

## Plans de sondage pour les indices des prix à la consommation

Alan H. Dorfman, Janice Lent, Sylvia G. Leaver et Edward Wegman<sup>1</sup>

### Résumé

L'échantillonnage en vue d'estimer un indice des prix à la consommation (IPC) est assez compliqué et requiert généralement la combinaison de données provenant d'au moins deux enquêtes, l'une donnant les prix et l'autre, la pondération par les dépenses. Deux approches fondamentalement différentes du processus d'échantillonnage – l'échantillonnage probabiliste et l'échantillonnage par choix raisonné – ont été vivement recommandées et sont utilisées par divers pays en vue de recueillir les données sur les prix. En construisant un petit « univers » d'achats et de prix à partir de données scannées sur les céréales, puis en simulant diverses méthodes d'échantillonnage et d'estimation, nous comparons les résultats de deux approches du plan de sondage et de l'estimation, à savoir l'approche probabiliste adoptée aux États-Unis et l'approche par choix raisonné adoptée au Royaume-Uni. Pour la même quantité d'information recueillie, mais avec l'utilisation d'estimateurs différents, les méthodes du Royaume-Uni semblent offrir une meilleure exactitude globale du ciblage d'un indice superlatif des prix à la consommation basé sur la population.

Mots clés : Indice élémentaire; échantillonnage avec probabilité proportionnelle à la taille; échantillonnage par choix raisonné; données scannées; indice superlatif.

### 1. Contexte

Du début à la fin, l'échantillonnage en vue d'établir un indice des prix à la consommation (IPC) représente l'une des entreprises d'échantillonnage les plus compliquées. La population cible est difficile à définir, le domaine approprié des articles est débattu, les définitions des ingrédients bruts, c'est-à-dire les prix, les quantités et les articles, sont ambiguës et mises en doute. L'estimateur ultime – l'estimateur de l'IPC d'ensemble – repose sur des données provenant d'au moins deux enquêtes, l'une donnant les prix et l'autre, les « pondérations ». Sous le niveau des « articles composites » (ou « strates d'articles »), c'est-à-dire des groupes d'articles dont il est supposé que le mouvement des prix est homogène, il n'existe habituellement pas de moyen rentable de tenir à jour les poids d'échantillonnage. Le débat se poursuit donc quant au choix approprié parmi divers estimateurs simples de la variation du prix pour les catégories d'articles, c'est-à-dire les « indices élémentaires ». La méthode appropriée d'agrégation de ces variations des prix, au moyen des pondérations, fait également l'objet de débats.

On distingue deux grandes approches de l'échantillonnage en vue de relever les prix : l'échantillonnage probabiliste et l'échantillonnage par choix raisonné, ou échantillonnage au jugé. L'approche d'échantillonnage la plus généralement reconnue consiste habituellement à injecter un élément aléatoire dans le processus de sondage et à se fonder sur cet élément aléatoire pour faire des inférences au sujet des caractéristiques de population que l'on veut étudier, c'est-à-dire un échantillonnage probabiliste ou « basé sur un plan de sondage »; voir, par exemple, Särndal, Swensson et

Wretman (1992). Cette approche n'a pas toujours été tenue pour acquise. Au début du XX<sup>e</sup> siècle, l'échantillonnage « par choix raisonné » ou « représentatif » était considéré comme une option viable, voire préférable. Plus récemment, l'école de pensée « prédictionniste » de Royall a contesté les hypothèses basées sur le plan de sondage; voir, par exemple, Valliant, Dorfman et Royall (2000).

Aux États-Unis, toutes les enquêtes reliées à l'IPC sont réalisées en utilisant des méthodes d'échantillonnage probabilistes complexes. Ailleurs dans le monde, la plupart des IPC sont construits à partir d'échantillons au jugé, dans lesquels les spécialistes des différentes strates d'articles choisissent des catégories plus ou moins générales ou sélectives d'articles pour lesquelles des agents de terrain relèvent les prix. La raison fondamentale de cette approche est la difficulté qu'il y a à obtenir toutes les données sur la pléthore d'articles vendus et sur les endroits où ils sont vendus nécessaires pour que l'échantillonnage probabiliste soit faisable.

L'aspect intéressant est que fort peu d'évaluations de l'exactitude relative des diverses approches d'échantillonnage ont été faites. En réalité, il n'est pas certain que ce genre d'évaluation soit réalisable. Même pour les pays les plus petits, l'indice des prix sous-jacent calculé en se basant sur la population totale d'articles, ou indice des prix de population, englobe un si grand nombre de transactions portant sur un si grand nombre d'articles dans un si grand nombre d'emplacements qu'il est inaccessible. De surcroît, la population d'articles sur le marché varie constamment, ce qui complique l'application des formules d'indice de population classiques. Alors, comment peut-on juger de la

1. Alan H. Dorfman, Office of Survey Methods Research et Sylvia G. Leaver, Office of Prices and Living Conditions, U.S. Bureau of Labor Statistics, 2 Massachusetts Ave NE, Washington, D.C., États-Unis, 20212; Janice Lent, U.S. Bureau of Transportation Statistics, 400 7<sup>th</sup> Street, SW, Washington, D.C., États-Unis, 20590; Edward Wegman, Center for Computational Statistics, George Mason University, Fairfax, VA, États-Unis, 22030.

justesse relative de diverses estimations basées sur un échantillon par rapport à la « valeur réelle »? De surcroît, dans la plupart des cas, on ne dispose même pas d'information échantillonnale pour l'un des éléments clés de l'indice de population, à savoir les *quantités* d'articles vendus, de sorte que même la construction d'une population artificielle à partir de données d'échantillon en vue d'une évaluation n'a pas été possible.

La disponibilité relativement récente de données *scannées*, aux États-Unis et ailleurs, offre une occasion sans précédent d'évaluer les approches d'échantillonnage et les estimateurs. Ces données comprennent les prix *et* les quantités, habituellement sur une base hebdomadaire, de *tous* les articles vendus dans une catégorie donnée dans un grand échantillon de points de vente dotés de scanners. Ce genre de données peuvent être utilisées pour construire des populations réalistes de transactions pour lesquelles l'indice des prix réel est *connu*. Nous pouvons alors utiliser diverses méthodes pour échantillonner cette population, construire différentes estimations d'indice d'intérêt et comparer les résultats aux paramètres de population connus. Une étude de ce type, décrite par de Haan, Opperdoes et Schut (1999), semble indiquer que l'« échantillonnage avec seuil d'inclusion » (*cutoff sampling*), c'est-à-dire l'échantillonnage des quelques articles les plus importants (en ce qui a trait aux recettes générées) dans la population, donne de meilleurs résultats que deux grandes approches probabilistes, à savoir l'échantillonnage aléatoire simple (*eas*) et l'échantillonnage avec probabilité proportionnelle à la taille (*ppt*) (où la mesure de taille est, de nouveau, les recettes).

L'une des difficultés que pose toute étude faisant ce genre de comparaison est la nécessité de maintenir des « règles du jeu équitables ». Si l'une des méthodes d'échantillonnage s'appuie, par exemple, sur des données (de population) susceptibles, en réalité, ne pas être disponibles en pratique, tandis qu'une autre ne le fait pas, la comparaison des méthodes est sérieusement minée. De même, si une méthode ne fournit qu'un seul échantillon ou quelques échantillons, et qu'une autre en fournit des milliers, des précautions particulières doivent être prises pour les comparer; en effet, ce genre de comparaison pourrait nécessiter d'importantes restrictions. Étant donné la complexité des méthodes d'échantillonnage et d'estimation utilisées pour le calcul des indices de prix, il n'est pas étonnant que ces difficultés et de nombreuses autres compliquent les expériences conçues en vue de comparer diverses méthodes.

Idéalement, pour comparer les approches, disons, de deux pays, nous imiterions entièrement le processus complexe d'échantillonnage et d'estimation de chacun et nous évaluerions les coûts. Le même budget serait affecté aux deux processus et nous serions capables, au moyen

d'une mesure préétablie et équitable, d'évaluer la justesse de chaque estimation par rapport à un indice cible connu.

Le présent article porte sur deux études, une grande étude principale et une étude de suivi secondaire plus petite.

L'étude principale est décrite aux sections 2 à 4. La section 2 donne la construction de la population cible. La section 3 expose les méthodologies des « États-Unis » et du « Royaume-Uni », et fournit les renseignements sur les simulations. Aucun effort n'est fait pour évaluer les coûts relatifs (ce qui nous écarte de la situation idéale), mais les approches concurrentes sont rendues aussi égales que possible en ce qui a trait à l'information utilisée. Les résultats, qui donnent l'avantage à l'approche britannique, sont présentés à la section 4.

L'étude de suivi, décrite à la section 5, a pour but d'essayer de dégager les effets des diverses composantes des deux approches, en particulier la méthode d'échantillonnage et la formule de l'indice élémentaire. La section 6 comprend un résumé final et une discussion.

*Note sur les indices cibles.* La littérature sur les indices de prix contient des myriades de formules pour calculer la variation des prix d'une période à une autre. Divers indices sont compatibles avec différentes hypothèses quant au comportement d'achat du consommateur « moyen » en réponse à la variation des prix. Les indices à « panier de consommation fixe », les formules fréquemment employées de Laspeyres et celles, moins fréquemment utilisées, de Paasche, sont compatibles avec l'hypothèse selon laquelle les consommateurs continuent d'acheter les mêmes articles en même quantité quelle que soit la variation des prix relatifs. L'indice de Laspeyres projette les quantités de la période 1 (« période de base ») sur la période 2 (« période courante »), tandis que celui de Paasche applique les quantités de la période 2 à la période 1. L'indice géométrique (ou « de Jevons » ou « *moyenne géométrique* »), habituellement pondéré par les parts des dépenses à la période de base, suppose que le consommateur ajuste les quantités qu'il achète de telle façon que la part des dépenses pour chaque article demeure constante au cours du temps. Les formules des indices « superlatifs » de Fisher, de Törnqvist et de Walsh, qui reposent sur des données sur les quantités (ou parts des dépenses) pour les deux périodes, ne nécessitent pas ces hypothèses. Les formules de ces indices, avec les exposants  $y$  représentant la période de base,  $y + 1$ , la période courante et  $i$ , l'article acheté, sont données à l'annexe A.

Le débat concernant l'indice d'ensemble cible se résume habituellement à choisir entre l'indice de Laspeyres et l'un des indices superlatifs. La plupart des pays optent pour un indice cible de Laspeyres, mais de solides arguments peuvent être présentés (Diewert 1997) en faveur du choix d'un indice superlatif comme cible (habituellement celui de

Fisher ou de Törnqvist), même si la formule de l'estimateur ne ressemble pas à l'une des formules d'indice superlatif de population. Compte tenu de la forme des agrégats élémentaires utilisés aux États-Unis – *moyenne géométrique* – et du fait que des travaux de recherche antérieurs (Dorfman, Leaver et Lent 1999) ont indiqué que le niveau le plus faible d'estimation peut avoir un effet important, la *moyenne géométrique* pondérée sera incluse parmi les indices cibles possibles. Les cibles pour un domaine particulier sont calculées en se basant sur les prix et sur les quantités de tous les articles compris dans le domaine, conformément aux formules qui figurent à l'annexe A (agrégation à un degré des prix et des quantités).

*Nota :* Ces formules donnent l'illusion d'être simples, mais nécessitent la notation de la section 3 pour leur développement complet. Donc, dans une formule telle que celle de l'indice de Fisher  $F$  (que nous choisirons comme cible dans le corps de l'étude aux sections 2 à 4), «  $i$  » représente un article  $i$  appartenant à une petite catégorie  $c$  (un « ANE » (article de niveau d'entrée) ou « article représentatif » – voir la section 3), où  $c$  est elle-même un sous-ensemble d'une catégorie plus grande. En outre, l'article  $i$  est vendu dans un point de vente particulier  $j$ , classifié comme faisant partie d'une chaîne particulière  $k$ , et situé dans une région géographique échantillonnée particulière, l'unité primaire d'échantillonnage (*upe*) $l$ . Donc, dans le cas de l'indice de population global, l'expression pour une somme  $\sum_i$  est en fait une abréviation de  $\sum_{l=1}^3 \sum_{k=1}^8 \sum_{j \in (k,l)} \sum_{C=1}^4 \sum_{h \in C} \sum_{c \in h} \sum_{i \in (j,c)}$ ; une remarque semblable s'applique à  $\prod_i$ . Brièvement, il existe des sommes et des produits sur la *totalité* des articles dans la population. Les contractions de ce développement complet donneront les indices de population pour les diverses catégories  $C$ , etc.

## 2. La population pour l'étude principale

La source des données sur lesquelles porte l'étude est un ensemble de données scannées sur les céréales pour petit déjeuner couvrant la période de 1995 à 2000 dans trois zones séparées, mais contiguës, d'une grande région métropolitaine. Le U.S. Bureau of Labor Statistics a acheté l'ensemble de données à la société A.C. Nielsen en vue de déterminer s'il est possible d'intégrer des données scannées dans l'IPC des États-Unis; voir Richardson (2000).

Des « populations » artificielles ont été tirées à partir de ces données par la méthode décrite plus loin. Donc, l'étude a pour champ d'observation un univers en apparence limité, celui des céréales, dans un domaine géographique relativement restreint. Toutefois, même cet univers limité permet d'observer des tendances des prix assez divergentes au cours de la période de six ans. Donc, bien que nous ne puissions pas généraliser nos résultats, de façon simple, aux

indices de prix globaux couvrant une vaste gamme hétérogène de produits, nous arriverons peut-être à dégager d'importants éclaircissements quant aux effets de diverses méthodes d'échantillonnage et au comportement d'estimateurs particuliers.

L'ensemble de données portant sur une période de six années nous a permis d'établir des tendances des prix d'assez long terme. Afin que le volume de données demeure raisonnable et pour éviter les complications de la saisonnalité, nous nous sommes limités aux données de février. Pour février de l'année  $y$ , pour chaque article (c'est-à-dire chaque combinaison particulière  $k$  de marque, type, format) dans un point de vente particulier, nous avons combiné les données sur les prix et les quantités pour une période de quatre semaines  $t$  en un prix et une quantité uniques pour un mois, en utilisant la somme des quantités  $q_k^{\text{Feb},y} = \sum_{t \in \text{Feb},y} q_k^t$  vendues durant le mois comme valeur de la quantité et la *valeur unitaire*  $p_k^{\text{Feb},y} = \sum_{t \in \text{Feb},y} q_k^t p_k^t / \sum_{t \in \text{Feb},y} q_k^t$  comme prix. Les valeurs unitaires calculées sur de courtes périodes (par exemple, un mois) correspondant peut-être au sens le plus significatif du prix « moyen » d'un article particulier. L'utilisation des valeurs unitaires lisse les données et les réduit à un volume plus raisonnable; pour une discussion des circonstances sous lesquelles l'utilisation des valeurs unitaires est ou non appropriée, voir Balk (1999).

