Distr.
GÉNÉRALE

CES/AC.49/2001/3
19 juin 2001

FRANÇAIS
Original: ANGLAIS

COMMISSION DE STATISTIQUE et COMMISSION ÉCONOMIQUE POUR L'EUROPE

CONFÉRENCE DES STATISTICIENS EUROPÉENS

<u>Réunion commune CEE/OIT sur les indices des prix à la consommation</u>
(Genève, 1er et 2 novembre 2001)

ORGANISATION INTERNATIONALE DU TRAVAIL (OIT)

RÉGRESSIONS HÉDONISTES: UNE MÉTHODE FONDÉE SUR LA THÉORIE DU CONSOMMATEUR

<u>Communication sollicitée présentée par l'Université de la Colombie britannique (Canada)*</u>

Résumé

Une régression hédoniste permet d'estimer le prix de différents modèles d'un produit (ou de différentes variantes d'un service) en fonction des caractéristiques de ce produit. La théorie économique sur laquelle repose une régression hédoniste est extrêmement complexe. Dans le présent document on s'appuie sur une méthode très simple fondée sur la théorie du consommateur afin de déduire une famille de formes fonctionnelles utilisables pour une régression hédoniste. La principale hypothèse simplificatrice est que, tous les consommateurs ont la même fonction d'utilité hédoniste, qui décrit la façon dont ils évaluent différents modèles dotés de caractéristiques différentes. On suppose que cette

---

* Document établi par Erwin Diewert, Département d'économie et NBER, Université de la Colombie britannique et NBER, Vancouver (Canada). L'auteur remercie Paul Armknecht, Bert Balk, Ernst Berndt, Jeff Bernstein, Angus Deaton, Robert Feenstra, Dennis Fixler, Robert Gillingham, Alice Nakamura, Richard Schmalensee, Mick Silver, Yrjö Vartia et Kam Yu pour leurs utiles observations et le Conseil de recherches en sciences humaines du Canada pour son aide financière.

GE.01-31408 (F)

fonction d'utilité est séparable des autres biens, ce qui constitue la deuxième grande hypothèse simplificatrice. Dans le document, on examine également différentes formes fonctionnelles pour la fonction d'utilité hédoniste du point de vue de leurs propriétés en matière de flexibilité, c'est-à-dire la qualité de leur approximation des formes fonctionnelles arbitraires. Le document relève que les régressions hédonistes permettent d'estimer le prix des modèles selon une fonction linéaire des caractéristiques n'est pas compatible avec la démarche fondée sur le consommateur qui est adoptée dans le document. Enfin, ce document permet de comparer les techniques d'appariement des modèles traditionnellement utilisées par les instituts de statistique pour prendre en compte les changements qualitatifs, d'une part, avec la méthode de régression hédoniste, d'autre part, et indique dans quelles conditions les deux méthodes ont de bonnes chances de coïncider.

Mots clefs

Régression hédoniste, formes fonctionnelles flexibles, théorie du consommateur, caractéristiques, changement qualitatif, modèles correspondants, indice des prix à la consommation.

Codes de classification du *Journal of Economic Literature*

C23, C43, C51, D11, D12, E31.

Introduction

1. Initialement, le présent document visait à formuler des observations sur l'article de Silver et Heravi (2001), document très utile et très intéressant s'inscrivant dans la tradition instaurée par Silver (1995), qui a été le premier à utiliser systématiquement des données obtenues par lecture optique pour construire des indices. Dans cet article de Silver et Heravi, les auteurs présentent un ensemble énorme de données sur pratiquement toutes les ventes de machines à laver qui ont eu lieu au Royaume-Uni pendant les 12 mois de l'année 1998. Ils utilisent ces données détaillées sur les prix et les quantités vendues, ainsi que des informations sur les caractéristiques de chaque machine, afin de calculer divers indices de prix mensuels agrégés pour les machines à laver, en tenant compte des problèmes résultant de l'évolution de la qualité de celles-ci. En particulier, les auteurs examinent trois grands types de méthode pour l'estimation de prix ajustés en fonction de la qualité en utilisant des données obtenues par lecture optique:

- La technique habituelle de régression hédoniste recourant à des variables indicatrices et à des séries chronologiques, qui n'utilise pas de données quantitatives sur les ventes de modèles;

- Les techniques d'appariement des modèles selon lesquelles on utilise le prix unitaire et les quantités vendues de modèles correspondants au cours de chacune des deux périodes faisant l'objet de la comparaison (ensuite, la théorie usuelle des indices permet d'agréger ces prix et quantités de base); et

- Une méthode hédoniste exacte reposant sur les travaux de Feenstra (1995).

2. Les auteurs ont également utilisé leur base de données scannées sur les machines à laver en vue de reproduire les techniques d'échantillonnage utilisées par les instituts de statistique.

3. Il est remarquable de constater que pratiquement tous[1] les indices de prix qu'ils ont calculés font apparaître une très forte baisse des prix des machines à laver ajustés en fonction de la qualité, soit d'environ 6 à 10 % sur l'année. La plupart de leurs indices mettent en évidence une baisse d'environ 8 à 10 % du prix agrégé des machines à laver. Dans l'indice des prix de détail du Royaume-Uni, les machines à laver appartiennent à la catégorie des appareils électriques, qui comprend des appareils très divers, notamment des fers à repasser, des grille-pain, des réfrigérateurs, etc. De janvier à décembre 1998, l'élément «appareils électriques» de l'indice des prix à la consommation est passé de 98,6 à 98,0, soit une baisse de 0,6 point. Il se peut que les prix des appareils électriques autres que les machines à laver aient suffisamment augmenté au cours de cette période pour annuler l'importante baisse apparente du prix des machines à laver, mais cela me paraît assez peu probable. Nous nous trouvons donc en face d'un phénomène un peu mystérieux: comment se fait-il que les études de l'évolution des prix reposant sur des données obtenues par lecture optique et une régression hédoniste indiquent, en moyenne, des augmentations de prix beaucoup plus faibles que celles qui ressortent des indices officiels correspondants qui incluent le produit considéré[2]? L'une des explications de ce mystère (s'il s'agit bien d'un mystère) pourrait être la suivante: à un certain moment, l'institut de statistique établit un échantillon de modèles dont les prix doivent être recueillis jusqu'à la période suivante d'établissement d'un échantillon. À moins que certains de ces modèles ne disparaissent, aucun autre modèle n'est ajouté à l'échantillon. Ce qui peut se passer, c'est que de nouveaux modèles sont lancés sur le marché pendant la période qui sépare la constitution des échantillons. Ces nouveaux modèles bénéficient du progrès technique et leur prix (ajusté en fonction de la qualité) a tendance à être inférieur à celui des modèles que l'institut de statistique suit. Théoriquement, les fabricants des modèles démodés devraient baisser leurs prix pour faire face à la concurrence nouvelle mais il se peut qu'au lieu de cela ils arrêtent simplement la production de ces modèles démodés, sans modifier leurs prix (ou en ne les diminuant pas de façon suffisante). Cependant, tant que le dernier exemplaire de ce modèle démodé est en vente, l'institut de statistique continue à suivre l'évolution de son prix, qui n'est plus représentatif du marché[3]. Si un modèle disparaît, il est possible que le modèle de remplacement choisi par l'institut de statistique ne soit pas introduit à un prix ajusté en fonction de la qualité suffisamment bas[4], étant donné que l'utilisation de régressions hédonistes est loin d'être fréquente dans les instituts de statistique. Ces deux facteurs peuvent contribuer à expliquer pourquoi la méthode fondée sur la régression hédoniste a tendance à faire apparaître, sur les marchés en mutation rapide, des taux d'augmentation de prix inférieurs aux taux obtenus par les instituts de statistique.

4. Un autre facteur peut contribuer à expliquer pourquoi les études portant sur des données obtenues par lecture optique qui ont recours à la méthode d'appariement des modèles obtiennent des taux d'augmentation de prix inférieurs (ou des taux plus élevés de baisse des prix, comme dans le cas des machines à laver) que ceux obtenus par les instituts de statistique. Si l'on examine la liste des modèles au moment de l'établissement de l'échantillon, il faut s'attendre à ce que certains d'entre eux remportent un grand succès sur le marché, étant donné que leur prix est plus avantageux compte tenu de la qualité[5]. Les consommateurs achèteront un nombre de plus en plus élevé de ces modèles, ce qui permettra à leurs fabricants d'en réduire les prix, car les coûts fixes de production par unité baisseront à mesure que leur part de marché s'accroîtra.

Si on calcule le prix du marché global en retenant l'ensemble des modèles, au moyen d'un indice superlatif reposant sur la lecture optique des données, on augmente avec le temps le facteur de pondération lié à la quantité pour ces modèles à succès caractérisés par des baisses rapides de prix, ce qui se traduit par une mesure de la variation de prix globale inférieure à celle qui est obtenue par l'institut de statistique, étant donné que celui-ci agrégera les prix de l'échantillon en utilisant des facteurs de pondération fixes[6].

5. Fondamentalement, je n'ai rien à redire à l'article de Silver et Heravi (2001); je considère qu'ils ont accompli un travail remarquable.

6. Comme je n'ai pas de critique fondamentale à formuler à l'égard de cet article, la question est de savoir sur quoi la suite du présent document portera. En fait, je traiterai des diverses questions méthodologiques que les auteurs n'ont pu aborder, par manque de place[7].

7. Par conséquent, dans la deuxième section de ce document (Un regard neuf sur la théorie des indices de prix hédonistes), j'examine l'article classique de Sherwin Rosen (1974) sur la méthode hédoniste, en vue de concevoir un modèle beaucoup plus simple que celui qu'il a établi. En particulier, j'émets un nombre suffisant d'hypothèses simplificatrices pour ramener le modèle très général de Rosen aux dimensions du modèle habituel de régression hédoniste à variables indicatrices dans la série chronologique utilisé par Silver et Heravi. Les hypothèses nécessaires pour parvenir à ce modèle simple sont très restrictives et il faut espérer qu'ultérieurement d'autres chercheurs trouveront des moyens d'assouplir certaines de ces hypothèses. Il convient de mentionner que j'utilise une approche traditionnelle fondée sur la demande des consommateurs pour résoudre des problèmes que pose l'institution d'un cadre économétrique permettant d'estimer les préférences hédonistes; en d'autres termes, je n'essaie pas de modéliser l'offre du fabricant sur le marché[8]. Un autre but important de la deuxième section est d'indiquer pourquoi les modèles de régression hédoniste linéaire (dans lesquels la variable dépendante est le prix du modèle et la variable indicatrice temporelle est incorporée dans la régression de façon linéaire) ont peu de chances d'être compatibles avec la théorie microéconomique.

8. Dans la troisième section (Questions relatives aux formes fonctionnelles), nous examinons les problèmes posés par le choix d'une forme fonctionnelle pour la régression hédoniste. Les questions suivantes sont notamment examinées:

- Une comparaison entre les trois formes fonctionnelles les plus couramment utilisées pour les régressions hédonistes.

- La façon dont il est possible d'utiliser les techniques de régression hédoniste pour modéliser le choix de la taille du conditionnement.

- Faut-il choisir des formes fonctionnelles flexibles pour établir des régressions hédonistes?

- Convient-il d'utiliser des formes fonctionnelles non paramétriques?

9. Silver et Heravi (2001) ont fait observer qu'il existe un lien entre les techniques d'appariement des modèles utilisés pour procéder à des ajustements de la qualité et les techniques de régression hédoniste: essentiellement, la méthode hédoniste permet d'utiliser

des informations sur des observations non concordantes, tandis que si l'on a recours à la méthode d'appariement des modèles, il faut écarter les informations sur les modèles qui font soudainement leur apparition sur le marché ou disparaissent brusquement. Triplett (2001) a également examiné le lien qui existe entre les deux méthodes, dans un excellent aperçu des travaux sur la régression hédoniste. L'un des résultats les plus intéressants auxquels Triplett parvient consiste à définir un ensemble de conditions dans lesquelles un modèle de régression hédoniste donne les mêmes résultats qu'un modèle d'appariement. Dans la quatrième section (Régressions hédonistes et méthodes traditionnelles d'ajustement de la qualité), nous généralisons ce résultat pour l'étendre à un ensemble plus général de modèles de régression que celui qui a été examiné par Triplett et nous élargissons ses résultats portant sur deux périodes aux cas où il y a un grand nombre de périodes.

10. Une des caractéristiques de l'article de Silver et Heravi est qu'ils utilisent les informations sur le nombre de modèles vendus en plus des informations habituelles sur le prix et les caractéristiques des modèles qui sont utilisées traditionnellement pour l'établissement de régressions hédonistes. Dans la cinquième section (Régressions hédonistes et utilisation d'un facteur de pondération fondés sur les quantités), nous examinons certaines des questions qui se posent lors de régressions hédonistes effectuées quand des informations sur les ventes sont disponibles.

11. La sixième section (Indices hédonistes exacts) présente des informations sur la méthode de l'indice hédoniste exact de Feenstra (1995), qui est utilisée par Silver et Heravi. Nous concluons qu'il ne semble pas vraiment nécessaire d'utiliser la méthode de Feenstra si l'on est disposé à faire les hypothèses simplificatrices que nous présentons dans la deuxième section.

12. Dans la septième section (Évolution des goûts et fonction d'utilité hédoniste), nous généralisons notre modèle hédoniste présenté dans la deuxième section à une situation plus générale dans laquelle des régressions hédonistes entièrement distinctes sont effectuées au cours de chaque période, au lieu d'effectuer une seule grande régression hédoniste portant sur toutes les périodes.

Enfin, la huitième section présente des ébauches de conclusions.

Un regard neuf sur la théorie des indices de prix hédonistes

13. Dans une optique pragmatique, les modèles de régression hédoniste permettent d'estimer le prix d'une unité d'un produit (appelé «modèle») en fonction des caractéristiques du modèle et d'une variable indicatrice temporelle. On suppose qu'il est possible de recueillir un échantillon de prix de modèles sur plusieurs périodes et d'établir un vecteur des caractéristiques des modèles. Une question théorique intéressante se pose à cet égard: pouvons-nous fournir une interprétation microéconomique de la fonction des caractéristiques dans le membre de droite de la régression?

14. Rosen (1974) le fait dans son article classique sur la méthode hédoniste. Cependant, son modèle économique est extrêmement complexe. Dans la présente section, nous apporterons à son modèle[9], deux modifications importantes:

- Nous supposerons que chaque consommateur a la même *fonction de sous-utilité séparable*, $f(z_1,\ldots,z_N)$, qui permet aux consommateurs de retirer la sous-utilité $Z = f(z)$ de l'achat d'une unité de la marchandise hédoniste complexe dont le vecteur de caractéristiques est $z \equiv (z_1,\ldots,z_N)$[10].

