



HAL
open science

Consommation marchande et contraintes non monétaires au Canada (1969-2008)

Julien Boelaert, François Gardes

► **To cite this version:**

Julien Boelaert, François Gardes. Consommation marchande et contraintes non monétaires au Canada (1969-2008). 2013. halshs-00922775

HAL Id: halshs-00922775

<https://shs.hal.science/halshs-00922775v1>

Submitted on 30 Dec 2013

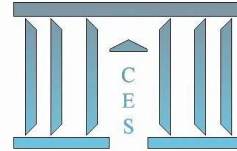
HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Documents de Travail du Centre d'Économie de la Sorbonne

C
E
S
W
o
r
k
i
n
g
P
a
p
e
r
s



Consommation marchande et contraintes non monétaires au Canada (1969-2008)

Julien BOELAERT, François GARDES

2013.82



Consommation marchande et contraintes non monétaires au Canada (1969-2008)

Julien Boelaert, François Gardes *

12 décembre 2013

Résumé - Nous construisons un pseudo-panel long à partir de huit enquêtes canadiennes de budgets de ménages (1969-2008) afin d'estimer des fonctions de demande sur longue période. La différence entre les estimations en dimension transversale et temporelle permet de révéler les contraintes non monétaires qui influent sur les comportements de consommation des ménages, et leur évolution. Nous trouvons des évolutions différenciées de ces contraintes selon les postes de consommation, mais une tendance commune à la réduction au cours du temps de la différenciation des contraintes selon le revenu.

Mots-clé : consommation, QUAIDS, pseudo-panel, contraintes non monétaires, prix virtuels

Abstract - A long pseudo panel is built from eight Canadian household budget surveys (1969-2008) in order to estimate demand functions on long periods. The difference between estimations in the cross-sectional and time dimensions allow the identification of non monetary constraints that influence consumption behaviour and its evolution. We find that while the evolution of constraints varies across consumption categories, there is a common tendency towards smaller differences in constraints across income levels.

Keywords : consumption, QUAIDS, pseudo panel, non monetary constraints, shadow prices

JEL : C33, D12

*Université Paris 1 Panthéon-Sorbonne

Centre d'Economie de la Sorbonne - Axe Economie Mathématique et Jeux

Maison des Sciences Economiques

106-112 boulevard de l'Hôpital, 75013 Paris

Auteur correspondant : julien.boelaert@gmail.com

Ce travail a bénéficié du financement du contrat ANR MALDI 06-BLAN-0140. Nous remercions Simon Langlois, professeur au département de sociologie de l'Université Laval (Québec), qui nous a donné accès aux données utilisées dans cette étude, ainsi que les participants au colloque 2011 de la Société Canadienne de Sciences Economiques et aux 30èmes Journées de Microéconomie Appliquée (2013) pour leurs commentaires.

1 Introduction

Le présent travail trouve son origine dans les travaux de Theil *et al* ([22], [21], [20]) sur les régularités internationales de la consommation des ménages, qui ont établi sur données agrégées par pays l'existence de lois de consommation stables et communes à tous les pays, ainsi que des résultats de convergence internationale des comportements de consommation. Dans le but de prolonger ces travaux sur données individuelles, nous nous sommes penchés dans une étude précédente [6] sur la convergence des structures de consommation des ménages canadiens depuis les années 1970, et avons trouvé que les structures moyennes de consommation par catégorie sociale (ou par quintile de revenu) ont eu tendance à se rapprocher de 1969 à 2008. Ce résultat, conforme à ceux de Theil, est étonnant lorsque l'on s'interroge sur ses causes, puisqu'au cours de cette période les revenus moyens par classe se sont écartés - du moins en termes absolus, l'écart relatif restant constant. Puisqu'en outre les prix s'imposent identiques à toutes les catégories sociales (c'est le propre d'une étude intranationale, par opposition à une étude internationale dans laquelle les prix relatifs peuvent tendre à s'harmoniser entre pays), il semble que les facteurs économiques monétaires ne peuvent pas expliquer ce mouvement de rapprochement.

C'est pour cette raison que nous nous tournons ici vers l'étude des déterminants non monétaires de la convergence des structures de consommation. En particulier, la méthode que nous présentons permet de révéler s'il existe des contraintes non monétaires corrélées au niveau de revenu. Les décisions économiques réelles sont affectées par une multitude de contraintes non monétaires : temps disponible pour la consommation d'un bien ou service, facilité d'accès à l'offre d'un type de biens, acceptabilité sociale d'un type de consommation... Parmi ces contraintes, certaines ne sont pas différenciées selon le revenu ; celles-là ne nous intéressent pas ici, puisque si elles sont distribuées uniformément à travers les catégories de revenu elles ne sauraient avoir de rôle dans la convergence entre catégories. En revanche, les contraintes corrélées au revenu ont pu évoluer au cours de la période, et ainsi influencer sur le rapprochement des comportements de consommation. En effet si les contraintes non monétaires sont moins différenciées qu'auparavant selon le niveau de revenu, il semble normal, si les préférences restent identiques dans toute la population, que les comportements de consommation des catégories sociales se soient rapprochés.

L'article est organisé comme suit : la deuxième section pose le cadre théorique de notre analyse, la troisième section expose le modèle et les données utilisés, la quatrième section traite des choix méthodologiques et des difficultés d'estimation, la cinquième section détaille les résultats d'estimations et en donne une discussion.