Pour la présente étude, la population de céréales pour petit déjeuner a été subdivisée en quatre groupes :

1. céréales chaudes ( $C$ )
2. céréales « sucrées » ( $S$ )
3. céréales « fruitées » ( $F$ )
4. céréales « ordinaires » ( $O$ ), c'est-à-dire les céréales froides ne rentrant pas dans les catégories (2) et (3).

Pour chaque groupe, pour chaque paire d'années successives, nous avons calculé les indices superlatifs et non superlatifs en utilisant les combinaisons article-point de vente disponibles les deux années. En pratique, il est généralement difficile d'obtenir des appariements parfaits de période en période, et il est important de trouver des moyens de faire face à ce problème en découvrant des substituts pour les produits originaux ou d'autres façons; la présente étude ne tient pas compte de ce problème particulier.

Nous avons calculé les indices de long terme (1995 à 2000) directement, ainsi que par enchaînement des indices annuels. En outre, nous avons calculé les indices fondés sur les articles « de base », c'est-à-dire ceux disponibles chacune des six années. D'une année sur l'autre, les articles de base représentaient entre 53 % et 61 % des articles disponibles durant une année particulière pour la comparaison d'une année à l'autre; les dépenses de base représentaient entre 83 % et 91 % des dépenses totales au titre de tous les

articles (de base et autres). Au cours de la période de 1995 à 2000, il y avait 326 articles de base et, en tout, 848 articles distincts.

Les valeurs des indices annuels de population sont représentées aux figures 1 à 5. La figure 1 donne les valeurs de l'indice  $I^{y,y+1}$  pour les céréales sucrées basé sur l'ensemble des articles vendus dans les magasins durant les années  $y$  et  $y+1$ , pour (février de)  $y = 1995, \dots, 1999$  (l'indice « d'ensemble »). Les valeurs sont présentées pour cinq indices, y compris l'indice de Paasche  $P$  et, en raison de son intérêt théorique, un indice de valeur unitaire  $U$ , le ratio des prix moyens pondérés par les quantités, la moyenne étant calculée sur tous les types d'articles et tous les points de vente. La figure 2 donne les résultats des mêmes calculs, mais en se limitant aux articles « de base ». Les figures 1 et 2 sont presque identiques et la ressemblance entre les indices calculés en utilisant tous les articles (indice d'ensemble) et ceux obtenus en utilisant uniquement les articles de base est vérifiée pour les autres catégories de céréales également. Les figures 3 à 5 donnent les résultats relatifs aux indices calculés pour les articles de base dans le cas des céréales chaudes, fruitées et ordinaires. Pour tout indice, les figures révèlent d'importants écarts entre les catégories de céréales. Les tendances des prix des quatre grands groupes sont assez différentes : celle de  $C$  est à la hausse, celle de  $S$  est fortement à la baisse, celle de  $F$  est moyennement à la baisse et celle de  $O$  est moyennement à la hausse.

Le tableau 1 donne les indices directs de long terme (1995 à 2000) et les indices enchaînés pour « l'ensemble des articles » et pour les « articles de base ». (« L'ensemble des articles » pour la construction d'un indice entre deux années données comprend les articles/points de vente pour lesquels les quantités vendues sont positives les deux années.) De nouveau, l'écart est très faible entre les valeurs obtenues pour les « articles de base » et « l'ensemble des articles », mais prononcé entre les catégories de céréales. Les résultats enchaînés et directs sont proches des indices superlatifs, mais ont tendance à diverger dans le cas de la *moyenne géométrique*, et des indices de Laspeyres et de Paasche. Les

indices de valeur unitaire enchaînés et directs sont proches et, en fait, le dernier serait identique à l'indice enchaîné fondé sur les articles de base, si, par souci de commodité, les indices annuels n'avaient pas été basés uniquement sur les combinaisons article-point de vente disponibles pour les deux années.

À part une certaine oscillation de la position de l'indice de valeur unitaire, nous observons un classement manifeste des indices en fonction de la formule, le même pour toutes les catégories de céréales, qui peut se résumer comme suit : 1) Les indices superlatifs diffèrent assez peu les uns des autres, résultat qui mérite d'être souligné étant donné le degré de variabilité due aux « ventes » des prix relatifs et des quantités liés aux combinaisons article-point de vente. 2) Les indices non superlatifs classiques diffèrent fortement les uns des autres et des indices superlatifs, la *moyenne géométrique*, pondérée par les dépenses à la première période, étant beaucoup plus *élevée* que les indices superlatifs, l'indice de Laspeyres étant encore plus élevé et l'indice de Paasche étant (sans surprise) nettement plus faible. Ces résultats donnent à penser que, dans l'univers des céréales, non seulement la quantité, mais aussi la part des dépenses, a tendance à diminuer à la période 2 pour un article dont le prix augmente fortement durant cette période. 3) L'indice de valeur unitaire est faible également, mais, sauf dans le cas des céréales chaudes, s'approche davantage des indices superlatifs que les indices non superlatifs classiques. 4) À la lumière de faits présentés plus loin dans l'article, et à la recommandation d'un examinateur, nous avons inclus deux indices non fondés sur les quantités dans ce tableau (mais pas dans les figures) : l'indice de *Dutot*, qui est un simple ratio des prix moyens (RM) – voir l'annexe A, et une *moyenne géométrique non pondérée* (c'est-à-dire pondérée par une constante); tous deux sont habituellement réservés au calcul des indices au niveau élémentaire. Les résultats sont surprenants : en ce qui a trait à l'approximation des indices superlatifs, ils donnent d'aussi bons, voire de meilleurs, résultats que les indices non superlatifs basés sur les quantités traditionnelles, à peu près à égalité avec l'indice de valeur unitaire.

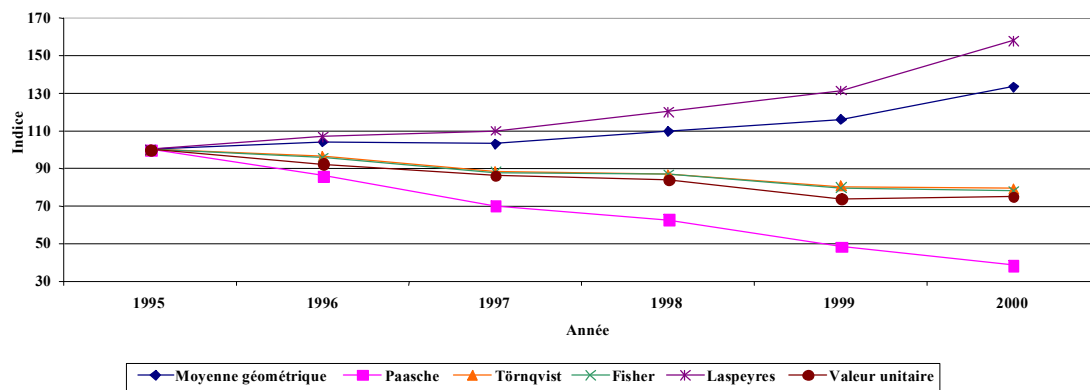


Figure 1. Indices de population cibles enchaînés annuellement pour les indices de février à février pour l'ensemble des céréales sucrées, 1995 = 100.

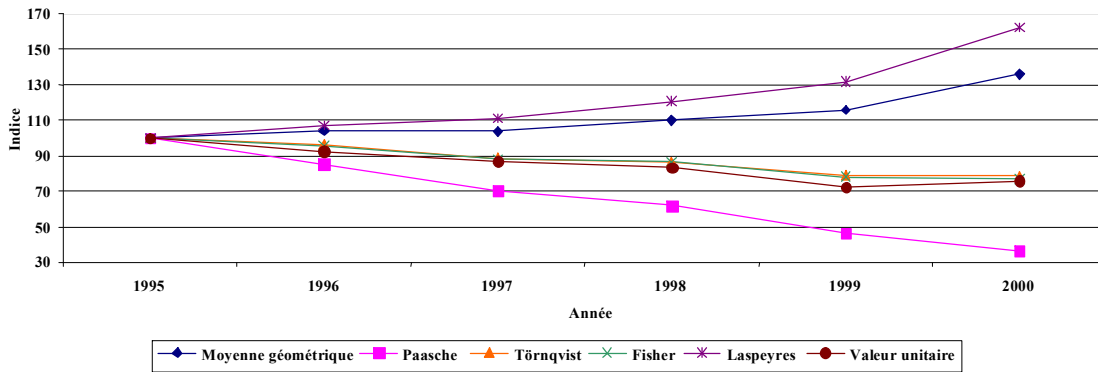


Figure 2. Indices de population cibles enchaînés annuellement pour les indices de février à février pour les céréales sucrées de base, 1995 = 100.

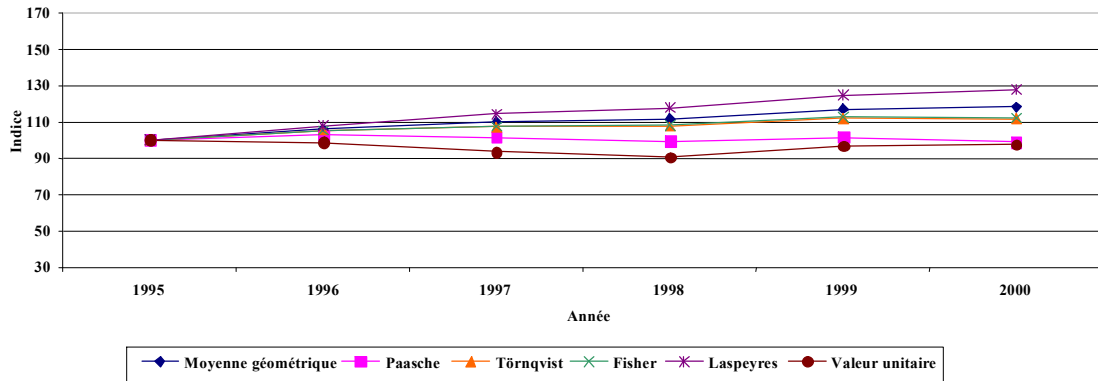


Figure 3. Indices de population cibles enchaînés annuellement pour les indices de février à février pour l'ensemble des céréales chaudes, 1995 = 100.

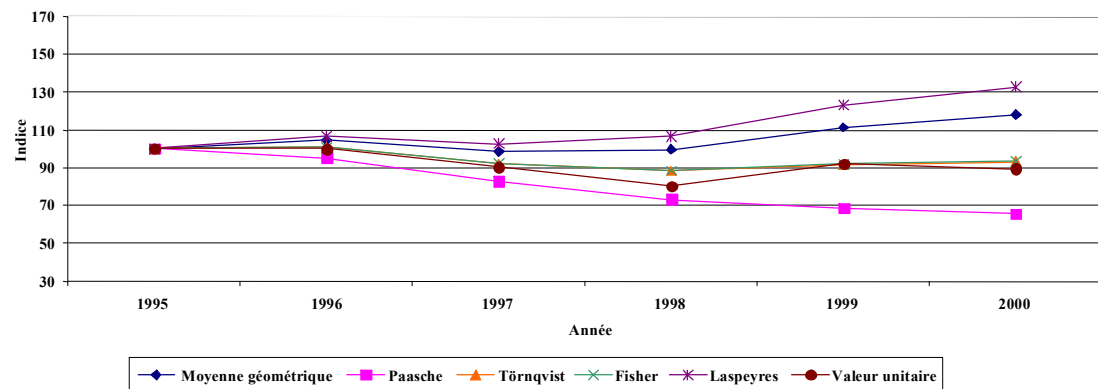


Figure 4. Indices de population cibles enchaînés annuellement pour les indices de février à février pour l'ensemble des céréales fruitées, 1995 = 100.

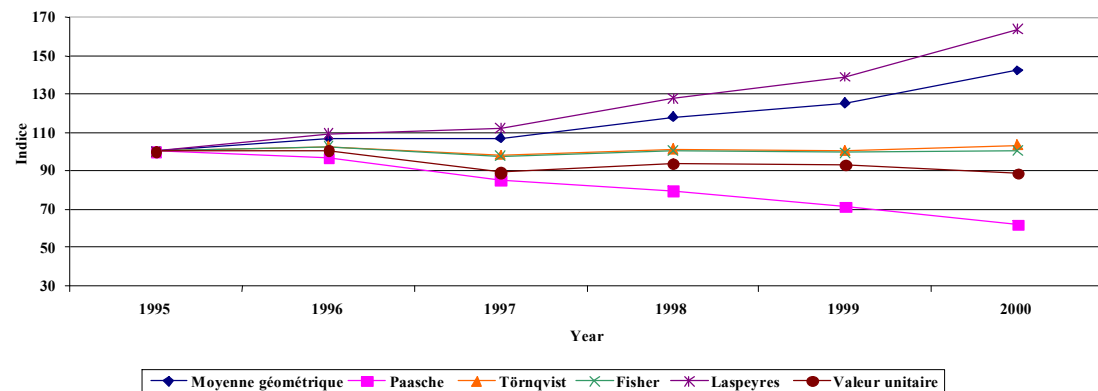


Figure 5. Indices de population cibles enchaînés annuellement pour les indices de février à février pour l'ensemble des céréales ordinaires, 1995 = 100.

