantités C1 et C2 de chacun des deux biens.

Une façon commode et très habituelle de représenter les préférences du consommateur consiste à lui attribuer une fonction d'utilité : à chaque panier, on associe un nombre, fonction des quantités de chaque bien consommé. La question de la préférence revient alors simplement à comparer les utilités respectives des différents paniers. Il est habituel d'effectuer certaines hypothèses sur cette fonction. En particulier, les courbes d'utilité constante définies dans le plan (C1, C2) doivent être concaves (cf. graphique ci-après donnant l'exemple de deux courbes d'utilité constante, correspondant à une même fonction d'utilité U). Cela signifie que lorsque l'on a beaucoup du bien 2 et peu du bien 1, il suffit d'une petite augmentation du bien 1 pour compenser une certaine diminution du bien 2, et inversement.



Une question fondamentale dans la théorie du consommateur consiste à définir quel sera le panier (C1*, C2*) de consommation choisi par le consommateur face à la donnée du prix de chaque bien et de la somme maximale dont il dispose. Cette contrainte de revenu correspond graphiquement au fait que l'on doit se situer sur la droite reliant les points sur les axes correspondant au fait de dépenser tout ce revenu à la consommation d'un seul produit. Dès lors, la recherche de l'utilité maximale conduit au point de tangence entre la droite de revenu et une courbe à utilité constante (cf. graphique ci-dessus).

Lorsque les prix évoluent, la droite de revenu change puisque les paniers de consommation possibles ne sont plus les mêmes, et l'utilité s'en trouve a priori modifiée. Par exemple, si les prix montent, l'utilité obtenue sera plus faible et pour conserver l'utilité de départ il faudra augmenter le revenu, de R_0 à R_1. Par définition, l'indice de prix à utilité constante sera égal à cette nécessaire évolution du revenu, R_1/R_0, pour conserver l'utilité. Cet indice intègre la possibilité qu'a le consommateur d'adapter sa structure de consommation à l'évolution des prix.

Par contre, l'indice de Laspeyres raisonne à structure de consommation fixe (celle de la situation initiale) puisqu'on peut l'analyser comme l'évolution du revenu nécessaire pour acheter le même panier avec les nouveaux prix. Cette absence de substituabilité entre les consommations des deux biens fait que l'on s'écarte de l'optimum du consommateur. Dès lors, en modifiant sa structure de consommation il pourrait obtenir une utilité supérieure à celle obtenue avec son panier fixe. On en déduit qu'il pourrait maintenir son utilité avec un revenu plus faible et qu'ainsi l'indice de Laspeyres est supérieur à l'indice à utilité constante. Une autre façon de voir consiste à dire que l'indice de Laspeyres surestime l'évolution des prix parce qu'il sous-pondère les produits dont le prix évolue le plus faiblement (sous l'hypothèse que leur part dans la consommation augmente). Un raisonnement de même nature montre que l'indice de Paasche est supérieur à l'indice à utilité constante, naturellement lorsque celui-ci existe, ce qui suppose que l'on se place sous les hypothèses de la théorie du consommateur.

3. Les propriétés d'agrégation

Le problème

On se place dans le cas ou l'on veut agréger selon un arbre d'agrégation, défini par une nomenclature.

Exemple de l'indice de la production industrielle (cf. tableau tiré d'un Informations Rapides de l'Insee, en page 48) : la nomenclature est la nomenclature économique de synthèse (NES) ; les pondérations sont les valeurs ajoutées de l'activité pour l'année de base (2000).

Pour calculer l'indice global, il y a plusieurs possibilités : soit calculer directement l'indice à partir de l'ensemble des postes au niveau fin de la nomenclature (ex : C1 Habillement cuir) ; soit partir des indices intermédiaires (ex : EC Biens de consommation) en les pondérant par les valeurs ajoutées correspondantes c'est à dire en faisant la somme des poids des activités concernées. Est-il indifférent de choisir l'une ou l'autre méthode ?

Propriétés

Les indices de **Laspeyres** et de **Paasche** vérifient les propriétés d'agrégation, c'est à dire que **les différentes méthodes sont mathématiquement équivalentes** (du fait de la structure de moyenne de ces indices).

Exemple numérique (cf. tableau page 48, extrait d'un Informations Rapides relatif à l'Indice de la Production Industrielle-IPI, qui est un indice de Laspeyres) : l'indice Industrie de février 2004 (100,5) peut être indifféremment calculé :
- Soit comme moyenne des 17 postes (dont C1 Habillement cuir, 58.6, avec une pondération de 190) ;
- Soit comme moyenne des 6 postes agrégés (dont EC Biens de consommation, 101.8, avec une pondération de 1257 = 190+341+499+226, à l'arrondi près).

Cette propriété est non seulement satisfaisante en soi, mais est importante en matière de diffusion lorsque l'on a des nomenclatures emboîtées comme pour l'IPI (la nomenclature complète comporte de nombreux niveaux).

Attention : cette propriété n'est pas du tout générale. En particulier, l'indice de Fisher ne la vérifie pas (tous les indices intermédiaires peuvent être égaux sans que l'indice global, calculé directement, soit égal à cette valeur commune. Cf. page 40).

4. Homogénéité-hétérogénéité

4.1 Exemple ferroviaire

Le tableau « Indices synthétiques », page 45, fait apparaître certains points qui méritent d'être analysés :

- Le prix du fret diminue sensiblement ;
- Pour les voyageurs, les quantités augmentent beaucoup alors que simultanément les prix sont plutôt en hausse.

Le calcul d'indices statistiques, destinés à résumer une information complexe, ne dispense pas d'examiner cette complexité et surtout de connaître un peu le domaine sur lequel portent les indices.

Le **transport ferroviaire de marchandises** comporte trois segments, correspondant à trois marchés différents :

- le wagon isolé : il est le plus coûteux à mettre en œuvre (triage) ; il est très fortement concurrencé par la route (prix, délais) ; en conséquence il est en déclin.

- Le train entier : ce créneau traditionnel résiste mieux mais perd tout de même des parts de marché face à la route ;

- Le transport combiné (caissons transportés par fer avant d'être livrés par la route) : coût ferroviaire plus faible ; créneau plus étroit (entre grandes villes suffisamment éloignées) mais en plein développement.

Le **transport de voyageurs** comporte deux segments :

- les trains classiques : ce segment est en déclin car il est concurrencé principalement par l'automobile mais aussi par l'avion ; et bien sûr il subit le transfert vers les TGV.

- Les TGV : ils sont en plein développement au cours de cette période, car ils offrent une qualité de service nettement supérieure alors que la politique tarifaire de la SNCF est attractive (hausse de prix limitée par rapport au train classique).

Ces éléments économiques conduisent à enrichir le tableau de données examinées. Le tableau « Homogénéité-hétérogénéité », en page 46, permet alors de comprendre les problèmes soulevés :

- Le prix du fret diminue sensiblement : il s'agit d'un **effet de structure**, le segment le plus cher (wagon isolé) étant en déclin, alors que le segment le moins cher (transport combiné) est en expansion. On a ici un « **effet qualité** », les différentes tonnes-km ne correspondant pas à la même prestation.

- Pour les voyageurs, prix et quantités augmentent simultanément : il y a là encore un effet qualité, la part de km-passagers effectués en TGV augmentant rapidement.

Le tableau Homogénéité-hétérogénéité permet de mesurer l'impact des choix suivants :

- les tonnes-km des différents segments du fret sont-elles comparables (homogénéité) ou bien non comparables (hétérogénéité)?
- Les voyageurs-km par TGV ou train classique sont-ils comparables (homogénéité) ou bien non comparables (hétérogénéité)?

Suivant que l'on choisi comparabilité ou non comparabilité :

- l'indice 2000/1994 de volume du fret est de 103.8 (il s'agit alors d'un simple indice de quantité) ou 94.9 (en choisissant un Paasche) !
- L'indice de prix correspondant est de 89.0 (il s'agit alors d'un prix moyen) ou de 97.4 (en choisissant un Laspeyres).
- L'indice global de Laspeyres des prix est de 100.0 avec 1 segment fret + 1 segment voyageurs ou avec 3 fret + 2 voyageurs : les deux segmentations se compensent. Mais on peut calculer L = 97.1 avec 1 segment fret + 2 voyageurs.

Quel calcul privilégier ?

Sur cet exemple, il semble préférable de segmenter pour mieux capter la réalité économique (mais qu'il faut connaître un peu). Les exemples suivants montreront que cette conclusion n'est pas universelle. **Il convient avant tout de se demander ce que l'on cherche à mesurer.**

Même dans l'exemple ferroviaire, la question n'est pas close. On peut imaginer des segmentations supplémentaires. Ainsi, pour le trafic voyageurs : TGV/corail/autres, 1ère classe/2ème classe, heures de pointe/heures creuses, etc.

4.2 Exemple d'un effet de structure qu'il ne faut pas corriger

La production de betterave à sucre va nous fournir un exemple où, contrairement à ce que pourrait laisser penser l'exemple ferroviaire, il est important de ne pas corriger les effets de structure (cf. tableau « Production de betterave à sucre », page 47).

La réalité économique est la suivante : il y a 3 segments caractérisés par des prix (versés aux producteurs) différents, mais un produit physique unique. Les deux premiers segments reçoivent des aides mais sont contingentés (quotas A et B) ; le troisième n'est pas soumis à quota mais le prix est celui du marché.

Pour évaluer la production, l'indice pertinent pour établir des comptes est dès lors celui du prix moyen car l'homogénéité physique des produits doit conduire à ce que l'évolution en volume est égale à celle des quantités.

Or, l'année 1998 montre que le partage volume-prix dépend complètement de ce choix : les 3 prix baissent, et donc l'indice de Paasche (ou de Laspeyres) baisse aussi, mais le prix moyen augmente car les quantités hors quota, moins payées, diminuent beaucoup.

4.3 D'autres exemples

L'indice du coût de la construction (ICC, cf. [7])

- l'effet de structure « régions regroupées» par exemple est corrigé (des logements identiques en Ile de France ou en Bretagne sont considérés comme différents).
- par contre, ce n'est pas le cas pour des différences de financement (aidé ou non).

Cet exemple montre que les choix à effectuer :

- ont une composante technique : s'il n'y avait aucune raison pour que les prix (ou leurs évolutions) soient significativement différents en Ile de France et en Bretagne, la question ne se poserait pas.
- relèvent aussi de considération d'opportunité : que veut-on mesurer ? (non correction de l'effet de structure « financement » mais correction de l'effet « région »).

L'indice des prix à la consommation (IPC, cf. [8])

Reprenons l'exemple indiqué en introduction : l'achat d'un même bien dans une grande surface et celui dans une épicerie doivent-ils être considérés comme différents ? Répondre à cette question suppose de répondre aux deux questions suivantes :

- Que veut-on mesurer ? Le service rendu n'est pas tout à fait le même ; mais le mouvement historique vers les grandes surfaces doit-il être neutre sur l'évolution de l'indice des prix (et donc du pouvoir d'achat)?

- Que peut-on mesurer ? Il est en fait plus difficile de calculer un prix moyen (supposant la comparabilité) qu'un indice synthétique (supposant la non comparabilité) car il faut prendre en compte de façon fine la déformation de la structure des lieux de ventes, par produits.

C'est la non comparabilité qui a été retenue. L'effet a été estimé au milieu des années 1990 à 0,4% par an (indice > évolution du prix moyen ; cf. Economie et statistique n°285-286).