2 Cadre théorique

Notre point de départ est une analyse en termes de coûts complets. Le coût véritable de la consommation d'un bien ou service ne se limite pas au prix monétaire, *ie* au prix de marché. A ce prix monétaire s'ajoute en effet une foule de coûts liés par exemple à

la technologie de production domestique, à la valorisation du temps investi dans l'activité de consommation, à la structure familiale, aux contraintes de liquidité, aux normes sociales de consommation. L'analyse en coûts complets n'est pas nouvelle. L'exemple le plus commun est le modèle d'allocation du temps de Becker [3] dans lequel le prix complet de la consommation d'un bien est la somme de son prix monétaire et de son coût en temps, le temps étant valorisé au taux de salaire de l'agent. Deux agents dont le taux de salaire diffère seront donc confrontés à des prix complets différents pour la consommation d'un même bien, à prix monétaire et temps de consommation constants. On trouve un autre exemple d'analyse en coûts complets chez Barten [2], qui montre que la structure familiale des ménages peut se traduire en termes de différences de prix relatifs, et que par exemple "un litre de lait coûte relativement plus cher pour une famille nombreuse". La structure familiale d'un ménage détermine donc en partie le système de prix complets qui s'impose à lui.

La relation que nous établissons entre caractéristiques du ménages et coûts non-monétaires s'inspire de l'analyse de Neary et Roberts [18], qui montrent comment sous hypothèse de convexité, continuité et stricte monotonie des préférences, toute contrainte de rationnement d'un ménage peut se traduire par des prix virtuels pour les biens contraints (les prix virtuels étant définis comme les prix qui conduiraient un ménage non contraint à se comporter de la même manière que s'il était soumis à un vecteur de contraintes de rationnement). Les différentes contraintes qui s'imposent aux ménages au titre de leur structure familiale, de leur place dans l'échelle des revenus, et de leur technologie de production domestique pose un problème dans l'analyse des fonctions de demande en ce qu'elles introduisent une grande hétérogénéité entre ménages, ce qui complique l'analyse de leur choix. La traduction des contraintes de rationnement en prix virtuels à la Neary-Roberts présente le grand avantage de résumer cette hétérogénéité de caractéristiques inter-ménages à de simples différences de prix.

Gardes *et al* [13] montrent qu'il est possible de révéler les prix virtuels à partir de l'estimation d'un même modèle de demande dans deux dimensions, transversale et temporelle. En effet, en suivant Mundlak [17], dans un modèle de panel linéaire à effets fixes, les paramètres estimés en dimension transversale sont biaisés en présence d'une corrélation entre l'effet fixe et la transformation between d'une partie des variables exogènes. Une telle corrélation peut provenir de l'existence de variables latentes permanentes (telles la génération, le milieu social d'origine, le niveau d'éducation) corrélées à la fois avec l'effet fixe et les moyennes individuelles des variables exogènes. L'estimation du même modèle dans la dimension temporelle ne souffrira pas de ce biais. Formellement, Mundlak pose le modèle :

$$y_{it} = X_{it}\beta + Z_{it}\alpha_i + u_{it}$$

où β est le vecteur de paramètres d'intérêt, les Z_{it} sont des vecteurs de variables qualitatives, les u_{it} sont des erreurs idiosyncratiques ($u_{it} \sim (0, \sigma^2), i.i.d.$), et α_i est un vecteur d'effets. Si α_i est un effet fixe lié aux moyennes par individu des variables X (transformation between) par la spécification suivante :

$$\alpha_i = X_i\nu + w_i.$$

où ν est non nul et $w_i \sim (0, \omega^2), i.i.d$, alors en notant b_w et b_b respectivement les estimateurs within et between de β , Mundlak montre que b_w est BLUE, alors que l'estimateur between est biaisé :

$$E(b_b|X, Z) = \beta + \nu$$

et que la différence entre les deux estimateurs donne un estimateur BLUE de ν :

$$\hat{\nu} = b_b - b_w$$

L'originalité de l'argument de Gardes *et al* réside dans l'interprétation de cette différence d'estimateurs en termes de prix virtuels dans le cas d'une équation de demande. En effet, le paramètre ν ci-dessus représente le lien entre les effets fixes individuels et les régresseurs X ; il résume donc le lien entre les régresseurs et des variables latentes qui influent sur le comportement de consommation. Ces variables latentes sont composées de ressources non monétaires et de contraintes de rationnement des ménages, et peuvent donc être exprimées par des prix virtuels à la Neary-Roberts. La différence entre les estimations dans les dimensions temporelle et transversale permet alors de révéler les différences de prix virtuels¹ entre agents. Plus spécifiquement, si on considère pour le bien i , le ménage h à la période t le modèle de demande

$$x_{iht} = Z_{ht} \cdot \beta_i + \sum_j p_j \cdot \gamma_{ij} + u_{iht}$$

où x_i est le coefficient budgétaire du bien i , h et t sont respectivement des indices de ménage et de période, Z est un ensemble de caractéristiques, p_j le prix logarithmique du bien j , et le terme d'erreur se décompose en un terme spécifique individuel et un terme idiosyncratique ($u_{iht} = \alpha_{ih} + \epsilon_{iht}$), alors en considérant uniquement l'effet sur le prix du bien i d'un sous-ensemble $\{z_1, z_2, \dots, z_K\}$ des variables Z , on obtient, pour un coefficient de prix γ_{ii} fixé,

$$d\pi_i = \sum_{k=1}^K \left(\frac{d\pi_i}{dz_i^k} \right) dz_i^k = \sum_{k=1}^K \left(\frac{\beta_i^{k(CS)} - \beta_i^{k(TS)}}{\gamma_{ii}} \right) dz_i^k \quad (1)$$

avec π_i prix virtuel logarithmique du bien i , et $\beta_i^{k(CS)}$ et $\beta_i^{k(TS)}$ respectivement les estimateurs en dimension transversale et temporelle du coefficient du régresseur Z_k pour le bien i .²