**Tableau 1**  
Indices directs et enchaînés pour 1995 à 2000

		Moyenne géométrique	Paasche	Törnqvist	Fisher	Laspeyres	Valeur unitaire
Chaudes	Direct	1,1176	1,0253	1,0847	1,0891	1,1569	0,9576
	Enchaîné, ensemble des articles	1,1801	0,9874	1,1159	1,1216	1,2742	0,9453
	Enchaîné, articles de base	1,1804	0,9865	1,1160	1,1221	1,2763	0,9759
Sucrées	Direct	0,8855	0,6739	0,7913	0,7898	0,9257	0,7417
	Enchaîné, ensemble des articles	1,3341	0,3825	0,7925	0,7771	1,5786	0,7506
	Enchaîné, articles de base	1,3591	0,3661	0,7849	0,7704	1,6212	0,7585
Fruitées	Direct	0,9716	0,8676	0,9319	0,9296	0,9960	0,8932
	Enchaîné, ensemble des articles	1,2202	0,6849	0,9661	0,9696	1,3728	0,9308
	Enchaîné, articles de base	1,1808	0,6557	0,9320	0,9328	1,3269	0,8950
Ordinaires	Direct	1,0811	0,8641	1,0045	0,9816	1,1150	0,8554
	Enchaîné, ensemble des articles	1,3969	0,6330	1,0333	1,0053	1,5965	0,8935
	Enchaîné, articles de base	1,4234	0,6175	1,0353	1,0054	1,6370	0,8879

**Tableau 1**  
Indices directs et enchaînés pour 1995 à 2000

		Moyenne géométrique*	Paasche	Törnqvist	Fisher	Laspeyres	Valeur unitaire	RM	Moyenne géométrique +
Chaudes	Direct	1,1176	1,0253	1,0847	1,0891	1,1569	0,9576	1,1192	1,0949
	Enchaîné, ensemble des articles	1,1801	0,9874	1,1159	1,1216	1,2742	0,9453	1,1395	1,1128
	Enchaîné, articles de base	1,1804	0,9865	1,1160	1,1221	1,2763	0,9759	1,1374	1,1151
Sucrées	Direct	0,8855	0,6739	0,7913	0,7898	0,9257	0,7417	0,8817	0,8702
	Enchaîné, ensemble des articles	1,3341	0,3825	0,7925	0,7771	1,5786	0,7506	0,9124	0,9010
	Enchaîné, articles de base	1,3591	0,3661	0,7849	0,7704	1,6212	0,7585	0,8984	0,8894
Fruitées	Direct	0,9716	0,8676	0,9319	0,9296	0,9960	0,8932	0,9815	0,9726
	Enchaîné, ensemble des articles	1,2202	0,6849	0,9661	0,9696	1,3728	0,9308	1,0263	1,0165
	Enchaîné, articles de base	1,1808	0,6557	0,9320	0,9328	1,3269	0,8950	0,9935	0,9820
Ordinaires	Direct	1,0811	0,8641	1,0045	0,9816	1,1150	0,8554	1,0620	1,0511
	Enchaîné, ensemble des articles	1,3969	0,6330	1,0333	1,0053	1,5965	0,8935	1,0642	1,0572
	Enchaîné, articles de base	1,4234	0,6175	1,0353	1,0054	1,6370	0,8879	1,0653	1,0571

\* Pondérée par les dépenses à la période de base.

+ Non pondérée.

D'après les examens préliminaires, et pour simplifier, nous limitons nos investigations plus approfondies aux données sur les articles de base. Afin d'étudier l'exactitude relative des échantillonnages probabiliste et par choix raisonné, tels qu'ils sont appliqués en pratique pour établir les IPC, nous nous sommes efforcés d'approximer les plans d'échantillonnage utilisés aux États-Unis et au Royaume-Uni, qui représentent l'échantillonnage probabiliste et l'échantillonnage par choix raisonné, respectivement. Dans les deux cas, nous avons eu la chance de disposer d'information détaillée sur les processus d'enquête complexes grâce

à des manuels et à des contacts avec les organismes respectifs. L'idée fondamentale consistait à procéder à l'échantillonnage répété d'une population donnée, par exemple, les transactions de base effectuées en 1995 et en 1996. Chaque « passage-machine » était une combinaison d'activités d'échantillonnage et d'estimation exécutées conformément aux méthodes d'un pays ou de l'autre. Il ne faut pas oublier que nous voulions comparer les mérites des *méthodologies* et non évaluer le succès avec lequel les États-Unis et le Royaume-Uni estiment les paramètres de leur population cible.

Les quatre groupes de population « naturels » décrits plus haut, qui sont appelés « grands groupes » au Royaume-Uni et « catégories de dépenses » aux États-Unis, ont été divisés en sous-groupes moins agrégés. En pratique, les sous-groupes seraient définis en fonction de types de produits. L'une des raisons, outre tout intérêt intrinsèque que l'on pourrait avoir pour ces produits, est que les sous-groupes ainsi formés ont tendance à être homogènes en ce qui a trait aux tendances des prix. Aux fins de la présente étude par simulation, nous définissons par conséquent les sous-groupes de la façon suivante :

- 1) Nous avons calculé la variation de prix de long terme pour chacun des 326 articles compris dans les données de base, en utilisant les indices de valeur unitaire pour les articles (sur l'ensemble des points de vente) pour 2000 comparativement à 1995.
- 2) Nous avons ajouté un bruit à ces indices, trié les articles dans chaque grand groupe en fonction de leur valeur de l'indice perturbé, et regroupé les articles adjacents. Le regroupement des articles dont les indices de long terme étaient proches avait pour but de rendre les sous-groupes homogènes, et l'ajout d'un bruit a été fait de sorte que l'homogénéité soit raisonnablement imparfaite.

Le tableau 2 donne la structure des articles de la population qui a été construite, y compris la nomenclature utilisée dans les deux pays, le nombre de groupes à chaque niveau d'agrégation et le symbole correspondant à chaque niveau de classification utilisé dans le présent article. L'« article représentatif » est le niveau d'agrégation le plus faible auquel un indice est produit aux Royaume-Uni. Il correspond à l'article de niveau d'entrée ou ANE (en anglais *Entry Level Item* ou ELI) des États-Unis, qui est en fait un ensemble d'articles similaires ou connexes. Aux États-Unis, les indices sont produits pour les catégories obtenues au niveau directement supérieur d'agrégation, c'est-à-dire le niveau de la « strate d'articles », mais ces catégories sont encore subdivisées en fonction des régions géographiques dans lesquelles les articles sont vendus. Notons qu'il existe 2 ou 3 strates d'articles/sections  $h$  dans une catégorie/un grand groupe  $C$ , 3 ANE/articles représentatifs  $c$  par strate d'articles/section  $h$  (sauf dans un cas où il y en a 2), et 10 ou 11 articles/variétés  $i$  dans chaque ANE/article représentatif  $c$ . (*Nota* : une catégorie réelle du Royaume-Uni pourrait être plus grande ou plus petite que la catégorie correspondante des États-Unis; par exemple, en règle générale, l'ANE comprend probablement plus de sortes d'articles spécifiques que l'article représentatif. Nous avons donc dû forcer l'équivalence pour assurer que la même quantité d'information soit utilisée dans chaque approche. Cet ajustement n'aura pas d'incidence sur nos conclusions

en ce qui concerne les mérites relatifs des méthodes de base utilisées dans les deux pays.

**Tableau 2**  
Structure de population de l'« univers des céréales » : articles

R.-U.	É.-U.	Nombre de groupes	Symbole
Grand groupe	Catégorie de dépenses	4	$C$
Section	Strate d'articles	10	$h$
Article représentatif	Article de niveau d'entrée (ANE)	29	$c$
Variété	Article	326	$i$

En plus de la structure fondée sur les articles, chaque population de transactions possède une structure « spatiale » caractéristique de l'endroit où a été vendu un article. Cette structure est résumée au tableau 3. Les points de vente appartiennent à des chaînes (par exemple, Safeway, Kroger), qui recoupent les trois unités primaires d'échantillonnage géographique des États-Unis à partir desquelles les données sur les céréales ont été recueillies. (Dans la terminologie du Royaume-Uni, les chaînes sont appelées « *magasins multiples* ».) Les points de vente appartenant à une chaîne donnée ont un propriétaire commun, à l'exception de la « chaîne 8 », qui est un groupe « fourre-tout » composé des points de vente  $n$  appartenant *pas* à une grande chaîne (il pourrait y avoir certaines « mini-chaînes »). Lors de l'appariement de cette « structure basée sur les chaînes » à la classification des points de vente utilisée pour l'échantillonnage au Royaume-Uni, la chaîne 8 a été considérée comme un ensemble de « magasins indépendants » (le terme utilisé pour désigner les magasins appartenant à un propriétaire indépendant au Royaume-Uni). La chaîne 4, qui semble présenter la plus grande homogénéité des prix sur l'ensemble des points de vente, a été considérée comme un « magasin multiple pour lequel le relevé des prix est centralisé » (centrally collected multiple), expression utilisée au Royaume-Uni pour les groupes de points de vente dont les prix sont contrôlés centralement. Chaque chaîne restante a été considérée comme une chaîne dont le relevé des prix n'est pas centralisé. Les méthodes de collecte et d'estimation pour ces trois types de chaînes sont données plus loin dans la description des méthodes appliquées au Royaume-Uni.

Donc, la population est constituée de  $N^{95} \approx 20\,000$  enregistrements pour les indices de 1995–1996, chaque enregistrement représentant l'achat d'un article  $i$  dans un point de vente  $j$ . Sont reliés à chaque article/point de vente son UPE/région  $l$ , sa chaîne/son type de magasin  $k$ , le point de vente/magasin  $j$ , l'article/la variété  $i$ , l'ANE/article représentatif  $c$ , la strate d'articles/section  $h$ , la catégorie de dépenses/le grand groupe  $C$ , et  $p^y$ ,  $q^y$ ,  $p^{y+1}$ , et  $q^{y+1}$ , les prix et quantités (en onces) des articles vendus (en février des) deux années en question. Nous avons utilisé ce fichier de population (appelé simplement « le fichier »

dans la suite) pour simuler toutes les phases des opérations aux États-Unis et au Royaume-Uni.

### 3. Méthodes d'échantillonnage simulées

Les méthodes d'échantillonnage compliquées que nous avons utilisées pour simuler les approches adoptées aux États-Unis et au Royaume-Uni sont modélisées d'après les pratiques respectives de ces deux pays. Ces pratiques évoluent au cours du temps et présentent même des variantes à un point particulier dans le temps. Notre objectif n'était pas de déterminer quel pays utilise la meilleure méthode, ni d'englober toutes les variantes. Nous cherchions plutôt à comparer deux modes distincts d'échantillonnage, en tenant compte de la gamme de complexités que postulent ces modes. Le lecteur que cela intéresse trouvera une description de la construction de l'IPC des États-Unis dans le *BLS Handbook of Methods* (2005), chapitre 17. Pour l'indice des prix de détail du Royaume-Uni, nous nous sommes fondés sur le document intitulé *The Retail Prices Index Technical Manual* (1998). Une description des pratiques plus récentes adoptées au Royaume-Uni peut être consultée dans le *Consumer Price Indexes Technical Manual* (2005).

#### 3.1 Méthodes d'échantillonnage aux États-Unis

Nous commencerons par décrire les méthodes d'échantillonnage appliquées aux États-Unis, qui nécessitent trois enquêtes avec échantillonnage probabiliste, à savoir 1) une enquête-ménages, appelée Consumer Expenditure Survey (CEX), pour estimer la répartition des dépenses des ménages entre diverses catégories de biens, 2) une deuxième enquête-ménages, appelée Point of Purchase Survey (POPS) pour estimer, dans chaque groupe d'articles, les montants relatifs dépensés dans divers points de vente, et 3) une enquête auprès des points de vente, grâce à laquelle sont sélectionnés des articles individuels dont le prix est relevé. Dans les trois cas, l'échantillonnage pour la simulation est aléatoire avec remise (quoique l'échantillonnage employé en pratique soit nettement plus compliqué). Les deux premières enquêtes sont fondées sur des échantillons aléatoires simples et la

dernière, sur un échantillon sélectionné avec probabilité proportionnelle à la taille ( $ppt$ ), où les mesures de taille sont fonction des dépenses estimées d'après la CEX et la POPS. L'échantillon de la troisième enquête est un ensemble d'articles compris dans les combinaisons point de vente/ANE.

##### 3.1.1 CEX (enquête-ménages)

L'objectif est d'estimer  $E_{lc}$ , c'est-à-dire les dépenses brutes des ménages au titre de l'ANE  $c$  dans l'UPE  $l$ . Nous avons procédé à un échantillonnage aléatoire simple avec remise (*esasar*) à partir du fichier décrit plus haut, dans les UPE, de manière à obtenir des estimations à facteur d'extension sans biais

$$\hat{E}_{lc}^{95} = \frac{N_l^{95}}{n_{xl}} \sum_{j \in l \cap s(xl)} \sum_{i \in c \cap s(xl)} E_{jci}^{95}$$

où  $E_{jci}^y = q_{jci}^y p_{jci}^y$ ,  $N_l^{95}$  est la taille de population (nombre d'enregistrements pour l'upe  $l$  en 1995–1996) et  $n_{xl}$  est la taille de l'échantillon CEX  $s(xl)$  dans l'UPE  $l$ , choisies de façon à concorder avec les tailles réelles d'échantillon de la CEX américaine et d'obtenir des coefficients de variation des estimations s'approchant de ceux obtenus dans le cas de la CEX américaine réelle; le  $x$  dans  $s(xl)$  et dans  $n_{xl}$  sert simplement à faire la distinction entre l'enquête CEX et l'enquête POPS (dont la notation correspondante est «  $p$  »; voir plus loin) ou l'enquête sur les prix. Cette « imitation de la CEX » est une version simplifiée de l'enquête réelle. Notre méthodologie reposait sur l'hypothèse tacite que tous les clients d'un point de vente donné achètent les articles dans les mêmes proportions; elle ne tenait pas compte de l'erreur de mesure inévitable dans toute enquête réelle sur les dépenses, et (pour 1995–1996) elle était trop récente : les données réelles de la CEX sont souvent antérieures de plusieurs années aux enquêtes auprès des points de vente pour lesquelles elles sont utilisées. Toutefois, puisque les « données des ménages » recueillies ont aussi été utilisées dans les méthodes correspondantes du Royaume-Uni (voir plus loin), la version simplifiée suffit pour la comparaison souhaitée des méthodologies.