On voit sur cet exemple que la maxime « si aucun prix ne bouge, l'indice ne bouge pas » n'est ici d'aucun secours, malgré son évidence dans une approche axiomatique. En l'occurrence, la question est de savoir si l'indice ne doit pas bouger lorsque les prix dans les différnets types de commerce restent constants mais que la part des achats entre ces différents commerces évolue. La fixité du prix d'un produit donné dépend en fait du point de vue que l'on adopte.

La mesure de l'évolution des salaires.

Une palette d'indicateurs différents sont envisageables : de l'indice de salaire moyen, utile en terme budgétaire, à l'indice correspondant à une perception individuelle, raisonnant à emploi, qualification et ancienneté donnés. Il faut dès lors savoir quel sens on veut donner à la phrase : les salaires évoluent plus/moins vite dans la fonction publique que dans le secteur privé. C'est poser là la fameuse question du GVT (glissement, vieillesse, technicité). Il n'y a pas de réponse unique.

En conclusion du thème hétérogénéité-homogénéité, on peut souligner les points suivants :

- Il s'agit d'un problème essentiel : les résultats obtenus peuvent dépendre de façon cruciale des choix effectués.

- Il doit être examiné en amont puisqu'il s'agit de savoir ce que l'on va retenir comme indices élémentaires.

- Il est nécessaire de disposer d'une certaine connaissance du domaine en question. On ne peut pas plaquer la méthodologie des indices sans connaître un peu le domaine auquel on l'applique.

- Les choix à effectuer relèvent avant tout de ce que l'on veut mesurer, mais ils sont indissociables de certains aspects techniques.

5. Séries temporelles et chaînage

Il est très fréquent de s'intéresser non pas à l'évolution d'une grandeur entre deux dates mais sur toute une période. Par exemple, on souhaite connaître l'évolution du volume de consommation des ménages tout au long de la période 1990 et 2000. Ce genre de question conduit à des difficultés nouvelles par rapport à ce que l'on a présenté jusque là.

- Une fois choisi le type d'indice (par exemple, un indice de Laspeyres) sa non transitivité conduirait à présenter les résultats du calcul dans un tableau à double entrées : à l'intersection de la ligne correspondant à l'année n avec la colonne correspondant à l'année m, on aurait l'indice d'évolution entre n et m.
- Cette présentation aurait l'inconvénient majeur de ne pas pouvoir se traduire par une courbe et permettrait difficilement l'analyse.

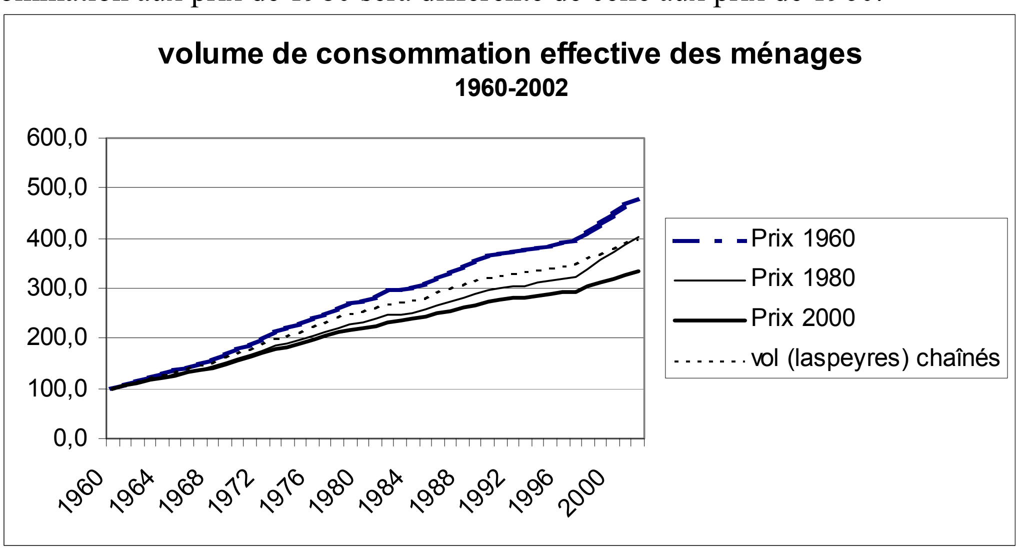
En pratique, deux solutions classiques peuvent être adoptées. Elles sont illustrées par l'exemple de la consommation des ménages, en volume, entre 1960 et 2000 (cf. graphique et tableau ci-dessous).

a) Calculer des indices à base fixe.

Cela revient à fixer n, à calculer des indices $I_{t/n}$, à tracer la courbe correspondante et à calculer, si besoin est, des évolutions entre deux dates t et t' par $I_{t'/t} = I_{t'/n} / I_{t/n}$.

Cette solution est très pratique mais présente des inconvénients :

- Sur la période [n, m], la structure de pondération peut évoluer beaucoup. Ainsi, avec des indices aux prix de 1960, les évolutions de 1995 à 2000 seront calculées avec la structure de prix de 1960 ! Il est d'ailleurs en soi curieux de faire intervenir une année 1960 extérieure à la période examinée.
- Il y a autant de résultats différents que de choix de n : l'évolution du volume de consommation aux prix de 1980 sera différente de celle aux prix de 1960.



Source : comptes nationaux de l'Insee et calculs de l'auteur.

Le tableau ci-dessous donne par ailleurs les évolutions moyennes annuelles correspondantes, par décennies et sur l'ensemble de la période.

	60-70	**70-80**	**80-90**	**90-2000**	**60-2000**
Prix 1960	1,059	1,045	1,029	1,020	1,038
Prix 1980	1,047	1,039	1,025	1,023	1,034
Prix 2000	1,045	1,036	1,022	1,015	1,029
Prix chaînés	1,055	1,041	1,024	1,016	1,034

Source : comptes nationaux de l'Insee et calculs de l'auteur.

Le graphique montre bien que la croissance de la consommation des ménages est plus ou moins vive suivant la méthode retenue : entre 1960 et 2000, la consommation a été multipliée par 4,5 aux prix de 1960 mais seulement par 3,2 aux prix de 2000, soit un écart de près d'un point de croissance par an (cf. tableau). Les résultats aux prix de 1980 et avec le chaînage (cf. ci-dessous), sont intermédiaires et assez voisins. Mais graphique et tableau montrent que, même dans ce cas, les évaluations par décennies sont sensiblement différentes : par exemple pour la dernière décennie, +2,3% par an avec des prix de 1980 contre +1,6% avec le chaînage.

b) Calculer des séries chaînées

La méthode

Elle consiste à calculer les évolutions d'une date à la suivante et à forcer la transitivité. On pose ainsi :

$$\mathbf{Ich_{t+1} = I\,ch_t \;\; x \;\; I_{t+1/t}}$$

où I peut être un indice de Laspeyres, de Paasche, de Fisher, etc. On définit ainsi un indice chaîne de Laspeyres (par exemple), ou encore un « Laspeyres chaîné ».

Par exemple, on aura $I\ vol_{2000/1960} = I_{61/60}$ x $I_{62/61}$ x …x $I_{2000/1999}$. Chacun des termes $I_{t+1/t}$ est appelé un maillon (de la chaîne).

Propriétés

- Ich est un indice transitif.
- Lch (prix) x Pch (vol) = I (valeur) = Pch (prix) x Lch (vol). On peut donc indifféremment chaîner les prix puis diviser la valeur par l'indice chaîné ou bien chaîner directement les volumes.

Le choix entre chaînage et base fixe, et dans ce cas celui de l'année de base, peut être important, notamment sur longue période comme l'ont montré graphique et tableau ci-dessus, relatifs à la consommation. Il n'y a pas de « bonne » année de référence sur longue période. Dans le cas présent, les comptables nationaux privilégient la série chaînée.

Considérations (plus ou moins) théoriques

La justification principale du chaînage vient du fait qu'il permet de prendre en compte la déformation de la structure de l'économie (prix relatifs par exemple) tout au long de la période en question. Cet argument a cependant ses limites et dans certains cas le chaînage n'est pas souhaitable : lors d'un retour à la situation initiale, le chaînage n'assure pas de revenir à l'indice 100 comme le ferait un indice à base fixe (cf. page 37).

Il est possible de donner quelques idées complémentaires, plus théoriques :

- Il existe un certain nombre de cas théoriques où le choix du chaînage est parfaitement justifié, mais les conditions sont très restrictives. Enoncés (trop) brièvement, le premier concerne les cas où il existe une fonction d'utilité homogène : d'après le théorème de Hulten, l'indice de Divisia est l'indice à utilité constante (cf. [9]). Le deuxième cas est celui où les quantités relatives sont fonctions -uniquement- des prix relatifs (cas de deux produits, cf.[6]).

- En définitive (cf. [10]), il y a l'idée que le chaînage, en faisant intervenir les dates intermédiaires, permet de prendre en compte le lien qui existe entre les variables prix et quantités (approche économique de la théorie des indices). S'il n'y avait aucun lien entre ces variables, il n'y aurait aucune justification à faire intervenir des dates intermédiaires (approche axiomatique de la théorie des indices).

Considérations pratiques

- Le chaînage introduit quelques complexités par rapport aux indices à base fixe : modification des pondérations à chaque période ; destruction des égalités comptables (cf partie « partage volume-prix »).

- Sur longue période, ou même dès que la structure de pondération se déforme, le chaînage peut être utile, voire indispensable.

- Il suppose cependant de respecter certaines règles de prudence : le chaînage est (en pratique) justifié lorsqu'à des dates intermédiaires correspondent des situations « intermédiaires ». Il faut donc se méfier du chaînage lorsque les séries sont très irrégulières.

Une autre utilisation du chaînage : les ruptures de séries

Le chaînage est souvent utilisé pour éviter des ruptures de séries (c'est à dire des marches d'escalier qui ne correspondent pas à la réalité que l'on veut décrire) :

- Ce peut être le cas lorsque la source statistique a changé ou bien contient en son sein une telle rupture de série, lorsque le champ couvert par une série a changé, etc. ou encore lorsque l'on a découvert une erreur passée.

- L’idée est de privilégier alors les évolutions (un indice ne dit rien sur le niveau lui-même). Pour cela on calcule l’évolution de l’indice (aussi proprement que possible) que l’on chaînera à la valeur précédente.

- Exemple d’un indice à base fixe : on a publié 113,6 en période t ; on calcule 115,2 en période t+1 ; mais on se rend compte que l’on aurait dû publier 114,1 en t. S’il est trop tard pour corriger le 113,6, on calculera l’évolution 115,2 / 114,1 que l’on appliquera à l’indice « faux » 113,6 pour obtenir 114,7.

Un cas important est celui où l’on dispose de séries d’indices relativement courtes alors que l’on voudrait constituer une série longue. C’est souvent le cas lorsque l’on prend les séries publiées par l’Insee (cf. changements de base, ci-après). Il est alors fréquent, faute d’information complémentaire, de chaîner les séries, en général en partant de la plus récente. Dans le jargon, on dit souvent que l’on raboute ou que l’on raccorde les séries.