- La sous-utilité que le consommateur retire de la consommation de Z unités de la marchandise hédoniste est combinée avec la consommation de X unités d'une «autre» marchandise composite pour donner au consommateur une utilité globale $u = U^t(X,Z)$ au cours de la période t, où U^t est la fonction de «macro»-utilité au cours de la période t. Rosen (1974; 38) a normalisé le prix de X pour qu'il soit égal à l'unité. Nous ne le faisons pas mais utilisons un prix explicite pour la période t, p^t, pour une unité de la marchandise de consommation générale X.

15. Nous commençons par examiner l'ensemble de combinaisons de X et Z qui peut donner le niveau d'utilité du consommateur au cours de la période t, u^t. Il s'agit de l'ensemble $\{(X,Z) : U^t(X,Z) = u^t\}$, qui est bien entendu la courbe d'indifférence du consommateur pour la période t par rapport aux combinaisons équivalentes de la marchandise de consommation générale X et de la marchandise hédoniste Z. On résout maintenant l'équation $U^t(X,Z) = u^t$ pour X en fonction de u^t et de Z; on a donc[11]

(1) $X = g^t(u^t,Z)$.

Nous supposerons que la pente de cette courbe d'indifférence est dirigée vers le bas et, en fait, nous émettrons l'hypothèse plus forte que g^t est différentiable par rapport à Z et que

(2) $\partial g^t(u^t,Z)/\partial Z<0$.

Soit p^t et P^t les prix d'une unité de X et de Z, respectivement, au cours de la période t. Le *problème de minimisation de la dépense du consommateur au cours de la période t* peut être défini comme suit:

(3) $\min_{X,Z} \{p^tX + P^tZ : X = g^t(u^t,Z)\} = \min_Z \{p^tg^t(u^t,Z) + P^tZ\}$.

La condition nécessaire de premier ordre pour Z pour résoudre l'équation (3) est:

(4) $p^t\, \partial g^t(u^t,Z)/\partial Z + P^t = 0$.

Il est maintenant possible de réécrire l'équation (4) pour obtenir le prix de l'agrégat hédoniste P^t en fonction du niveau d'utilité u^t de la période t et du prix de la consommation générale p^t:

(5) $P^t = - p^t\, \partial g^t(u^t,Z)/\partial Z > 0$.

où l'inégalité découle de l'hypothèse (2) ci-dessus. Nous interprétons maintenant le membre de droite de l'équation (5) comme la *fonction de prix que le consommateur consent à payer au cours de la période t* $w^t(Z,u^t,p^t)$:

(6) $w^t(Z,u^t,p^t) \equiv - p^t\, \partial g^t(u^t,Z)/\partial Z$.

Ainsi, à mesure que nous parcourons vers le bas la courbe d'indifférence du consommateur pour la période t, pour chaque point (indexé par Z) sur cette courbe, (6) nous indique la somme que le consommateur serait disposé à payer *par unité de Z* pour rester sur la même courbe d'indifférence qui est indexée par le niveau d'utilité u^t.

La *fonction de valeur du consentement à payer de la période t* v^t peut maintenant être définie comme la quantité de Z consommés multipliée par le prix que le consommateur consent à payer par unité correspondante, $w^t(Z,u^t,p^t)$:

(7) $v^t(Z,u^t,p^t) \equiv Z\, w^t(Z,u^t,p^t) = -\, Z\, p^t\, \partial g^t(u^t,Z)/\partial Z$

où la dernière égalité découle de l'utilisation de l'équation (6). La fonction v^t est l'homologue de la *fonction de valeur* de Rosen (1974; 38); elle nous indique le montant que le consommateur consent à payer pour consommer Z unités.

16. L'ensemble des formules algébriques qui précèdent a une interprétation qui est indépendante du modèle hédoniste; il s'agit simplement d'un exposé de la façon de déterminer le prix que le consommateur est disposé à payer et une fonction de valeur en utilisant les préférences du consommateur définies pour deux marchandises. Cependant, nous supposons maintenant que le consommateur a une *fonction de sous-utilité séparable* $f(z_1,...,z_N)$ qui indique la sous-utilité $Z = f(z)$ que le consommateur retire de l'achat d'une unité de la marchandise hédoniste complexe[12], dont le vecteur de caractéristiques est $z \equiv (z_1,...,z_N)$. Il convient de relever que nous avons supposé que la fonction f ne varie pas en fonction du temps[13]. Nous supposons maintenant que la fonction d'utilité de la période t du consommateur est $U^t(X, f(z))$. Les formules algébriques qui précèdent au sujet du consentement à payer restent valables. En particulier, notre nouvelle *fonction de prix que le consommateur consent à payer* pour la période t, en ce qui concerne un modèle donné ayant les caractéristiques $z = (z_1,...,z_n)$, est:

(8) $w^t(f(z),u^t,p^t) \equiv -\, p^t\, \partial g^t(u^t,f(z))/\partial Z$

Notre nouvelle *fonction de valeur du consentement à payer* pour la période t (qui est le montant que le consommateur est disposé à payer pour obtenir les services offerts par un modèle ayant le vecteur de caractéristiques z) est:

(9) $v^t(f(z),u^t,p^t) \equiv f(z)\, w^t(f(z),u^t,p^t) = -\, f(z)\, p^t\, \partial g^t(u^t,f(z))/\partial Z$

Supposons maintenant que K^t modèles soient disponibles pour le consommateur pendant la période t, au cours de laquelle le modèle k se vend au prix unitaire de $P_k{}^t$ et a le vecteur de caractéristiques $z_k{}^t \equiv (z_{1k}{}^t,...,z_{Nk}{}^t)$ pour $k = 1, 2,..., K^t$. Si le consommateur achète une unité du modèle k pendant la période t, alors nous pouvons identifier le prix du modèle $P_k{}^t$ à la valeur de consentement à payer appropriée définie par (9), où z est remplacé par $z_k{}^t$; en d'autres termes, les équations suivantes devraient être valables:

(10) $P_k{}^t = -\, f(z_k{}^t)\, p^t\, \partial g^t(u^t,f(z_k{}^t))/\partial Z\,;$ $\qquad t = 1,..., T;\ k = 1,..., K^t.$

17. Quelle est la signification de l'hypothèse de séparabilité? Supposons que la marchandise hédoniste est une voiture et qu'il n'existe que trois caractéristiques: le nombre de sièges, la consommation de carburant et le nombre de chevaux. Dans l'hypothèse de la séparabilité,

le consommateur peut doser à sa guise ces trois caractéristiques et déterminer l'utilité de toute voiture ayant un dosage quelconque de ces trois caractéristiques *indépendamment de ses autres choix* de marchandises. En particulier, le classement des modèles de voiture en fonction de leur utilité est indépendant du nombre d'enfants du consommateur et du prix de l'essence. Bien entendu, l'hypothèse de la séparabilité a peu de chances de se vérifier intégralement dans la réalité mais, pour que notre modèle soit maniable, il faut émettre cette hypothèse assez restrictive.

18. Un autre aspect de notre modèle nécessite des explications complémentaires. Nous supposons explicitement que les consommateurs *ne peuvent acheter des unités fractionnelles de chaque modèle*; ils ne peuvent acheter qu'une quantité de chaque modèle représenté par un entier non négatif; en d'autres termes, nous postulons explicitement des *indivisibilités* en ce qui concerne l'offre. En conséquence, au cours de chaque période, il n'existe qu'un nombre fini de modèles de la marchandise hédoniste, de sorte que, alors que l'on suppose que le consommateur a des préférences continues portant sur l'ensemble des combinaisons possibles des caractéristiques $(z_1,..., z_N)$, au cours de chaque période, il n'existe qu'un nombre fini de modèles distincts disponibles sur le marché.

19. Maintenant, nous précisons davantage notre modèle. Nous supposons que *chaque* consommateur a la *même* fonction de sous-utilité hédoniste f(z)[14] et que le consommateur i a la fonction de *macro-utilité de courbe d'indifférence linéaire* suivante pendant la période t:

(11) $g_i^t(u_i^t,Z) \equiv - a^t Z + b_i^t u_i^t$; $\quad t = 1,..., T; i = 1,..., I$

où a^t et b_i^t sont des constantes positives. Ainsi, pour chaque période t et chaque consommateur i, la courbe d'indifférence pour la période t entre les combinaisons de X et Z est linéaire, la pente constante $- a^t$ étant la même pour tous les consommateurs[15]. Cependant, il convient de noter que nous laissons cette pente varier dans le temps. On dérive maintenant l'équation (11) par rapport à Z et on substitue cette dérivée partielle dans (10). Les équations qui résultent sont les suivantes[16]:

(12) $P_k^t = p^t a^t f(z_k^t)$; $\quad t = 1,..., T; k = 1,..., K^t$.

On définit maintenant le *prix agrégé d'une unité de Z pendant la période t* comme suit[17]:

(13) $\rho_t \equiv p^t a^t$; $\quad t = 1,..., T$

et on utilise cette identité dans l'équation (12), afin d'obtenir notre *système de base d'équations hédonistes*[18]:

(14) $P_k^t = \rho_t f(z_k^t)$; $\quad t = 1,…, T; k = 1,…, K^t$.

20. Il nous suffit maintenant de postuler une forme fonctionnelle pour la fonction de sous-utilité hédoniste f et d'ajouter une spécification stochastique à l'équation (14) pour disposer maintenant de notre modèle de régression hédoniste de base. Les paramètres inconnus de f, ainsi que les paramètres de prix hédonistes pour la période ρ_t peuvent ensuite être estimés[19].

21. Il est possible de généraliser le modèle présenté plus haut tout en obtenant le même modèle (14) si nous remplaçons l'«autre» marchandise composite X par h (x), où x est un vecteur de consommation et h est une fonction d'agrégation linéairement homogène, croissante et concave. Avec ces nouvelles hypothèses, nous obtenons, au lieu des équations (12), les équations suivantes:

(15) $P_k^t = c(p^t)a^t f(z_k^t)$; $t = 1,...,T$; $k = 1,...,K^t$,

où p^t est maintenant le vecteur de prix pour les x marchandises pendant la période t, et c est le coût unitaire ou la fonction de dépense qui est duale de h[20]. On définit maintenant ρ_t comme $c(p^t)a^t$ et on obtient encore le système de base d'équations hédonistes (14).

22. Les équations (14) ont une propriété qui a de bonnes chances d'être présente dans des modèles plus complexes et plus réalistes du choix du consommateur. Cette propriété est que les prix des modèles pendant la période t sont *homogènes de degré un* dans le niveau général de prix p^t. En conséquence, si p^t est remplacé par λp^t pour tout $\lambda > 0$ (on peut songer à un brusque accès d'hyperinflation, où λ est élevé), alors les équations (12) et (14) impliquent que les prix des modèles devraient devenir λP_k^t. Il est à noter que cette propriété d'homogénéité *ne se vérifie pas* pour le *modèle hédoniste additif* suivant:

(16) $P_k^t = \rho_t + f(z_k^t)$; $t = 1,...,T$; $k = 1,...,K^t$.

23. En conséquence, je suis porté à exclure la mise en œuvre de régressions hédonistes reposant sur le modèle linéaire (16) pour des raisons de principe. Il est à noter que les modèles hédonistes où on prend le logarithme du prix de modèle, P_k^t, comme variable dépendante ont tendance à être compatibles avec nos équations hédonistes de base (14), tandis que des modèles linéaires tels que ceux de l'équation (16) ne sont pas compatibles avec les propriétés d'homogénéité linéaire normale impliquées par la théorie microéconomique.

Passons maintenant à l'examen de certains des problèmes que pose le choix d'une forme fonctionnelle pour la fonction de sous-utilité hédoniste f(z)[21].

Questions relatives aux formes fonctionnelles

<u>Formes fonctionnelles fréquemment utilisées</u>

24. Les trois formes fonctionnelles les plus couramment utilisées dans les articles scientifiques sur la régression hédoniste sont les formes log-log, semi-log et linéaire[22]. Nous examinons successivement chacune de ces formes.

Dans le *modèle log-log*, la fonction d'agrégation hédoniste f est définie par son logarithme de la façon suivante:

(17) $\ln f(z_1,...,z_N) \equiv \alpha_0 + \sum_{n=1}^{N} \alpha_n \ln z_n$

où α_n sont les paramètres inconnus à estimer. Si nous prenons les logarithmes des deux membres de l'équation (14), en utilisant l'équation (17) et en ajoutant les termes d'erreur ε_k^t, nous obtenons le modèle de régression hédoniste suivant:

(18) $\ln P_k^{\,t} = \beta_t + \alpha_0 + \sum_{n=1}^{N} \alpha_n \ln z_{nk}^{\,t} + \varepsilon_k^{\,t}$; $t = 1,...,T$; $k = 1,...,K^t$

où $\beta_t \equiv \ln \rho_t$ avec $t = 1,...,T$. Pour identifier tous les paramètres, nous devons normaliser β_t et α_0. Généralement, nous posons $\beta_1 = 0$, ce qui est équivalent à $a^1p^1 = 1$. Si nous voulons imposer une homogénéité linéaire (ou des rendements d'échelle constants) à la fonction de sous-utilité hédoniste f(z), nous pouvons le faire en imposant $\sum_{n=1}^{N} \alpha_n = 1$.

Dans le *modèle semi-log*, le logarithme de la fonction hédoniste f(z) est défini comme suit:

(19) $\ln f(z_1,...,z_N) \equiv \alpha_0 + \sum_{n=1}^{N} \alpha_n z_n$.

Si nous prenons le logarithme des deux membres de l'équation (14), en utilisant l'équation (18) et en ajoutant les termes d'erreur $\varepsilon_k^{\,t}$, nous obtenons le modèle de régression hédoniste suivant:

(20) $\ln P_k^{\,t} = \beta_t + \alpha_0 + \sum_{n=1}^{N} \alpha_n z_{nk}^{\,t} + \varepsilon_k^{\,t}$; $t = 1,...,T$; $k = 1,...,K^t$

où $\beta_t \equiv \ln \rho_t$ pour $t = 1,...,T$. Ici encore, pour identifier tous les paramètres, nous devons normaliser β_t et α_0, de manière que $\beta_1 = 0$, ce qui est équivalent à $a^1p^1 = 1$.

Le modèle semi-log présente un *inconvénient* par rapport au modèle log-log: il n'est pas possible d'imposer des rendements d'échelle constants à la fonction hédoniste semi-log f(z)[23]. En revanche, le modèle semi-log présente un *avantage* par rapport au modèle log-log, étant donné que le modèle semi-log peut prendre en charge les situations dans lesquelles une ou plusieurs caractéristiques $z_{nk}^{\,t}$ sont égales à zéro, ce que ne peut faire le modèle log-log. Cela est important lorsque de nouvelles caractéristiques font leur apparition sur le marché au cours de la période d'échantillonnage.