Ce raisonnement fournit une explication au constat empirique de différence entre élasticités transversales et temporelles de la consommation, constat établi sur des données de plusieurs pays. Par exemple, Gardes *et al* [13] et Cardoso et Gardes [8] trouvent, pour la France, les Etats-Unis et la Pologne, une différence significative pour les deux

1. La définition des prix virtuels dans Gardes *et al* [13] et ici n'est pas la même que celle de Neary-Roberts. Alors que chez ces derniers les prix virtuels correspondent aux prix complets, pour nous le prix virtuel π est la partie non monétaire du prix complet, selon une spécification logarithmique $\log(p^C) = \log(p^M) + \pi$

2. Pour un exposé plus détaillé, voir Gardes *et al* [13] ou Gardes [12].

tiers des postes de consommation entre les élasticités-revenus estimées dans les dimensions transversale et temporelle. Ils trouvent en particulier que, pour chaque pays, les élasticités-revenu en dimension transversale sont affectées d'un biais négatif pour l'alimentation à domicile et d'un biais positif pour l'alimentation à l'extérieur. Ce qui, en liant les prix virtuels à une approche de production domestique à la Becker [3] (intégrant le temps dans le coût d'une consommation), peut s'expliquer par une croissance de la composante temporelle du prix complet de la consommation alimentaire avec le revenu relatif du ménage.

L'estimation des prix virtuels par cette méthode présente un intérêt particulier pour l'étude de la convergence des comportements de consommation, puisqu'elle permet de rendre compte des variations de conditions de choix des ménages, poste par poste, pour chaque période considérée. On peut alors étudier l'évolution des contraintes au cours du temps, et observer dans quelle mesure les conditions de choix s'harmonisent entre classes sociales.

3 Modèle et données

Pour étudier les variations des prix virtuels au Canada, nous partons d'un modèle QUAIDS (Banks, Blundell, Lewbel [1]), extension quadratique de l'Almost Ideal Demand System de Deaton et Muellbauer [11]. Dans le cadre de données de panel, pour un ménage h à l'époque t la part budgétaire w_{iht} allouée à la consommation du bien i s'écrit :

$$w_{iht} = \alpha_{iht} + \sum_{j=1}^n \gamma_{ij} \ln p_{jht} + \beta_i \ln \left(\frac{m_{ht}}{a(\mathbf{p}_{ht})} \right) + \frac{\lambda_i}{b(\mathbf{p}_{ht})} \left(\ln \left(\frac{m_{ht}}{a(\mathbf{p}_{ht})} \right) \right)^2 + \epsilon_{iht} \quad (2)$$

avec

$$\ln a(\mathbf{p}_{ht}) = \alpha_0 + \sum_{i=1}^n \alpha_{iht} \ln p_{iht} + \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n \gamma_{ij} \ln p_{iht} \ln p_{jht}$$

$$b(\mathbf{p}_{ht}) = \prod_{i=1}^n p_{iht}^{\beta_i}$$

où p_{iht} représente le prix du bien i , \mathbf{p}_{ht} est le vecteur des p_{iht} , et m_{ht} est la dépense totale du ménage. Nous suivons Pashardes [19] en prenant pour α_0 une valeur légèrement inférieure à la dépense totale logarithmique minimale observée dans l'échantillon (nous appliquons un coefficient de 0,9). Nous décomposons les α_{iht} en un terme constant par ménage (qui comprend une constante et un effet spécifique de ménage) et un terme variant selon la composition du ménage :

$$\alpha_{iht} = \delta_{ih}^0 + \delta_i^1 \cdot \text{nombre.adultes}_{ht} + \delta_i^2 \cdot \text{proportion.enfants}_{ht} \quad (3)$$

Cette décomposition permet de prendre en compte l'hétérogénéité des structures familiales, y compris dans le terme déflateur de la dépense totale, $\ln a(\mathbf{p}_{ht})$.

La méthode de révélation des prix virtuels présentée ci-dessus nécessite des données structurées en panel, puisqu'elle demande d'estimer des paramètres à la fois dans la dimension transversale (biaisée) et dans la dimension temporelle (non biaisée). Or nous ne disposons pour le Canada que de données d'enquêtes répétées (décrites ci-dessous) qui ne présentent pas a priori cette structure ; le problème est résolu par la construction d'un pseudo-panel. La pseudo-panélistation, initiée par Deaton [10], consiste à donner une structure de panel à des enquêtes successives en coupe transversale, en regroupant les individus de chaque enquête par cohortes (une cohorte est un groupe défini par des critères stables au cours du temps), à calculer pour toutes les variables des moyennes par cohorte, puis à traiter les cohortes obtenues comme des individus que l'on suit d'une enquête sur l'autre. L'intuition de la méthode de Deaton est simple : s'il existe dans la population une relation linéaire au niveau individuel, on retrouvera cette relation au niveau des moyennes par cohorte.