6. Le partage volume-prix

a) Partage de l'indice de valeur en indices de prix et de volume

Lorsque l'on s'intéresse à la fois aux prix et aux volumes, il est généralement demandé d'avoir l'égalité :

Indice (élémentaire) de valeur
=
indice (synthétique) de prix x indice (synthétique) de volume

Cette égalité n'est pas vérifiée pour tous les types d'indices (rappel). Les possibilités les plus classiques sont **L(p) et P(q)** d'une part, et **P(p) et L(q)** de l'autre. Mais l'utilisation du **Fisher** à la fois pour les prix et les volumes est une très bonne solution.

Imposer l'égalité ci-dessus est naturelle mais cela implique que tout ce qui n'est pas considéré comme constitutif d'une évolution de prix est considéré comme constitutif de volume, et réciproquement. On rappellera ici l'exemple de l'indice des prix à la consommation : l'effet « lieu d'achat » n'est pas constitutif de prix ce qui signifie que le fait que l'achat des mêmes biens se soit effectué de plus en plus en grande surface a tendu à faire baisser l'indice de volume correspondant (dans les comptes nationaux c'est bien en déflatant les valeurs par l'IPC qu'est calculé le volume de consommation).

b) Volumes aux prix d'une année donnée

Il est souvent intéressant de calculer des volumes (et non des indices de volume) aux prix d'une année donnée. Par définition on pose (comme souvent, au facteur 100 près) :

$$\textbf{volume en t au prix de n = valeur en t / } I_{t/n}\textbf{(prix)}$$

Le volume est alors exprimé en unité monétaire. Mais il est essentiel de ne pas confondre :

- un volume au prix de l'année 2000 par exemple,
- une valeur exprimée en euros constants de 2000.

Dans le premier cas, le déflateur est l'indice de prix du produit concerné. Dans le deuxième, le déflateur est un indice de prix général (souvent l'IPC ou le prix du PIB) de façon à pouvoir comparer des euros à des dates différentes.

Propriétés

- Le volume de l'année n aux prix de n n'est rien d'autre que la valeur.

- Avec une base fixe (année n), l'indice de volume entre t et t' est égal au rapport des deux volumes aux prix de la même année n.

- Mais attention, le rapport de deux volumes exprimés aux prix d'années différentes n'a pas de signification. En particulier, lorsqu'on appelle de façon trop elliptique « volume » le volume aux prix de l'année précédente, le rapport des volumes de deux années ne représente pas l'indice de volume entre les deux dates !

En partie pour pallier cet inconvénient, les comptables nationaux annuels publient des volumes chaînés, en euros 2000 par exemple. Il sont calculés en prenant comme déflateur l'indice de prix chaîné annuellement entre t et 2000. Ces volumes vérifient alors, comme les indices à base fixe, le fait que l'indice de volume (chaîné) entre t et t' est égal au rapport des deux volumes (chaînés).

c) Partage volume-prix et égalités comptables

A partir du moment où l'on a défini des volumes en euros, il est naturel de vouloir les ajouter, ce que les indices ne permettent pas en eux-mêmes. Par exemple, on peut additionner le volume (au prix de 1995) de trafic fret avec le volume de trafic voyageurs (toujours aux prix de 1995). Le résultat est-il égal au volume global de trafic ferroviaire (au prix de 1995), sachant que celui-ci est déjà défini par : valeur globale / indice de prix global ?

Ce problème n'a qu'une seule solution : des indices de Paasche pour les prix, ce qui correspond à des indices de Laspeyres pour les volumes. C'est ce qui est utilisé en Europe pour les comptes nationaux (dans les calculs aux prix d'une année fixe ou aux prix de l'année précédente).

En pratique, il est alors équivalent, pour agréger les indices de volume, d'additionner les différents volumes et de calculer l'indice correspondant ou bien de calculer directement l'indice de volume à partir d'un Laspeyres.

Mais attention : le chaînage détruit cette additivité. Le chaînage de la somme n'est pas égal à la somme des chaînages. Il y a un « écart de chaînage » dû à des effets de structures (il a ainsi une origine « purement » mathématique et n'a rien à voir avec l'incertitude des données).

De façon générale, le chaînage détruit les égalités comptables. C'est l'inconvénient majeur de l'utilisation des indices chaînés en comptabilité nationale ([cf. 11]).

d) Remarque sur les indices de valeurs

Contrairement aux indices de volume ou de prix, les indices de valeur (laquelle s'exprime de façon homogène en unité monétaire) ne relèvent **a priori** pas des indices synthétiques mais des indices élémentaires : il suffit de poser $I_{t'/t} = val_{t'}/val_t$.

En pratique, on peut cependant être amené à les calculer sous la forme d'indices synthétiques. Il est alors naturel de choisir un indice du type Laspeyres (vu comme moyenne arithmétique des indices élémentaires où les pondérations sont les parts en valeurs). En effet, la valeur totale (notée Val) étant exprimée comme somme des valeurs correspondant aux différentes composantes (notées val^i), on peut écrire :

$$\frac{Val_1}{Val_0} = \frac{\sum_i val_1^i}{\sum_i val_0^i} = \sum_i \frac{val_0^i}{\sum_j val_0^j} \frac{val_1^i}{val_0^i}$$

On reconnaît l'expression un indice de Laspeyres, les indices élémentaires étant relatifs aux valeurs des différentes composantes.

L'**exemple de l'indice de chiffre d'affaires** de l'Insee permet de l'illustrer. Il s'agit d'un indice conjoncturel mensuel mesurant l'activité des secteurs industriels, commerciaux et des services marchands, basé sur les données fiscales liées au reversement de la TVA par les entreprises. Le calcul de cet indice comme un indice synthétique est justifié par le fait :

- que ces indices sont calculés et publiés à différents niveaux de la nomenclature d'activité, suivant un arbre d'agrégation ;
- qu'au niveau le plus fin (NAF 700) l'indice (élémentaire) est calculé à partir d'un échantillon d'entreprises qui permet de disposer d'évolutions mais pas de niveaux ;
- que les pondérations doivent dès lors être établies par ailleurs (elles sont calculées pour l'année de base à partir des statistiques structurelles d'entreprises, sur un champ exhaustif).

7. Base et changement de base

Période de base

On se place dans le cas, très courant, où l'on veut établir une série. En vue de construire un indice synthétique, tous les indices élémentaires doivent être établis en valant 1 (ou 100) à la même période. Cette période commune est appelée période de base. Attention : l'évolution de l'indice synthétique obtenu dépend du choix de la période de base.

Dans le cas le plus simple (et le plus fréquent) d'un indice de Laspeyres, les indices élémentaires sont pondérés par les valeurs de la période de base.

Changements de base

Lorsque l'on s'éloigne de la période de base, les déformations structurelles (prix relatifs par exemple) font progressivement perdre à l'indice calculé sa pertinence. Il faut alors changer de (période de) base. On peut remarquer que le chaînage correspond en quelque sorte à un changement de base à chaque période (en prenant comme période de base la période qui précède immédiatement).

Pour un indice de Laspeyres, le changement de base correspond à une actualisation des pondérations. Mais un changement de base peut-être l'occasion d'introduire d'autres modifications : choix des séries, nouveaux concepts, nouvelle nomenclature, sources statistiques, extension du champ, etc.

Le choix de la période de base peut être important. Il faut éviter de prendre une période atypique, donnant par exemple à un indice élémentaire un poids inhabituel. Ce choix est en fait souvent contraint (année ronde, décision européenne, etc.). Il est parfois pris comme pondération une moyenne centrée sur la période affichée (ex : moyenne 1999-200-2001 pour une base « 2000 »).

Etablissement de séries longues

Le cas type est celui d'un indice de Laspeyres avec changement de base tous les 5 ans, avec un certain recouvrement entre les séries établies dans deux bases successives (par exemple, base 95 établie jusqu'en 2002, base 2000 établie à partir de 2000). Pour établir une série longue, deux cas de figure sont possibles :

a) Une véritable rétropolation

Elle consiste à recalculer sur le passé les séries d'une façon homogène avec la dernière base. C'est une opération très lourde, qui ne peut (en général) pas être effectuée par l'utilisateur lui-même.

b) Un simple raccordement de série.

Il s'agit ici de chaîner les évolutions, en choisissant une période pivot sur la plage de recouvrement. Cela peut être fait par le producteur mais aussi par l'utilisateur. Vis à vis des pondérations, il est probablement préférable de laisser ainsi celles de l'époque plutôt que de rétropoler en fixant sur le passé les pondérations d'aujourd'hui. Par contre, pour les changements de concept, champ, etc. un raccordement est un pis-aller à une rétropolation. En cas de changement de nomenclature, un raccordement est souvent impossible, sauf à un niveau très agrégé.

Période de référence

Lorsque tous les calculs sont effectués, il est possible d'afficher la série de l'indice synthétique obtenue en fixant l'indice égal à 100 pour une période quelconque. Par exemple, l'IPC est actuellement en base 1998. On peut souhaiter (le producteur ou plus souvent l'utilisateur) l'afficher en prenant 100 en 2000.

Ceci n'apporte aucune information nouvelle (toutes les évolutions sont identiques) et ne change pas la période de base. On dit alors que l'on a un indice base 98 mais en référence 100 en 2000 (on note parfois « 100 en 2000 » voire « 2000 = 100 »).

Il est essentiel de distinguer un changement de base et un changement de référence. Malheureusement, le vocabulaire en la matière n'est pas complètement codifié. On emploie souvent base n dans le sens précédent de base et référence n. Beaucoup plus gênant, il arrive d'entendre « changement de base » pour « changement de référence », ce qui est trompeur et doit être banni.

8. Le choix du type d'indice

On cherchera ici à aider celui qui doit construire un indice synthétique à choisir le type d'indice (Laspeyres, Paasche, etc.). Il s'agit d'un mini guide pratique qui vise délibérément la simplicité. On peut naturellement s'écarter de ces « règles », mais en ayant conscience que cela réclame généralement une certaine expertise.

Les trois indices classiques sont le Laspeyres, le Paasche et le Fisher. Pour les applications courantes, on peut sans inconvénient se restreindre à ces 3 types d'indices, éventuellement chaînés.

Le Laspeyres est de loin le plus utilisé et peut être considéré comme l'indice classique par excellence. La raison de ce succès tient à la fois à son interprétation simple (panier de consommation), à ses propriétés d'agrégation et au fait qu'il permet de calculer une série temporelle sans changer les pondérations (dont la détermination peut être coûteuse).

L'indice de Fisher est le meilleur des trois. Son utilisation est notamment recommandée lorsque les évolutions de structure (les pondérations en valeur ou les prix relatifs) sont très différentes en début et en fin de période. Son utilisation est cependant assez réduite à la fois car il apparaît plus compliqué (argument peu valide) et qu'il nécessite, à travers l'indice de Paasche, de calculer des pondérations aux diverses périodes (ce qui peut être très coûteux). Par ailleurs il pose des problèmes d'agrégation.

L'indice de Paasche est le moins intéressant des trois : il a l'inconvénient pratique du Fisher sans en avoir l'avantage théorique. Son utilisation est restreinte à des cas spécifiques, soit théoriques (partage volume-prix) soit pratiques (Indice du coût de la construction). Utilisé en série, il peut poser des problèmes d'interprétation : l'évolution que l'on en déduit entre deux périodes successives peut être paradoxale du fait que les pondérations sont différentes entre les deux dates (on peut avoir une baisse de chaque indice élémentaire et une hausse de l'indice synthétique. Cf page 40).