Dans le modèle linéaire, la fonction hédoniste f(z) est une fonction linéaire simple des caractéristiques:

(21) $f(z_1,...,z_N) \equiv \alpha_0 + \sum_{n=1}^{N} \alpha_n z_n$.

Si on utilise l'identité (21), dans l'équation (14) et si l'on ajoute les termes d'erreur $\varepsilon_k^{\,t}$, on obtient le modèle de régression hédoniste suivant:

(22) $P_k^{\,t} = \rho_t\,[\alpha_0 + \sum_{n=1}^{N} \alpha_n z_{nk}^{\,t}] + \varepsilon_k^{\,t}$; $t = 1,...,T$; $k = 1,...,K^t$.

25. Ici encore, pour identifier tous les paramètres, nous devons normaliser ρ_t et α_n, de manière que $\rho_1 = 0$, ce qui est équivalent à $a^1p^1 = 1$. Malheureusement, l'équation (22) correspond à un *modèle de régression non linéaire* tandis que les modèles log-log et semi-log antérieurs étaient des *modèles de régression linéaire*. Il est possible d'imposer des rendements d'échelle constants à la fonction hédoniste linéaire en posant $\alpha_0 = 0$. Le modèle (22) peut lui aussi prendre en charge sans difficulté l'introduction sur le marché de nouvelles caractéristiques.

26. Il apparaît qu'aucun des trois modèles (18), (20) et (22) n'est réellement meilleur que les deux autres; chacun des trois modèles a au moins un avantage par rapport aux deux autres.

27. En raison de la non-linéacité de l'équation (22), ce modèle n'a pas permis d'estimer très fréquemment, à supposer qu'il l'ait jamais été, alors que le modèle suivant très semblable l'a été à de très nombreuses reprises:

(23) $P_k^t = \rho_t + \alpha_0 + \sum_{n=1}^{N} \alpha_n z_{nk}^t + \varepsilon_k^t$; $\quad t = 1,...,T$; $k = 1,...,K^t$.

28. Comme cela a été indiqué dans la section précédente, le modèle linéaire (23) a peu de chances d'être compatible avec la théorie microéconomique, de sorte que nous ne pouvons recommander son utilisation.

<u>Régressions hédonistes et problème de la taille du conditionnement</u>

29. Pour de nombreuses marchandises, le prix baisse à mesure que le volume acheté augmente. Comment peut-on modéliser ce phénomène en utilisant une régression hédoniste?

30. Supposons que le vecteur de caractéristiques $z \equiv (z1,...,zN)$ est scalaire, de sorte que $N = 1$ et que la quantité caractéristique unique z1 est la taille du conditionnement; en d'autres termes, il s'agit de la quantité d'une marchandise homogène qui est contenue dans le conditionnement vendu. Dans ce cas, il est naturel de considérer que la fonction de sous-utilité hédoniste $f(z_1)$ est une fonction continue monotone non décroissante à une seule variable avec $f(0) = 0$. Nous supprimons l'indice 1 dans les développements qui suivent.

31. Une hypothèse simple pour la fonction $f(z)$ est qu'elle soit une fonction linéaire par sections, une fonction continue ou une fonction *spline linéaire*. Dans le cas de trois segments linéaires, le système d'équations permettant l'estimation (14) serait donné par le système suivant après ajout des erreurs à l'équation (14): pour $t = 1,\ldots,T$, nous avons:

$$
\begin{aligned}
(24)\quad P_k^t &= \rho_t\,\alpha_1 z_k^t + \varepsilon_k^t && \text{si } 0 \le z_k^t \le z_1^* \\
&= \rho_t\,[\alpha_1 z_1^* + \alpha_2\{z_k^t - z_1^*\}] + \varepsilon_k^t && \text{si } z_1^* \le z_k^t \le z_2^* \\
&= \rho_t\,[\alpha_1 z_1^* + \alpha_2\{z_2^* - z_1^*\} + \alpha_3\{z_k^t - z_2^*\}] + \varepsilon_k^t && \text{si } z_2^* \le z_k^t.
\end{aligned}
$$

32. Les tailles de conditionnement prédéterminées z_1^* et z_2^*, lorsque nous passons d'un segment linéaire à l'autre, sont appelées «seuils». Les paramètres inconnus à estimer sont $\rho_1,...,\rho_T$, α_1, α_2 et α_3. Comme d'habitude, ils ne peuvent être tous identifiés, de sorte qu'il est nécessaire d'imposer une normalisation telle que $\rho_1 = 1$.

33. Le système d'équations permettant l'estimation (24) soulève deux difficultés:

- La régression est non linéaire dans les paramètres inconnus.

- Les coefficients estimés α_1, α_2 et α_3 doivent être non négatifs[24]. Si une régression initiale donne une valeur α_i négative, il est alors possible d'effectuer à nouveau la régression en remplaçant α_i, par $(\alpha_i)^2$.

Nous examinons maintenant les propriétés de flexibilité d'une fonction de sous-utilité hédoniste supposée $f(z)$.

Questions de flexibilité

34. Dans la théorie de la demande normale du consommateur, nous voulons habituellement que la forme fonctionnelle de la fonction d'utilité du consommateur (ou d'une quelconque de ses représentations duales) soit flexible; en d'autres termes, nous souhaitons que notre forme fonctionnelle supposée puisse être une approximation d'une fonction arbitraire d'utilité, deux fois continûment différentiable, au deuxième ordre[25]. Dans les articles scientifiques sur les régressions hédonistes, il n'est généralement pas exigé que la forme fonctionnelle de la fonction d'utilité soit flexible[26]. Par exemple, les formes fonctionnelles examinées ci-dessus dans la section intitulée «Formes fonctionnelles fréquemment utilisées» ne peuvent fournir qu'une approximation linéaire et non quadratique. Si des formes fonctionnelles flexibles n'ont pas été utilisées plus largement dans les articles sur la méthode hédoniste, cela est probablement dû au problème de multicolinéarité; en d'autres termes, si nous essayons d'estimer une fonction de sous-utilité hédoniste f(z) qui soit capable de fournir une approximation au second ordre, alors elle peut comporter trop de paramètres inconnus pour pouvoir être estimée de façon exacte[27]. Néanmoins, il peut être utile d'examiner les inconvénients et les avantages de l'utilisation d'autres formes fonctionnelles flexibles dans le contexte hédoniste.

35. Pour la première forme fonctionnelle flexible pour f(z), examinons la *forme fonctionnelle translog* suivante[28], qui généralise notre précédente fonction d'agrégation hédoniste log-log définie par l'équation (17) ci-dessus:

$$(25)\quad \ln f(z_1,\ldots,z_N) \equiv \alpha_0 + \sum_{n=1}^{N} \alpha_n \ln z_n + (1/2) \sum_{i=1}^{N} \sum_{j=1}^{N} \alpha_{ij} \ln z_i \ln z_j$$

où le α_n et le α_{ij} sont les paramètres inconnus à estimer. Si nous prenons les logarithmes des deux membres de l'équation (14), en utilisant l'équation (25) et en ajoutant les termes d'erreur $\varepsilon_k{}^t$, nous obtenons le *modèle de régression hédoniste* translog suivant:

$$(26)\quad \ln P_k{}^t = \beta_t + \alpha_0 + \sum_{n=1}^{N} \alpha_n \ln z_{nk}{}^t + (1/2) \sum_{i=1}^{N} \sum_{j=1}^{N} \alpha_{ij} \ln z_{ik}{}^t \ln z_{jk}{}^t + \varepsilon_k{}^t\ ;$$
$$\alpha_{ij} = \alpha_{ji}\ ;\ t = 1,\ldots,T\ ;\ k = 1,\ldots,K^t$$

où $\beta_t \equiv \ln \rho_t$ pour t = 1,...,T. Pour identifier tous les paramètres, nous avons besoin de normaliser β_t et α_0. Généralement, nous posons $\beta_1 = 0$, ce qui est équivalent à $a^1p^1 = 1$. Si nous voulons imposer une homogénéité linéaire (ou des rendements d'échelle constants) à la fonction de sous-utilité hédoniste f(z), nous pouvons le faire en imposant $\sum_{n=1}^{N} \alpha_n = 1$ et en appliquant les restrictions $\sum_{j=1}^{N} \alpha_{ij} = 0$ pour i = 1,...,N. Bien entendu, le modèle translog (26) contient le modèle log-log (18) comme cas particulier[29].

36. Le modèle hédoniste translog (26) a deux propriétés intéressantes:

- Le membre de droite de l'équation (26) est linéaire en fonction des paramètres inconnus, de sorte qu'il est possible d'utiliser les techniques de régression linéaire pour les estimer.

- Il est aisé d'imposer des rendements d'échelle constants à la fonction d'utilité hédoniste translog f(z) sans supprimer la flexibilité de la forme fonctionnelle.

37. Le principal inconvénient du modèle hédoniste translog est que, comme le modèle log-log, il ne peut prendre en charge le problème du nombre nul de caractéristiques.

38. Pour notre deuxième forme fonctionnelle flexible, examinons la généralisation suivante de la fonction d'utilité hédoniste semi-log (19):

$$(27)\quad \ln f(z_1,...,z_N) \equiv \alpha_0 + \Sigma_{n=1}{}^N \alpha_n z_n + (1/2)\, \Sigma_{i=1}{}^N \Sigma_{j=1}{}^N \alpha_{ij}\, z_i z_j$$

où α_n et α_{ij} sont les paramètres inconnus à estimer. Si nous prenons les logarithmes des deux membres de l'équation (14), en utilisant l'équation (27) et ajoutant les termes d'erreur $\varepsilon_k{}^t$, nous obtenons le *modèle de régression hédoniste quadratique semi-log* suivant:

$$(28)\quad \ln P_k{}^t = \beta_t + \alpha_0 + \Sigma_{n=1}{}^N \alpha_n z_{nk}{}^t + (1/2)\, \Sigma_{i=1}{}^N \Sigma_{j=1}{}^N \alpha_{ij}\, z_{ik}{}^t z_{jk}{}^t + \varepsilon_k{}^t\ ;$$

$$t = 1,...,T\ ;\ k = 1,...,K^t$$

où $\beta_t \equiv \ln \rho_t$ pour t = 1,..., T. Ici encore, pour identifier tous les paramètres, nous devons normaliser β_t et α_0, de manière que $\beta_1 = 0$, ce qui est équivalent à $a^1p^1 = 1$.

39. Le modèle quadratique semi-log présente un *inconvénient* par rapport au modèle translog: il n'est pas possible d'imposer des rendements d'échelle constants à la fonction hédoniste quadratique semi-log f(z). Ces deux modèles ont l'avantage d'être linéaires en fonction des paramètres inconnus. Cependant, le modèle quadratique semi-log présente un *avantage* par rapport au modèle translog: il peut prendre en charge des situations dans lesquelles une ou plusieurs caractéristiques $z_{nk}{}^t$ sont égales à zéro, tandis que le modèle translog ne peut le faire. Ceci est important lorsque de nouvelles caractéristiques font leur apparition sur le marché au cours de la période d'échantillonnage.

40. Pour la troisième forme fonctionnelle flexible pour la fonction d'utilité hédoniste f(z), examinons la *forme fonctionnelle linéaire généralisée*[30] suivante:

$$(29)\quad f(z_1,...,z_N) \equiv \alpha_0 + \Sigma_{n=1}{}^N \alpha_n (z_n)^{1/2} + (1/2)\, \Sigma_{i=1}{}^N \Sigma_{j=1}{}^N \alpha_{ij} (z_i)^{1/2} (z_j)^{1/2}$$

où α_n et α_{ij} sont les paramètres inconnus à estimer. Il convient de noter que l'équation (29) généralise notre forme fonctionnelle linéaire (21)[31]. Si nous utilisons l'identité (29) dans l'équation (14) et si nous ajoutons les termes d'erreur $\varepsilon_k{}^t$, nous obtenons le *modèle de régression hédoniste linéaire généralisé* suivant:

$$(30)\quad P_k{}^t = \rho_t\, [\alpha_0 + \Sigma_{n=1}{}^N \alpha_n (z_{nk}{}^t)^{1/2} + (1/2)\, \Sigma_{i=1}{}^N \Sigma_{j=1}{}^N \alpha_{ij} (z_{ik}{}^t)^{1/2} (z_{jk}{}^t)^{1/2}] + \varepsilon_k{}^t;$$

$$t = 1,..., T\ ;\ k = 1,..., K^t.$$

41. Ici encore, pour identifier tous les paramètres, nous devons normaliser ρ_t, α_n et α_{ij} de manière que $\rho_1 = 0$, ce qui est équivalent à $a^1p^1 = 1$. Malheureusement, l'équation (30) correspond à un *modèle de régression non linéaire*, alors que les précédents modèles quadratiques translog et semi-log étaient des *modèles de régression linéaires.* Il est possible d'imposer des rendements d'échelle constants à la fonction hédoniste linéaire généralisée en posant $\alpha_n = 0$ pour n = 0,1,...,N. En outre, le modèle (22) peut aisément prendre en charge l'apparition de caractéristiques nouvelles sur le marché.

42. Comme c'était le cas dans la section intitulée «Formes fonctionnelles fréquemment utilisées», aucun des trois modèles de régression hédoniste flexibles présentés dans la présente section n'est réellement meilleur que les deux autres. Les modèles (26) et (28) ont l'avantage d'être des modèles de régression linéaires, tandis que le modèle (30) est non linéaire. Le modèle (26) ne peut facilement prendre en charge l'apparition de caractéristiques nouvelles pendant la période d'observation, contrairement aux modèles (28) et (30). Il est facile d'imposer des rendements d'échelle constants en ce qui concerne les caractéristiques dans les modèles (26) et (30), alors que cela n'est pas possible dans le cas du modèle (28). Chacun des trois modèles a donc deux caractéristiques favorables et une caractéristique défavorable.

Formes fonctionnelles non paramétriques

43. Il est possible d'aborder le problème de la forme fonctionnelle d'une manière non paramétrique en utilisant les *techniques généralisées des variables indicatrices*[32].