**Tableau 3**  
Structure de population de l'« univers des céréales » : points de vente

R.-U.	É.-U.	N <sup>bre</sup>	Symbole
Région	Unité d'échantillonnage	primaire 3	$l$
Type de magasin :	Indépendant	Chaîne 8	$k$
	Multiple :	Chaîne 4	
	{ Relevé central Non central }	Chaînes 1 à 3; 5 à 7	
Magasin	Point de vente	~300	$j$



Les dépenses de plus haut niveau ont été estimées par simple sommation. Par exemple, le total, étendu à l'ensemble des UPE, dans un ANE donné  $c$  est estimé par  $\hat{E}_c^{95} = \sum_l \hat{E}_{lc}^{95}$ , etc. En tout, nous avons tiré 500 échantillons CEX, chacun produisant un ensemble correspondant d'estimations des dépenses.

### 3.1.2 POPS (enquête-ménages)

L'objectif de cette enquête est d'estimer la distribution des dépenses dans différents points de vente pour des catégories particulières de biens. Ces catégories pourraient être des ANE ou des groupes d'ANE; ici, nous supposons qu'il s'agit d'ANE. La TPOPS (Telephone Point of Purchase Survey) réelle des États-Unis est, comme son nom l'indique, réalisée par téléphone, selon un plan avec renouvellement de l'échantillon tous les quatre ans. Nous nous sommes efforcés, comme nous l'avons fait dans le cas de la CEX, de faire concorder les propriétés statistiques de notre procédure avec celles de la TPOPS réelle, mais il s'est avéré que faire correspondre les tailles d'échantillon dans notre fichier de 20 000 enregistrements nous aurait donné des fractions d'échantillonnage dans les UPE plus grandes qu'il n'était souhaitable. Par conséquent, nous avons réduit les tailles d'échantillon de moitié, de sorte que notre « imitation de POPS » devrait avoir une précision d'environ  $1/\sqrt{2}$  celle de la TPOPS réelle. De nouveau, cette modification n'aura pas d'incidence sur les conclusions de l'étude, parce que nous avons utilisé des données identiques pour la construction de l'enquête britannique. Nous avons tiré des échantillons  $s(pl)$  de taille  $n_{pl}$  par *earstar* et procédé à l'estimation au moyen de l'estimateur à facteur d'extension :

$$\tilde{E}_{lci}^y = \frac{N_l^y}{n_{pl}} \sum_{i \in c \cap s(pl)} E_{lci}^y.$$

Puisque les données de la POPS ont tendance à être plus à jour que celles de la CEX, nous choisissons  $y$  comme année de base de l'indice, 1995 pour 1995–1996, mais 1996 pour 1996–1997, etc. Nous avons exécuté 500 passages machines et obtenu 500 ensembles d'estimations, qui ont chacun été apparié à une réalisation de la CEX.

### 3.1.3 Échantillonnage des points de vente

Pour chaque année  $y$ , la sélection des articles pour lesquels les prix doivent être relevés comprend les étapes suivantes :

- Pour chaque UPE  $l$ , et chacune des dix strates d'articles  $h$ , nous sélectionnons deux ANE  $c$  par échantillonnage avec probabilité proportionnelle à la taille avec remise (*pptar*), avec la mesure de taille  $\hat{E}_{lc}^{95}$  dérivée de la CEX.

- Pour chaque ANE  $c$  sélectionné, nous tirons huit points de vente  $j$  par échantillonnage *pptar*, en utilisant comme mesure de taille les estimations des dépenses d'après la POPS  $\tilde{E}_{lji}^y$ . Donc, en tout, nous obtenons 160 paires ANE-point de vente par UPE, et un nombre total de 480, avec un certain degré de répétition éventuel.
- Dans chaque groupe point de vente-ANE ( $j, c$ ), nous « allons » (comme l'agent de terrain irait littéralement) dans le point de vente et « dressons la liste » de tous les articles appartenant à l'ANE et des dépenses correspondantes à la première période  $E_{lji}^y$ , et, au moyen de cette base de sondage dans le point de vente, nous sélectionnons un article par échantillonnage *ppt*.

Pour chaque article ainsi sélectionné, nous enregistrons les prix  $p_{lji}^y$ ,  $y=1, 2$ . Donc, nous notons que tous les aspects de l'échantillonnage des points de vente sont *ppt* avec remise, en fonction des estimations des dépenses provenant de l'une ou l'autre des deux enquêtes-ménages ou provenant directement du magasin sélectionné. De nouveau, nous avons exécuté 500 passages machines, chacun correspondant à une réalisation CEX/POPS unique.

## 3.2 Estimation aux États-Unis

Les « agrégats élémentaires »  $\hat{I}_{lh}^{y, y+1}$ , c'est-à-dire des estimations d'indice au niveau de l'UPE  $\times$  strate d'articles, sont les éléments à partir desquels est construit l'IPC. Dans la plupart des IPC partout dans le monde, les indices de niveau le plus faible sont des moyennes non pondérées d'une sorte ou l'autre, comme l'estimateur RM du Royaume-Uni décrit plus loin, et les données sur les dépenses ne sont utilisées que pour agréger ces indices à des niveaux plus élevés. Aux États-Unis, les indices élémentaires sont essentiellement des estimateurs d'Horvitz-Thomson fondés explicitement ou implicitement sur les estimations des dépenses provenant de la CEX ainsi que de la POPS. Ces dernières années, pour la plupart des strates d'articles, les États-Unis ont adopté la formule de la *moyenne géométrique* (voir l'annexe A), de sorte que les estimations à ce niveau prennent la forme

$$\hat{I}_{lh}^{y, y+1} = \prod_{\substack{j \in l, \\ i \in c \in h, \\ (i, j) \in s}} \left( \frac{p_{lji}^{y+1}}{p_{lji}^y} \right)^{s_{lji}},$$

où

$$s_{lji} = \frac{w_{lji}}{\sum_{\substack{j \in l, c \in h, i \in c \\ (i, j) \in s}} w_{lji}},$$

avec

$$w_{ljhci} = \frac{\tilde{E}_{lc} \hat{E}_{lh}}{\hat{E}_{lc}} w_{ljhci},$$

$j \in l, i \in c \in h$  et  $(i, j) \in s$ . Notons que les poids ne sont pas particuliers au  $i^e$  article; nous omettons les indices supérieurs de période par souci de brièveté. Les poids ne sont pas simplement égaux à l'inverse du nombre  $n_{lh}$  d'articles échantillonnés dans  $lh$ , comme des considérations d'absence de biais d'échantillonnage pourraient porter à le croire (Balk 2003), parce que les probabilités d'échantillonnage ne reflètent pas les dépenses exactes en articles à la période de base; voir le *BLS Handbook of Methods* (2005).

Puis, les indices élémentaires sont agrégés en utilisant les dépenses estimées d'après la CEX conformément à la formule de Laspeyres, par exemple

$$\hat{I}_h^{y, y+1} = \frac{\sum_l \hat{E}_{lh} \hat{I}_h^{y, y+1}}{\sum_l \hat{E}_{lh}}$$

pour obtenir l'indice pour une strate d'articles donnée  $h$ , sur l'ensemble des UPE.

### 3.3 Méthodes d'échantillonnage du Royaume-Uni

Au Royaume-Uni, comme aux États-Unis, la méthode d'estimation comporte la combinaison de trois composante : 1) une enquête-ménages, appelée Family Expenditure Survey (FES), pour estimer les montants consacrés à l'achat de divers groupes d'articles, 2) une enquête-magasins, appelée Annual Retailing Inquiry (ARI), pour obtenir des renseignements sur les dépenses par section et type de magasin et 3) une enquête auprès des points de vente des magasins, pour sélectionner des articles pour l'établissement des prix.

#### 3.3.1 FES (enquête-ménages)

L'objectif est d'estimer les dépenses  $E_{..c}$  pour des articles représentatifs  $c$ , et les dépenses  $E_{l..h}$  pour des combinaisons région-section. Aux fins de la présente étude, nous supposons qu'il existe une concordance passage-machine par passage-machine entre les données pour la CEX des États-Unis et pour la FES du Royaume-Uni, de sorte que nous avons, de nouveau, 500 ensembles de données FES. Notons que le Royaume-Uni ne cherche pas à obtenir les estimations plus détaillées  $E_{l..c}$  que visent les États-Unis.

#### 3.3.2 Annual Retailing Inquiry (enquête-magasins)

L'objectif est d'obtenir des estimations des dépenses  $\tilde{E}_{kh}$ , par section et type de magasin. Cet objectif est considérablement plus général que l'obtention d'estimations point de vente (magasin) par ANE (article représentatif) visées par la POPS des États-Unis. Nous utilisons, pour construire les estimations ARI, pour chacun des 500 passages, les mêmes données que celles utilisées pour construire les estimations POPS pour la simulation de l'IPC des États-Unis.

#### 3.3.3 Échantillonnage des points de vente

La sélection des articles pour lesquels les prix doivent être relevés comprend les étapes suivantes :

- Un « échantillon au jugé » d'articles représentatifs  $c$  est sélectionné dans chaque section  $h$ . Dans la présente étude (uniquement pour permettre la simulation), dans chaque section, nous sélectionnons les deux articles représentatifs ayant les valeurs les plus grandes de  $\hat{E}_{hc}$ . Il convient de souligner deux différences par rapport à l'étape (a) correspondante de la méthode américaine : i) la sélection est uniforme sur l'ensemble des régions  $l$ ; ii) la sélection n'est pas aléatoire et, en particulier, ne permet pas la sélection répétée des articles représentatifs. (La sélection répétée peut avoir lieu dans la méthode américaine simulée, à cause de l'échantillonnage avec remise des ANE dans les strates d'articles.)
- Les économistes de terrain choisissent les magasins dans une localité particulière dans laquelle le prix d'un article représentatif doit être établi. Traditionnellement, cela se faisait par *eassr*, après que l'économiste de terrain ait construit une base de sondage des magasins appropriés. Plus récemment, la sélection a été faite par échantillonnage *ppt*, où la mesure de taille est la superficie consacrée dans le magasin au type de biens que l'article représentatif représente. Les économistes de terrain ne tirent pas d'échantillons d'articles « dont le relevé des prix est centralisé » : dans le cas d'un très grand magasin multiple, le prix d'un article est relevé auprès du bureau central de ce magasin et est considéré comme représentatif du prix de l'article dans tous les magasins faisant partie du multiple. Dans la présente étude, nous avons procédé de la façon suivante : pour chaque région  $l$  et chaque article représentatif  $c$ , nous avons sélectionné huit magasins comme il suit :
  - 4 parmi les magasins multiples à relevé de prix non centralisé (chaînes 1, 2, 3, 5, 6, 7)

- 1 provenant d'un magasin multiple à relevé de prix centralisé (chaîne 4)
- 3 parmi les magasins indépendants (chaîne 8)

Dans chaque cas, pour simplifier, nous avons procédé à l'échantillonnage aléatoire simple (*eas*) sans remise des magasins pour lesquels les dépenses étaient positives pour l'article représentatif. Le nombre de magasins dans la simulation pour le Royaume-Uni (8 par article représentatif de chaque région) coïncide avec le nombre de « points de vente » dans la simulation pour les États-Unis; il existe 160 paires magasin-article représentatif par région, soit 480 en tout. Soulignons les différences qui suivent par rapport à la méthode appliquée aux États-Unis :

1. L'information sur le type de magasin est utilisée pour la stratification (et jouera un rôle dans l'estimation décrite plus loin). Cette information est disponible dans l'échantillon des États-Unis, mais est passée outre au profit de la méthode d'échantillonnage *ppt*.
  2. Pour le Royaume-Uni, nous permettons de l'information sur la présence ou l'absence de l'article représentatif spécifique  $c$  (équivalent à l'ANE) dans la liste de magasins avant l'échantillonnage, tandis que pour les États-Unis, on ne connaît effectivement que l'existence d'un certain ANE dans la strate d'articles donnés. (Cela sous-entend des mises en correspondance multipliées de catégories ANE à POPS, ce qui était typiquement le cas jusqu'à récemment aux États-Unis; la version courante des appariements de catégories ANE à TPOPS (telephone point of purchase survey) est 1 à 1; autrement dit, une base de sondage de points de vente est construite pour chaque ANE.)
- c) Traditionnellement, pour chaque élément représentatif  $c$ , dans un magasin particulier, l'économiste de terrain sélectionne la variété  $i$  qu'il ou elle considère comme dominant les ventes, c'est-à-dire un échantillonnage au jugé de la variété achetée la plus systématiquement. Nous formalisons cela de la façon suivante :
1. Pour une paire magasin-article représentatif donnée ( $j$ ,  $c$ ), nous dressons la liste de toutes les variétés  $i$ .
  2. Pour chaque variété, nous trouvons la quantité minimale  $q_i^* = \text{Min}(q_i^y, q_i^{y+1})$  sur deux années.

3. Nous échantillonnons la variété  $i$  avec  $\text{Max}\{q_i^*\}$ .

Naturellement, ce processus requiert plus d'information que n'en posséderait un économiste de terrain à la première période (et, de nouveau, n'est pas utilisé dans la méthode d'échantillonnage américaine décrite plus haut), et peut être considéré comme offrant un substitut pour l'évaluation par l'économiste de terrain de la continuité relative des biens vendus.

*Nota* : Il est commode d'utiliser l'expression échantillonnage *maxminq* pour désigner la combinaison de la sélection d'un point de vente par *eassr* comme en (b) et d'un article dans le magasin comme en (c).

### 3.4 Estimation au Royaume-Uni

Pour le Royaume-Uni, les agrégats élémentaires ont été calculés au moyen d'une formule de ratio des moyennes (RM) dans chaque cellule de classification croisée définie par la région, le type de magasin et l'article représentatif. Il s'agit fondamentalement d'une estimation non pondérée donnée, pour les magasins indépendants, par

$$\hat{I}_{lkhc}^{y,y+1} = \frac{\frac{1}{n} \sum_{i \in c, j \in k} P_{ljhci}^{y+1}}{\frac{1}{n} \sum_{i \in c, j \in k} P_{ljhci}^y}$$

Dans le cas des magasins multiples, une version pondérée de la formule susmentionnée est utilisée avec les dépenses par type de magasin, estimées d'après l'ARI, qui fournissent les poids relatifs des magasins multiples à relevé des prix centralisé par opposition à non centralisé.