Pour faire un partage volume-prix, on prend généralement des prix Paasche et des volumes Laspeyres. Si l'on n'a pas besoin d'égalité comptable en volume, une solution Fisher-Fisher est techniquement préférable, bien que moins fréquente.

En série longue, l'utilisation d'un indice à base fixe est à éviter et le chaînage est a priori la bonne réponse (mais il supprime les égalités comptables). Il doit cependant être évité lorsque les séries sont très heurtées. La solution la plus « passe-partout » est alors d'utiliser un Fisher. Une alternative consiste à effectuer des changements de base suffisamment fréquents (avec nécessité de chaîner les différentes périodes pour reconstituer une série longue).

9. Construction des indices élémentaires

Le choix des séries.

Très souvent, une « série élémentaire » doit représenter l'évolution, d'un prix par exemple, pour un poste aussi fin que possible d'une nomenclature. « Aussi fin que possible » signifie que :

- On est capable d'affecter une pondération à ce poste de la nomenclature.

- On est incapable d'affecter des pondérations à une désagrégation de ce poste.

Exemple pour l'IPC (on parle alors de « variété », cf.[8] page 76) : les imperméables pour homme. La connaissance des achats correspondants est déjà approximative. Ce poste contient pourtant une certaine diversité de produits.

Pour construire ces « indices élémentaires », au sens où c'est à partir d'eux que l'on appliquera les techniques d'agrégation classiques (par utilisation d'un indice de Laspeyres par exemple), il faut suivre un certain nombre de produits bien précis, censés représenter le poste en question. On parle parfois de « séries témoins ».

Deux cas de figure sont distingués classiquement :

- Soit le poste est considéré comme homogène (377 cas sur 909 pour l'IPC de la base 90). L'indice élémentaire est alors pris comme l'évolution du prix moyen.

- Soit le poste est considéré comme hétérogène (532 cas sur 909). L'évolution du prix moyen ne pas être considérée comme un indice de prix. Faute de pondération, on prend alors la moyenne géométrique des évolutions de chaque série témoin (la moyenne géométrique ayant de meilleures propriétés que la moyenne arithmétique, notamment lorsque les prix fluctuent. Voir encadré « Utilisation de la moyenne géométrique »).

Au delà de cet aspect méthodologique concernant le calcul de l'indice lui même, le choix des séries peut être une opération très lourde. Ainsi, pour l'IPC (base 1998), 130 000 séries sont suivies. Elles sont définies à partir d'un plan de sondage stratifié par type de produit (plus de 1000 variétés), agglomérations (106) et type de point de vente. Il faut par ailleurs ajouter 40 000 séries de type « tarif ».

Les produits nouveaux ou ayant des caractéristiques nouvelles**.**

L'établissement d'indices de prix d'un ensemble de produits se heurte au fait que les caractéristiques des produits évoluent au cours du temps, que certains produits nouveaux sont introduits sur le marché alors que d'autres disparaissent. On peut citer comme exemples de produits véritablement nouveaux plus ou moins récents les micro-ordinateurs ou le téléphone portable. Plus fréquemment, ce sont les caractéristiques de produits existants qui changent : les performances des micro-ordinateurs sont sans cesse croissantes par exemple. Mais au-delà de ces

Utilisation de la moyenne géométrique

Dans le cas des postes hétérogènes (cf. texte), la formule traditionnelle il y a encore 10 ou 15 ans était de calculer l'indice comme la moyenne arithmétique des indices des séries témoins (moyenne simple faute de pouvoir disposer de pondérations à ce niveau de détail). Il est aujourd'hui admis au niveau français et international qu'il est préférable d'utiliser la moyenne géométrique (simple). Plusieurs arguments peuvent être avancés dont les deux suivants.

1) Au niveau de détail auquel on se place, il n'est pas rare qu'il y ait des fluctuations de prix. Or, dans ce cas, la moyenne arithmétique présente des inconvénients qui peuvent être importants et que la moyenne géométrique permet d'éviter. Pour comprendre l'origine de cette question qui n'est pas que thèorique, prenons le cas d'école suivant : les prix de deux produits valent alternativement p et p' de la façon suivante : suivant les périodes d'observation (les mois par exemple) le prix du produits 1 vaut alternativement p, p', p…alors que celui du produit 2 vaut p', p, p'…de telle sorte qu'à une période donnée on ait toujours un produit à p et l'autre à p'. Puisque l'on ne dispose pas de pondérations, les deux produits doivent jouer le même rôle dans notre calcul et la situation doit être considérée comme équivalente pour toutes les périodes.

La moyenne géométrique rend bien compte de cette équivalence : entre la période initiale (p, p') et une période (p', p), le carré de cet indice vaut (p'/p)x(p/p') =1. Il en est bien sûr de même entre la période initiale et une période (p, p') identique à la période initiale. Par contre, avec la moyenne arithmétique, entre la période initiale et une période (p', p) l'indice vaut 1/2(p'/p+p/p') qui est supérieur à 1 si p n'est pas égal à p' (la moyenne arithmétique est supérieure à la moyenne géométrique).

Si l'on chaîne les évolutions d'une période à la suivante, on vérifie aisément que la moyenne géométrique continue à donner un indice toujours égal à 1. Par contre, avec la moyenne arithmétique, l'indice s'écarte de plus en plus de l'unité alors même que toutes les deux périodes la situation se reproduit à l'identique. En effet, entre une période quelconque et la suivante, l'indice est toujours égal à 1/2(p'/p+p/p').

2) Le deuxième argument tient à la prise en compte ou non d'une forte substituabilité entre les produits. Au niveau de finesse de la nomenclature de produits à laquelle on se place, il existe en effet des possibilités importantes de substitution d'un produit à un produit voisin en fonction des prix relatifs. Il est donc important que l'indice choisi soit cohérent avec cette réalité économique. Or, la moyenne arithmétique peut s'analyser comme un indice de Laspeyres particulier (ceux-ci pouvant s'écrire comme moyenne arithmétique pondérée d'indices élémentaires) et donc correspond à une structure de consommation fixe. Par contre, la théorie des indices à utilité constante montre que la moyenne géométrique (pondérée) est égale à l'indice à utilité constante (le « bon » indice) dès lors que la fonction d'utilité du consommateur correspond à une élasticité de substitution de 1, ce qui correspond à un ordre de grandeur qui n'est pas irréaliste en pratique à ce niveau de détail des produits.

cas que chacun a en tête, les caractéristiques fines des produits évoluent très souvent sans que l'on y porte forcément attention.
Ainsi, pour l'IPC (cf. [8] page 95), le renouvellement des produits a concerné 46% de l'échantillon en 1997. Cette question peut donc avoir une importance pratique considérable.

Sur le plan méthodologique, la question, lorsqu'un produit ayant des caractéristiques nouvelles en remplace un autre est d'effectuer un partage entre ce qui relève d'une variation de la qualité du produit (l'effet qualité) et ce qui relève d'un effet de prix pur qui seul doit être pris en compte lorsque l'on veut établir un indice de prix. Plusieurs méthodes sont envisageables (et pratiquées) suivant l'analyse que l'on fait et les moyens dont on dispose.

Commençons par les deux cas polaires, dans lesquels on considère que l'écart de prix constaté entre l'ancien produit et le nouveau correspond entièrement soit à un prix pur, soit au seul effet qualité.

a) La comparaison directe (ou remplacement « en équivalent » pour l'IPC)

On considère ici qu'il n'y a pas d'effet qualité et que la variation de prix est une variation de prix pure. Cette solution a bien sûr comme principal avantage la simplicité puisque l'on fait comme s'il n'y avait pas eu de changement de produit. En pratique, elle est assez largement utilisée lorque les produits ont des caractéristiques suffisamment voisines (par exemple : changement mineur de conditionnement). Il convient cependant d'être très vigilant sur le caractère « mineur » de ces changements. Ainsi, les prix à la production (IPPAP) des pommes de terre d'une variété bien définie peut varier très fortement suivant que la livraison s'effectue par 10, 25 ou 100 kg, lavée ou non, etc.

b) Le chaînage (ou remplacement en « dissemblable pur » pour l'IPC, ou encore « méthode des modèles appariés »).

On considère à l'inverse que la variation de prix observée est entièrement due à un effet qualité. Le principe est alors de ne calculer d'indice qu'à partir des seuls produits communs aux deux périodes et de chaîner les évolutions de prix. Cette méthode ne pose pas non plus de problème opérationnel et elle est largement utilisée.

Elle présente cependant plusieurs inconvénients :
-elle ne peut s'appliquer telle quelle que si pendant une petite période les deux produits sont simultanément sur le marché (cf. ci-après dans le cas contraire).
- l'hypothèse de départ est plus ou moins valide : un changement de prix peut s'effectuer à l'occasion d'une modification des caractéristiques du produit.

L'impact du choix d'un tel traitement peut être dans certains cas important. Certains produits ont en effet des caractéristiques qui évoluent très rapidement. C'est le cas des micro-ordinateurs (mais aussi des automobiles par exemple). Et la baisse de prix de la micro-informatique passe souvent par l'introduction de micro ayant des caractéristiques meilleures. Le remplacement « en dissemblable pur » peut ici biaiser largement le résultat du calcul (cf. d).

Lorsque ces deux cas polaires paraissent trop frustes ou bien lorsque l'on souhaiterait effectuer un chaînage mais que les produits ancien et nouveau ne coexistent pas, il convient alors d'« inventer » un prix fictif pour la période où le produit n'existait pas encore.

c) Dans certains cas, cela peut être fait d'une façon simple et largement automatique. On utilisera l'évolution des prix d'un produit comparable ou bien en affectant l'évolution de l'ensemble du poste de la nomenclature comprenant ce produit (l'IPC parle de « dissemblable corrigé »).

d) Il est parfois nécessaire d'effectuer une évaluation explicite du prix fictif à une période antérieure à l'apparition du produit. Des méthodes plus ou moins complexes sont en pratiques utilisées :

- On peut corriger à la marge le prix des produits en fonctions du changement de leurs caractéristiques. Ainsi, pour l'automobile lorsqu'une option passe en série, l'IPC corrige de la moitié du prix qui était demandé pour l'option (cf. [8] page 101). Ce coefficient conventionnel s'appuie sur l'idée qu'en passant en série une caractéristique (climatisation par exemple) coûte moins cher à fabriquer et que d'autre part une partie des consommateurs n'étaient pas prêts à payer le prix correspondant à l'option (ce qui traduit une utilité moindre).

- Dans certains cas, il est souhaitable de partir des caractéristiques des produits elles-mêmes. Ces méthodes dites « hédoniques » s'appuient sur l'étude économétrique du lien constaté entre prix et caractéristiques. L'exemple emblématique des méthodes hédoniques est celui des micro-ordinateurs. Les caractéristiques principales sont la vitesse d'horloge et la capacité (en Go), chaque marque donnant lieu à une équation différente. L'écart par rapport à la méthode du chaînage peut être important lorsque les produits évoluent très vite.

L'introduction des méthodes hédoniques à l'Insee est intervenue au début des années 90 pour calculer l'indice de prix de vente industriel (appelé aujourd'hui indice de prix à la production) des micro-ordinateurs. L'étude indiquait qu'en prenant 100 en 88 T1 (1er trimestre 1998), l'indice (hédonique) pour 91 T2 devait être égal à 61,6 et non 70,9 comme avec la méthode basée sur les produits communs à deux périodes successives. Notons que, bien que l'IPP (ex IPVI) soit un indice à base fixe (rebasé tous les cinq ans), l'évolution rapide de la micro-informatique a conduit à actualiser les estimations économétriques chaques mois et donc à effectuer un chaînage mensuel des prix hédoniques.