44. Supposons que deux caractéristiques seulement présentent de l'importance pour les modèles disponibles sur le marché pendant les périodes $t = 1,\ldots,T$. Supposons également qu'il n'y a que I configurations de la première caractéristique et J configurations de la seconde caractéristique pendant la période d'échantillonnage, où I et J sont des entiers supérieurs à un[33]. Supposons en outre qu'au cours de la période t nous ayons $K_{ij}{}^t$ observations dont la première caractéristique est dans le groupe i et la seconde caractéristique dans le groupe j. Représentons la k-ième observation pendant la période t dans ce groupement i, j sous la forme $z_{ijk}{}^t = (z_{1ijk}{}^t, z_{2ijk}{}^t)$. Pour cette configuration de caractéristiques, nous définissons l'utilité hédoniste correspondante comme suit:

$$(31) \quad f(z_{ijk}{}^t) \equiv \alpha_{ij}; \qquad t = 1,\ldots,T;\ i = 1,\ldots,I;\ j = 1,\ldots,J;\ k = 1,\ldots,K_{ij}{}^t.$$

Soit $P_{ijk}{}^t$ le prix au cours de la période t pour l'observation k pour laquelle les caractéristiques du modèle rangent celui-ci dans le groupement i, j de modèles. Si nous utilisons l'identité (31) dans l'équation (14) et si nous ajoutons le terme d'erreur $\varepsilon_{ijk}{}^t$, nous obtenons le *modèle de régression hédoniste à variables indicatrices généralisé (non linéaire)* suivant:

$$(32) \quad P_{ijk}{}^t = \rho_t\alpha_{ij} + \varepsilon_{ijk}{}^t; \qquad t = 1,\ldots,T;\ i = 1,\ldots,I;\ j = 1,\ldots,J;\ k = 1,\ldots,K_{ij}{}^t.$$

Comme précédemment, il n'est pas possible d'identifier tous les paramètres ρ_t pour $t = 1,\ldots,T$ et α_{ij} pour $i = 1,\ldots,I$ et $j = 1,\ldots,J$, de sorte qu'il est nécessaire de leur imposer la normalisation $\rho_1 = 1$.

45. Le modèle de régression hédoniste (32) est non linéaire. Cependant, dans ce cas, nous pouvons reparamétriser notre modèle théorique de façon à obtenir un modèle de régression linéaire. Supposons que nous prenions les logarithmes des deux membres de l'équation (31). Alors, en définissant $\ln\alpha_{ij}$ comme étant γ_{ij}, nous obtenons:

$$(33) \quad \ln f(z_{ijk}{}^t) \equiv \gamma_{ij}; \qquad t = 1,\ldots,T;\ i = 1,\ldots,I;\ j = 1,\ldots,J;\ k = 1,\ldots,K_{ij}{}^t.$$

En utilisant l'identité (33) dans l'équation (14) après avoir pris les logarithmes des deux membres de l'équation (14) et en ajoutant le terme d'erreur $\varepsilon_{ijk}{}^t$, on obtient le *modèle de régression hédoniste à variables indicatrices généralisé linéaire* suivant:

(34) $\ln P_{ijk}{}^t = \beta_t + \gamma_{ij} + \varepsilon_{ijk}{}^t$; $t = 1,...,T;\ i = 1,...,I;\ j = 1,...,J;\ k = 1,...,K_{ij}{}^t$

où $\beta_t \equiv \ln\rho_t$ pour $t = 1,\ldots,T$. Ici encore, il n'est pas possible d'identifier tous les paramètres β_t pour $t = 1,\ldots,T$ et γ_{ij} pour $i = 1,\ldots,I$ et $j = 1,\ldots,J$, de sorte qu'il est nécessaire d'imposer aux paramètres une normalisation telle que $\beta_1 = 0$, ce qui correspond à $\rho_1 = 1$.

Lequel des deux modèles de régression hédoniste à variables indicatrices généralisés (32) et (34) est le «meilleur»? Manifestement, ils ont tous deux exactement le même contenu économique mais bien entendu leurs spécifications stochastiques sont différentes. En conséquence, nous devons examiner les propriétés statistiques des résidus dans les deux modèles pour déterminer lequel est le meilleur[34]. Cependant, si l'on n'examine pas les résidus, le modèle de régression linéaire (34) est beaucoup plus facile à mettre en œuvre que le modèle non linéaire (32), en particulier pour les grands ensembles de données.

46. Les modèles de régression hédoniste à variables indicatrices généralisés linéaires (32) et (34) présentent deux importants avantages par rapport aux modèles à forme fonctionnelle flexible traditionnels énumérés précédemment:

- Les modèles à variables indicatrices (32) et (34) sont entièrement non paramétriques et, par conséquent, beaucoup plus flexibles que les formes fonctionnelles flexibles traditionnelles.

- Les modèles à variables indicatrices peuvent aisément prendre en charge les espaces de caractéristiques discrètes.

47. Cependant, les régressions hédonistes à variables indicatrices présentent également certains inconvénients.

- Il peut exister un nombre énorme de paramètres à estimer, en particulier s'il y a un grand nombre de caractéristiques distinctes.

- Si nous essayons de réduire le nombre de paramètres en retenant un nombre d'intervalles de classe plus petit pour chaque caractéristique, nous introduisons une plus grande variance dans l'estimation de nos coefficients.

- Le nombre de cellules de classement varie en fonction des enquêteurs; en d'autres termes, différentes spécifications hédonistes à variables indicatrices faites par différents opérateurs hédonistes se traduiront par un choix de I et de J différents, ce qui empêchera la reproductibilité dans les modèles[35].

- Si j reste constant, les coefficients α_{ij} et γ_{ij} devraient augmenter (ou au moins ne pas diminuer) lorsque i augmente de 1 à I[36]. De même, si i reste constant, les coefficients α_{ij} et γ_{ij} devraient augmenter (ou au moins ne pas diminuer) lorsque j augmente de 1 à J. Les modèles de régression (32) et (34) ne tiennent pas compte de ces restrictions et il peut être difficile de les imposer[37].

48. Néanmoins, je pense que ces techniques de régression hédoniste à variables indicatrices généralisées sont très prometteuses. Ces modèles, ainsi que d'autres modèles non paramétriques, méritent d'être examinés attentivement par les responsables de la recherche appliquée.

Régressions hédonistes et méthodes traditionnelles d'ajustement de la qualité

49. Silver et Heravi (2001) ont démontré comment on peut réinterpréter les techniques traditionnelles d'appariement des modèles pour l'ajustement de la qualité dans le contexte des modèles de régression hédoniste. Triplett (2001) et Koskimäki et Vartia (2001; 9) ont également obtenu certains résultats dans ce domaine. Dans la présente section, nous examinons deux des résultats de Triplett.

50. Supposons que les équations de régression hédoniste (14) sont valables pendant la période t et que nous souhaitons comparer la qualité du modèle 1 avec la qualité du modèle 2. Les deux premières équations des équations (14) impliquent que l'utilité de la variété 2 par rapport à celle de la variété 1 est donnée par la formule

(35) $f(z_2^t)/f(z_1^t) = [P_2^t/\rho_t]/[P_1^t/\rho_t] = P_2^t / P_1^t$;

en d'autres termes, l'utilité ou la valeur intrinsèque, pour le consommateur, du modèle 2 par rapport à l'utilité du modèle 1 est le rapport des prix P_2^t / P_1^t. Par conséquent, dans ce cas, un ajustement de la qualité qui sort du cadre d'un modèle de régression hédoniste est équivalent à un ajustement de la qualité réalisé selon la technique «traditionnelle» des instituts de statistique, qui consiste à utiliser le *rapport des prix observés* des deux marchandises au cours de la même période comme un indicateur de la qualité relative des deux marchandises[38].

51. Dans un deuxième exemple, qui indique les moyens d'établir un lien entre les régressions hédonistes et les techniques traditionnelles d'ajustement de la qualité utilisées par les instituts de statistique, Triplett (2001) a montré que, dans certaines conditions, la méthode habituelle d'appariement des modèles pour calculer globalement les variations de prix d'une période à la suivante (à l'aide de moyennes géométriques) donnait les mêmes résultats que ceux obtenus à l'aide d'un modèle de régression hédoniste[39]. Nous examinons maintenant le résultat obtenu par Triplett dans un cadre un peu plus général.

Rappelons-nous nos équations de modèle de régression hédoniste (14). Supposons en outre que le logarithme de f(z) est une fonction linéaire des J paramètres inconnus, $\alpha_1,...,\alpha_J$; nous avons donc:

(36) $\ln f(z_k^t) \equiv \alpha_1 + \sum_{j=2}^J x_j(z_k^t)\, \alpha_j$; $\qquad t = 1,...,T;\ k = 1,...,K^t$

où les fonctions $x_j(z_k^t)$ sont connues. Il est à noter que nous avons supposé que $x_1(z_k^t) \equiv 1$; autrement dit, nous avons supposé que la forme fonctionnelle pour f(z) contient un terme constant. On prend maintenant les logarithmes des deux membres des équations (14), on y utilise l'identité (36) et on ajoute les termes stochastiques ε_k^t pour obtenir le système suivant d'équations de régression:

(37) $\ln P_k^t = \beta_t + \alpha_1 + \sum_{j=2}^J x_j(z_k^t)\,\alpha_j + \varepsilon_k^t$; $\qquad t = 1,...,T;\ k = 1,...,K^t$

où, comme de coutume, nous avons défini $\beta_t \equiv \ln \rho_t$ pour t = 1,...,T. Une normalisation est nécessaire pour pouvoir identifier tous les paramètres dans (37). Nous choisissons la normalisation $\rho_1 = 1$, qui se traduit par la normalisation suivante:

(38) $\beta_1 = 0$.

En utilisant la notation matricielle, nous pouvons exprimer les équations (37) pour la période t comme suit:

(39) $y^t = 1^t\,\beta_t + X^t\,\alpha + \varepsilon^t$; $\qquad t = 1,...,T$

où $y^t \equiv [\ln P_1^t,...,\ln P_{Kt}^t]'$ est un vecteur de logarithmes de prix de modèles pour la période t (le symbole indiquant la transposée du vecteur qui précède), β_t est le paramètre scalaire $\ln\rho_t$, 1^t est un vecteur colonne consistant à K^t, composantes de valeur un, X^t est une matrice K^t x J de variables exogènes, $\alpha \equiv [\alpha_1,...,\alpha_J]'$ est un vecteur colonne de paramètres qui déterminent la fonction de sous-utilité hédoniste et $\varepsilon^t \equiv [\varepsilon_1^t,...,\varepsilon_{Kt}^t]'$ est un vecteur colonne de perturbations au cours de la période t. On réécrit maintenant le système d'équations (39) sous la forme suivante:

(40) $y = W\gamma + \varepsilon$

où $y' \equiv [y^{1'},...,y^{T'}]$, $\varepsilon' \equiv [\varepsilon^{1'},...,\varepsilon^{T'}]$, $\gamma' \equiv [\beta_2, \beta_3,..., \beta_T,\alpha_1,...,\alpha_J]$, et la matrice W est une matrice assez compliquée, que l'on construit en utilisant les vecteurs colonnes 1^t et les matrices K^t x J X^t pour t = 1,…,T[40].

Le vecteur des estimateurs des moindres carrés pour les composantes de γ est

(41) $\gamma^* \equiv (W'W)^{-1}W'y$.

On définit le vecteur des résidus des moindres carrés e comme suit:

(42) $e \equiv y - W\gamma^* = y - W(W'W)^{-1}W'y$.

Il est bien connu que le vecteur des résidus des moindres carrés e est orthogonal aux colonnes de W; en d'autres termes, nous avons:

(43) $W'e = W'[y - W(W'W)^{-1}W'y] = W'y - W'y = 0_{T-1+J}'$

où 0_{T-1+J} est un vecteur de zéros de dimension T−1+J. On prémultiplie maintenant les deux membres de $e \equiv y - W\gamma^*$ par les transposées des premières colonnes T–1 de W. En utilisant l'équation (43), nous obtenons les équations suivantes:

(44) $0 = 1^{t\prime}y^t - 1^{t\prime}1^t\,\beta_t^* - 1^{t\prime}X^t\,\alpha^*$; $\qquad t = 2,3,...,T$

où β_t^* est l'estimateur des moindres carrés pour β_t et $\alpha^* \equiv [\alpha_1^*,...,\alpha_J^*]'$ est le vecteur des estimateurs des moindres carrés pour $\alpha \equiv [\alpha_1,...,\alpha_J]'$. Maintenant, la colonne T dans W correspond au terme constant α_1 et est par conséquent un vecteur de composantes à valeur un.

On prémultiplie les deux membres de l'équation (42) par cette colonne et, en utilisant l'équation (43), on obtient l'équation suivante:

(45) $0 = \sum_{t=1}^{T} 1^{t\prime} y^t - \sum_{t=2}^{T} 1^{t\prime} 1^t \beta_t^* - \sum_{t=1}^{T} 1^{t\prime} X^t \alpha^*.$

On utilise les équations (44) dans l'équation (45) pour obtenir l'équation suivante:

(46) $1^{t\prime} y^1 = 1^{t\prime} X^1 \alpha^*.$

Comme $1^{t\prime} 1^t = K^t$ (le nombre de prix de modèles collectés pendant la période t), nous pouvons reformuler les équations (44) de la façon suivante:

(47) $\beta_t^* = (1/K^t) \sum_{k=1}^{Kt} y_k^t - (1/K^t)\, 1^{t\prime} X^t \alpha^*;$ $t = 2,3,\ldots,T.$

52. Le β_t^* donné par le membre de droite de (47) peut être interprété de façon intéressante comme une moyenne arithmétique du vecteur des prix logarithmiques pendant la période t ajustés en fonction de la qualité $y^t - X^t\alpha^*$. Cependant, un résultat très intéressant découle de l'utilisation des équations (46) et (47) si nous supposons que l'échantillon de prix des modèles *correspond* à celui de l'ensemble T de périodes (de sorte que les prix, au cours de chaque période, concernent exactement les mêmes modèles). Si l'échantillon correspond, chaque matrice X^t est exactement la même (et tous les K^t ont une taille d'échantillon commune K). Si la matrice commune X^t est la matrice X de dimension K x T–1+J, alors en utilisant les équations (46) et (47), on obtient la formule suivante pour β_t^*:

(48) $\beta_t^* = (1/K) \sum_{k=1}^{K} y_k^t - (1/K) \sum_{k=1}^{K} y_k^1;$ $t = 2,3,\ldots,T.$

En conséquence, dans le cas d'échantillons correspondants, si l'on prend l'exponentielle de β_t^* comme estimateur de ρ_t, en se rappelant que $y_k^t \equiv \ln P_k^t$, nous avons:

(49) $\rho_t^* \equiv [\prod_{k=1}^{K} P_k^t]^{1/K} / [\prod_{k=1}^{K} P_k^1]^{1/K} = [\prod_{k=1}^{K} (P_k^t/P_k^1)]^{1/K};$ $t = 2,3,\ldots,T;$

c'est-à-dire que, dans le modèle d'appariement la méthode de régression hédoniste *donne exactement le même résultat pour la variation globale des prix de la période 1 à la période t que celui que nous obtiendrions en prenant la moyenne géométrique des rapports des prix obtenus selon la méthode de l'appariement des modèles pour les deux périodes considérées*.

Triplett a indiqué que ce résultat était vrai dans le cas T = 2 et dans l'hypothèse où f était la fonction d'utilité hédoniste log-log décrite ci-dessus dans la section «Formes fonctionnelles fréquemment utilisées».