Un indice à l'échelle du pays pour les articles représentatifs  $c$  dans l'échantillon (agrégés sur les types de magasin  $k$  et les régions  $l$ ) est alors calculé à l'aide d'un estimateur de type Laspeyres :

$$\hat{I}_c^{y,y+1} = \sum_l \sum_k \tilde{w}_{lkhc} \hat{I}_{lkhc}^{y,y+1},$$

où  $c \in h$ , et  $\tilde{w}_{lkhc}$  est fondé sur  $\tilde{E}_{kh}^y$  provenant de l'ARI et  $\hat{E}_{lh}^{95}$  provenant de la FES (l'utilisation de ces périodes fait en sorte que l'information utilisée est la même pour les États-Unis et le Royaume-Uni). Une agrégation supplémentaire (sur les articles représentatifs  $c$ ) est faite en utilisant  $\hat{E}_{hc}^y$ , etc. provenant de la FES.

### 3.5 Comparaison

Le tableau 4 donne une comparaison sommaire des méthodes appliquées aux États-Unis et au Royaume-Uni que nous avons considérées. La caractéristique prédominante de

la méthode américaine est l'échantillonnage et l'estimation probabiliste, typiquement *pptar*, et celle de la méthode britannique est l'échantillonnage sélectif, en choisissant l'article ou la catégorie que l'on estime avoir le plus d'importance selon le montant des dépenses ou la quantité vendue. Les méthodes de formation des agrégats élémentaires diffèrent et les poids utilisés pour l'agrégation au Royaume-Uni sont estimés à un niveau un peu moins fin de détail aux étapes inférieures.

Le tableau 5 résume ce que l'on pourrait considérer comme étant les points forts et les points faibles des méthodes américaine et britannique. Par avantage de « force brute », que nous attribuons à l'approche britannique, nous

entendons l'exploitation d'une combinaison de deux facteurs qui jouent souvent un rôle dans l'établissement des prix et la construction des indices des prix. En premier lieu, les leaders du marché ont tendance à dicter le prix; par exemple, s'ils augmentent ou réduisent fortement les prix, leurs concurrents moins importants vendant des biens similaires peuvent penser qu'il est nécessaire ou justifié de suivre leur exemple. En deuxième lieu, même si la tendance des prix varie entre biens semblables, les principaux vendeurs domineront vraisemblablement l'indice des prix en raison du montant élevé des dépenses, autrement dit, à cause de leur pondération conséquemment élevée.

**Tableau 4**  
Comparaison sommaire des méthodes appliquées aux États-Unis et au Royaume-Uni

	É.-U.	R.-U.
Enquête sur les dépenses des ménages	$\hat{E}_{lc}^{95}$	$\hat{E}_{.c}^{95}, \hat{E}_{lh}^{95}$
Dépenses par point de vente/catégorie	Ménages (POPS) $\tilde{E}_{jc}^y$	Enquête auprès des magasins (ARI) $\tilde{E}_{kh}^y$
Choix des catégories d'articles	2 ANE <i>c</i> /strate d'articles <i>h</i> /UPE <i>l</i> <i>pptar</i> ( $\hat{E}_{lc}^{95} / \hat{E}_{lh}^{95}$ )	2 articles représentatifs <i>c</i> /section <i>h</i> /région <i>l</i> <i>Les plus importants</i> ( $\hat{E}_{.c}^{95} / \hat{E}_{.h}^{95}$ )
Choix des points de vente	8 points de vente <i>j</i> /ANE <i>c</i> × UPE <i>l</i> <i>pptar</i> ( $\tilde{E}_{jc}^y / \tilde{E}_{l,c}^y$ )	8 points de vente <i>j</i> /article représentatif <i>c</i> × région <i>l</i> – <i>eas</i> avec type de magasin <i>k</i> , $E_{jc}^y > 0$
Article dans un point de vente/une catégorie	1 article <i>i</i> / <i>jc</i> <i>ppt</i> ( $E_{jci}^y / E_{jc}^y$ )	1 variété <i>i</i> / <i>jc</i> $\max[\text{Min}(q_{ji}^y, q_{ji}^{y+1})]$
Indice élémentaire	$\hat{I}_{lh}^{y,y+1} = \prod_{\substack{j \in l, \\ i \in c \in h \\ (i,j) \in s}} \left( \frac{P_{ljhci}^{y+1}}{P_{ljhci}^y} \right)^{S_{ljhci}}$	$\hat{I}_{lkhc}^{y,y+1} = \frac{\frac{1}{n} \sum_{i \in c, j \in k} P_{ljhci}^{y+1}}{\frac{1}{n} \sum_{i \in c, j \in k} P_{ljhci}^y}$
Niveau plus élevé d'agrégation	$\hat{I}_h^{y,y+1} = \frac{\sum_l \hat{E}_{lh} \hat{I}_{lh}^{y,y+1}}{\sum_l \hat{E}_{lh}}$	$\hat{I}_c = \sum_l \sum_k \hat{W}_{lkhc} \hat{I}_{lkhc}$ $\hat{W}_{lkhc} = f(\tilde{E}_{.kh}, \hat{E}_{l,h})$

**Tableau 5**  
Comparaison des approches américaine et britannique

Points forts	Points faibles
É.-U. Recueil plus d'information Plus grande utilisation de l'information Satisfait la théorie classique de l'échantillonnage Donne des estimateurs pondérés des estimations régionales (UPE) au niveau le plus faible Procédure opérationnelle plus normalisée	Répétition éventuelle lors de la sélection Ne tient pas compte de la stratification des magasins (c'est-à-dire, de la classification en chaînes)
R.-U. S'appuie sur le principe de « force brute » Stratification des points de vente  L'enquête sur le terrain auprès des magasins s'appuie sur diverses sources	Ensemble disparate de coefficients de pondération Inconsistante dans le cas de l'agrégation des prix relevés de façon centralisée?  Estimateur non pondéré et en apparence arbitraire au niveau le plus faible.

#### 4. Résultats de l'étude principale

Le tableau 6 donne les indices comparant 1995 à 1996 pour la population 1) dans son ensemble (les trois domaines combinés), 2) ventilée en classes/grands groupes et 3) ventilée encore davantage en strates d'articles/sections. Quatre indices, qui pourraient être considérés comme les cibles de l'estimation, sont donnés. À cet égard, rappelons la discussion sur les cibles qui figurent à la section 1.

Le tableau 7 donne les moyennes, variances et erreurs quadratiques moyennes correspondantes pour les estimations américaines et britanniques, où l'erreur quadratique moyenne est calculée par rapport aux indices de Fisher. Nous faisons les constatations suivantes :

- 1) Pour l'ensemble des articles, les catégories et les strates d'articles, les estimations américaines semblent s'approcher de la *moyenne géométrique*  $G$ . Cela confirme ce que d'autres travaux nous avaient fait soupçonner (Dorfman et coll., 1999), à savoir que le niveau le plus faible d'agrégation est dominant (nous avons utilisé une formule de Laspeyres pour les niveaux d'agrégation plus élevés). Le fait que  $G$  se situe entre l'indice de Laspeyres et l'indice cible superlatif donne certaines preuves que le passage des États-Unis à cette méthode d'agrégation élémentaire représente un pas dans la bonne direction.
- 2) Il ne semble exister aucune relation d'ordre claire entre les estimations britanniques au niveau de la *section* et les cibles correspondantes; par exemple, l'indice pour la section 11 est plus élevé que la cible  $L$ , tandis que l'indice pour la section 12 est plus faible que les indices superlatifs, *etc.* Toutefois, lorsque nous poursuivons l'agrégation jusqu'aux niveaux du grand groupe et de l'ensemble des articles, les estimations commencent manifestement à s'approcher des indices superlatifs  $F$  ou  $T$ . (Dalén (1998) a noté un résultat similaire lors de l'agrégation d'échantillons avec seuil d'inclusion.)
- 3) Si nous choisissons l'indice de Fisher comme cible, *même au* niveau de la section, la racine carrée de l'erreur quadratique moyenne de l'estimateur du Royaume-Uni est nettement plus faible que celle de l'estimateur des États-Unis. Étant donné la nature relativement restreinte du plan d'échantillonnage du Royaume-Uni, il n'est pas étonnant que l'estimateur de ce pays présente une variance plus faible, mais sa forme ne porterait pas à penser qu'il donne une approximation sans biais d'un indice de Fisher. Néanmoins, nos résultats laissent entendre que, du

moins pour une population d'achats tels que ceux utilisés dans l'étude, les méthodes par choix raisonné, à « force brute », du Royaume-Uni (et de nombreux autres pays) donnent de bons résultats.

Des résultats similaires ont été obtenus pour les paires successives d'années jusqu'à 1999–2000. La figure 6 donne la *moyenne géométrique* et l'indice de Fisher d'une année sur l'autre pour l'ensemble des articles pour cinq paires d'années, ainsi que les moyennes sur l'ensemble des échantillons des estimateurs américain et britannique correspondant. (Il convient de souligner la différence d'échelle entre la figure 6 et les figures 1 à 5.) Il est facile de voir que l'estimateur américain a tendance à suivre la *moyenne géométrique* de population. L'estimateur britannique, qui suit l'indice de Fisher, a tendance à surestimer les prix dans les années les plus récentes, bien qu'il s'approche nettement plus de l'indice de Fisher que de la *moyenne géométrique* de population. Il convient de souligner que nous avons utilisé des données sur les dépenses de plus en plus périmées, à savoir les données pour 1995, pour l'échantillonnage et l'estimation. Il se peut que les données périmées sur les dépenses aient une incidence plus grande sur les estimations britanniques que sur les estimations américaines, peut-être en nous menant à suréchantillonner les articles représentatifs coûteux ou à nous concentrer sur certains groupes de magasins qui pratiquent des prix de plus en plus élevés.

Les résultats pour les catégories (« chaudes », *etc.*) étaient forts semblables pour les États-Unis relativement à la *moyenne géométrique* et ne sont pas présentés. La figure 7 donne la différence entre les estimations britanniques moyennes d'une année sur l'autre et l'indice de Fisher, pour chacune des quatre catégories. Son examen révèle que la tendance à la surestimation dans les années les plus récentes affecte les quatre catégories.

Dans l'ensemble, les estimateurs du Royaume-Uni fournissent de meilleures estimations de l'indice superlatif cible de Fisher que ceux des États-Unis. Le tableau 8 donne le ratio de la racine carrée de l'erreur quadratique moyenne du Royaume-Uni à celle des États-Unis, pour les cinq paires d'années, pour l'ensemble des articles, pour les groupes et pour les sections. Il contient quelques valeurs anormales, notamment dans les indices 1998–1999 où, pour la section 2 de « chaudes » et, par conséquent, pour la catégorie « chaudes » complète, les estimations britanniques sont appréciablement moins bonnes. Cependant, en général, les méthodes du Royaume-Uni produisent de nettement meilleures estimations. Cela est attribuable, en partie, à une structure d'échantillonnage plus stricte (principalement parce que l'échantillonnage par choix raisonné/avec seuil d'inclusion est sensiblement plus contraignant que la sélection aléatoire de l'ensemble d'articles qui peuvent entrer dans l'échantillon), qui produit, ce qui n'est pas

étonnant, moins de variance. Néanmoins, cela est également dû, en partie, à une tendance étonnante des estimateurs britanniques à cibler les indices de Fisher correspondants, ce qui réduit le biais. Puisque les estimateurs britanniques ne

ressemblent pas formellement à l'indice de Fisher, les raisons de leur tendance à approximer cet indice méritent d'être étudiées plus en profondeur. Nous nous penchons sur cette question à la section suivante.

**Tableau 6**  
Indices 1995-1996 cibles possibles

Description	Moyenne géométrique	Törnqvist	Fisher	Laspeyres
Tous les articles	1,053	1,002	0,997	1,079
Catégories/Grands groupes				
1 – Chaudes	1,058	1,052	1,052	1,078
2 – Sucrées	1,042	0,964	0,956	1,072
3 – Fruitées	1,044	1,007	1,007	1,067
4 – Ordinaires	1,069	1,027	1,027	1,092
Strates d'articles/Sections				
Chaudes – 11	1,043	1,044	1,044	1,057
Chaudes – 12	1,073	1,059	1,058	1,097
Sucrées – 21	1,003	0,917	0,910	1,034
Sucrées – 22	1,063	0,982	0,972	1,093
Sucrées – 23	1,093	1,052	1,054	1,119
Fruitées – 31	0,977	0,955	0,950	0,985
Fruitées – 32	1,165	1,110	1,116	1,204
Ordinaires – 41	1,067	1,021	1,021	1,094
Ordinaires – 42	1,030	0,996	0,996	1,050
Ordinaires – 43	1,104	1,063	1,062	1,125

**Tableau 7**  
Résultats des simulations pour les indices 1995-1996

Description	Indice cible	É.-U.			R.-U.		
		Moyenne	Écart-type	REQM	Moyenne	Écart-type	REQM
Tous les articles	0,997	1,057	0,016	0,062	1,002	0,011	0,012
Catégories/Grands groupes							
1 – Chaudes	1,052	1,059	0,031	0,032	1,045	0,022	0,023
2 – Sucrées	0,956	1,046	0,030	0,095	0,971	0,023	0,027
3 – Fruitées	1,007	1,053	0,035	0,058	0,986	0,027	0,034
4 – Ordinaires	1,027	1,072	0,025	0,051	1,025	0,016	0,016
Strates d'articles/Sections							
Chaudes – 11	1,044	1,045	0,035	0,035	1,064	0,025	0,032
Chaudes – 12	1,058	1,072	0,049	0,051	1,027	0,035	0,047
Sucrées – 21	0,910	1,004	0,050	0,106	0,850	0,045	0,074
Sucrées – 22	0,972	1,070	0,051	0,111	1,089	0,030	0,121
Sucrées – 23	1,054	1,095	0,044	0,060	1,026	0,027	0,039
Fruitées – 31	0,950	0,979	0,020	0,035	0,932	0,020	0,027
Fruitées – 32	1,116	1,178	0,084	0,104	1,077	0,059	0,071
Ordinaires – 41	1,021	1,069	0,050	0,070	1,060	0,030	0,049
Ordinaires – 42	0,996	1,033	0,035	0,051	0,987	0,031	0,032
Ordinaires – 43	1,062	1,107	0,042	0,061	1,028	0,023	0,041

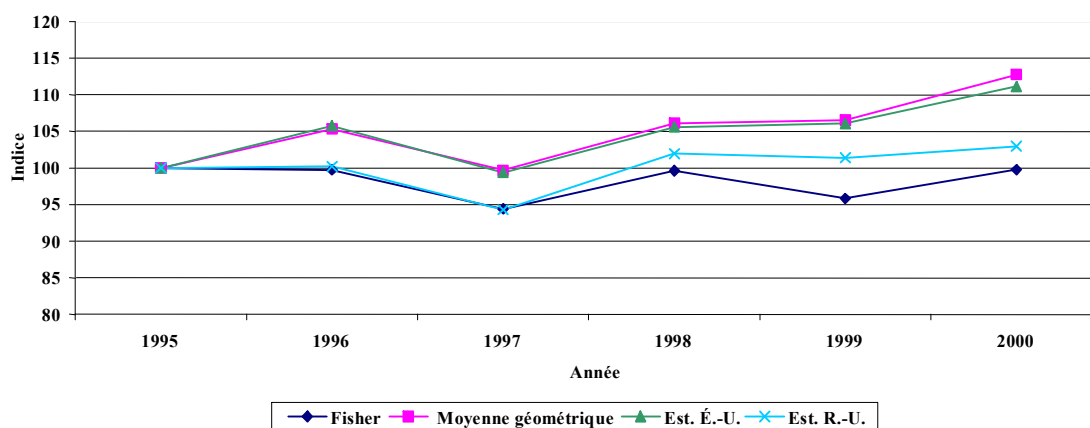


Figure 6. Indices cibles et estimations pour l'ensemble des céréales, indices et estimations de l'indice de février à février, 1995 = 100.