Ces méthodes hédoniques sont toutefois complexes, coûteuses à mettre au point et surtout à maintenir. Seule une mutualisation des travaux au niveau international permettra d'étendre quelque peu leur utilisation.

e) L'introduction dans l'indice des produits entièrement nouveaux, ne correspondant au remplacement d'aucun produit plus ou moins semblable, est beaucoup plus rare. Il est habituel de considérer que l'apparition sur le marché d'un tel produit n'a d'effet sur l'indice des prix que par l'évolution de son prix. Son introduction dans l'indice se fait nécessairement avec un certain retard, en pratique à un moment où la pondération qu'il aurait dans l'indice n'est plus négligeable (et a souvent tendance à se stabiliser). L'inconvénient d'un tel retard doit donc être relativisé dès lors que les pondérations sont revues suffisamment souvent (par chaînage ou rebasements fréquents).

10. Quelques points particuliers

a) **Les paniers variables**

Pour les produits saisonniers, comme par exemple les fruits (indice de prix à la consommation ou, plus encore, à la production), l'importance relative des produits dépend fortement de la période de l'année. De plus certains produits peuvent même être absents à certaines périodes. Il est alors assez classique d'introduire un « panier variable » qui permet de traiter ces cas, non sans certains inconvénients.

Techniquement

La pondération des différents fruits (par exemple) sera variable pour chacun des mois de l'année. Il y aura donc 12 coefficients par produit au lieu d'un seul. Mais la pondération de l'ensemble des fruits sera maintenue constante (sinon les postes non saisonniers auraient eux aussi des pondérations fonction du mois).

L'indice du mois m de l'année n peut être calculé à partir de l'évolution du prix depuis le même mois m de l'année de base. Mais il faut ensuite multiplier cette évolution par l'indice de ce mois m de l'année de base. Pour établir ce dernier, on fixe en général à 100 la moyenne de l'année de base pour chaque produit, ce qui conduit à définir aussi l'indice global avec une référence 100 en moyenne pour l'année de base.

Les inconvénients

On peut malheureusement en citer plusieurs :

- Une certaine complexité technique.
- Certains problèmes doivent être résolus en marge. En particulier, les aléas peuvent faire que le panier prévu corresponde mal au panier réel. Par exemple, il peut y avoir un décalage d'un mois dans les récoltes ou bien des évolutions dans les importations.
- Surtout, l'introduction d'un panier variable introduit une difficultés d'analyse des indices obtenus. L'évolution d'un mois sur l'autre de l'indice agrégé ne signifie pas grand chose, puisqu'elle résulte de l'évolution de prix de deux paniers différents. Seuls les glissements annuels ont une signification claire.

Malgré tout, les paniers variables restent nécessaires lorsque les phénomènes sont très saisonniers.

b) Les indices corrigés des variations saisonnières et des jours ouvrables (CVS-CJO)

La technique de désaisonnalisation (et de correction des jours ouvrables) n'est pas une question d'indices. Mais de nombreux indices conjoncturels sont publiés en CVS (ou CVS-CJO). Le choix de calculer des indices CVS doit être conditionné par une double affirmation :

L’utilisation des indices

D’une façon générale, les indices peuvent être utilisés à plusieurs fins :
- Ils peuvent servir à l’élaboration de certaines statistiques : les comptes nationaux utilisent par exemple l’IPC et les IPP pour effectuer des partages volume-prix, l’IPI pour les comptes trimestriels ou annuels provisoires, etc.
- La publication des indices constitue une information socio-économique utile à la fois au débat social et à la conduite de politiques économiques ou sociales : l’IPCH (IPC harmonisé au niveau européen) constitue ainsi un indicateur essentiel pour la Banque Centrale Européenne, chargée de la politique monétaire européenne. Et l’économie est de plus en plus réactive aux informations conjoncturelles : l’impact sur les marchés boursiers de certains indicateurs (IPI par exemple) peut être important.
- Les indices ont parfois un rôle réglementaire important : bon nombre de loyers sont indexés aujourd’hui sur l’indice du coût de la construction (bientôt sur une moyenne pondérée de trois indices, dont l’ICC), l’IPC intervient dans l’évolution du SMIC, sert à indexer les pensions alimentaires et l’un des critères de Maastricht est basé sur l’IPCH.
- Enfin, de très nombreux contrats privés sont indexés sur des indices publiés par l’Insee : IPP (ex IPVI), ICC ou indices des comptes nationaux par exemple.

Ces différentes utilisations ne représentent pas à proprement parler des problèmes méthodologiques. Elles constituent néanmoins pour le producteur d’indices un contexte qu’il doit connaître et dont il doit tenir compte. On ne donnera qu’un aperçu des questions qui se posent à ce sujet :
- publication : de nombreux indices doivent être publiés au journal officiel et diffusés à une date (et à une heure) indiquée à l’avance, et des précautions de confidentialité strictes doivent prises jusque là.
- transparence des méthodes et des sources utilisées : elle est d’autant plus nécessaire que l’indice est utilisé. Cette transparence est de plus en plus une obligation internationale, dans le cadre de l’Union européenne notamment, laquelle recherche et impose de plus en plus souvent la comparabilité entre Etats membres des principaux indicateurs, fixe leurs dates de disponibilité, etc.
- continuité des séries : il faut avoir en tête que tout arrêt de publication d’une indice pose un problème difficile aux parties prenantes de contrats utilisant cet indice à des fins d’indexation. Il en est de même, à un degrès moindre, en cas de changement de nom, de nomenclature, d’année de référence, etc. L’expérience montre que la diversité des indices utilisés dans des contrat privés est plus importante que les producteurs ne pensent.
- élaboration d’indicateurs complémentaires : il est parfois nécessaire de publier officiellement un indicateur dérivé, comme la moyenne glissante de quatre indices (trimestriels) du coût de la construction.

- Il y a bien une saisonnalité réelle. Faute de quoi on ne fait qu’introduire du « bruit ».

- Ce que l’on veut mesurer correspond bien à un indice désaisonnalisé (ou corrigé des jours ouvrables) et non brut.

La réponse à chacune de ces deux questions peut être délicate. De plus, et indépendamment de la technique CVS-CJO elle-même, il se pose des questions spécifiques au fait qu'elle s'applique à l'élaboration d'indices. Ainsi, la CVS de la somme n'étant pas égale à pas la somme des séries CVS, la question se pose de savoir à quel niveau d'agrégation il faut désaisonnaliser. Deux choix sont possibles :

- Désaisonnaliser à tous les niveaux, au risque de voir apparaître des paradoxes du type : toutes les composantes baissent mais l'agrégation augmente.

- Désaisonnaliser (uniquement) à un niveau fin (par exemple le niveau diffusé le plus fin) et effectuer ensuite les agrégations.

Par ailleurs, le produit de séries CVS n'est pas égal à la série CVS du produit, ce qui peut poser des problèmes de partage de la valeur en volume-prix. Pour résoudre ce problème, une solution courante consiste à calculer l'une des trois séries « CVS » comme résultantes des deux autres, calculées directement. Ainsi, pour l'établissement des comptes nationaux trimestriels, les indicateurs de prix et de volume (respectivement de prix et de valeur dans certains cas) sont désaisonnalisés, et les séries en valeur (resp. de volume) sont calculées à partir de celles en prix et en volume (resp. de prix et de valeur).

c) Les contributions à la variation de l'indice

Il peut être utile de connaître (et donc de définir) la part, que l'on appellera contribution, de chaque composante ou produit i dans l'évolution de l'indice global. Pour cela on cherche à écrire le taux de croissance de l'indice global comme une somme de termes ne faisant intervenir chacun qu'un seul produit. Dans le cas courant d'un indice de Laspeyres, une telle décomposition s'obtient de façon naturelle en partant de l'écriture $I_t = \Sigma\ w_i\ I^i_t$ où w_i est la pondération de la composante i dans l'indice global (la somme des w_i étant égale à 1).

On peut écrire sans difficulté :
$$\frac{I_t - I_{t-1}}{I_{t-1}} = \sum_i w_i \left(\frac{I^i_t - I^i_{t-1}}{I_{t-1}} \right) = \sum_i w_i\ \frac{I^i_{t-1}}{I_{t-1}}\ \frac{I^i_t - I^i_{t-1}}{I^i_{t-1}} \qquad (1)$$

et l'on est donc amené à poser :

$$C_i(t) = w_i\ \frac{I^i_{t-1}}{I_{t-1}}\ \frac{I^i_t - I^i_{t-1}}{I^i_{t-1}}$$

La contribution de la composante i au taux de croissance de l'indice synthétique est donc le produit de 3 termes :

- Le taux de croissance de l'indice de la composante i : $(I^i_t - I^i_{t-1}) / I^i_{t-1}$.
- La pondération de la composante i dans l'indice agrégé : w_i qui représente la part en valeur de la composante i, lors de la période de base.
-L'évolution entre la période de base et la période t-1 de l'indice de la composante i, relativement à celle de l'indice global : I^i_{t-1} / I_{t-1}.

L'intervention de ce dernier terme est moins intuitive que celle des deux premiers. Elle s'analyse cependant clairement. On peut en effet lire la formule (1) de la façon suivante :

- Le taux de croissance de l'indice agrégé est la moyenne des taux de croissance des différentes composantes, pondérées par le produit des deux autres termes.

- Ce produit correspond aux parts en valeurs w_i (relatives à la période de base) actualisées par l'évolution I^i_{t-1} / I_{t-1} des prix relatifs (s'il s'agit d'un indice de prix) jusqu'en t-1 et non par celle des valeurs elles-mêmes comme dans le chaînage. Ceci s'explique naturellement si l'on pense à l'expression du Laspeyres pour laquelle on fait varier les prix en maintenant les quantités constantes. On voit ainsi qu'un indice de Laspeyres à base fixe peut mal vieillir.

11. Exemples d'indices officiels

Il s'agit ici de donner principalement des exemples montrant la variété des solutions adoptées, de laquelle se dégage cependant une tendance à l'utilisation privilégiée des indices de Laspeyres, chaînés ou non, et du couple Laspeyres-Paasche pour un partage volume-prix.

Indices de prix à la consommation (IPC, cf. [8] et note d'actualisation sur internet pour la base 1998) : indices mensuels destinés à mesurer l'évolution des prix de la consommation des ménages (21 000 indices élémentaires basés sur 160 000 relevés mensuels et 40 000 tarifs); Laspeyres ; chaînage annuel (pondérations fixes pendant 12 mois égales à la consommation des ménages tirées des comptes nationaux annuels) ; existence d'un panier variable et de méthode hédonique; indice d'ensemble également en CVS.

Indices de la production industrielle (IPI, cf. [12]) : indices mensuels destinés à mesurer l'activité en volume des branches industrielles (en base 95, 671 séries -témoin couvrant 261 postes de la nomenclature d'activité); Laspeyres à base fixe; changement de base tous les 5 ans ; pondérations = valeur ajoutée tirée des comptes nationaux ; brut et CVS-CJO.