L'estimation d'un modèle linéaire sur pseudo-panel n'est pas sans poser certains problèmes par rapport à un panel véritable, et nécessite d'adapter légèrement la méthode d'estimation. Deaton (*op cit*) considère que les moyennes calculées par cohortes sont entachées d'erreurs de mesure en raison du nombre limité d'observations, et propose un estimateur en conséquence. Cependant Verbeek et Nijman [25] montrent que cet estimateur n'est pas convergent pour un nombre fini de périodes lorsque les effets spécifiques sont corrélés avec les variables explicatives, et proposent un nouvel estimateur convergent. Verbeek et Nijman considèrent un modèle linéaire exprimant un y en fonction d'un vecteur-ligne de régresseurs x ; dans le pseudo-panel on ne dispose que des moyennes construites \bar{x}_{ct} et \bar{y}_{ct} (moyennes pour la cohorte c à l'instant t), qui sont des observations entachées d'erreurs des vraies moyennes par cohorte. En considérant que les erreurs de mesure suivent une distribution normale multivariée centrée, et en notant σ le vecteur des covariances entre les erreurs portant sur y et celles portant sur les x , et Ω la matrice de variance-covariance des erreurs de mesure des x (σ et Ω sont facilement estimés à partir des données individuelles), l'estimateur de Verbeek et Nijman, paramétré par un $\alpha \in [0, 1]$, est défini comme suit :

$$\hat{\beta}(\alpha) = (M_{xx} - \alpha\Omega)^{-1}(m_{xy} - \alpha\sigma) \quad (4)$$

avec

$$M_{xx} = \frac{1}{CT} \sum_{c=1}^C \sum_{t=1}^T \left(\bar{x}_{ct} - (1/T) \sum_{s=1}^T \bar{x}_{cs} \right)' \left(\bar{x}_{ct} - (1/T) \sum_{s=1}^T \bar{x}_{cs} \right)$$

$$m_{xy} = \frac{1}{CT} \sum_{c=1}^C \sum_{t=1}^T \left(\bar{x}_{ct} - (1/T) \sum_{s=1}^T \bar{x}_{cs} \right)' \left(\bar{y}_{ct} - (1/T) \sum_{s=1}^T \bar{y}_{cs} \right)$$

Cet estimateur généralise l'estimateur de Deaton [10] (que l'on obtient pour le cas particulier $\alpha = 1$), et converge quand le nombre de cohortes $C \rightarrow \infty$ pour un nombre fini T de périodes si l'on prend $\alpha = (T - 1)/T$. Dans un article précédent [24], Verbeek et Nijman ont montré que le problème des erreurs de mesure des moyennes de cohortes est négligeable si d'une part les cohortes contiennent chaque année un grand nombre

d'observations (100 à 200) et si d'autre part les vraies moyennes par cohorte évoluent suffisamment au fil du temps. Dans ce cas les estimateurs standards de modèles à effets fixes (within en différences premières) sont valides, ce qui simplifie considérablement l'estimation. Cependant nos données ne remplissent pas la première de ces conditions, puisqu'environ la moitié des couples cohorte-année d'observation contiennent moins de 100 observations. Il nous faut donc employer l'estimateur de Verbeek et Nijman (4).

Les données analysées proviennent des microdonnées des enquêtes de dépenses des ménages de Statistique Canada, compilées et harmonisées par S. Langlois à l'Université Laval (Québec). Pour les années 1969, 1978, 1982, 1986 et 1992 nous utilisons l'Enquête de Dépenses des Familles, et pour années de 1996, 2004 et 2008 son successeur l'Enquête de Dépense des Ménages. L'Enquête de Dépense des Ménages a également été menée chaque année entre 1997 et 2003 et entre 2005 et 2008 mais nous choisissons de ne pas utiliser ces données; d'une part parce que les enquêtes de 1997 à 2003 ne comprennent pas la variable "niveau d'éducation de la personne de référence", variable essentielle à la construction du pseudo-panel, d'autre part pour garder une périodicité relativement stable et ne pas accorder un poids plus important à la décennie 2000 dans les estimations. La construction du pseudo-panel nous a également obligés à retirer de l'échantillon les ménages dont la province de résidence n'est pas renseignée. Afin d'éliminer les données aberrantes nous avons effectué un filtrage des ménages présentant une dépense totale négative ou des postes de dépense négatifs, et ceux dont le centile de dépense nette s'éloigne de plus de 30 du centile de revenu. Ces opérations retirent pour chaque enquête moins de 10% de l'échantillon. Il reste finalement 14 266 ménages en 1969, 8 786 en 1978, 10 010 en 1982, 9 467 en 1986, 8 624 en 1992, 9 219 en 1996, 12 999 en 2004 et 8820 en 2008. Les prix sont tirés des données d'Indices des Prix à la Consommation de Statistique Canada, et les prix des différentes provinces sont harmonisés à l'aide des Indices Comparatifs des Niveaux de Prix à la Consommation de 2008. Nous avons agrégé les données selon les douze postes de la nomenclature fonctionnelle COICOP du Système de Comptabilité Nationale des Nations Unies [23]. Cependant les colinéarités de prix nous ont conduit à regrouper plusieurs postes; nous gardons finalement 8 postes, décrits dans le tableau 1.