53. J'estime que l'article de Silver et Heravi (2001) et le Manuel de Triplett (2001) sont tous deux très utiles, dans la mesure où ils indiquent très clairement que les techniques traditionnelles d'appariement des modèles pour l'ajustement de la qualité donnent des résultats très proches de ceux de la méthode de régression hédoniste. Cette correspondance entre les deux méthodes devrait contribuer à démythifier dans une certaine mesure les méthodes hédonistes. En outre, comme le soulignent si bien Silver et Heravi que Triplett, l'avantage qu'il y a, du point de vue statistique, à utiliser la méthode de régression hédoniste plutôt que la méthode d'appariement des modèles augmente à mesure que la concordance entre les échantillons des modèles diminue.

Autrement dit, la technique hédoniste utilise toutes les informations relatives aux modèles pour les deux périodes considérées, tandis que la méthode d'appariement des modèles ne peut par définition utiliser que les informations disponibles sur les modèles présents sur le marché pendant les deux périodes.

Régressions hédonistes et utilisation de facteurs de pondération des quantités

54. L'étude de Silver et Heravi (2001) fondée sur la régression hédoniste est assez unique dans son genre, puisque ses auteurs disposaient non seulement de données sur les prix et les caractéristiques des machines à laver vendues au Royaume-Uni en 1998, mais également sur les ventes de chaque modèle. La question abordée dans la présente section est la suivante: dans quelle mesure les données sur les quantités doivent-elles être utilisées dans une étude fondée sur la régression hédoniste?

55. Nous commençons par examiner un modèle très simple dans lequel il n'existe qu'une seule variété sur le marché au cours de la période t, mais pour lequel on dispose de K observations sur les prix, P_k^t au cours de la période t, ainsi que des quantités correspondantes vendues à chacun de ces prix, q_k^t. Compte tenu de ces hypothèses, nos équations de régression hédoniste de base (14) pour la période t deviennent:

(50) $P_k^t = \rho_t f(z_k^t) = \rho_t$; $\qquad k = 1,2,...,K$

où nous supposons que $f(z_k^t) = 1$, puisque l'ensemble des K transactions portent sur le même modèle.

56. Lorsqu'on examine l'équation (50), on constate que ρ_t peut être interprété comme une sorte de moyenne des K prix sur le marché observés pendant la période t, P_k^t. La *fréquence relative* à laquelle le prix P_k^t est observé sur le marché pendant la période t peut être définie comme suit:

(52) $\theta_k^t \equiv q_k^t / \sum_{i=1}^K q_i^t$.

La *valeur moyenne de la distribution discrète des prix au cours de la période t* est

(53) $\rho_t^* \equiv \sum_{k=1}^K \theta_k^t P_k^t = \sum_{k=1}^K q_k^t P_k^t / \sum_{i=1}^K q_k^t$ $\qquad$ avec l'équation (52).

57. Il est à noter que le troisième membre de l'équation (53) a la *valeur unitaire*. En conséquence, il est possible d'utiliser des données quantitatives sur les ventes d'un modèle pour établir un *prix moyen représentatif* pour ce modèle sur une période, ce prix représentatif étant un prix moyen global pondéré pour l'ensemble des ventes pour le modèle considéré ou la valeur unitaire[41].

58. Comment pouvons-nous obtenir l'estimateur de la valeur unitaire pour le prix représentatif sur la période t ρ_t en utilisant une régression hédoniste? Il y a au moins deux façons de le faire.

59. Examinons l'équation k dans le système d'équations de prix (50). Comme q_k^t ventes ont été conclues à ce prix pendant la période t, on peut reprendre l'équation $P_k^t = \rho_t$ un certain nombre de fois, q_k^t fois pour être précis. Soit 1_k un vecteur de dimension q_k^t. En utilisant la notation

vectorielle, nous pourrions reformuler le système d'équations (50), en répétant chaque prix P_k^t autant de fois qu'une vente a eu lieu à ce prix au cours de la période t; nous avons alors:

(54) $1_k P_k^t = 1_k \rho_t$; $\qquad$ k = 1,2,...,K.

60. Ajoutons maintenant les termes d'erreur dans chacune des équations de l'équation (54) et calculons l'estimateur des moindres carrés pour la régression linéaire obtenue. Il en résulte que l'estimateur est l'estimateur de la valeur unitaire ρ_t^* défini par l'équation (53).

61. La deuxième façon d'obtenir l'estimateur de la valeur unitaire pour le prix représentatif sur la période ρ_t au moyen d'une régression hédoniste consiste à multiplier les deux membres de l'équation k dans les équations (50) par la racine carrée du nombre de ventes de modèle k pendant la période t, $(q_k^t)^{1/2}$, et à ajouter ensuite un terme d'erreur ε_k^t. Nous obtenons alors le système d'équation suivant:

(55) $(q_k^t)^{1/2} P_k^t = (q_k^t)^{1/2} \rho_t + \varepsilon_k^t$; $\qquad$ k = 1,2,...,K.

62. Il est à noter que les variables du membre de gauche dans l'équation (55) sont connues. On traite maintenant cette équation comme une régression linéaire pour estimer le paramètre inconnu ρ_t. Il peut être vérifié que l'estimateur des moindres carrés pour ρ_t est l'estimateur de la valeur unitaire ρ_t^* défini par l'équation (53)[42]. Nous pouvons donc utiliser une régression hédoniste des moindres carrés pondérée pour obtenir un prix de modèle moyen plus représentatif pour la période t.

63. Les développements qui précèdent peuvent contribuer à expliquer pourquoi Silver et Heravi (2001) utilisent des régressions hédonistes pondérées en fonction des ventes dans leurs modèles de régression. L'utilisation de régressions pondérées en fonction des quantités vendues diminue l'influence de prix non représentatifs[43] et devrait permettre de mieux mesurer la tendance centrale de la distribution de prix de modèles ajustés en fonction de la qualité; en d'autres termes, l'utilisation de facteurs de pondération des quantités vendues devrait permettre d'obtenir des estimations plus exactes des paramètres ρ_t dans les équations (14).

Indices hédonistes exacts

64. Silver et Heravi (2001) ont fourni un travail considérable en vue d'évaluer deux des limites de Feenstra (1995) pour l'indice hédoniste exact. Plus haut, dans la deuxième section (Un regard neuf sur la théorie des indices de prix hédonistes), nous avons fait de fortes hypothèses simplificatrices concernant la structure des préférences des consommateurs, hypothèses qui étaient assez différentes de celles faites par Feenstra. Dans la présente section, nous examinons les conséquences de nos hypothèses dans l'optique de la construction d'indices hédonistes exacts[44].

65. Rappelons-nous nos équations hédonistes de base (14): $P_k^t = \rho_t f(z_k^t)$ pour t = 1,...,T et k = 1,...,K^t. Nous supposons que le prix P_k^t est le prix moyen pour tous les modèles de type k vendus pendant la période t et nous supposons que q_k^t est le nombre d'unités vendues du modèle k pendant cette période t. On se rappellera que le nombre de modèles présents sur le marché pendant la période t s'élevait à K^t.

66. Dans la présente section, nous supposerons qu’il existe K modèles sur le marché pendant l’ensemble des T périodes de notre période d’échantillonnage. Si un modèle k donné ne s’est pas vendu du tout au cours de la période t, alors nous supposerons que P_k^t et q_k^t sont tous deux égal à zéro. Compte tenu de ces conventions, la *valeur totale des achats des consommateurs au cours de la période t* est égale à:

(56) $\sum_{k=1}^{K} P_k^t q_k^t = \sum_{k=1}^{K} \rho_t f(z_k) q_k^t;$ $t = 1,...,T.$

67. La fonction de sous-utilité hédoniste f a accompli l’ensemble de la tâche difficile dans notre modèle, en convertissant l’utilité fournie par le modèle k au cours de la période t en une utilité «normalisée» $f(z_k)$ qui, en tant que nombre cardinal, peut être comparée dans les différents modèles. Ensuite, pour chaque modèle k, nous effectuons simplement la multiplication par le nombre total d’unités vendues au cours de la période t, q_k^t, afin d’obtenir *la quantité totale de la marchandise hédoniste présente sur le marché pendant la période t*, qui peut être indiquée par Q_t. Nous avons donc[45]:

(57) $Q_t \equiv \sum_{k=1}^{K} f(z_k) q_k^t;$ $t = 1,...,T.$

68. Le prix agrégé correspondant de la marchandise hédoniste est ρ_t. En conséquence, dans notre modèle très simplifié, *le prix et la quantité exacts agrégés pour la période t* en ce qui concerne la marchandise hédoniste sont ρ_t et Q_t définis par (57), qu’il est aisé de calculer, pour autant que nous ayons estimé les paramètres de la régression hédoniste (14) et que nous disposions de données sur les quantités vendues au cours de chaque période, c’est-à-dire q_k^t[46].

69. Lorsque ρ_t et Q_t ont été déterminés pour $t = 1,\ldots,T$, ces estimations de prix et de quantités agrégés pour la marchandise hédoniste peuvent être combinées avec les prix et quantités agrégés de marchandises non hédonistes à l’aide de la théorie normale des indices.

70. Nous terminons cette section en examinant un autre aspect de l’article de Silver et Heravi, à savoir leur utilisation d’*indices superlatifs dans les modèles d’appariement*. Un *indice des prix dans les modèles d’appariement* pour la marchandise hédoniste de la période t à la période t+1 est construit de la façon suivante. Soit I(t, t+1) l’ensemble de modèles k qui sont vendus au cours des deux périodes t et t+1. Alors, *les indices de prix de Laspeyres et de Paasche des modèles correspondants* de la période t à la période t+1, soit P_L et P_p, respectivement, sont:

(58) $P_L^t \equiv [\sum_{k\in I(t,\,t+1)} P_k^{t+1} q_k^t]/[\sum_{k\in I(t,\,t+1)} P_k^t q_k^t];$

(59) $P_P^t \equiv [\sum_{k\in I(t,\,t+1)} P_k^{t+1} q_k^{t+1}]/[\sum_{k\in I(t,\,t+1)} P_k^t q_k^{t+1}].$

71. Dans les indices des modèles correspondants ci-dessus, nous ne comparons que les modèles qui ont été vendus au cours des deux périodes considérées. En conséquence, nous négligeons certaines informations sur les prix dont nous disposons (c’est-à-dire sur les prix des modèles seulement présents au cours de l’une des deux périodes). L’*indice de prix idéal de Fisher superlatif des modèles correspondants* de la période t à la période t+1 est $P_F^t \equiv [P_L^t P_P^t]^{1/2}$; autrement dit, il s’agit de la racine carrée du produit des indices de Laspeyres et de Paasche des modèles correspondants. Il est maintenant possible de comparer la mesure de Fisher de l’évolution des prix des modèles correspondants de la période t à la période t+1, (P_F^t), à la mesure correspondante du changement de prix agrégé que nous pourrions obtenir dans notre

modèle hédoniste, soit ρ_{t+1}/ρ_t. Nous espérons que ces mesures de variation des prix seront très semblables, en particulier si la proportion de modèles correspondants est élevée pour chaque période (comme c'est le cas en ce qui concerne les données de Silver et Heravi). Silver et Heravi (2001) font cette comparaison dans leurs modèles hédonistes et constatent que l'indice de Fisher des modèles correspondants est inférieur d'environ 2 % à ceux des modèles hédonistes en ce qui concerne leurs données relatives aux machines à laver au Royaume-Uni en 1998. Il semble très possible que cet écart relativement élevé soit dû au fait que les formes fonctionnelles hédonistes de Silver et Heravi ne permettent d'obtenir qu'une approximation au premier ordre des préférences hédonistes arbitraires, tandis que les indices superlatifs donnent une approximation au deuxième ordre et que les effets de substitution sont donc plus importants pour les indices des prix superlatifs des modèles correspondants[47].

72. Une conséquence importante de l'article de Silver et Heravi se dégage donc: il n'est pas nécessaire d'entreprendre une étude hédoniste

- si des données détaillées sur le prix et la quantité vendue sont disponibles pour chaque modèle; et

- si, entre des périodes consécutives, le nombre de nouveaux modèles et de modèles qui ont disparu est faible, de sorte que la correspondance est relativement élevée.

73. Nous examinons maintenant notre dernier thème: les problèmes supplémentaires qui se posent si nous assouplissons l'hypothèse selon laquelle la fonction de sous-utilité hédoniste f(z) ne varie pas dans le temps.

Évolution des goûts et fonction d'utilité hédoniste

74. Plusieurs économistes ont suggéré qu'il existe de bonnes raisons pour que la fonction d'utilité hédoniste f(z) présentée dans la deuxième section (Un regard neuf sur les indices de prix hédonistes) dépende du temps t[48]. Dans la présente section, nous examinons les modifications qu'il faut apporter à notre modèle hédoniste de base présenté dans la deuxième section si nous remplaçons notre fonction d'utilité hédoniste invariante dans le temps f(z) par une autre fonction qui varie dans le temps, par exemple $f^t(z)$[49].

75. Si nous remplaçons la fonction f(z) utilisée dans la deuxième section par $f^t(z)$ et faisons les mêmes autres hypothèses que dans cette dernière, nous constatons que nos équations (14) sont maintenant remplacées par les équations suivantes:

(60) $P_k^{\,t} = \rho_t\, f^t(z_k^{\,t})$; $\qquad t = 1,\ldots,T\ ;\ k = 1,\ldots,K^t.$

76. Jusqu'à présent, peu de changements ont été apportés au modèle présenté dans la deuxième section, qui postulait une fonction de sous-utilité hédoniste invariante dans le temps f(z), si ce n'est que notre nouvelle fonction de sous-utilité $f^t(z)$ comprendra naturellement certains paramètres dépendants du temps. Cependant, notre nouveau modèle (60) est lié à un autre changement important. Il faut se rappeler que, dans les modèles invariants dans le temps dont il a été question dans la section 3 (Questions relatives aux formes fonctionnelles), *une seule* normalisation des paramètres, telle que $\rho_1 = 1$, était nécessaire. Dans notre nouveau cadre dépendant du temps, une normalisation des paramètres de (60) est nécessaire pour

chaque période; autrement dit, *T normalisations des paramètres sont nécessaires au lieu d'une seule*, pour identifier les paramètres ρ_1 et α qui caractérisent $f^t(z)$.

77. La façon la plus simple d'obtenir les normalisations nécessaires consiste à faire l'hypothèse selon laquelle un *modèle de référence* ayant les caractéristiques $z^* \equiv (z_1^*,...,z_N^*)$ donne au consommateur la *même utilité* pendant toutes les périodes d'échantillonnage. Si nous choisissons que ce niveau d'utilité de référence est l'unité, alors cette hypothèse se traduit par l'imposition des restrictions suivantes aux paramètres de $f^t(z)$:

(61) $f^t(z^*) = 1;$ $t = 1,...,T.$

Les équations (60) et (61) deviennent maintenant notre système de base d'équations de régression hédoniste qui remplace notre ancien système (14) plus la normalisation $\rho_1 = 1$[50].