Indices de chiffre d'affaires (ex CA3) : indices destinées à mesurer l'activité des secteurs industrie-commerce-services en valeur (cf. ci-dessus). Les indices élémentaires sont calculés au niveau NAF 700, à partir d'échantillons d'entreprises dans les déclaration fiscales TVA ; des maillons annuels sont calculés à échantillon constant sur deux années puis chaînés. Les indices synthétiques sont calculés comme indices de Laspeyres à base fixe, les pondérations étant les chiffres d'affaires de l'année de base déduits des statistiques structurelles d'entreprises. Indices bruts et CVS-CJO.

Indices de prix des produits agricoles à la production (IPPAP) : indices mensuels ; Laspeyres à base fixe ; pondération tirées des comptes agricoles ; changement de base tous les 5 ans ; existence d'un panier variable ; brut et CVS (en fait une minorité des séries font l'objet de CVS).

Indices de prix à la production (IPP, ex IPVI) : indices mensuels ; Laspeyres à base fixe ; changement de base tous les 5 ans ; existence de méthodes hédoniques.

Indices des **comptes nationaux annuels** français (et européens) utilisés pour les agrégations : volume = Laspeyres et prix = Paasche ; chaînage annuel ; uniquement brut.

Indice des **comptes nationaux trimestriels** français : volume = Laspeyres et prix = Paasche ; base fixe ; changement de base tous les 5 ans ; brut et CVS-CJO.

Indices des **comptes nationaux américains** : indices de volume et de prix = Fisher chaînés.

Indice du coût de la construction (ICC, cf.[7]) : trimestriel ; méthodologie complexe et originale mais indice s'apparentant à un Paasche : on calcul un volume en valorisant la production (de logements neufs) du trimestre courant à l'aide d'un bordereau de prix fixe ; chaînage annuel pour prendre en compte l'évolution de la structure des logements (individuel, collectif, etc.).

L'indice des prix à la consommation surestime-t-il ou sous-estime-t-il l'évolution des prix ?

L'importance de l'IPC est telle que de façon récurrente la question de l'existence d'un certain biais dans la mesure de l'évolution des prix fait l'objet d'un débat. On illustrera ce propos en rappelant que dans un passé assez récent, la possibilité d'une sous-évaluation ou d'une sur-estimation s'est posée. Même s'il n'est pas question ici d'arbitrer ces débats, on fera référence aux analyses détaillées qui ont été publiées par l'Insee.

1) Une éventuelle sous-estimation de l'indice est régulièrement avancée du fait de son champ qui exclut, pour des raisons conceptuelles, l'acquisition des logements et n'impute pas de loyer aux propriétaires occupants. Mais la question d'une sous-évaluation de l'IPC a émérgé d'une façon spécifique à l'occasion du passage à l'euro : les médias se sont faits l'échos d'un sentiment plus ou moins général des ménages comme quoi le passage à l'euro avait donné lieu à des hausses de prix que l'IPC n'aurait que partiellement reflétées. Dans son édition 2003-2004 de « L'économie française », l'Insee a publié un dossier sur « L'inflation au moment du passage à l'euro » qui estime à 0,26% l'effet du passage à l'euro sur 2001-2002 (0,10% pour le mois de janvier 2002), ce qui n'est pas de nature à provoquer un sentiment de hausse important. L'analyse effectuée conduit à penser que les ménages accordent psychologiquement un poids trop fort aux prix des biens ou services qu'ils achètent fréquemment, produits qui ont justement connus de fortes hausses dont l'origine n'est pas toujours liée au passage à l'euro.

2) En décembre 1996, la « Commission Boskin » rend un rapport, à la demande du Sénat américain, sur une possible sur-estimation de l'inflation au Etats-Unis. Le retentissement de ce rapport en Europe provient à la fois de l'ampleur du biais estimé (1,1% par an) et par le fait que certains arguments méthodologiques s'appliquent aussi potentiellement aux indices européens. La commission Boskin estime que plusieurs problèmes insuffisamment résolus tendent à sur-estimer l'indice américain :
- En ce qui concerne la prise en compte des produits nouveaux ou des changements de caractéristiques des produits, la sous-estimation de l'amélioration de la qualité conduirait à biaiser la mesure de l'inflation.
- Une prise en compte insuffisante des substitutions possibles dans la consommation des différents produits entrainerait une sur-estimation de l'indice par rapport à une vision en terme d'utilité pour le consommateur.
- Les baisses de prix liées aux nouveaux circuits de distribution (hard-discounters par exemple) seraient mal prises en compte.

L'Insee, en 1997, a fait paraître dans sa revue Economie et statistique (cf. [13]) une étude serrée pour examiner ces critiques dans le cas des méthodes et des données françaises. La conclusion est nuancée : les arguments avancés par la commission Boskin traduisent bien une présomption de sur-estimation, mais qui serait dans le cas français d'une ampleur nettement plus faible.

Indices de valeur unitaire du commerce extérieur (IVU) : indices mensuels ; à un niveau extrêmement fin (mais parfois hétérogène), il s'agit de prix moyens. L'agrégation se fait à l'aide d'indices de Paasche à base fixe ; changement de base tous les 5 ans. Les indices de volume étant calculés en divisant les valeurs par les IVU, il s'agit en fait d'un partage volume-prix.

Bibliographie

[1] « Indices statistiques. Quels outils pour quelles mesures ?», J. Vacher, Insee Méthodes, n° 15, 1991.

[2] « The making of index numbers » I. Fisher. Boston. Houghton Mifflin. 1922

[3] " Généralité sur les indices" M. Lévy. Polycopié de l'Ensae. 1967.

[4] "Exact and superlatif index numbers", W.E. Diewert, Journal of Econometrics, 4, 1976.

[5] « Les indices à utilité constante : une référence pour mesurer l'évolution des prix », F. Magnien et J.Pougnard, Économie et Statistique, n° 335, 2000.

[6] « Les indices additifs : une autre approche de la théorie des indices et de l'étude du chaînage », J.-P. Berthier, Courrier des statistiques, n° 104, Insee, décembre 2002.

[7] « L'indice du coût de la construction » V. Le Calonnec. Insee méthodes n°17. 1992.

[8] « Pour comprendre l'indice des prix. Edition 1998. » Insee méthodes n°81-82, janvier 1999.

[9] "Divisia index numbers", C.R. Hulten, Econometrica, 41, 1973, n° 6, pp. 1017-25.

[10] « Le chaînage des indices : entre nécessité pratique et justification théorique », J.P. Berthier, Courrier des statistiques N°108. Décembre 2003.

[11] « Réflexions sur les différentes notions de volume dans les comptes nationaux : comptes aux prix d'une année fixe ou aux prix de l'année précédente, séries chaînées », J.P. Berthier. Document de travail G 2002 /08 de l'Insee.

[12] « L'indice de la production industrielle. Sources et méthodes de la base 95 » J. Harnois. Insee méthodes n°104. 2003.

[13] « L'indice des prix à la consommation surestime-t-il l'inflation ? » F. Lequiller. Economie et statistique, pages 3 à 32. Insee 1997.

ANNEXE 1
Quelques exemples pathologiques

1. Evolution paradoxale d'une série calculée avec un indice de Paasche

Considérons le cas très simple suivant, à deux produits et trois périodes.

	Période 0	Période 1	Période 2
Prix produit 1	100	120	125
Prix produit 2	100	80	85
Pondération (en valeur) produit 1		0,5	0,3
Pondération (en valeur) produit 2		0,5	0,7

La série correspondante de l'indice de prix, calculée avec un Paasche ($P_{t/0}$) est la suivante :

	Période 0	Période 1	Période 2
$P_{t/0}$	100	96	94

Ainsi, entre la période 1 et la période 2, l'indice global diminue, alors que chacune des composantes (les deux indices élémentaires) augmente : le prix 1 passe de 120 à 125 et le prix 2 de 80 à 85.

Ce paradoxe est rendu possible par le fait que l'évolution d'un Paasche correspond au quotient de deux moyennes (harmoniques) pondérées par des coefficients différents au numérateur et au dénominateur, puisque pour un Paasche les pondérations correspondent à la période courante. Avec un indice de Laspeyres un tel phénomène, difficile à expliquer à l'occasion de la publication d'indices officiels, ne peut pas se produire puisque les pondérations restent fixes.

Le trait a été forcé dans l'exemple numérique ci-dessus. Mais le problème est loin d'être uniquement théorique, cette situation se produisant réellement de temps en temps.

2. Problème d'agrégation d'un Fisher

Considérons l'exemple ci-dessous, avec une nomenclature comportant deux postes A et B, ce dernier se subdivisant lui-même en deux, B1 et B2.

	prix période 0	prix période 1	quantité période 0	quantité période 1
Poste A	1	1,1	1	1
Poste B1	1	1	1	4,75
Poste B2	1	1,3	1	1

On peut calculer l'indice de Fisher des prix du poste B puis l'indice global à partir des postes A et B ou bien directement à partir des trois postes A, B1 et B2. On obtient les résultats suivants (en reprenant aussi l'indice du poste A).

	Indice de Fisher
Poste A	110,0
Poste B	110,0
Global à partir de A et B	110,0
Global à partir de A, B1 et B2	109,6

La propriété d'agrégation n'est donc pas vérifée. Cet exemple théorique montre même que l'indice de chaque poste du premier niveau de nomenclature peut être identique (1,100) sans que l'indice global soit pour autant égal à cette valeur commune.

3. Problème lié au chaînage

Le chaînage est potentiellement mauvais lorsque la situation finale revient à la situation intiale après s'en être écartée. L'exemple classique est de considérer le chaînage d'un indice de Laspeyres avec les données suivantes :

	Période 0	Période1	Période 2
Prix produit 1	1	2	1
Prix produit 2	2	1	2
Pondération (en valeur) produit 1	0,5	0,5	0,5
Pondération (en valeur) produit 2	0,5	0,5	0,5

On obtient que $L_{1/0} = L_{2/1} = 1,25$ d'où un indice chaîné de $Lch_{2/0} = 1,25 \times 1,25 = 1,56$ environ, alors l'indice de Laspeyres $L_{2/0}$ est bien égal à 1 puisque chacun des deux indices élémentaires est lui-même égal à l'unité. La remarque suivante est également classique : si l'on considère un phénomène cyclique, où la situation en t+2 est identique à celle en t et où le cycle est décrit par le tableau ci-dessus, on obtient alors $L_{2n/0} = (1,25)^{2n}$ qui tend vers l'infini avec n alors que la situation en 2n est identique à celle en 0.

En fait, cet exemple n'est pas aussi destructeur qu'il n'y paraît. La configuration des données permet en effet d'espérer que c'est l'indice de Laspeyres lui-même qui est incapable de fournir un bon résumé des données. Et il est vrai que l'utilisation d'un indice de Fisher supprime le problème comme le montre le tableau suivant :

	t1/t0	t2/t1	Chaînage
Laspeyres	1,25	1,25	$(1,25)^2$
Paasche	0,8	0,8	$(0,8)^2$
Fisher	1	1	1

On peut donc, à ce stade, penser que ce n'est pas le principe même du chaînage qui est en cause, mais le choix du type d'indice. Malheureusement, cet espoir doit être abandonné. Pour le montrer, il faut cependant prendre un exemple un peu plus compliqué (avec notamment une période de plus).