Nous construisons un pseudo-panel regroupant les enquêtes de 1969, 1978, 1982, 1986, 1992, 1996, 2004 et 2008. Nous formons des cohortes selon trois critères: la décennie de naissance du chef de ménage³, la province de résidence (5 groupes de provinces: Colombie-Britannique, Québec, Ontario, l'ensemble "provinces des prairies" regroupant l'Alberta, le Manitoba, le Nunavut, le Saskatchewan, les Territoires du Nord-Ouest et le Yukon, et l'ensemble "provinces de l'Atlantique" regroupant l'Île-du-Prince-Edouard, le Nouveau-Brunswick, la Nouvelle-Ecosse et la province de Terre-Neuve-et-Labrador) et le niveau d'éducation du chef de ménages (5 niveaux d'éducation). Après regroupement des cohortes proches à trop faible effectif et élimination de certaines cohortes pour obtenir des pseudo-panels cylindrés (nous ne retenons que les ménages dont la personne de référence

3. Un problème de pseudo-panélisation est posé par l'évolution de la définition du "chef de ménage" ou "personne de référence" au fil des enquêtes. La règle de stabilité de l'appartenance aux cohortes n'est donc pas strictement respectée.

TABLE 1: Nomenclature fonctionnelle utilisée.

notation	Intitulés COICOP des postes
alim	Alimentation à domicile et boissons non alcoolisées Boissons alcoolisées et tabac
habi	Habillement et chaussures
loge	Logement, eau, électricité, gaz et autres combustibles Meubles, équipement ménager et entretien routinier du logement Communication
tran	Transports
lois	Loisirs et culture
educ	Education
rest	Restaurants et hôtels
dive	Biens et services divers Santé

est née entre 1920 et 1949), nous obtenons 55 cohortes (soit 440 points).

La date de naissance de la personne de référence du ménage est un critère classique de construction de pseudo-panels, puisque cette information est presque toujours disponible et que par définition elle ne varie pas avec le temps. Le niveau d'éducation est un critère moins classique mais utilisé notamment par Gardes *et al* [13], et valide si l'on considère que le niveau d'éducation est acquis une fois pour toute en début de vie. Le troisième critère que nous avons retenu, la province de résidence, est beaucoup plus discutable puisqu'il ne s'agit pas d'un critère naturellement invariable dans le temps (un ménage peut déménager d'une province à une autre entre deux périodes). Cependant si l'on construit le pseudo-panel sans intégrer le critère de province, il faut utiliser les prix nationaux, qui varient bien trop peu (seulement avec l'année d'enquête) et présentent des problèmes de multicolinéarité. L'identification des paramètres de prix dans l'équation 2 nous oblige donc à retenir la province de résidence comme troisième critère de construction des cohortes, et donc à considérer les déménagements inter-provinciaux comme négligeables⁴. Notons enfin que le pseudo-panel long que nous avons construit présente l'inconvénient de ne suivre que les ménages dont la personne de référence est née entre 1920 et 1949, de sorte que pour les observations de 2008 les personnes de références retenues ont toutes plus de 60 ans. Nos résultats sont donc probablement affectés par des effets de cycle de vie, et semblent difficilement généralisables à la population canadienne dans son ensemble.

Nous estimons le système QUAIDS sans prendre en compte les poids d'échantillonnage de Statistique Canada. Ce choix est justifié par Davesies et D'Haultfoeuille [9],

4. Ou encore que ces déménagements sont aléatoires et non corrélés avec les autres variables, hypothèse probablement trop forte si l'on considère par exemple le cas des jeunes diplômés attirés par les provinces productrices de pétrole.

puisque l'estimation en pseudo-panels est un cas particulier de régression par variables instrumentales, que les variables de construction des cellules sont celles qui influent sur l'échantillonnage, et que dans ce cas "les estimateurs pondérés et non pondérés [sont] convergents, et l'estimateur non pondéré sera plus précis."

A partir de l'estimation du système QUAIDS (équation 2) dans les deux dimensions transversale et temporelle sur le pseudo-panel, nous sommes en mesure de calculer les variations de prix virtuels d'après l'équation 1. Nous nous intéressons uniquement aux variations de prix virtuels liées aux variations des deux termes de dépense totale, $\ln(m_{ht}/a(\mathbf{p}_{ht}))$ et $(\ln(m_{ht}/a(\mathbf{p}_{ht})))^2$. Nous obtenons :

$$d\pi_i = \frac{\partial \pi_i}{\partial \ln(m/a(\mathbf{p}))} \cdot d \ln(m/a(\mathbf{p})) + \frac{\partial \pi_i}{\partial \ln(m/a(\mathbf{p}))^2} \cdot d \ln(m/a(\mathbf{p}))^2$$

et comme

$$d \ln(m/a(\mathbf{p}))^2 = 2 \ln(m/a(\mathbf{p})) \cdot d \ln(m/a(\mathbf{p}))$$

on a :

$$d\pi_i = \frac{\varphi_i}{\gamma_{ii}} \cdot d \ln(m/a(\mathbf{p})) \quad (5)$$

où

$$\varphi_i = \left((\beta_i^{(CS)} - \beta_i^{(TS)}) + \left(\frac{\lambda_i^{(CS)}}{b(\mathbf{p})^{(CS)}} - \frac{\lambda_i^{(TS)}}{b(\mathbf{p})^{(TS)}} \right) \cdot 2 \ln(m/a(\mathbf{p})) \right) \quad (6)$$