78. Comment devons-nous choisir la forme fonctionnelle de $f^t(z)$? Il existe bien entendu de nombreuses possibilités. Cependant, la plus simple (qui est du reste celle qui a été choisie par Silver et Heravi) consiste à laisser les paramètres α_n que nous avons définis pour diverses formes fonctionnelles dans la troisième section (Questions relatives aux formes fonctionnelles) dépendre de t; en d'autres termes, les α_n définis dans la troisième section sont remplacés par des α_n^t et chaque ensemble de paramètres relatifs à la période t est estimé au moyen d'une régression hédoniste qui utilise *uniquement* les données relatives aux prix et aux caractéristiques pour la période t[51]. Nous laissons au lecteur le soin d'étudier les détails des modifications à apporter: réécrire les formules algébriques figurant dans la troisième section, changer les α_n en α_n^t et effectuer les normalisations (61) au lieu de notre ancienne normalisation $\rho_1 = 1$.

79. Jusqu'ici, pas de problème. Il semble que nous ayons considérablement généralisé notre ancien modèle hédoniste «statique», pratiquement sans inconvénient. Cependant, il existe un inconvénient caché. Notre nouveau système d'équations de régression (60) et (61) est en général *non invariant à l'égard du choix du modèle de référence à vecteur de caractéristiques* z^*. Ainsi, si nous choisissons un modèle de référence différent ayant un vecteur de caractéristiques $z^{**} \neq z^*$ et remplaçons les normalisations (61) par

(62) $f^t(z^{**}) = 1;$ $t = 1,\ldots, T.$

alors, en général, les nouvelles estimations de prix de marchandises hédonistes agrégées ρ_1 s'en trouvent modifiées. Par conséquent, si l'on postule une fonction d'utilité hédoniste dépendante du temps, il en résulte l'inconvénient d'un *manque d'invariance* dans les prix relatifs de la marchandise hédoniste agrégée dans le temps, à l'égard de nos normalisations de fonctions d'utilité (61) ou (62).

80. Ce manque d'invariance de notre ρ_1 estimé ne pose pas nécessairement de problème aux instituts de statistique, pour autant que nous puissions nous mettre d'accord sur un choix «raisonnable» en ce qui concerne le modèle de référence caractérisé par le vecteur de caractéristiques z^*, car ce qui importe pour les instituts de statistique, c'est d'obtenir des estimations «raisonnables» et reproductibles pour les prix des marchandises hédonistes agrégées. Sur la base d'une analyse de ce problème dans Silver (1999b; 47), on peut suggérer,

à titre préliminaire, de prendre z^* comme vecteur de caractéristiques moyen pondéré en fonction des ventes des modèles qui ont fait leur apparition au cours de la période d'observation:

(63) $z^* \equiv \sum_{k=1}^{K} \sum_{t=1}^{T} q_k^t z_k / \sum_{k=1}^{K} \sum_{t=1}^{T} q_k^t$

où nous sommes revenus à la notation utilisée dans la sixième section (Indices hédonistes exacts); en d'autres termes, K est le nombre total de modèles distincts qui ont été vendus pendant l'ensemble des T périodes dans notre échantillon et q_k^t est le nombre de modèles qui ont le vecteur de caractéristiques z_k qui ont été vendus pendant la période t[52].

81. Ainsi, une fois que nous avons choisi nos formes fonctionnelles pour $f^t(z)$ et ajouté les termes stochastiques à (60), les équations (60) et (61) et la définition (63) spécifient complètement notre nouveau cadre de régression hédoniste. Bien entendu, nous recommandons toujours que des pondérations en fonction des quantités (si elles sont disponibles) soient utilisées dans l'estimation économétrique pour les raisons exposées dans la section 5 (Régressions hédonistes et utilisation de coefficients de pondération fondés sur la quantité); on se rappellera les équations (55) plus haut.

82. Cependant, si le nombre de périodes pour notre échantillon T est élevé, alors il existe un risque que le vecteur de caractéristiques z^* défini par (63) ne soit pas très représentatif pour une période quelconque ou deux périodes consécutives quelconques. En conséquence, nous suggérons maintenant une autre méthode pour normaliser ou rendre comparables les fonctions d'utilité hédoniste dépendantes du temps $f^t(z)$ qui feront face à ce problème de manque de représentativité. Pour chaque période t, on définit z^{t*} comme étant le vecteur de caractéristiques moyen pondéré en fonction du nombre de ventes des modèles qui ont fait leur apparition sur le marché pendant la période t:

(64) $z^{t*} \equiv \sum_{k=1}^{K} q_k^t z_k / \sum_{k=1}^{K} q_k^t$; $\qquad t = 1,..., T.$

On se rappellera nos équations de régression hédoniste de base (60), $P_k^t = \rho_t f^t(z_k^t)$. On procède maintenant aux normalisations suivantes:

(65) $\rho_t = 1$; $\qquad t = 1,..., T.$

En supposant que les paramètres des fonctions d'utilité hédoniste de la période $f^t(z)$ ont été estimés, nous pouvons maintenant définir les indices de prix hédonistes des types Laspeyres, Paasche[53] et Fisher de la période t à la période t+1 respectivement comme suit:

(66) $P_L^{t,t+1} \equiv f^{t+1}(z^{t*})/f^t(z^{t*})$; $\qquad t =1,..., T-1;$

(67) $P_P^{t,t+1} \equiv f^{t+1}(z^{t+1*})/f^t(z^{t+1*})$; $\qquad t =1,..., T-1;$

(68) $P_F^{t,t+1} \equiv [P_L^{t,t+1} P_P^{t,t+1}]^{1/2}$; $\qquad t =1,..., T-1.$

83. L'indice de prix hédoniste de type Fisher est celui que nous préférons. On peut constater que les indices de Laspeyres et de Paasche définis par (66) et (67) peuvent être très étroitement liés aux indices de frontières supérieure et inférieure de Feenstra qui correspondent à son indice vrai [et cette méthode hédoniste exacte superlative est utilisée par Silver et Heravi (2001)], selon la forme fonctionnelle qui est choisie pour f^t.

84. Une fois que les paramètres qui caractérisent les fonctions d'utilité hédoniste dépendantes du temps $f^t(z)$ ont été estimés, de même que les prix de la marchandise hédoniste de la période t agrégés associés ρ_t[54], alors nous pouvons définir la demande agrégée de la période t pour la marchandise hédoniste de la façon suivante[55]:

(69) $Q_t \equiv \sum_{k=1}^{K} f^t(z_k)q_k^t$; $\quad t = 1,..., T.$

85. Le modèle ci-dessus est la *méthode directe* que nous proposons pour former les prix et quantités agrégés exacts de la période t, ρ_t et Q_t, pour la marchandise hédoniste.

86. Il est possible d'utiliser les résultats de régressions hédonistes d'une autre façon, plus *indirecte*, tout en appliquant la théorie normale des indices, afin de construire des indices de prix et de quantités agrégés pour la marchandise hédoniste [56]. On se rappellera les équations (58) et (59) qui figurent dans la section précédente et définissent des indices de prix de Laspeyres et de Paasche des modèles correspondants pour des modèles hédonistes allant de la période t à la période t+1. Ces indices présentent l'inconvénient d'écarter des informations sur les modèles qui ne sont vendus que pendant l'une des deux périodes considérées. Une façon d'utiliser ces informations écartées consiste à utiliser les régressions hédonistes en vue d'*estimer* les prix manquants[57].

87. On suppose que le modèle k n'était pas disponible ou ne s'est pas vendu pendant la période t (de sorte que $q_k^t = 0$) mais qu'il s'est vendu pendant la période t+1 (de sorte que P_k^{t+1} et q_k^{t+1} sont positifs). Le problème est que nous ne disposons pas de prix P_k^t pour ce modèle pendant la période t au cours de laquelle il ne s'est pas vendu. Cependant, pour la période t+1, notre équation de régression hédoniste pour ce modèle se présente comme suit (si on néglige le terme d'erreur):

(70) $P_k^{t+1} = \rho_{t+1} f^{t+1}(z_k).$

88. Nous pouvons maintenant utiliser la fonction d'utilité hédoniste estimée pour la période t+1, f^{t+1}, et le prix agrégé estimé pour la période t pour la marchandise hédoniste, ρ_t, afin de définir un *prix estimé pour le modèle k pendant la période t*, de la façon suivante:

(71)
$$\begin{aligned} P_k^{t*} &\equiv \rho_t f^{t+1}(z_k) \\ &= \rho_t\,[P_k^{t+1}/\rho_{t+1}] && \text{en utilisant (70)} \\ &= [\rho_t/\rho_{t+1}]\,P_k^{t+1}. \end{aligned}$$

89. Par conséquent, le prix estimé du modèle k pendant la période t, P_k^{t*}, est égal au prix observé du modèle k pendant la période t+1, P_k^{t+1}, multiplié par la réciproque du taux estimé de variation globale du prix de la marchandise hédoniste de la période t à la période t+1, reproduit à la formule $[\rho_t/\rho_{t+1}]$.

90. On suppose maintenant que le modèle k s'est vendu pendant la période t (de sorte que P_k^t et q_k^t sont positifs) mais que ce modèle k soit a disparu, soit ne s'est pas vendu pendant la période t+1 (de sorte que q_k^{t+1} est égal à 0). Le problème qui se pose est que nous n'avons pas de prix P_k^{t+1} pour ce modèle pendant la période t+1, au cours de laquelle il ne s'est pas vendu.

Cependant, pour la période t, notre équation de régression hédoniste pour le modèle k est la suivante (si on néglige le terme d'erreur):

(72) $P_k^t = \rho_t f^t(z_k)$.

91. Nous pouvons maintenant utiliser la fonction d'utilité hédoniste estimée pour la période t, f^t, et le prix agrégé estimé pour la période t+1 pour la marchandise hédoniste, ρ_{t+1}, afin de définir un *prix estimé pour le modèle k pendant la période t+1*, de la façon suivante:

$$
\begin{aligned}
(73)\quad P_k^{t+1*} &\equiv \rho_{t+1} f^t(z_k) \\
&= \rho_{t+1}\,[P_k^t/\rho_t] \qquad \text{en utilisant (72)} \\
&= [\rho_{t+1}/\rho_t]\,P_k^t.
\end{aligned}
$$

92. Par conséquent, le prix estimé pour le modèle k pendant la période t+1, P_k^{t+1*}, est égal au prix observé du modèle k pendant la période t, P_k^t, multiplié par le taux estimé de variation globale du prix de la marchandise hédoniste de la période t à la période t+1, $[\rho_{t+1}/\rho_t]$[58].

93. Nous pouvons maintenant utiliser les prix estimés définis par (71) et (73) afin d'obtenir des informations sur les prix et les quantités de *tous* les modèles qui étaient disponibles au cours d'une des deux périodes t et t+1 ou des deux périodes et nous pouvons donc calculer les *indices de prix de Laspeyres et de Paasche exactement vérifiés* suivants:

(74) $P_L^t \equiv [\sum_{k=1}^K P_k^{t+1} q_k^t]/[\sum_{k=1}^K P_k^t q_k^t]$;

(75) $P_P^t \equiv [\sum_{k=1}^K P_k^{t+1} q_k^{t+1}]/[\sum_{k=1}^K P_k^t q_k^{t+1}]$

où nous utilisons le prix estimé P_k^{t*} défini par (71) au lieu du P_k^t if $q_k^t = 0$ manquant mais, comme q_k^{t+1} est positif, nous utiliserons le prix imputé P_k^{t+1*} défini par (73) au lieu du $P_k^{t+`}$ manquant si $q_k^{t+1} = 0$ mais q_k^t est positif[59]. Si l'on compare nos nouveaux indices de prix de Laspeyres et de Paasche définis par (74) et (75) à nos anciens indices de prix de Laspeyres et de Paasche des modèles correspondants définis par (58) et (59), on peut constater que nos nouveaux indices n'écartent aucune information pertinente sur les prix et les quantités et donc s'attendre à ce qu'ils soient plus «exacts» dans un certain sens.

Conclusions

94. L'article de Silver et Heravi (2001) et les développements que nous lui avons consacrés plus haut permettent de tirer un certain nombre d'ébauches de conclusions:

- Les techniques traditionnelles d'indices superlatifs qui agrègent des données relatives à des modèles correspondants peuvent donner plus ou moins les mêmes résultats qu'une méthode hédoniste, pour autant que le taux de correspondance entre les échantillons soit relativement élevé;

- Il est difficile de justifier les régressions hédonistes linéaires en s'appuyant sur des considérations théoriques (du moins sur la base de notre approche très simplifiée des régressions hédonistes) et, en conséquence, il faut les éviter si possible;

- Si des régressions hédonistes entièrement exemptes de contraintes sont appliquées aux données de chaque période, alors il faut veiller à choisir soigneusement un modèle de référence qui permette de comparer l'utilité de la marchandise hédoniste entre périodes. En particulier, les estimations de variation de prix agrégés de la marchandise hédoniste ne sont généralement pas invariantes à l'égard du choix de modèle de référence;

- L'utilisation de coefficients de pondération fondés sur les quantités dans les modèles de régression hédoniste est fortement recommandée, si elle est possible;

- Dans certaines conditions, si les modèles sont disponibles au cours de chacune des périodes, alors la démarche de régression hédoniste produit exactement le même résultat que la démarche utilisée traditionnellement par les instituts de statistique pour calculer un indice élémentaire;

- Nous ne sommes pas parvenus à un consensus sur ce que devrait être la «meilleure pratique» en matière de spécification de régression hédoniste, mais des considérations relatives à des formes fonctionnelles flexibles devraient probablement être prises en compte lors de discussions consacrées à ce problème.

Notes

[1] L'unique exception était constituée par un indice de valeur unitaire qui était le prix moyen de l'ensemble des machines à laver sans ajustement pour l'évolution de la composition de l'ensemble de machines. Cet indice non ajusté en fonction de la qualité faisait apparaître une baisse de seulement 1 % sur l'année. Il est particulièrement intéressant de relever que la méthode hédoniste exacte de Feenstra (1995) débouche sur des résultats pratiquement identiques à ceux qui résultent des autres méthodes.

[2] Voir l'article de Diewert (1998) pour un examen des études reposant sur des données obtenues par lecture optique jusqu'alors.

[3] Si cette hypothèse est exacte, les modèles anciens devraient avoir tendance à présenter des résidus positifs dans les régressions hédonistes. Berndt, Griliches et Rappaport (1995; 264), Kokoski, Moulton et Zieschang (1999; 155) et Koskimäki et Vartia (2001; 4) ont trouvé des données qui étayent cette hypothèse pour les ordinateurs de bureau, les légumes frais et les ordinateurs, respectivement.

[4] En outre, lorsqu'un modèle disparaît, il est courant que les instituts de statistique demandent aux personnes chargées de recueillir des prix de rechercher le modèle le plus proche du modèle dépassé, c'est-à-dire un modèle qui ne tardera pas lui-même à être obsolète.