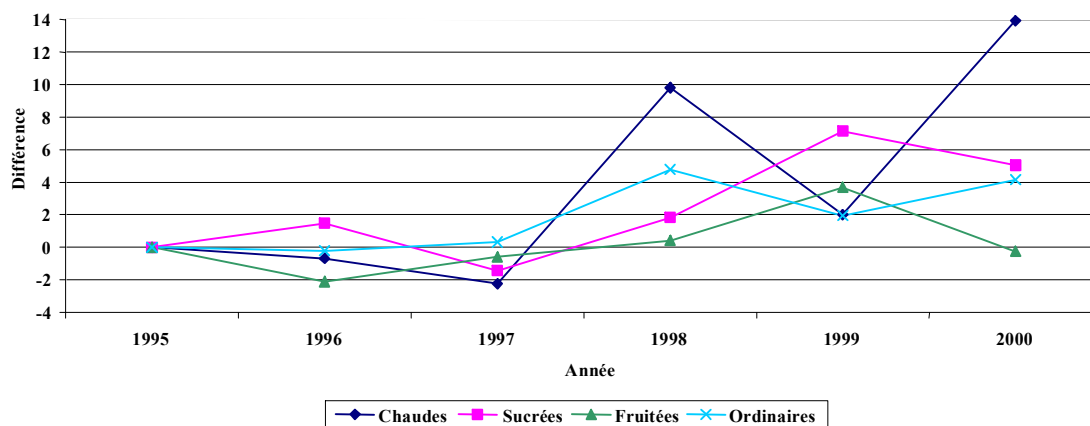


Figure 7. Différences entre les estimations britanniques et les indices de population de Fisher, indices et estimations de l'indice de février à février, 1995 = 100.

Tableau 8  
Ratios de la REQM du R.-U. à la REQM des É.-U.

Description	1995-1996	1996-1997	1997-1998	1998-1999	1999-2000
Tous les articles	0,196	0,192	0,419	0,548	0,288
Catégories/Grands groupes					
1 - Chaudes	0,713	0,517	0,483	1,437	0,589
2 - Sucrées	0,286	0,336	0,314	0,522	0,282
3 - Fruitées	0,595	0,508	0,308	0,501	0,405
4 - Ordinaires	0,310	0,297	0,777	0,319	0,404
Strates d'articles/Sections					
Chaudes - 11	0,923	1,066	0,682	0,529	0,508
Chaudes - 12	0,920	0,850	1,169	1,860	0,842
Sucrées - 21	0,702	0,392	0,421	0,595	0,330
Sucrées - 22	1,092	0,426	0,380	0,341	0,365
Sucrées - 23	0,650	0,455	0,448	0,925	0,851
Fruitées - 31	0,778	1,059	0,637	0,581	0,618
Fruitées - 32	0,683	0,809	0,314	0,457	0,356
Ordinaires - 41	0,709	0,623	0,494	0,567	0,317
Ordinaires - 42	0,642	0,511	1,117	1,092	1,005
Ordinaires - 43	0,678	0,839	0,641	0,815	0,701

## 5. Étude de suivi

Les approches du Royaume-Uni et des États-Unis diffèrent à quatre égards : 1) la structure de stratification, en particulier l'utilisation au Royaume-Uni de différentes strates de magasins et, dans une certaine mesure, de l'échantillonnage centralisé, 2) la structure d'agrégation et de pondération, 3) le mode d'échantillonnage à divers degrés et 4) la formule des agrégats élémentaires. Il est donc difficile de déterminer dans quelle mesure chaque aspect contribue au mérite relatif des méthodes américaine et britannique de construction de l'indice. En particulier, comme nous l'avons mentionné à la dernière section, la raison pour laquelle l'estimateur de l'indice du Royaume-Uni a tendance à cibler les indices superlatifs, surtout au niveau plus élevé d'agrégation, demeure un peu mystérieuse.

Dans l'étude de suivi, nous nous concentrons sur le niveau le plus faible de construction de l'indice, c'est-à-dire sur (3), le niveau magasin-article représentatif (ANE) d'échantillonnage et sur (4), les formules des indices élémentaires. Nous comparons les avantages relatifs des diverses options, en prenant comme cibles les indices élémentaires à l'intérieur des régions. L'agrégation en vue d'obtenir des indices de niveau plus élevé sera exécutée uniformément pour toutes les options de niveau plus faible considérées, en utilisant les parts réelles des dépenses de population. L'importance de la méthode de construction des indices élémentaires est généralement reconnue; voir Diewert (2004) et les références, ainsi que Dorfman et coll. (1999). L'exemple exposé à l'annexe B, conjugué aux résultats présentés au tableau 9, illustre l'effet décisif que le niveau le plus faible de construction de l'indice a sur l'indice dans son ensemble.

Donc, une source vraisemblablement importante de la différence entre les résultats donnés par les méthodes américaine et britannique tient à l'estimation sur échantillon des indices élémentaires de population. Mais cela laisse ouverte la question de savoir si les écarts sont dus à des différences entre les méthodes d'échantillonnage ou entre les formules utilisées pour l'estimation, ou les deux. Donc, nous cherchons à déterminer 1) comment l'échantillonnage au jugé (ici, échantillonnage avec seuil d'inclusion basé sur *maxminq*) se compare à l'échantillonnage probabiliste représenté par *pptar*, en maintenant fixe l'estimateur des indices élémentaires, et 2) comment les estimateurs des indices élémentaires se comparent lorsque nous maintenons fixe la méthode d'échantillonnage. Il sera également intéressant de déterminer ce qui se passe lorsque l'échantillonnage *maxminq* est fondé sur des données provenant de la période de base et de la période *précédente*, plutôt que de la période de base et la période courante.

### 5.1 Méthodes d'échantillonnage et estimateurs au niveau élémentaire

Pour explorer ces questions, nous avons exécuté d'autres études par simulation. Nous nous sommes servis des mêmes données sur les céréales que celles utilisées pour l'étude principale (mois de février successifs), mais nous nous limitons aux magasins indépendants, *chaîne 8*. Nous avons procédé ainsi pour que l'étude soit plus facile à gérer, mais aussi parce que, pour les autres chaînes, les estimateurs des indices élémentaires utilisés au Royaume-Uni étaient plus compliqués que le simple indice de *Dutot*. En outre, il est raisonnable de s'attendre à ce que le comportement des prix soit le plus hétérogène dans cette chaîne, de sorte que les différences intrinsèques seront plus évidentes. La chaîne 8 est la plus grande des chaînes étudiées, comprenant chaque année environ 30 % de l'ensemble de la population, soit environ 6 000 enregistrements.

La structure de base est restée la même : 3 *upe*, 4 grands groupes/catégories de dépenses (chaudes, sucrées, fruitées et ordinaires), 10 sections/strates d'articles, et 29 articles représentatifs/*ANE*. Pour chaque *ANE/article représentatif*, 3 points de vente (un article par point de vente) ont été sélectionnés, par opposition à 10 dans le cas de l'étude principale. Pour étudier l'approche *maxminq* basée sur des périodes antérieures, nous avons réduit les cinq ensembles de données originaux, contenant chacun les données sur les prix et les quantités pour une paire d'années (1995–1996, 1996–1997, *etc.*) afin de n'inclure que les articles permettant un « rétro-appariement »; c'est-à-dire l'appariement sur trois années pour comparer les prix des articles dans les points de vente pour 1995/1996/1997, 1996/1997/1998, *etc.* Environ 90 % des enregistrements de la chaîne 8 ont permis un rétro-appariement (en ce qui concerne les résultats qui suivent, il vaut probablement la peine de souligner que la réduction de l'échantillon pourrait influencer de façon disproportionnée le *maxminq*). Nous déplaçons notre attention de l'indice de Fisher vers l'indice superlatif de Walsh, grâce à une suggestion astucieuse d'un examinateur, dont nous discutons à l'annexe C.

Nous avons utilisé trois estimateurs pour les indices élémentaires : le ratio des moyennes arithmétiques (RM) (le *Dutot*), la *moyenne géométrique* non pondérée (aussi appelée indice de Jevons) et la moyenne des ratios (*MR*). Dans l'échantillonnage *ppt* des points de vente, puis dans l'échantillonnage des articles dans les points de vente, nous avons supposé que la variable de taille (dépenses) était connue (au lieu d'être estimée, comme dans l'étude principale). Outre l'échantillonnage *ppt* avec remise (comme dans l'approche américaine), et *maxminq*, nous avons également étudié l'échantillonnage *ppt* sans remise, parce que nous soupçonnions qu'il serait moins variable que la version avec remise.



Pour chaque mode d'échantillonnage, dans chaque combinaison *upe/ANE*, nous avons tiré 500 échantillons. Nous avons calculé l'erreur quadratique moyenne des estimations par rapport à un indice cible de Walsh au niveau de l'*ANE*. Les moyennes des *eqm* sur l'ensemble des *ANE* ont été calculées pour chaque mode d'échantillonnage/estimation, dans chaque *upe*.

Le tableau 10 donne le ratio de ces moyennes par rapport à l'*eqm* moyenne pour la combinaison *maxminq/Dutot*. Pour chaque estimateur, pour chaque *upe*, à une exception près (*upe* 3, 1999–2000), *maxminq* donne l'*eqm* la plus faible, souvent avec une marge appréciable. L'échantillonnage *ppt* sans remise est la deuxième des solutions les meilleures. Si nous maintenons la méthode d'échantillonnage fixe (en

comparant les lignes 1, 4 et 7, puis 2, 5 et 8, *etc.* au tableau 10), nous constatons qu'à quelques exceptions près, les résultats sont meilleurs pour le *Dutot* que pour la *moyenne géométrique*, qui donne de meilleurs résultats que le *RM*. Ces résultats donnent à penser que 1) *maxminq* est meilleur que *ppt*(dép.), et que *ppt*(dép.) est meilleur que *pptar*(dép.) et que 2) le *Dutot* est plus efficace que la *moyenne géométrique*, et que la *moyenne géométrique* est plus efficace que la moyenne des ratios. Il existe une synergie favorable entre l'échantillonnage *maxminq* et l'indice de *Dutot*. Nous avons également étudié les biais et les variances, et les résultats (non présentés) avaient tendance à suivre le même profil.

**Tableau 9**  
Indices de population 1995-1996, chaîne 8

Description	Laspeyres	Moyenne géométrique*	Fisher	Walsh	Laspeyres de l'indice élémentaire de Walsh
Tous les articles	1,129	1,091	1,028	1,030	1,040
Catégories/Grands groupes					
1 – Chaudes	1,161	1,115	1,080	1,082	1,084
2 – Sucrées	1,129	1,088	1,007	1,012	1,025
3 – Fruitées	1,084	1,054	0,997	1,005	1,015
4 – Ordinaires	1,135	1,101	1,046	1,042	1,050
Strates d'articles/Sections					
Chaudes – 11	1,157	1,117	1,088	1,089	1,090
Chaudes – 12	1,164	1,113	1,072	1,075	1,079
Sucrées – 21	1,086	1,045	0,962	0,970	0,992
Sucrées – 22	1,187	1,142	1,055	1,056	1,058
Sucrées – 23	1,117	1,091	1,034	1,039	1,043
Fruitées – 31	1,003	0,992	0,949	0,965	0,966
Fruitées – 32	1,228	1,172	1,100	1,091	1,102
Ordinaires – 41	1,212	1,161	1,091	1,080	1,090
Ordinaires – 42	1,048	1,030	0,997	0,997	0,998
Ordinaires – 43	1,136	1,107	1,048	1,046	1,056

\* Pondérée par les dépenses à la période de base.

**Tableau 10**  
Moyenne standardisée de l'erreur quadratique moyenne relative sur l'ensemble des ANE, populations réduites, chaîne 8

Estimateur/méthode d'échantillonnage	<i>upe</i> 2				<i>upe</i> 3				<i>upe</i> 4			
	96–97	97–98	98–99	99–00	96–97	97–98	98–99	99–00	96–97	97–98	98–99	99–00
<i>Dutot/maxminq</i> (R.-U.)	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00
<i>Dutot/pptsr</i>	1,73	1,70	1,68	1,91	1,23	1,82	1,35	2,24	1,22	1,06	1,12	0,93
<i>Dutot/pptar</i>	2,13	2,10	1,91	2,14	1,42	2,10	1,46	2,67	1,45	1,23	1,36	1,07
<i>Moyenne géométrique/maxminq</i>	1,20	1,16	1,16	1,06	1,06	1,14	1,08	1,05	1,10	1,11	1,12	0,96
<i>Moyenne géométrique/pptsr</i>	2,08	1,88	1,98	2,27	1,33	1,94	1,47	2,59	1,33	1,09	1,28	0,97
<i>Moyenne géométrique/pptar</i> (É.-U.)	2,49	2,29	2,18	2,53	1,58	2,23	1,58	3,09	1,59	1,30	1,52	1,12
<i>RM/maxminq</i>	1,42	1,32	1,31	1,14	1,24	1,03	1,30	1,05	1,11	1,20	1,21	1,07
<i>RM/pptsr</i>	2,81	2,35	2,49	2,85	1,70	2,31	1,77	3,43	1,57	1,30	1,42	1,17
<i>RM/pptar</i>	3,23	2,77	2,66	3,08	2,03	2,58	1,87	3,96	1,83	1,49	1,66	1,30
<i>Dutot/maxminq</i> , <i>q</i> antérieures	1,12	1,19	1,19	1,41	1,56	1,42	1,69	1,51	1,20	1,02	0,85	1,48

Le fait que l'indice d'échantillon de *Dutot* puisse cibler l'indice de population de Walsh (et donc, indirectement, tout indice superlatif) lorsque les vendeurs les plus importants sont systématiquement échantillonnés est, selon nous, le résultat d'un mécanisme très simple, de « force brute » : dans la mesure où l'indice de Walsh peut être représenté par un petit échantillon d'articles, il est représenté le mieux par ceux pour lesquels les quantités sont régulièrement les plus grandes, et ce sont ces articles que le scénario d'échantillonnage *maxminq* fournit presque toujours. À l'annexe C, nous discutons d'une autre explication des bonnes propriétés de la combinaison *maxminq/Dutot*.