et les indices (CS) et (TS) correspondent respectivement aux estimations en dimensions transversale et temporelle. Dans le cas où les coefficients de prix γ_{ii} ne sont pas correctement estimés, on peut observer les variations de prix virtuels en étudiant uniquement les variations de φ_i , puisque les γ_{ii} sont constants lorsque l'on applique le modèle à l'ensemble des données. Dans ce cas seules des comparaisons poste par poste sont possible, puisque les variations des prix virtuels entre postes dépendent des grandeurs γ_{ii} , qui vraisemblablement varient selon les postes. Notons que puisque φ_i est une fonction linéaire de la dépense totale logarithmique, les prix virtuels logarithmiques sont une fonction quadratique de la dépense totale logarithmique - particularité qui vient de l'expression quadratique de QUAIDS. En intégrant l'expression 5, on obtient :

$$\pi_i = \frac{\Phi_i}{-\gamma_{ii}}$$

avec

$$\Phi_i = - \left(C + (\beta_i^{(CS)} - \beta_i^{(TS)}) \cdot \ln(m/a(\mathbf{p})) + \left(\frac{\lambda_i^{(CS)}}{b(\mathbf{p})^{(CS)}} - \frac{\lambda_i^{(TS)}}{b(\mathbf{p})^{(TS)}} \right) \cdot \ln(m/a(\mathbf{p}))^2 \right) \quad (7)$$

où C est une constante d'intégration. En l'absence de paramètres γ_{ii} , la représentation des Φ_i donne une idée des variations de prix virtuels sur chaque période. Puisque d'une

part C est inconnu, et que d'autre part les prix virtuels sont le quotient de Φ_i et de paramètres γ_{ii} (que l'on traite comme des inconnues s'ils sont mal estimés), *l'analyse des Φ_i ne porte pas sur leur valeur, mais uniquement sur leurs variations*. Pour faciliter la représentation graphique, nous normalisons les Φ_i de sorte que leur minimum soit 0 pour chaque poste. Notons que nous prenons pour Φ_i l'opposé d'une primitive de φ_i , pour qu'ils reflètent mieux les prix virtuels (les paramètres de prix γ_{ii} étant habituellement négatifs).

4 Méthodes et difficultés d'estimation

Nous estimons le modèle QUAIDS explicité en 2 sur le pseudo-panel canadien, dans les deux dimensions transversale et temporelle. Le système de demande présentant des termes non linéaires, on l'estime habituellement à l'aide d'une méthode itérative décrite par Blundell et Robin [5], en recalculant à chaque itération les $\ln a(\mathbf{p})$ et $b(\mathbf{p})$ à partir des paramètres estimés à l'itération précédente, jusqu'à convergence. Nous initialisons les indices de prix $\ln a(\mathbf{p})$ et $b(\mathbf{p})$ en utilisant l'indice de Stone :

$$\ln a(\mathbf{p}) = \sum_i \bar{w}_i \log(p_i)$$

$$b(\mathbf{p}) = \prod_i p_i^{\bar{w}_i}$$

où p_i représente le prix du poste i et \bar{w}_i représente le coefficient budgétaire moyen du poste i dans l'échantillon.

La spécification retenue en 2 est quelque peu restrictive puisqu'elle ne permet qu'une variation limitée des α_{iht} , et aucune hétérogénéité des paramètres β_i et λ_i (voir Blundell *et al* [4] pour une spécification moins restrictive dans le cadre de l'AIDS). Elle nous semble cependant préférable dans le cadre de notre pseudo-panel, puisque le nombre de points d'observation est relativement faible, et que l'inclusion de termes d'hétérogénéité dans chaque équation fait augmenter rapidement le nombre de paramètres à estimer puisque chaque paramètre doit être estimé pour chaque poste de dépense (la présente spécification nécessite l'estimation de 96 paramètres en dimension temporelle).

Lorsque le QUAIDS est estimé sur un pool de données d'enquêtes, les termes de dépense totale logarithmique sont instrumentés pour répondre au problème d'endogénéité (voir [1] et [5]). Nous considérons ici que la pseudo-panélisation élimine ce problème puisqu'elle ne considère que les moyennes par cohortes des coefficients budgétaires et de la dépense totale logarithmique ; les termes de dépense totale logarithmique ne sont donc pas instrumentés. En outre puisque les régresseurs sont les mêmes dans toutes les équations du modèle, l'estimation est menée équation par équation (voir Zellner [26]). Enfin, les différences dans le nombre d'observations par cellule du pseudo-panel amènent un problème d'hétéroscédasticité des erreurs. Nous corrigeons ce problème en multipliant les variables correspondant à chaque cellule par la moyenne de la racine du nombre d'individus de cette cellule au cours du temps (voir Gardes *et al* [13]).