[5] Ces modèles devraient avoir des résidus négatifs lors de la constitution de l'échantillon dans une régression hédoniste.

[6] Cette observation est présentée par Berndt et Rappaport (2001). Cependant, il est intéressant de noter que tant Silver et Heravi que Berndt, Griliches et Rappaport (1995) constatent que cette erreur de pondération était relativement faible dans leurs études sur les machines à laver et les ordinateurs, lorsqu'ils ont comparé des indices superlatifs reposant sur la méthode d'appariement des modèles avec les résultats de régressions hédonistes non pondérées. Berndt, Griliches et Rappaport (1995) ont conclu que cette erreur de pondération pour les ordinateurs s'élevait à environ 0,7 point par an.

[7] Je tiens à signaler qu'un grand nombre de questions méthodologiques sont examinées de façon plus complète dans un document connexe, qui porte sur les postes de télévision au Royaume-Uni et non sur les machines à laver; voir Silver (1999b).

[8] Je suis donc l'exemple de Muellbauer (1974; 977), lorsqu'il écrit qu'il n'a pas honte de reconnaître que «sa démarche est délibérément unilatérale, dans la mesure où seule la demande est prise en compte. …L'objet de cet article est par conséquent assez différent de celui du récent article publié par Sherwin Rosen. L'offre et les problèmes de simultanéité qui peuvent surgir sont laissés de côté».

[9] Nous avons utilisé la notation de Rosen, qui est quelque peu différente de celle utilisée par Silver et Heravi.

[10] Nous ne supposons pas que tous les modèles possibles sont disponibles sur le marché. En fait, nous supposerons que seul un ensemble fini de modèles est disponible au cours de chaque période. En revanche, nous supposons que le consommateur a des préférences relatives qui portent sur tous les modèles possibles, étant entendu que chaque modèle est indexé par son vecteur de caractéristiques, $z = (z_1,\ldots,z_N)$. En conséquence, chaque consommateur préfère un modèle potentiel dont le vecteur de caractéristiques est $z^1 = (z_1{}^1,\ldots,z_N{}^1)$ à un autre modèle potentiel dont le vecteur de caractéristiques est $z^2 = (z_1{}^2,\ldots,z_N{}^2)$ si et seulement si $f(z^1) > f(z^2)$.

[11] Si la courbe d'indifférence pour la période t croise les deux axes, alors $g^t(u^t,Z)$ ne sera défini que pour un ensemble de valeurs de Z non négatives jusqu'à une limite supérieure.

[12] Si un consommateur achète par exemple deux unités d'un modèle à un prix P qui a les caractéristiques $z_1,...,z_N$, nous pouvons alors modéliser cette situation en introduisant un modèle artificiel qui se vend au prix de 2P, et a les caractéristiques $2z_1,...,2z_N$. Ainsi, la surface hédoniste $Z = f(z)$ n'englobe que les modèles les plus performants, y compris les modèles artificiels.

[13] Nous ne supposons pas que f(z) est une fonction quasi concave ou concave de z. Dans la théorie de la demande normale du consommateur, on peut supposer que f(z) est quasi concave sans perte de généralité, étant donné que des contraintes budgétaires linéaires et l'hypothèse de divisibilité parfaite impliquent que les courbes d'indifférence «efficaces» englobent les ensembles convexes. Cependant, comme Rosen (1974; 37-38) le fait observer, dans le cas des marchandises hédonistes, les différentes caractéristiques ne peuvent être séparées les unes des autres. En outre, on ne peut postuler une divisibilité parfaite et toutes les combinaisons possibles de caractéristiques ne sont pas disponibles sur le marché. En conséquence, les hypothèses habituelles faites dans la théorie de la demande normale du consommateur ne sont pas satisfaites dans le contexte hédoniste. Il est également à noter que, bien que nous posions comme hypothèse la progressivité des fonctions macro $g^t(u,Z)$ (l'existence de la dérivée partielle $\partial g^t(u,Z)/\partial Z$), nous n'imposons aucune restriction en matière de progressivité à la fonction de sous-utilité hédoniste f(z).

[14] Cette hypothèse est très forte et requiert certaines justifications. Elle est entièrement analogue à l'hypothèse selon laquelle les consommateurs ont les mêmes préférences homothétiques concernant, par exemple, les aliments. Bien que cette hypothèse ne soit pas justifiée dans certains cas, aux fins de la construction d'un indice des prix des aliments, elle est suffisante parce que ce qui nous intéresse avant tout c'est d'appréhender les effets de substitution dans le prix agrégé des aliments à mesure que les prix relatifs des éléments des aliments varient. De façon analogue, nous souhaitons déterminer comment le consommateur «moyen» évalue une vitesse plus grande d'un ordinateur par rapport à une mémoire accrue; en d'autres termes, nous nous intéressons principalement aux effets de substitution hédoniste.

[15] Nous n'avons pas besoin d'une courbe d'indifférence linéaire globalement mais seulement localement pour une certaine gamme d'achats. Autrement dit, nous pouvons considérer que la courbe d'indifférence linéaire est une approximation au premier ordre d'une courbe d'indifférence non linéaire.

[16] En comparant l'équation (12) avec l'équation (10), on peut constater que les hypothèses simplificatrices (11) nous ont permis de supprimer les termes $\partial g^t(u_i^t, f(z_k^t))/\partial Z$, qui dépendent des courbes d'indifférence des différents consommateurs entre la marchandise hédoniste et les autres marchandises. Si nous disposions, pour les différents ménages, des données sur la consommation de marchandises hédonistes et d'autres marchandises, nous pourrions utiliser les techniques relatives à la demande normale du consommateur pour estimer les paramètres qui caractérisent ces courbes d'indifférence.

[17] Nous avons substitué des indices aux exposants conformément aux conventions relatives aux paramètres dans les modèles de régression; en d'autres termes, les constantes ρ_t seront des paramètres de régression dans la suite du texte. Il convient également de noter que ρ_t est le prix de l'«autre» marchandise p^t multiplié par le paramètre pente a^t de la période t. Nous devons laisser ce paramètre pente varier dans le temps pour pouvoir modéliser la demande de marchandises hédonistes de haute technologie, dont le prix a baissé par rapport aux «autres» marchandises; autrement dit, nous considérons que a^t baisse dans le temps pour les marchandises de haute technologie.

[18] En fin de compte, notre modèle de base est très semblable à l'un des modèles hédonistes de Muellbauer (1974; 988-989); voir en particulier son équation (32).

[19] Il est possible d'adapter la théorie exposée plus haut et de lui donner une interprétation conforme à la théorie du producteur. L'homologue du problème de minimisation des dépenses (3) devient maintenant le problème de maximisation du bénéfice suivant: $\max_{X,Z} \{P^tZ - w^tX : X = g^t(k^t, Z)\}$, où Z est la production hédoniste et P^t est un prix au cours de la période t pour une unité de la production hédoniste, w^t est le prix au cours de la période t d'un facteur de production variable et X est la quantité utilisée de celui-ci, k^t est la quantité au cours de la période t d'un facteur fixe (le capital, par exemple) et g^t est la fonction des besoins en facteurs de l'entreprise. Si l'on suppose que $Z = f(z)$, nous obtenons l'homologue de l'équation (10) suivant conformément à la théorie du producteur: $P_k^t = f(z_k^t)\partial g^t(k^t, f(z_k^t))/\partial Z$. L'homologue de l'hypothèse (11) est, pour l'entreprise i, $g_i^t(k_i^t, Z) \equiv a^tZ - b_i^tk_i^t$ et l'homologue de l'équation (12) devient $P_k^t = w^ta^tf(z_k^t)$. Cependant, les hypothèses du modèle reposant sur la théorie du producteur ne sont pas aussi plausibles que les hypothèses correspondantes du modèle reposant sur la théorie du consommateur. En particulier, il y a assez peu de chances pour que chaque producteur pratique le même prix agrégé au cours de la période t pour une unité de facteur de production variable w^t et il y a assez peu de chances pour que chaque entreprise présente sur le marché hédoniste ait le même paramètre de technologie a^t. Cependant, l'hypothèse essentielle qui ne se vérifiera généralement pas dans le contexte du producteur est que chaque *producteur est capable de produire la gamme complète de modèles hédonistes* tandis que, dans le contexte du consommateur, il est tout à fait plausible que tous les consommateurs aient la possibilité d'acheter et de consommer tous les modèles.

[20] On définit c par l'équation $c(p^t) \equiv \min_x \{p^t \bullet x : h(x) = 1\}$, où $p^t \bullet x$ désigne le produit intérieur entre les vecteurs p^t et x.

[21] Notre analyse s'inspire en grande partie de Triplett (2001) et de Berndt (1991; chap. 4).

[22] Voir Berndt (1991; chap. 4) pour des indications historiques concernant les premières utilisations de ces formes fonctionnelles.

[23] Dans certains cas, il est commode que la fonction d'utilité hédoniste soit du type de fonction d'utilité qui est postulé dans la théorie des indices, où l'on suppose habituellement que la fonction d'utilité est homogène de degré un, croissante et concave. Par exemple, si nous voulons utiliser le cadre hédoniste pour modéliser des *achats liés* (c'est-à-dire que deux marchandises sont vendues ensemble à un prix forfaitaire), alors la fonction d'utilité hédoniste devient une fonction d'utilité ordinaire $f(z_1,z_2)$, où z_1 et z_2 sont les quantités des deux marchandises réunies dans l'emballage. Dans cette situation, il peut être bien de supposer que f est homogène de degré un, auquel cas le prix d'un ensemble réuni consistant en z_1 et z_2 unités des deux marchandises est $c(p_1,p_2)f(z_1,z_2)$, où $c(p_1,p_2) \equiv \min_{z\text{'s}} \{p_1z_1 + p_2z_2 : f(z_1,z_2) = 1\}$ est la fonction de coût unitaire qui est duale de la fonction d'utilité f. Il existe de nombreuses autres applications pour lesquelles il serait utile que f soit une fonction linéairement homogène.

[24] Pakes (2001) est d'avis que nous ne devons pas nous attendre à ce que nos estimations de régression hédoniste satisfassent des restrictions de monotonicité reposant sur le comportement stratégique d'entreprises lorsque celles-ci lancent de nouveaux modèles. Cependant, pour des raisons de crédibilité, il est probable que les instituts de statistique souhaiteront imposer des restrictions de monotonicité.

[25] Voir Diewert (1974; 127-133) (1993; 158-164) pour des exemples de formes fonctionnelles flexibles.

[26] L'article récent de Yu (2001) constitue une exception à cet égard. L'analyse de Yu est similaire à la nôtre à bien des égards et est plus générale parfois.

[27] La situation dans la théorie de la demande normale du consommateur peut être plus propice à une estimation exacte des formes fonctionnelles flexibles, étant donné que nous disposons d'un *système* complet d'équations permettant l'estimation dans le contexte normal. Ainsi, s'il existe N marchandises et observations sur les prix et les quantités pour T périodes sur H ménages, nous aurons H(N-1)T degrés de liberté pour l'approche systémique habituelle destinée à l'estimation des préférences des consommateurs. Dans le cadre de la régression hédoniste, nous avons $K^1+K^2+...+K^T$ ou approximativement KT degrés de liberté, où K est le nombre moyen de modèles au cours de chaque période.

[28] Voir Christensen, Jorgenson et Lau (1975).

[29] Compte tenu de ce qui a été indiqué dans la deuxième section («Un regard neuf sur la théorie des indices de prix hédonistes»), la fonction translog f(z) ne doit répondre à aucune condition en matière de courbure.

[30] Voir Diewert (1971).

[31] Si nous supposons que α_n et α_{ij} pour $i\neq j$ sont tous deux égaux à zéro dans l'équation (29), nous obtenons l'équation (21).

---

[32] Le contenu de cette section est essentiellement équivalent à ce que les statisticiens appellent un modèle d'analyse de variance (un schéma double avec termes d'interaction); voir le chapitre 4 de Scheffé (1959).

[33] Nous pouvons aussi *regrouper* les observations de façon que tous les modèles ayant une quantité z_1 de la première caractéristique comprise entre 0 et z_1^* fassent partie du groupe 1, tous les modèles ayant une quantité z_1 de la première caractéristique comprise entre z_1^* et z_2^* fassent partie du groupe 2,…, et tous les modèles ayant une quantité z_1 de la première caractéristique comprise entre z_{1-1}^* et z_1^* fassent partie du groupe I. Nous regroupons de la même manière les modèles pour la seconde caractéristique. De cette façon, tout modèle k au cours de chaque période se trouve dans un des IJ groupements discrets de modèles.

[34] Pour le choix entre les équations (32) et (34), un autre élément doit être pris en considération. Les paramètres qui nous intéressent le plus sont ρ_t et non leurs logarithmes, β_t. Cependant, comme Berndt (1991; 127) l'a fait observer, «expliquer les variations du logarithme naturel du prix n'est pas la même chose qu'expliquer les variations de prix». En conséquence, tant Silver et Heravi (2001) que Triplett (2001) font observer que l'antilog de l'estimateur des moindres carrés pour β_t n'est pas un estimateur non biaisé de ρ_t avec la spécification stochastique habituelle et ils citent Goldberger (1968), qui propose une méthode permettant de corriger ce biais. Koskimäki et Vartia (2001; 15) abordent également ce problème. Ces considérations conduisent à privilégier l'estimation de l'équation (32) à celle de l'équation (34).

[35] La question de la reproductibilité est très importante pour les instituts de statistique.

[36] Nous appliquons la convention habituelle selon laquelle les différentes caractéristiques sont définies de façon qu'une plus grande quantité d'une caractéristique quelconque se traduise par une plus grande utilité pour le consommateur.

[37] Il est à noter qu'il existe des restrictions de monotonicité comparables, que les modèles hédonistes continus énumérés dans les sections «Formes fonctionnelles fréquemment utilisées» et «Questions de flexibilité» doivent aussi respecter, et qu'il sera difficile d'imposer ces conditions à ces modèles également.

[38] Ici, nous ne tenons pas compte des termes d'erreur dans les régressions hédonistes.

[39] Koskimäki et Vartia (2001 ; 9) donnent un résultat plus général, très semblable au résultat que nous obtenons dans la suite de l'exposé.

[40] Il est à noter que nous avons utilisé la normalisation (38) en vue d'éliminer le paramètre β_1 du vecteur paramétrique γ.