La moyenne des erreurs quadratiques moyennes a également été calculée pour la combinaison *maxminq/Dutot* en se fondant sur les valeurs antérieures de  $q$ , c'est-à-dire  $q_i^{y-1}$ ,  $q_i^y$ . Les résultats sont présentés à la dernière ligne du tableau 10. Nous observons un affaiblissement attendu comparativement à la combinaison *maxminq/Dutot* mise à jour, mais la comparaison des résultats à ceux d'autres options demeure favorable. Nous étudions cet aspect plus en profondeur à la sous-section 5.2.

## 5.2 Effet des quantités décalées sur l'échantillonnage *maxminq*

Afin de placer les résultats de la section 4 dans leur contexte, nous devons déterminer quel est l'effet de l'utilisation de valeurs décalées de  $q$  dans *maxminq*. La raison en est simple : si, à première vue, l'utilisation des quantités des périodes de base et courante semble le moyen évident de refléter la notion d'articles persistants utilisée au Royaume-Uni, cela implique néanmoins l'utilisation d'information (les quantités de la période courante) qui n'a pas été utilisée pour simuler l'échantillonnage utilisé aux États-Unis, ce qui pourrait donner un avantage injuste à la méthode du Royaume-Uni.

Par conséquent, nous comparons l'approche des États-Unis, c'est-à-dire *pptar* (avec taille variable basée sur les dépenses de la période) et la *moyenne géométrique* au niveau élémentaire à l'approche du Royaume-Uni représentée par *maxminq-Dutot*, mais en fondant ici *maxminq* sur les quantités  $q_{y-1}$  et  $q_y$ . Nous avons réduit légèrement les ensembles de données pour être certains d'obtenir des données concordantes pour les trois années consécutives. L'agrégation pour produire des indices de niveau plus élevé a été faite d'après les dépenses réelles de population pour les États-Unis, ainsi que le Royaume-Uni.

Le tableau 11 donne les résultats pour les indices calculés pour l'ensemble des céréales pour la chaîne 8, en comparant les biais, les écarts-types et les racines de l'erreur quadratique moyenne par rapport à l'indice de population de Walsh. Comme prévu, les résultats ne sont pas aussi bons que ceux obtenus en utilisant les valeurs courantes de  $q$ .

Néanmoins, en ce qui a trait aux trois mesures d'exactitude (biais, écart-type et racine de l'erreur quadratique moyenne), la combinaison *maxminq/Dutot* du Royaume-Uni continue de donner de meilleurs résultats que l'approche des États-Unis représentant l'échantillonnage probabiliste.

Pour les catégories plus fines, le tableau 12 donne les ratios des erreurs quadratiques moyennes obtenues sous la méthode du Royaume-Uni avec les valeurs décalées de  $q$  à celles obtenues sous la méthode des États-Unis. Bien qu'ils soient généralement plus élevés que ceux du tableau 8, ils donnent encore à penser que l'approche de l'échantillonnage par choix raisonné du Royaume-Uni est meilleure.

## 6. Discussion

Nous avons présenté une comparaison de deux approches fondamentalement différentes de l'échantillonnage et de l'inférence pour l'établissement d'un indice des prix à la consommation. La conclusion inévitable est que, dans la population que nous avons étudiée, l'approche « R.-U. », qui comporte une stratification plus stricte et, par-dessus tout, un échantillonnage au jugé dans les strates plus restrictif que l'échantillonnage probabiliste de l'approche « É.-U. », produit de meilleures estimations d'un indice superlatif cible.

Nous montrons qu'il en est ainsi, quel que soit l'estimateur de l'indice de prix de faible niveau (*Dutot*, ou *moyenne géométrique*, ou la moyenne des ratios) employé, bien que le *Dutot* (rapport des moyennes) donne les meilleurs résultats.

L'approche du Royaume-Uni est supérieure pour deux raisons : 1) son échantillonnage plus strict, limité aux articles sélectionnés (par exemple, voir le tableau 13 décrit à l'annexe C), mène, sans surprise, à une variance plus faible, constatation qui avait déjà été faite par de Haan et coll. (1999) et 2) les indices d'échantillon de *Dutot* ciblent les indices superlatifs sous échantillonnage du marché dominant, ce qui nous a surpris et a suscité l'étude décrite à la section 5. Par ailleurs, l'approche des États-Unis a donné un estimateur de l'indice pouvant être décrit comme étant sans biais, mais il était sans biais pour le « mauvais » indice de population basé sur la *moyenne géométrique* pondérée par les dépenses à la première période. Donc, il avait tendance à être considérablement plus élevé que l'indice superlatif cible, qu'il s'agisse de celui de Fisher, de Walsh ou de Törnqvist).

Si nous permettions aux tailles d'échantillon d'augmenter, nous pourrions nous attendre à ce que les variances des approches américaine et britannique diminuent l'une et l'autre, mais la variance du Royaume-Uni demeurerait plus faible. Le biais de l'estimateur des États-Unis pour l'indice superlatif cible ne serait pas affecté par l'accroissement de la taille d'échantillon, de sorte que l'erreur

quadratique moyenne relative de l'approche du Royaume-Uni deviendrait de plus en plus faible.

En pratique, évidemment, les quantités de la période 2 ne sont pas disponibles au moment de la sélection de l'échantillon (la période 1) et, dans le cadre de notre étude de suivi, nous donnons une certaine idée de la dégradation partielle qui résulte de l'utilisation des quantités antérieures : elle n'est pas suffisamment importante pour empêcher de conclure que l'approche du Royaume-Uni donne de meilleurs résultats. De surcroît, le jugement de l'économiste de terrain quant au meilleur vendeur pourrait être fondé sur des

données plus récentes que celles d'il y a un an. Donc, l'effet réel pourrait être compris entre ceux des versions décalées et non décalées de *maxminq* que nous avons utilisées. Toutefois, en pratique, les économistes de terrain des États-Unis pourraient échantillonner fréquemment des articles dans les points de vente d'après une estimation de la part des dépenses qui est réellement une moyenne lissée des parts des dépenses de la période de base *et* de la période récente. Cela pourrait atténuer le biais que nous avons observé dans nos simulations, où seules les dépenses de la période de base ont été utilisées pour l'échantillonnage dans les magasins.

**Tableau 11**

Biais, écart-type et racine de l'erreur quadratique moyenne (tous multipliés par 1 000), dans l'estimation de l'indice de population de Walsh pour l'ensemble des céréales, chaîne 8, fondée sur trois approches d'échantillonnage/estimation des indices élémentaires\*

	a) Biais				
	1995 – 1996	1996 – 1997	1997 – 1998	1998 – 1999	1999 – 2000
<i>Dutot/maxminq</i>	29	15	-13	33	2
<i>Dutot/maxminq, q antérieures</i>	-	46	32	82	36
<i>Moyenne géométrique/pptar</i>	78	62	66	82	66
	b) Écart-type				
	1995 – 1996	1996 – 1997	1997 – 1998	1998 – 1999	1999 – 2000
<i>Dutot/maxminq</i>	16	13	11	14	12
<i>Dutot/maxminq, q antérieures</i>	-	14	12	15	14
<i>Moyenne géométrique/pptar</i>	22	18	17	18	20
	c) Racine de l'erreur quadratique moyenne				
	1995 – 1996	1996 – 1997	1997 – 1998	1998 – 1999	1999 – 2000
<i>Dutot/maxminq</i>	33	20	17	36	12
<i>Dutot/maxminq, q antérieures</i>	-	48	34	83	39
<i>Moyenne géométrique/pptar</i>	80	65	68	84	68

\* Au niveau de l'ANE/article représentatif. Pour obtenir des estimations de l'indice global, nous avons agrégé les estimations des indices élémentaires en utilisant les dépenses de population connues.

**Tableau 12**

Ratios de la REQM du R.-U. à la REQM des É.-U., chaîne 8, indices cibles de Walsh : *maxminq* en utilisant les valeurs décalées de *q* et *Dutot* versus *pptar*(dépenses) et *moyenne géométrique*

Description	1996 – 1997	1997 – 1998	1998 – 1999	1999 – 2000
Tous les articles	0,748	0,498	0,993	0,567
Catégories/Grands groupes				
1 – Chaudes	1,539	0,495	1,280	0,765
2 – Sucrées	0,563	0,676	0,941	0,797
3 – Fruitées	0,409	0,323	0,463	0,852
4 – Ordinaires	0,915	0,560	1,164	0,359
Strates d'articles/Sections				
Chaudes – 11	0,748	0,607	0,660	0,657
Chaudes – 12	1,695	0,599	1,333	0,843
Sucrées – 21	0,757	0,593	1,136	0,924
Sucrées – 22	0,370	0,776	0,751	0,671
Sucrées – 23	0,479	0,785	0,796	0,508
Fruitées – 31	0,570	0,443	0,678	1,008
Fruitées – 32	0,526	0,350	0,277	0,674
Ordinaires – 41	1,167	0,509	1,395	0,397
Ordinaires – 42	0,623	0,411	0,918	0,624
Ordinaires – 43	0,919	1,171	0,668	0,560

**Tableau 13**  
Articles sélectionnés par *maxminq* et *ppt*  $\sqrt{q_y q_{y+1}}$  dans 500 échantillons

1995 – 1996, chaîne 8, upe 2 ANE 105											
<i>ppt</i>	articles sélectionnés	2889	2803	1564	2763	1558	2242	2344	2776	760	2850
	% d'échantillons dans lesquels sélectionnés	43,2	32,2	10,4	5,4	3,87	1,53	1,33	0,87	0,8	0,4
<i>maxminq</i>	articles sélectionnés	2889	2803								
	% d'échantillons dans lesquels sélectionnés	80,87	19,13								
1995 – 1996, chaîne 8, upe 3 ANE 401											
<i>ppt</i>	articles sélectionnés	1731	2378	2866	1742	2922	2375	2528	403	871	
	% d'échantillons dans lesquels sélectionnés	33,27	18,8	12,8	12,73	9,47	4,6	4,27	2,8	1,27	
<i>maxminq</i>	articles sélectionnés	2378	1731	2866	1742						
	% d'échantillons dans lesquels sélectionnés	46,27	24,47	15	14,27						
1999 – 2000, chaîne 8, upe 4 ANE 401											
<i>ppt</i>	articles sélectionnés	1731	2866	1742	2378	2922	2528	403			
	% d'échantillons dans lesquels sélectionnés	30,07	21,93	14,3	11,07	9,53	6,8	6,27			
<i>maxminq</i>	articles sélectionnés	1742	2866	2922	1731						
	% d'échantillons dans lesquels sélectionnés	34,27	30,87	18	16,87						

Il est généralement reconnu que les approches sans randomisation sont intrinsèquement moins coûteuses. Ainsi, le nombre de points de vente à visiter est habituellement plus faible et le relevé des prix dans les points de vente requiert moins de main-d'œuvre. Donc, pour un budget donné, nous pouvons nous attendre à ce que la supériorité de l'approche du Royaume-Uni par rapport à l'échantillonnage probabiliste appliqué aux États-Unis soit encore plus grande que ne le laisse entendre la présente étude.

Il serait utile d'étendre l'étude aux données scannées recueillies pour d'autres produits que les céréales. Les articles dont les fluctuations des prix sont plus importantes seraient particulièrement intéressants. Dans une certaine mesure, le bon comportement de l'approche *maxminq/Dutot* peut être relié à la justesse étonnante de l'indice de population de *Dutot* par rapport aux indices superlatifs (comme l'illustre le tableau 1). Dans quelle mesure cette justesse est-elle typique et, en son absence, le bon comportement d'échantillonnage persisterait-il?

Un dernier *avertissement*. En pratique, il serait souhaitable d'injecter une dose de randomisation à certains stades du processus d'échantillonnage et, en particulier, de se montrer légèrement prudent à l'égard de l'échantillonnage centralisé – non pas pour des raisons statistiques, mais pour assurer l'impartialité et l'apparence d'impartialité (Reinsdorf et Triplett 2005, section II; Royall 1976).

### Remerciements

Les opinions exprimées dans le présent article sont celles des auteurs et ne représentent pas la politique du Bureau of Labor Statistics ni du Bureau of Transportation Statistics des États-Unis. Les auteurs remercient David Richardson et Lyuba Rozental de leur avoir fourni les données sur les

céréales et de leur appui à point nommé, Sonja Mapes et Scott Pinkerton, pour leurs travaux sur la classification des céréales en catégories, et Mick Silver, Adrian Ball et Dawn Camus, de leur avoir fourni des explications et de la documentation sur les méthodes d'établissement de l'IPC du Royaume-Uni. Les auteurs remercient également trois examinateurs et un rédacteur adjoint de leurs commentaires constructifs et de leurs encouragements à étendre l'étude, ainsi que J. De Haan, M. Reinsdorf et B. Moulton, pour leurs suggestions utiles. Nous tenons tout spécialement à souligner les encouragements du regretté M.P. Singh dont les suggestions en tant que rédacteur en chef ont orienté l'évolution finale du présent article.