Pour l'estimation en dimension temporelle, nous remplaçons dans la méthode itérative décrite ci-dessus les MCO par l'estimateur de pseudo-panel de Verbeek et Nijman (équation 4). La convergence (définie comme la stabilité des $\ln a(\mathbf{p})$ et des $b(\mathbf{p})$ à 0.001 près entre deux itérations successives) est atteinte après 9 itérations. La méthode présente cependant une limite, puisqu'elle ne permet pas d'estimer convenablement la constante α_i de l'équation 2.⁵ C'est un problème important dans l'estimation de ces systèmes, puisque les termes α_i entrent dans le calcul du déflateur de la dépense totale $\ln a(\mathbf{p})$, qui à leur tour influencent les paramètres de dépense totale. Nous contournons partiellement le problème en décomposant, comme décrit plus haut, les α_i en une part constante par ménage - éliminée par l'estimateur - et une part variable qui, elle, est estimée. Nous utilisons alors cette part variable uniquement dans le calcul du $\ln a(\mathbf{p})$ (comme si l'on avait $\delta_{ih}^0 = 0$ dans l'équation 3). La méthode est imparfaite puisque les déflateurs ainsi calculés n'intègrent pas la totalité de α_i , mais nous n'avons pas trouvé de meilleure solution compatible avec le système QUAIDS. Une solution alternative consisterait à estimer les effets fixes individuels à partir des résidus de l'estimation temporelle, et ainsi récupérer l'intégralité du terme α_i . Cependant cette méthode engendrerait probablement des problèmes d'endogénéité, puisqu'elle consiste à utiliser des moyennes de résidus par cohortes comme régresseurs. Une autre possibilité consisterait à estimer une approximation linéaire du modèle AIDS (en prenant comme déflateur l'indice de Stone) et à appliquer la correction de Pashardes [19] pour les élasticités-revenu. Cette méthode présente en effet l'avantage de ne pas intégrer de paramètres estimés dans le déflateur, et ne pose donc pas de problème d'estimation en dimension temporelle ; cependant elle est sujette aux limites inhérentes au modèle AIDS (courbes d'Engel linéaires, prix virtuels fonctions linéaires de la dépense totale). La meilleure solution consiste probablement à estimer l'approximation linéaire du système EASI de Lewbel et Pendakur [15], qui présente le double avantage d'utiliser un simple déflateur de Stone et d'avoir des courbes d'Engel (et *a priori* des expressions de prix virtuels) extrêmement flexibles [travail en cours].

Pour l'estimation en dimension transversale, nous remplaçons les MCO par l'estimateur between dans la méthode itérative de Blundell et Robin. Là encore se pose le problème du calcul du déflateur $\ln a(\mathbf{p})$ dans la méthode itérative, puisque l'estimateur between, qui revient à effectuer l'ajustement sur les moyennes par individu par ménage de toutes les périodes, rend impossible l'identification des paramètres de prix (qui sont l'élément principal du déflateur $\ln a(\mathbf{p})$). Nous utilisons donc, pour l'estimation transversale, les déflateurs calculés à l'aide des paramètres estimés en dimension temporelle⁶. Ici encore, la solution est malheureusement imparfaite. Le même problème se pose pour toute estimation transversale de modèle AIDS ou QUAIDS lorsque les données de prix ne varient pas ou pas assez (prix nationaux ou par grande zone géographique).

5. Cette limite s'applique de la même façon aux autres méthodes d'estimation en dimension temporelle d'un système AIDS ou QUAIDS (estimateur within ou différences premières pour un panel classique), puisque ces méthodes reposent sur l'élimination de la constante pour l'estimation convergente des autres paramètres.

6. Nous avons essayé une méthode intermédiaire, consistant pour la dimension transversale à estimer par itération les $b(\mathbf{p})$, tout en conservant le $\ln a(\mathbf{p})$ estimé en dimension temporelle. Cependant cette procédure donnait des itérations divergentes sur notre pseudo-panel.

Les méthodes d'estimations retenues donnent lieu à des variances difficiles à calculer - le calcul demanderait de combiner les formules données par Blundell et Robin et celles de Verbeek et Nijman, entre autres. Nous optons donc pour une approximation des variances par bootstrap. Puisque les données ont une structure de panel et que le nombre d'années d'observation est faible, le bootstrap est effectué en tirant avec remise des cellules, et en prenant pour chaque cellule tirée les observations de toutes les années (voir Cameron et Trivedi [7], chapitre 11). Nous effectuons 1000 itérations de bootstrap - un nombre amplement suffisant d'après Cameron et Trivedi - en calculant simultanément les estimateurs temporels et transversaux pour chaque échantillon de bootstrap tiré. Cette procédure présente l'avantage de permettre le calcul direct des covariances des estimateurs temporels et transversaux, ce qui facilite grandement le calcul de statistiques de tests (comme le test d'Hausman). Elle présente cependant deux inconvénients. Le premier est posé par la possibilité de divergence dans les procédures d'estimations présentées ci-dessus, probablement due aux multicollinéarités de prix dans les échantillons de bootstrap. Nous faisons le choix de ne pas retenir les estimations non-convergentes. Ce choix entraîne un rejet de 10% des échantillons de bootstrap, ce qui biaise probablement les variances estimées. Le second problème vient du caractère imbriqué ("nested") des données, puisqu'au sein de chaque groupe année-province, les prix sont identiques pour toutes les observations. Ceci entraîne une sous-évaluation des variances estimées des paramètres de prix. Il est possible de corriger les effets de regroupement en utilisant une procédure bootstrap de panel (ou "block-bootstrap"), mais d'une part nous utilisons déjà cette procédure pour l'aspect panel de nos données, et d'autre part la procédure nécessite un grand nombre de clusters (groupes année-province), alors que nos données en présentent peu (5 provinces et 8 années d'enquête). La portée de ce second problème est cependant limitée, puisque même les variances sous-estimées obtenues par la procédure utilisée donnent des coefficients de prix pour la plupart non significatifs (voir table 4). La conscience du problème de sous-estimation nous porte seulement à croire qu'ils sont en réalité tous non significatifs.