[41] Il existe d'autres façons de pondérer les prix P_k^t. Par exemple, nous pourrions prendre la part des dépenses pour tous les modèles vendus au prix P_k^t pendant la période t, soit $s_k^t \equiv P_k^t q_k^t / \sum_{i=1}^K P_i^t q_i^t$ pour k = 1,...,K comme facteur de pondération pour P_k^t. Si l'on utilise ces facteurs de pondération, le prix moyen représentatif sur la période t devient $\rho_t^{**} \equiv \sum_{k=1}^K s_k^t P_k^t$. Il est à noter que, si nous divisons la valeur des transactions au cours de la période t, $\sum_{i=1}^K P_i^t q_i^t$ par ce prix, nous obtenons l'estimateur de quantité correspondant, $[\sum_{i=1}^K P_i^t q_i^t]^2 / \sum_{k=1}^K [P_k^t]^2 q_k^t$,

qui n'est pas facile à interpréter. D'autre part, si nous divisons la valeur des transactions au cours de la période t, $\sum_{i=1}^{K} P_i^t q_i^t$, par l'estimateur de la valeur unitaire du prix agrégé sur la période t ρ_t^* défini par l'équation (53), nous obtenons la simple somme des quantités vendues pendant la période t, $\sum_{k=1}^{K} q_k^t$, comme estimateur de la quantité correspondant. L'emploi de valeurs unitaires pour l'agrégation sur l'ensemble des transactions relatives à une marchandise identique pendant une période donnée pour obtenir un prix représentatif unique et une quantité pour la période considérée a été préconisé par Walsh (1901; 96) (1921; 88), Davies (1924; 187) et Diewert (1995; 20-24).

[42] Berndt (1991; 127) présente un argument économétrique similaire justifiant le modèle des moindres carrés pondérés (54) en termes d'un modèle incorporant des variances hétéroscédastiques pour le modèle non transformé.

[43] Griliches (1961) (1971; 5) a présenté cette observation il y a de nombreuses années.

[44] Nos hypothèses sont également très différentes de celles faites par Fixler et Zieschang (1992), qui ont adopté leur propre méthode de construction des indices hédonistes exacts.

[45] Ceci est équivalent à l'indice de quantité défini par Muellbauer (1974; 988) dans l'un de ses modèles hédonistes; voir son équation (30). Bien entendu, le fait de considérer ρ_t comme un prix pour l'agrégat de la quantité de marchandises hédonistes défini par l'équation (57) peut être justifié par le théorème d'agrégation de Hicks (1946; 312-313), puisque les prix des modèles $P_k^t = \rho_t f(z_k)$ ont tous le facteur commun de proportionnalité ρ_t.

[46] Si nous disposons de données pour q_k^t, alors il vaut mieux appliquer des régressions pondérées en fonction du nombre de ventes, comme indiqué dans la section précédente. Si nous ne disposons pas de données complètes sur les ventes des différents modèles mais si nous connaissons le nombre total de ventes pendant chaque période, nous pouvons appliquer le modèle de régression hédoniste (14) en utilisant un échantillon de prix de modèles et en divisant ensuite le nombre de ventes au cours de la période t par notre paramètre estimé ρ_t afin d'obtenir un estimateur pour Q_t.

[47] Cette interprétation est étayée par le fait que l'indice de Laspeyres des modèles correspondants était environ le même que les indices hédonistes calculés par Silver et Heravi. Cependant, d'autres facteurs sont en jeu et il se peut que cette «explication» soit incomplète.

[48] Plus précisément, Silver (1999a) et Pakes (2001) présentent des arguments très convaincants (reposant sur la théorie de l'organisation industrielle) en faveur de la thèse selon laquelle les coefficients de régression hédoniste qui sont estimés à l'aide des données relatives à la période t devraient dépendre de t. Griliches (1961) a également soutenu qu'il était peu probable que les coefficients de régression hédoniste soient constants au cours de périodes différentes.

[49] Avant de décrire de façon générale les fonctions d'agrégateurs hédonistes dépendantes du temps $f^t(z)$, nous tenons à attirer l'attention sur une méthode simple, conçue initialement par Court (1939) et Griliches (1961), qui tient compte de la dépendance du temps et n'exige pas de nouvelle méthode: on utilise simplement l'ancienne méthode fondée sur l'indépendance du temps, mais on restreint la régression à deux périodes consécutives. Cela fournit une mesure

de la variation de prix globale pour la marchandise hédoniste de la période t à la période t+1, par exemple. On effectue ensuite une autre régression hédoniste en utilisant uniquement les données relatives aux périodes t+1 et t+2, ce qui donne une mesure de l'évolution des prix de la période t+1 à la période t+2. Et ainsi de suite.

[50] Si nous définissons le prix estimé du modèle de référence pendant la période t comme P^{t*}, on constate, en utilisant (60) et (61), que $P^{t*} = \rho_t$ pour t = 1,…, T. Cela dit, dans la pratique, lorsque des régressions hédonistes non soumises à des restrictions pour la période t sont appliquées de façon isolée, les chercheurs omettent la variable indicatrice du temps et se bornent à estimer par régression, par exemple ln P_k^t sur ln $f^t(z_k^t)$, où les variables de régression du membre de droite ont un terme constant. Ensuite, le chercheur estime le prix agrégé de la période t de la marchandise hédoniste par l'identité $\rho^{t*} \equiv f^t(z^*)$, où z^* est un vecteur des caractéristiques de référence choisi pour des raisons de commodité. Cette procédure équivaut à notre procédure de variable indicatrice temporelle utilisant les normalisations (61).

[51] Si des données sur le nombre de ventes sont disponibles, nous recommandons l'utilisation de la démarche fondée sur une régression pondérée, qui a été expliquée dans la cinquième section (Régressions hédonistes et utilisation de coefficients de pondération fondées sur les quantités); il convient de se rappeler les équations (55). En outre, dans ce cas, si les modèles sont vendus à plus d'un prix au cours d'une période quelconque, alors nous pouvons pondérer chaque prix distinct en fonction du nombre de ventes correspondant ou simplement agréger sur l'ensemble des ventes du modèle spécifique k au cours de la période t, et faire de P_k^t le prix de la valeur unitaire sur l'ensemble de ces ventes. Dans les développements qui suivent, nous supposons que la deuxième option est retenue.

[52] Si des informations quantitatives sur les ventes de modèles, q_k^t, ne sont pas disponibles, alors on définit z^* comme une moyenne arithmétique non pondérée de z_k.

[53] Berndt, Griliches et Rappaport (1995; 262-263) et Berndt et Rappaport (2001) définissent les indices hédonistes des types Laspeyres et Paasche de cette façon. Cependant, l'idée de base remonte à Griliches (1971; 59) et à Dhrymes (1971; 111-112). On remarquera que (66) et (67) ne fonctionnent pas si le vecteur de caractéristiques de la période t est totalement différent du vecteur de caractéristiques de la période t+1. De même, des problèmes peuvent surgir si certaines caractéristiques sont zéro pour une période et non zéro pour une autre période; on se rappellera le problème du logarithme de zéro dont il a été question dans la troisième section (Questions relatives aux formes fonctionnelles).

[54] Selon notre deuxième méthode, dans laquelle nous posons que ρ_t est égal à l'unité, on définit $\rho_1 = 1$ et $\rho_{t+1} = \rho_t P_F^{t,t+1}$ pour t = 1,2,...,T−1, où l'indice-chaîne hédoniste de type Fisher $P_F^{t,t+1}$ est défini par (68). Selon cette deuxième méthode, une fois que les prix agrégés ρ_t ont été déterminés, nous obtenons les quantités agrégées Q_t comme valeurs déflatées, $\sum_{k=1}^{K} P_k^t q_k^t/\rho_t$, plutôt que d'utiliser les équations (69).

[55] Si des coefficients de pondération en fonction des quantités ne sont pas disponibles, nous ne pouvons calculer Q_t.

[56] Voir Moulton (1996; 170) pour une description de ces méthodes.

[57] Voir Armknecht et Maitland-Smith (1999) pour un très bon exposé des méthodes d'estimation.

[58] Je pense que la méthode décrite ici est compatible avec la méthode utilisée par Silver et Heravi pour obtenir des prix estimés pour les modèles manquants. Triplett (2001) décrit d'autres méthodes.

[59] Bien entendu, si tant q_k^t que q_k^{t+1} sont égaux à zéro, nous n'avons pas besoin d'estimateurs pour les prix manquants P_k^t et P_k^{t+1} pour calculer les indices de Laspeyres et de Paasche définis par (74) et (75).

Références

Armknecht, P.A. and Fenella-Maitland-Smith (1999), "Price Imputation and Other Techniques for Dealing with Missing Observations, Seasonality and Quality Change in Price Indices", pp. 25-49 in *Proceedings of the Measurement of Inflation Conference*, M. Silver and D. Fenwick (eds.), Cardiff University, August 31-September 1, London: Office for National Statistics.

Berndt, E.R. (1991), *The Practice of Econometrics: Classic and Contemporary*, Reading, MA: Addison-Wesley.

Berndt, E.R. and N.J. Rappaport (2001), "Price and Quality of Desktop and Mobile Personal Computers: A Quarter Century Historical Overview", *The American Economic Review*, forthcoming.

Berndt, E.R., Z. Griliches, and N.J. Rappaport (1995), "Econometric Estimates of Price Indexes for Personal Computers in the 1990's," *Journal of Econometrics* 68, 243-268.

Christensen, L.R., D.W. Jorgenson and L.J. Lau (1975), "Transcendental Logarithmic Utility Functions", *American Economic Review* 65, 367-383.

Court, A.T. (1939), "Hedonic Price Indexes with Automotive Examples", pp. 99-117 in *The Dynamics of Automobile Demand*, New York: General Motors Corporation.

Davies, G.R. (1924), "The Problem of a Standard Index Number Formula", *Journal of the American Statistical Association* 19, 180-188.

Diewert, W.E. (1971), "An Application of the Shephard Duality Theorem: A Generalized Leontief Production Function", *Journal of Political Economy* 79, 481-507.

Diewert, W.E. (1974), "Applications of Duality Theory", pp. 106-171 in *Frontiers of Quantitative Economics*, Volume 2, M.D. Intriligator and D.A. Kendrick (eds.), Amsterdam: North-Holland.

Diewert, W.E. (1993), "Duality Approaches to Microeconomic Theory", pp. 105-175 in *Essays in Index Number Theory*, Volume 1, W.E. Diewert and A.O. Nakamura (eds.), Amsterdam: North-Holland.

Diewert, W.E. (1995), "Axiomatic and Economic Approaches to Elementary Price Indexes", Discussion Paper 95-01, Department of Economics, University of British Columbia, Vancouver, Canada, V6T 1Z1. Available on the Web at: http://web.arts.ubc.ca/econ/diewert/hmpgdie.htm

Diewert, W.E. (1998), "Index Number Issues in the Consumer Price Index", *The Journal of Economic Perspectives*, 12, (1998), 47-58.

Dhrymes, P.J. (1971), "Price and Quality Changes in Consumer Capital Goods: An Empirical Study", pp. 88-149 in *Price Indexes and Quality Change*, Z. Griliches (ed.), Cambridge, MA: Harvard University Press.

Feenstra, R.C. (1995), “Exact Hedonic Price Indices”, *Review of Economics and Statistics* 77, 634-654.

Fixler, D. and K.D. Zieschang (1992), “Incorporating Ancillary Measures of Process and Quality Change into a Superlative Productivity Index”, *The Journal of Productivity Analysis* 2, 245-267.

Goldberger, A.A. (1968), “The Interpretation and Estimation of Cobb-Douglas Functions”, *Econometrica* 35, 464-472.

Griliches, Z. (1971), “Introduction: Hedonic Price Indexes Revisited”, pp. 3-15 in *Price Indexes and Quality Change*, Z. Griliches (ed.), Cambridge, MA: Harvard University Press.

Griliches, Z. (1961), “Hedonic Price Indexes for Automobiles: An Econometric Analysis of Quality Change”, in *The Price Statistics of the Federal Government*, G. Stigler (chairman): Washington D.C.: Government Printing Office. Reprinted as pp. 55-87 in *Price Indexes and Quality Change*, Z. Griliches (ed.), Cambridge, MA: Harvard University Press.

Hicks, J.R. (1946), *Value and Capital*, Second Edition, Oxford: Clarendon Press.

Kokoski, M.F., B.R. Moulton and K.D. Zieschang (1999), “Interarea Price Comparisons for Heterogeneous Goods and Several Levels of Commodity Aggregation”, pp. 123-166 in *International and Interarea Comparisons of Income, Output and Prices*, A. Heston and R.E. Lipsey (eds,), NBER Studies in Income and Wealth 61, Chicago: The University of Chicago Press.

Koskimäki, T. and Y. Vartia (2001), “Beyond Matched Pairs and Griliches-Type Hedonic Methods for Controlling Quality Changes in CPI Sub-Indices”, Paper presented at the 6th Ottawa Group Meeting, Canberra, Australia, April.

Moulton, B. (1996), “Bias in the Consumer Price Index: What is the Evidence?”, *Journal of Economic Perspectives* 10:4, 139-177.

Muellbauer, J. (1974), “Household Production Theory, Quality, and the ‘Hedonic Technique’”, *The American Economic Review* 64:6, 977-994.

Pakes, A. (2001), “Some Notes on Hedonic Price Indices, with an Application to PC’s”, paper presented at the NBER Productivity Program Meeting, March 16, Cambridge MA.

Scheffé, H. (1959), *The Analysis of Variance*, New York: John Wiley and Sons.

Silver, M. (1995), “Elementary Aggregates, Micro-Indices and Scanner Data: Some Issues in the Compilation of Consumer Prices”, *Review of Income and Wealth* 41, 427-438.

Silver, M. (1999a), “Bias in the Compilation of Consumer Price Indices When Different Models of an Item Coexist”, pp. 21-37 in *Proceedings of the Fourth Meeting of the International Working Group on Price Indices*, Washington D.C., April 22-24, 1998, W. Lane (ed.), Washington D.C,: U.S. Department of Labor, Bureau of Labor Statistics.

Silver, M. (1999b), “An Evaluation of the Use of Hedonic Regressions for Basic Components of Consumer Price Indices”, *The Review of Income and Wealth* 45:1, 41-56.

Silver, M. and S. Heravi (2001), “The Measurement of Quality-Adjusted Price Changes”, paper presented at the NBER Conference on Scanner Data and Price Indexes, September 15-16, 2000 at Arlington, Virginia, forthcoming in an CRIW-NBER Volume edited by R. Feenstra and M. Chapiro.

Rosen, S. (1974), “Hedonic Prices and Implicit Markets: Product Differentiation in Pure Competition”, *Journal of Political Economy* 82:1, 34-55.

Triplett, J. (2001), *Handbook on Quality Adjustment of Price Indexes for Information and Communication Technology Products*, forthcoming, Paris: OECD.

Walsh, C.M. (1901), *The Measurement of General Exchange Value*, New York: Macmillan and Company.

Walsh, C.M. (1921), *The Problem of Estimation*, London: P.S. King and Son.

Yu, K. (2001), “Trends in Internet Access Prices in Canada”, Paper presented at the 6th Ottawa Group Meeting, Canberra, Australia, April.