### Annexe A

#### Cibles – Indices de population

$$\text{Laspeyres } L = \frac{\sum_i q_i^y p_i^{y+1}}{\sum_i q_i^y p_i^y}$$

$$\text{Paasche } P = \frac{\sum_i q_i^{y+1} p_i^{y+1}}{\sum_i q_i^{y+1} p_i^y}$$

$$\text{Walsh } W = \frac{\sum_i \sqrt{q_i^y q_i^{y+1}} p_i^{y+1}}{\sum_i \sqrt{q_i^y q_i^{y+1}} p_i^y}$$

$$\text{Fisher } F = \left\{ \frac{\sum_i q_i^y p_i^{y+1}}{\sum_i q_i^y p_i^y} \frac{\sum_i q_i^{y+1} p_i^{y+1}}{\sum_i q_i^{y+1} p_i^y} \right\}^{1/2} = \sqrt{LP}$$

Törnqvist 
$$T = \prod_i \left( \frac{p_i^{y+1}}{p_i^y} \right)^{s_i^{y,y+1}},$$

où

$$s_i^{y,y+1} = \frac{1}{2} \left( \frac{p_i^y q_i^y}{\sum_i p_i^y q_i^y} + \frac{p_i^{y+1} q_i^{y+1}}{\sum_i p_i^{y+1} q_i^{y+1}} \right)$$

Moyenne géométrique

$$G = \prod_i \left( \frac{p_i^{y+1}}{p_i^y} \right),$$

où

$$w_i = s_i^y = \left( \frac{p_i^y q_i^y}{\sum_i p_i^y q_i^y} \right)$$

ou

$$w_i = 1/N$$

Valeur unitaire

$$U = \frac{\sum_i q_i^{y+1} p_i^{y+1} / \sum_i q_i^{y+1}}{\sum_i q_i^y p_i^y / \sum_i q_i^y}$$

Dutot

$$RM = \frac{\sum_i p_i^{y+1} / N}{\sum_i p_i^y / N}$$

(« rapport des moyennes arithmétiques » )

Moyenne des ratios

$$MR = \frac{\sum_i p_i^{y+1} / p_i^y}{N}$$

## Annexe B

### Exemple illustrant l'importance du niveau le plus faible d'agrégation

Nous présentons ici un exemple simple en vue d'illustrer l'importance de la méthode utilisée pour construire les indices élémentaires. Nous comparons les indices de population de Walsh aux indices résultant de l'agrégation des indices élémentaires de Walsh selon une formule de Laspeyres. La raison pour laquelle nous nous concentrons sur l'indice de Walsh est donnée à l'annexe C. L'indice de Walsh « pur » est

$$W = \frac{\sum_i \sqrt{q_i^y q_i^{y+1} p_i^{y+1}}}{\sum_i \sqrt{q_i^y q_i^{y+1} p_i^y}} = \sum \tilde{s}_h W_h^{y,y+1},$$

où  $W_h^{y,y+1}$  est le  $h^e$  indice élémentaire de Walsh et

$$\tilde{s}_h = \frac{\sum_{i \in h} \sqrt{q_i^y q_i^{y+1} p_i^y}}{\sum_i \sqrt{q_i^y q_i^{y+1} p_i^y}}$$

sont les poids d'agrégation de Walsh appropriés. Nous comparons à cela une agrégation selon Laspeyres d'indices élémentaires de Walsh (« ersatz de Walsh »),  $L_W^{y,y+1} = \sum \sum s_h W_h^{y,y+1}$ , où les  $s_h$  sont les poids standard à la période de base.

Les résultats sont donnés au tableau 9. Nous observons un écart perceptible entre l'indice de population réelle de Walsh et l'agrégat selon Laspeyres des indices élémentaires de Walsh, celui-ci ayant tendance à être un peu plus élevé. Cependant, ces différences sont du même ordre que celles entre ces indices et l'indice de Fisher. Elles sont faibles comparativement à l'écart entre la *moyenne géométrique* ou l'indice de Laspeyres et les indices superlatifs. Ce genre de résultat confirme qu'une procédure valable au niveau le plus faible est un élément essentiel de la construction d'un indice.

## Annexe C

### La combinaison *maxminq/Dutot*

Pourquoi la combinaison *maxminq/Dutot* donne-t-elle d'aussi bons résultats, paraissant donner lieu à une absence de biais pour les indices superlatifs?

Un examinateur nous a fait remarquer que l'échantillonnage *maxminq* ressemble considérablement à l'échantillonnage *ppt* avec taille variable  $\sqrt{q_i^y q_i^{y+1}}$ ; pour l'échantillonnage *pptsr* ( $\sqrt{q_i^y q_i^{y+1}}$ ), l'indice de *Dutot* est approximativement sans biais pour un indice cible de Walsh et, par conséquent, indirectement, pour tout autre indice superlatif.

En effet, pour l'espérance du numérateur de l'indice de *Dutot*, sous le scénario d'échantillonnage probabiliste, nous avons

$$\begin{aligned} E_\pi \left( \sum_{i \in S} p_i^{y+1} \right) &= E_\pi \left( \sum_{i' \in U} I_{i'} p_i^{y+1} \right) \\ &= \frac{n}{\sum_{i'} \sqrt{q_{i'}^y q_{i'}^{y+1}}} \sum_{i'} \sqrt{q_{i'}^y q_{i'}^{y+1}} p_{i'}^{y+1}, \end{aligned}$$

où  $E_{\pi}()$  signifie l'espérance par rapport au plan d'échantillonnage et  $I_{i'}$  est un indicateur aléatoire prenant la valeur de 1 ou 0, si  $i'$  est dans l'échantillon ou non. Nous obtenons une expression similaire pour le dénominateur. Le ratio de ces deux valeurs espérées est l'indice de Walsh. Par conséquent, à part le (léger) biais de ratio habituel, qui, comme nous pouvons le montrer, est généralement positif, l'indice de *Dutot* cible en effet l'indice de Walsh sous le scénario d'échantillonnage *ppt*.

Nous devons nous demander si les deux modes d'échantillonnage ont effectivement tendance à présenter un chevauchement appréciable en ce qui concerne les articles sélectionnés. Pour chaque passage-machine, pour chaque *upe l*, *ANE c*, trois articles ont été sélectionnés par échantillonnage *maxminq* ou par échantillonnage *pptsr* ( $\sqrt{q^y q^{y+1}}$ ) des articles compris dans *lc*. Le tableau 13 donne le pourcentage de fois (sur 500 passages) que des articles différents ont été sélectionnés dans l'échantillon, pour certains cas représentatifs sélectionnés arbitrairement. Nous concluons, pas entièrement sans nous étonner, que a) l'échantillonnage *ppt* produit une plus grande dispersion des articles sélectionnés, b) les articles sélectionnés par échantillonnage *maxminq* constituent un sous-ensemble de ceux sélectionnés par échantillonnage *ppt*, c) il existe une certaine corrélation des « articles dominants », c'est-à-dire de ceux qui ont le plus tendance à être sélectionnés par l'une ou l'autre méthode. Brièvement, les échantillonnages *maxminq* et *ppt* ( $\sqrt{q^y q^{y+1}}$ ) semblent être reliés, mais lâchement.

Afin de mieux comprendre la relation entre les deux méthodes d'échantillonnage, nous avons estimé le biais et l'erreur quadratique moyenne, par rapport à l'indice de population de Walsh, de l'indice de *Dutot* pour chaque ANE, pour l'échantillonnage *maxminq* ainsi que *ppt* ( $\sqrt{q^y q^{y+1}}$ ). Les estimations du biais et de l'EQM étaient fondées sur 500 passages pour chaque méthode d'échantillonnage. Les statistiques sommaires ont été calculées sur l'ensemble des ANE pour chaque paire d'années et chaque *upe*. Le tableau 14 donne le pourcentage d'ANE pour lequel les indices élémentaires de *Dutot* présentent un biais positif pour chaque mode d'échantillonnage. Comme prévu, l'échantillonnage *ppt* a tendance à produire un biais positif; nous constatons que l'échantillonnage *maxminq* donne des résultats tout aussi biaisés positivement que négativement.

Le tableau 15 (a) donne le pourcentage d'ANE pour lequel le biais absolu dû à l'utilisation de *maxminq* est plus important que celui du *ppt* ( $\sqrt{q^y q^{y+1}}$ ). À cet égard, l'échantillonnage *ppt* est meilleur. Cependant le tableau 15(b) donne le pourcentage d'ANE pour lequel *maxminq* a produit une plus grande erreur quadratique moyenne et, ici, *maxminq* donne de meilleurs résultats pour toutes les combinaisons période/*upe*, sauf deux. Nous considérons

l'erreur quadratique moyenne comme étant un critère plus décisif, surtout sachant la bidirectionnalité des biais produits par *maxminq*.

**Tableau 14**

Pourcentage d'ANE pour lequel l'indice de *Dutot* présente un biais positif pour un indice cible de Walsh pour deux scénarios d'échantillonnage

	<i>ppt</i> ( $\sqrt{q_y q_{y+1}}$ )			<i>maxminq</i>		
	<i>upe 2</i>	<i>upe 3</i>	<i>upe 4</i>	<i>upe 2</i>	<i>upe 3</i>	<i>upe 4</i>
1995 – 1996	75,0	86,2	75,9	64,3	61,1	61,1
1996 – 1997	60,7	72,4	65,5	53,6	65,5	51,8
1997 – 1998	65,5	75,9	78,6	41,4	27,6	42,9
1998 – 1999	72,4	75,9	70,4	48,3	75,9	40,8
1999 – 1900	89,7	72,4	75,9	48,3	20,7	44,9

**Tableau 15**

Pourcentage d'ANE pour lequel le biais et l'erreur quadratique moyenne de l'indice de *Dutot* pour un indice cible de Walsh sont plus faibles pour l'échantillonnage avec probabilité proportionnelle à la taille (taille variable =  $\sqrt{q_y q_{y+1}}$ ) que pour l'échantillonnage *maxminq*

	a) Biais de <i>ppt</i> plus faible			b) EQM de <i>ppt</i> plus faible		
	<i>upe 2</i>	<i>upe 3</i>	<i>upe 4</i>	<i>upe 2</i>	<i>upe 3</i>	<i>upe 4</i>
1995 – 1996	82,1	93,1	86,2	32,1	58,6	41,4
1996 – 1997	89,2	96,6	100,0	35,7	37,9	27,6
1997 – 1998	89,7	86,2	100,0	41,4	24,1	64,3
1998 – 1999	89,7	82,8	92,6	41,4	37,9	40,7
1999 – 1900	89,7	96,6	41,4	34,5	31,0	37,9

Nous concluons que les bons effets de l'échantillonnage *maxminq* combinés à l'estimateur de *Dutot* ne peuvent pas être expliqués par l'imitation approximative de l'échantillonnage *ppt*. Les comportements sont différents; et, dans l'ensemble, *maxminq* semble être un peu meilleur que *ppt* ( $\sqrt{q_y q_{y+1}}$ ).

Nous ne voyons aucune autre raison expliquant pourquoi l'indice d'échantillon de *Dutot* devrait cibler l'indice de population de Walsh lorsque les produits les mieux vendus sont échantillonnés systématiquement, à part le mécanisme de « force brute » : dans la mesure où l'indice de Walsh peut être représenté par un petit échantillon d'articles, il est mieux représenté par ceux pour lesquels les quantités sont systématiquement les plus grandes, et ces articles sont ceux que fournit le scénario d'échantillonnage *maxminq*.

## Bibliographie

- Balk, B. (1999). On the use of unit values as consumer price subindices. *Proceedings of the Fourth Meeting of the International Working Group on Price Indices*, BLS, Washington, D.C.
- Balk, B. (2003). Price indexes for elementary aggregates: The sampling approach. *Proceedings of the Seventh Meeting of the International Working Group on Price Indices (Ottawa Group)*, Paris.

- BLS Handbook of Methods* (2005). <http://stats.bls.gov/bls/descriptions.htm>.
- Consumer Price Indexes Technical Manual* (2005). Office for National Statistics, London, [http://www.statistics.gov.uk/downloads/theme\\_economy/CPI\\_Technical\\_Manual\\_2005.pdf](http://www.statistics.gov.uk/downloads/theme_economy/CPI_Technical_Manual_2005.pdf).
- De Haan, J., Opperdoes, E. et Schut, C. (1999). Le choix des produits pour l'indice des prix à la consommation : Le seuil d'inclusion par opposition au sondage probabiliste. *Techniques d'enquête*, 25, 1, 33-45.
- Dalén, J. (1998). Studies on the comparability of consumer price indices. *Revue Internationale de Statistique*, 66, 1, 83-113.
- Diewert, E. (1997). "Commentary" [sur 'Alternative Strategies for Aggregating Prices in the CPI' par M.D. Shapiro et D.W. Wilcox]. *Federal Reserve Bank of St. Louis Review*, 79, 3, 27-37.
- Diewert, E. (2004). Index number theory: Past progress and future challenges. Presented at the SSHRC Conference on Price Index Concepts and Measurement, Vancouver, Canada, à <http://www.econ.ubc.ca/diewert/concepts.pdf>.
- Dorfman, A.H., Leaver, S.G. et Lent, J. (1999). Some observations on price index estimators. *Proceedings of the Federal Committee on Statistical Methodology Research Conference, Monday B Sessions*, 56-65.
- Reinsdorf, M., et Triplett, J.E. (2005). A review of reviews: Ninety years of professional thinking about the consumer price index. A paraître, *Proceedings of the June 2004 NBER-CRIW Conference on Price Indexes*, Vancouver.
- The Retail Prices Index Technical Manual* (1998). (Éd. M. Baxter, The Stationary Office, London, à [http://www.statistics.gov.uk/downloads/theme\\_economy/RPI\\_TECHNICAL\\_MANUAL.pdf](http://www.statistics.gov.uk/downloads/theme_economy/RPI_TECHNICAL_MANUAL.pdf).
- Richardson, D.H. (2000). Scanner indexes for the CPI. *Proceedings of the Conference on Scanner Data and Price Indexes*, NBER, Cambridge, <http://www.nber.org/books/>.
- Royall, R.M. (1976). Current advances in sampling theory: Implications for human observational studies. *American Journal of Epidemiology*, 104, 463-473.
- Särndal, C.-E., Swensson, B. et Wretman, J. (1992). *Model assisted Survey Sampling*. Springer, New York.
- Valliant, R., Dorfman, A.H. et Royall, R.M. (2000). *Finite Population Sampling and Inference*. New York: John Wiley & Sons, Inc.