	Période 0	Période1	Période 2	Période 3
Prix produit 1	1	2	2	1
Prix produit 2	2	1	2	2
Pondération (en valeur) produit 1	0,6	0,3	0,5	0,6
Pondération (en valeur) produit 2	0,4	0,7	0,5	0,4

On obtient alors :

	t1/t0	t2/t1	t3/t2	Chaînage
Laspeyres	1,4	1,7	0,750	1,785
Paasche	0,645	1,333	0,625	0,538
Fisher	0,950	1,506	0,685	0,980

C'est donc bien le principe du chaînage lui-même qui pose parfois problème. L'espoir que si l'on chaînait en continu (indice de Divisia) on parviendrait à résoudre le problème doit également être abandonné, mais ceci dépasse le cadre ce document (cf. [6]).

Le dernier tableau ci-dessus, qui décrit des variations extrêmes, montre toutefois la robustesse de l'indice de Fisher. Le chaînage utilisant cet indice est en effet infiniment meilleur que le chaînage effectué à partir du Laspeyres ou du Paasche.

4. Problème spécifique aux variations de stocks

Les variations de stocks posent des problèmes spécifiques liés au fait qu'elles peuvent être positives ou négatives. L'exemple suivant montre qu'avec des données tout à fait banales, les résultats en terme d'indices peuvent être catastrophiques.

Plaçons nous dans le cadre des comptes nationaux et considérons deux variations de stocks (par exemple pour deux produits) que l'on notera VS1 et VS2. Le tableau suivant fournit les valeurs pour une période 0 et une période 1 ainsi que les volumes de la période 1 au prix de 0.

	VS1	VS2	VS totales
Valeur en 1	50	-51	-1
Volume en 1	49	-48	+1
Valeur en 0	48	-49	-1

L'évolution des prix des variations de stocks est de 2% environ (50/49) pour le produit 1 et d'environ 6% (-51/-48) pour le produit 2. L'indice de prix Paasche, conformément aux comptes nationaux de façon à conserver les égalités comptables en volume, est de -1/+1 = -1 soit une évolution de -200%. Ce résultat est (très largement) en dehors de la fourchette définie par les deux indices élémentaires (2% et 6%) car, si l'indice synthétique est bien une moyenne de ces deux indices, les pondérations sont l'une positive, l'autre négative.

Si l'on utilisait un indice de Laspeyres pour les prix, on obtiendrait +208%. Ces deux indices étant totalement contradictoires, on peut avoir le réflexe de se tourner vers le Fisher. Malheureusement, celui-ci n'est même pas défini puisque Laspeyres et Paasche sont de signes opposés.

La conclusion qui s'impose est que pour les variations de stocks, il convient de raisonner en différences : dans le cas présent, la stabilité des variations de stocks en valeur se décompose en +2 dû à un effet volume et -2 dû à un effet prix. Les +2 de l'effet volume peuvent éventuellement être convertis en contribution (à la croissance du PIB, par exemple, comme cela est fait classiquement en comptabilité nationale au niveau global). Mais il est tout de même bien désagréable de devoir changer de méthode et de discours spécialement pour les variations de stocks.

On peut ajouter que le problème se pose également, en théorie, avec la valeur ajoutée : le problème est formellement le même en remplaçant VS1 par la production et VS2 par « moins les consommations intermédiaires ». Il est cependant rare de se trouver dans une configuration réelle qui présente effectivement des problèmes de cette ampleur. Il faudrait pour cela que la valeur ajoutée soit suivant les cas positive ou négative, ce qui est heureusement très inhabituel à un niveau macroéconomique.

ANNEXE 2

TABLEAUX
de données utilisées dans le texte

Indices élémentaires

FRET	1994	1995	1996	1997	1998	1999	2000
Quantité (en milliards tonnes.km)	44,8	44,6	44,7	45,2	45,4	45,9	46,5
indices (100 en 1994)	100,0	99,6	99,8	100,9	101,3	102,5	103,8
Prix (en cent. d'euros)	3,71	3,64	3,53	3,45	3,43	3,36	3,30
indices (100 en 1994)	100,0	98,1	95,2	93,1	92,5	90,6	89,0
Valeur (en millions d'euros)	1661,71	1622,74	1578,94	1560,33	1557,62	1541,82	1535,18
indices (100 en 1994)	100,0	97,7	95,0	93,9	93,7	92,8	92,4

Sources : Prévisions 1994 de trafic SNCF

Indices synthétiques

FRET	**1994**	**1995**	**1996**	**1997**	**1998**	**1999**	**2000**
Quantité (en milliards tonnes.km)	44,8	44,6	44,7	45,2	45,4	45,9	46,5
indices (100 en 1994)	100,0	99,6	99,8	100,9	101,3	102,5	103,8
Prix (en cent. d'euros)	3,71	3,64	3,53	3,45	3,43	3,36	3,30
indices (100 en 1994)	100,0	98,1	95,2	93,1	92,5	90,6	89,0
Valeur (en millions d'euros)	1661,71	1622,74	1578,94	1560,33	1557,62	1541,82	1535,18
indices (100 en 1994)	100,0	97,7	95,0	93,9	93,7	92,8	92,4
VOYAGEURS	**1994**	**1995**	**1996**	**1997**	**1998**	**1999**	**2000**
Quantité (en milliards voyageurs.km)	43,4	46,2	48,6	50,9	51,4	51,8	52,3
indices (100 en 1994)	100,0	106,5	112,0	117,3	118,4	119,4	120,5
Prix (en cent. d'euros)	6,94	7,10	7,20	7,26	7,29	7,33	7,36
indices (100 en 1994)	100,0	102,3	103,7	104,6	105,1	105,5	106,1
Valeur (en millions d'euros)	3012,40	3280,30	3497,45	3694,44	3748,17	3794,63	3850,48
indices (100 en 1994)	100,0	108,9	116,1	122,6	124,4	126,0	127,8

Indices synthétiques	**1994**	**1995**	**1996**	**1997**	**1998**	**1999**	**2000**
Laspeyres volume (t/94)	100,0	104,0	107,6	111,5	112,4	113,3	114,6
Paasche volume (t/94)	100,0	104,1	107,9	111,9	112,8	113,9	115,2
L volume (t/t-1)		104,0	103,5	103,6	100,8	100,9	101,1
L volume chaîné	100,0	104,0	107,7	111,6	112,5	113,5	114,7

Source : prévisions 1994 de trafic SNCF

Homogénéité - hétérogénéité

FRET	**1994**	**1995**	**1996**	**1997**	**1998**	**1999**	**2000**
Quantité (en milliards tonnes.km)	44,8	44,6	44,7	45,2	45,4	45,9	46,5
train entier	22,5	22,3	22,0	21,8	21,6	21,4	21,2
transport combiné	8,6	9,0	10,2	11,3	11,7	13,0	14,2
wagon isolé	13,7	13,3	12,5	12,1	12,1	11,5	11,1
quantité (100 en 1994)	100,0	99,6	99,8	100,9	101,3	102,5	103,8
Paasche volume (100 en 1994)	100,0	98,7	96,8	96,2	96,3	95,3	94,9
Prix(moyen en cent. d'euros)	3,71	3,64	3,53	3,45	3,43	3,36	3,30
train entier	3,22	3,20	3,19	3,17	3,17	3,17	3,17
transport combiné	1,92	1,88	1,84	1,81	1,80	1,80	1,80
wagon isolé	5,64	5,56	5,52	5,49	5,47	5,47	5,47
prix moyen (100 en 1994)	100,0	98,1	95,2	93,1	92,5	90,6	89,0
Laspeyres prix (100 en 1994)	100,0	98,9	98,2	97,6	97,4	97,4	97,4
Valeur (en millions d'euros)	1661,71	1622,74	1578,94	1560,33	1557,62	1541,82	1535,18
train entier	723,75	713,92	700,96	691,26	684,92	678,58	672,24
transport combiné	165,19	168,76	188,15	205,00	210,47	233,86	255,44
wagon isolé	772,76	740,06	689,83	664,07	662,22	629,39	607,49
indice (100 en 1994)	100,0	97,7	95,0	93,9	93,7	92,8	92,4
VOYAGEURS	**1994**	**1995**	**1996**	**1997**	**1998**	**1999**	**2000**
Quantité (en milliards voyageurs.km)	43,4	46,2	48,6	50,9	51,4	51,8	52,3
TGV	21,4	25,8	29,1	32,2	33,1	33,9	34,7
Classique	22,0	20,4	19,5	18,7	18,3	17,9	17,6
quantité (100 en 1994)	100,0	106,5	112,0	117,3	118,4	119,4	120,5
Paasche volume (100 en 1994)	100,0	108,1	114,9	121,3	123,0	124,4	125,9
Prix (moyen en cent. d'euros)	6,94	7,10	7,20	7,26	7,29	7,33	7,36
TGV	7,71	7,82	7,88	7,90	7,93	7,96	7,99
Classique	6,19	6,19	6,17	6,16	6,14	6,13	6,13
prix moyen (100 en 1994)	100,0	102,3	103,7	104,6	105,1	105,5	106,1
Laspeyres prix (100 en 1994)	100,0	100,8	101,1	101,1	101,2	101,3	101,5
Valeur (en millions d'euros)	3012,40	3280,30	3497,45	3694,44	3748,17	3794,63	3850,48
TGV	1650,78	2017,72	2293,55	2542,79	2623,95	2697,71	2771,95
Classique	1361,62	1262,58	1203,90	1151,65	1124,22	1096,92	1078,53
indices (100 en 1994)	100,0	108,9	116,1	122,6	124,4	126,0	127,8

Indices synthétiques	**1994**	**1995**	**1996**	**1997**	**1998**	**1999**	**2000**
1 segment fret + 1 voy							
Laspeyres prix	100,0	100,8	100,7	100,5	100,6	100,2	100,0
1 fret + 2 voyageurs							
Laspeyres prix	100,0	99,8	99,0	98,2	98,1	97,5	97,1
3 fret + 1 voyageur							
Laspeyres prix	100,0	101,1	101,7	102,1	102,3	102,6	103,0
3 fret + 2 voyageurs							
Laspeyres prix	100,0	100,1	100,1	99,8	99,8	99,9	100,0

Source : prévisions 1994 de trafic SNCF

L'exemple de la betterave à sucre

PRODUCTION DE BETTERAVE A SUCRE			
Quantité	**1997**	**évolution**	**1998**
en milliers de tonnes			
- QUOTA A	15126	104,7	15834
- QUOTA B	4509	104,7	4721
- Hors quota	6876	62,8	4316
TOTAL BETT A SUCRE	**26511**	93,8	**24871**
Prix	**1997**	**évolution**	**1998**
en euros par tonne			
- QUOTA A	47,52	97,2	46,19
- QUOTA B	30,48	94,8	28,89
- Hors quota	14,07	62,3	8,77
Valeur	**1997**	**évolution**	**1998**
en milliers d'euros			
- QUOTA A	718788	101,8	731372
- QUOTA B	137434	99,2	136390
- Hors quota	96745	39,1	37851
TOTAL BETT A SUCRE	952967	95,0	905613
	Evolutions 1998/1997		
Indice de prix			
1 seul produit	**101,3**		
3 produits différents (Paasche)	**94,6**		
Indice de volume			
1 seul produit	**93,8**		
3 produits différents (Laspeyres)	**100,4**		
Indice de valeur			
1 seul produit	**95,0**		
3 produits différents	**95,0**		

Source : Comptes nationaux Insee

Exemple d'arbre d'agrégation

INDICE DE LA PRODUCTION INDUSTRIELLE (données cjo-cvs, base 100 en 2000)

		2004										
Identifiant NES et (MIG)	pondération (10000=EB...EH)	Fév.	Mars	Avril	Mai	Juin	Juil.	Août	Sept.	Oct.	Nov.	Déc.
INDUSTRIE (EB...EG)	**8617**	**100,5**	**100,9**	**100,4**	**100,9**	**101,7**	**101,0**	**99,4**	**102,2**	**101,2**	**101,6**	**102,2**
INDUSTRIE hors énergie et IAA (EC...EF)	**6710**	**99,0**	**99,0**	**99,4**	**99,9**	**100,9**	**100,4**	**98,2**	**101,7**	**100,6**	**100,5**	**101,3**
EB Industries agricoles et alimentaires	**760**	**102,3**	**102,3**	**102,3**	**101,5**	**103,0**	**102,9**	**101,8**	**101,5**	**101,0**	**102,8**	**101,3**
EC Biens de consommation	**1257**	**101,8**	**102,8**	**101,2**	**100,6**	**102,7**	**101,2**	**101,8**	**102,7**	**100,4**	**101,5**	**102,6**
C1 Habillement, cuir	190	58,6	59,0	55,9	52,3	60,2	57,3	55,8	57,5	51,4	53,7	54,9
C2 Produits de l'édition, imprimés ou reproduits	341	97,1	101,0	97,1	96,3	98,5	95,4	97,9	98,7	94,3	95,8	97,9
C3 Produits pharmaceutiques, de parfumerie et d'entretien	499	126,6	126,7	127,1	126,8	127,2	126,8	128,0	127,9	128,2	128,2	128,7
C4 Equipements du foyer	226	90,3	89,7	88,4	90,0	90,6	90,5	88,4	91,1	89,7	91,1	91,8
ED Industrie automobile	**784**	**107,4**	**109,8**	**109,2**	**112,1**	**115,5**	**115,7**	**96,9**	**118,4**	**114,1**	**115,9**	**116,9**
EE Biens d'équipement	**1644**	**100,0**	**98,9**	**101,0**	**100,5**	**101,0**	**101,0**	**101,9**	**103,3**	**102,8**	**102,3**	**102,8**
E1 Bateaux, avions, trains, motos	325	97,6	94,2	98,6	104,2	104,7	99,9	105,7	105,1	108,0	109,8	110,5
E2 Equipements mécaniques	740	104,2	104,1	105,2	103,6	105,3	104,9	104,8	108,0	105,8	106,6	106,6
E3 Equipements électriques et électroniques	579	96,1	94,7	97,0	94,5	93,3	96,7	96,2	96,2	96,1	92,5	93,7
EF Biens intermédiaires	**3026**	**95,2**	**94,7**	**95,3**	**96,2**	**96,4**	**95,7**	**95,0**	**96,1**	**96,0**	**95,1**	**96,0**
F1 Produits minéraux	333	98,8	97,8	98,3	98,7	98,6	97,6	96,6	98,7	97,8	98,4	99,1
F2 Produits de l'industrie textile	176	78,0	77,8	76,1	76,3	76,1	75,8	80,1	77,6	74,7	75,8	74,7
F3 Produits en bois, papier ou carton	311	99,8	98,6	100,1	98,7	99,2	98,9	99,1	99,4	99,9	100,6	100,8
F4 Produits chimiques, en caoutchouc ou plastique	907	98,2	97,1	97,1	99,3	99,9	101,0	97,1	100,4	101,8	98,2	98,4
F5 Métaux et produits métalliques	871	94,6	94,8	94,9	96,3	96,3	92,9	93,8	95,2	94,3	94,8	96,5
F6 Composants électriques et électroniques	428	90,7	91,2	94,2	93,5	94,0	94,4	95,0	92,4	91,8	90,6	92,4
EG Energie	**1146**	**108,2**	**111,4**	**105,1**	**106,1**	**105,1**	**103,6**	**104,8**	**105,4**	**104,9**	**107,1**	**108,3**
G1 Combustibles et carburants	168	102,8	101,3	104,0	91,7	93,4	98,2	95,6	98,4	98,4	91,7	97,9
G2 Eau, gaz, électricité	978	109,2	113,2	105,3	108,6	107,1	104,5	106,3	106,6	106,1	109,8	110,1
MIGS[(1)]												
(050) Biens d'investissement	2412	102,7	102,4	103,8	104,3	105,8	106,0	100,5	108,3	106,7	106,8	107,4
(040) Biens intermédiaires	3058	95,5	95,1	95,6	96,5	96,8	96,0	95,4	96,5	96,4	95,5	96,3
(060) Biens de consommation durables	242	88,3	89,8	88,0	90,0	90,7	89,2	86,8	90,5	88,4	90,9	92,3
(070) Biens de consommation non durables	1758	103,0	103,8	102,8	101,8	103,8	102,8	102,9	103,2	101,4	102,8	103,0
EH Construction	**1383**	**107,6**	**107,5**	**107,2**	**107,8**	**107,6**	**107,2**	**106,3**	**107,0**	**106,8**	**107,2**	**106,6**

Source : Insee

Série des Documents de Travail 'Méthodologie Statistique'

9601 : ‘Une méthode synthétique, robuste et efficace pour réaliser des estimations locales de population’
G. DECAUDIN, J.-C. LABAT

9602 : ‘Estimation de la précision d'un solde dans les enquêtes de conjoncture auprès des entreprises’
N. CARON, P. RAVALET, O. SAUTORY

9603 : ‘La procédure FREQ de SAS® - Tests d'indépendance et mesures d'association dans un tableau de contingence’
J. CONFAIS, Y. GRELET, M. LE GUEN

9604 : ‘Les principales techniques de correction de la non-réponse et les modèles associés’
N. CARON

9605 : ‘L'estimation du taux d'évolution des dépenses d'équipement dans l'enquête de conjoncture : analyse et voies d'amélioration’
P. RAVALET

9606 : ‘L'économétrie et l'étude des comportements. Présentation et mise en œuvre de modèles de régression qualitatifs. Les modèles univariés à résidus logistiques ou normaux (LOGIT, PROBIT)’
S. LOLLIVIER, M. MARPSAT, D. VERGER

9607 : ‘Enquêtes régionales sur les déplacements des ménages : l'expérience de Rhône-Alpes’
N. CARON, D. LE BLANC

9701 : ‘Une bonne petite enquête vaut-elle mieux qu'un mauvais recensement ?’
J.C. DEVILLE

9702 : ‘Modèles univariés et modèles de durée sur données individuelles’
S. LOLLIVIER

9703 : ‘Comparaison de deux estimateurs par le ratio stratifiés et application aux enquêtes auprès des entreprises’
N. CARON, J.C. DEVILLE
9704 : ‘La faisabilité d'une enquête auprès des ménages
1. au mois d'août. 2. à un rythme hebdomadaire’
C. LAGARENNE, C. THIESSET

9705 : ‘Méthodologie de l'enquête sur les déplacements dans l'agglomération toulousaine’
P. GIRARD

9801 : ‘Les logiciels de désaisonnalisation TRAMO & SEATS : philosophie, principes et mise en œuvre sous SAS’
K. ATTAL-TOUBERT, D. LADIRAY

9802 : ‘Estimation de variance pour des statistiques complexes : technique des résidus et de linéarisation’
J.C. DEVILLE

9803 : ‘Pour essayer d'en finir avec l'individu Kish’
J.C. DEVILLE

9804 : ‘Une nouvelle (encore une !) méthode de tirage à probabilités inégales’
J.C. DEVILLE

9805 : ‘Variance et estimation de variance en cas d'erreurs de mesure non corrélées ou de l'intrusion d'un individu Kish’
J.C. DEVILLE

9806 : ‘Estimation de précision de données issues d'enquêtes : document méthodologique sur le logiciel POULPE’
N. CARON, J.C. DEVILLE, O. SAUTORY

9807 : ‘Estimation de données régionales à l'aide de techniques d'analyse multidimensionnelle’
K. ATTAL-TOUBERT, O. SAUTORY

9808 : ‘Matrices de mobilité et calcul de la précision associée’
N. CARON, C. CHAMBAZ

9809 : ‘Echantillonnage et stratification : une étude empirique des gains de précision’
J. LE GUENNEC

9810 : ‘Le Kish : les problèmes de réalisation du tirage et de son extrapolation’
C. BERTHIER, N. CARON, B. NÉROS

9811 : ‘Vocabulaire statistique Français - Chinois - Anglais’
LIU Xiaoyue, CUI Bin

9901 : ‘Perte de précision liée au tirage d'un ou plusieurs individus Kish’
N. CARON

9902 : ‘Estimation de variance en présence de données imputées : un exemple à partir de l’enquête Panel Européen’
N. CARON

0001 : ‘L’économétrie et l’étude des comportements. Présentation et mise en oeuvre de modèles de régression qualitatifs. Les modèles univariés à résidus logistiques ou normaux (LOGIT, PROBIT) ’ (version actualisée)
S. LOLLIVIER, M. MARPSAT, D. VERGER

0002 : ‘Modèles structurels et variables explicatives endogènes’
Jean-Marc Robin INRA-LEA et CREST-INSEE.

0003 : ‘L’enquête 1997-1998 sur le devenir des personnes sorties du RMI- Une présentation de son déroulement’
D. ENEAU, D. GUILLEMOT

0004 : ‘Plus d’amis, plus proches? Essai de comparaison de deux enquêtes peu comparables’
O. GODECHOT

0005 : ‘Estimation dans les enquêtes répétées : Application à l’Enquête Emploi en Continu’
N. CARON, P. RAVALET

0006 : ‘Non-parametric approach to the cost-of-living index’
F. MAGNIEN, J. POUGNARD

0101 : ‘Diverses Macros SAS : Analyse exploratoire des données, Analyse des séries temporelles’
D. LADIRAY

0102 : ‘Econométrie linéaire des panels : une introduction’
T. MAGNAC

0201 : ‘Application des méthodes de calage à l'enquête EAE-Commerce’
N. CARON

0203 : ‘General principles for data editing in business surveys ans how to optimise it’
P. RIVIERE

0301 : ‘Les modèles logit polytomiques non ordonnés :théorie et applications’
C. AFSA ESSAFI

0401 : ‘Enquête sur le patrimoine des ménages - Synthèse des entretiens monographiques’
V. COHEN, C. DEMMER

0402 : ‘La macro SAS CUBE d'échantillonnage équilibré’
S. ROUSSEAU, F. TARDIEU

M0501 : ‘Correction de la non-réponse et calage de l’enquête Santé 2002’
N. CARON, S. ROUSSEAU

M0502 : ‘Correction de la non-réponse par repondération et par imputation’
N. CARON

M0503 : ‘Introduction à la pratique des indices statistiques - notes de cours’
J-P. BERTHIER

Série des Documents de Travail 'Méthodologie de Collecte'

C0201 : ‘Comportement face au risque et à l'avenir et accumulation patrimoniale - Bilan d'une expérimentation’
L. ARRONDEL, A. MASSON, D. VERGER

0202 : ‘Enquête Méthodologique Information et Vie Quotidienne - Tome 1 : bilan du test 1, novembre 2002’
L-A. VALLET, G. BONNET, J-C. EMIN, J. LEVASSEUR, T. ROCHER,
P. VRIGNAUD, X. D’HAULTFOEUILLE, F. MURAT, D. VERGER, P. ZAMORA