5 Résultats

Les tables 2 et 3 présentent les résultats d'estimation dans les dimensions temporelle et transversale. Les coefficients de dépense totale sont significativement différents de 0 en dimension temporelle (pour chaque poste, au moins un des deux paramètres de dépense totale est significatif au seuil de 5%), sauf pour les postes habillement, transports et éducation. En dimension transversale, tous les postes ont des paramètres de dépense totale significativement différents de 0, à l'exception du poste Education. Les estimations donnent quelques résultats étonnants, notamment le fait que le terme de dépense totale quadratique est significatif pour les dépenses d'alimentation à domicile, alors qu'on considère généralement que ce poste a une courbe d'Engel linéaire (voir Banks *et al* [1]). La table 4 reprend les paramètres de prix directs (γ_{ii}) estimés en dimension temporelle, et les écarts-types correspondants. Les γ_{ii} n'ont le signe attendu que pour la moitié des postes, mais comme indiqué plus haut la sous-évaluation des variances en bootstrap nous

Variable	alim	habi	loge	tran	lois	educ	rest	dive
prop_enfants	0.0338 (0.0331)	0.0255 (0.0180)	-0.1100 (0.0707)	0.0851 (0.0360)	-0.0106 (0.0200)	0.0224 (0.0122)	-0.0317 (0.0187)	-0.2427 (0.0399)
nb_adultes	0.0359 (0.0295)	0.0197 (0.0112)	-0.0214 (0.0516)	0.0225 (0.0207)	-0.0139 (0.0134)	-0.0027 (0.0097)	-0.0170 (0.0131)	-0.0277 (0.0209)
log_px_alim	-0.0608 (0.0485)	-0.0268 (0.0242)	-0.1379 (0.1138)	-0.0462 (0.0495)	0.0791 (0.0274)	0.0026 (0.0199)	0.0503 (0.0271)	0.2002 (0.0520)
log_px_habi	-0.1169 (0.2302)	-0.0258 (0.0721)	1.3883 (0.5857)	0.2117 (0.1714)	-0.3448 (0.1317)	-0.2405 (0.0961)	-0.2703 (0.1367)	-0.1113 (0.1625)
log_px_loge	-0.1340 (0.0730)	-0.0217 (0.0302)	0.2812 (0.1984)	0.0846 (0.0592)	-0.0870 (0.0417)	-0.0468 (0.0326)	-0.0841 (0.0381)	-0.1622 (0.0595)
log_px_tran	0.0799 (0.0395)	0.0475 (0.0151)	-0.3481 (0.0858)	-0.0819 (0.0300)	0.0362 (0.0194)	0.0483 (0.0137)	0.1045 (0.0223)	-0.0516 (0.0321)
log_px_lois	-0.0757 (0.0617)	-0.0437 (0.0252)	-0.2995 (0.1654)	0.0514 (0.0547)	0.0666 (0.0403)	0.0565 (0.0283)	0.0427 (0.0387)	0.0186 (0.0526)
log_px_educ	-0.0736 (0.0448)	-0.0041 (0.0176)	0.1589 (0.1050)	0.0538 (0.0346)	-0.0417 (0.0254)	-0.0240 (0.0191)	-0.0709 (0.0254)	-0.0289 (0.0358)
log_px_rest	0.1961 (0.1247)	0.0121 (0.0387)	-0.5823 (0.3412)	-0.1711 (0.0961)	0.1577 (0.0777)	0.1068 (0.0573)	0.1275 (0.0767)	-0.0324 (0.0865)
log_px_dive	-0.0386 (0.0493)	0.0137 (0.0255)	-0.1517 (0.0937)	0.0112 (0.0507)	-0.0013 (0.0247)	0.0302 (0.0170)	-0.0193 (0.0251)	0.0372 (0.0453)
log_m	0.1905 (0.0402)	-0.0004 (0.0156)	0.1298 (0.0895)	0.0303 (0.0303)	0.0515 (0.0220)	0.0062 (0.0163)	0.0530 (0.0216)	0.1044 (0.0313)
log_m_2	-0.0476 (0.0062)	0.0019 (0.0024)	-0.0251 (0.0127)	-0.0051 (0.0048)	-0.0008 (0.0032)	0.0028 (0.0023)	-0.0025 (0.0033)	-0.0039 (0.0048)

TABLE 2: Paramètres du modèle QUAIDS, estimés sans contraintes d'homogénéité ou de symétrie sur le pseudo-panel en dimension temporelle. Ecart-types entre parenthèses.

Source : microdonnées Statistique Canada, calculs de l'auteur.

Variable	alim	habi	loge	tran	lois	educ	rest	dive
_(Intercept)	0.5694 (0.0906)	0.2192 (0.0509)	-0.0064 (0.1239)	-0.0274 (0.0688)	-0.0776 (0.0330)	-0.0411 (0.0232)	0.0653 (0.0246)	0.2985 (0.0787)
prop_enfants	0.1515 (0.0262)	0.0629 (0.0134)	-0.0673 (0.0408)	0.0116 (0.0209)	-0.0301 (0.0112)	0.0112 (0.0059)	-0.0328 (0.0085)	-0.1069 (0.0308)
nb_adultes	0.0512 (0.0102)	0.0123 (0.0040)	-0.0147 (0.0109)	0.0111 (0.0058)	-0.0235 (0.0040)	-0.0015 (0.0024)	-0.0162 (0.0035)	-0.0187 (0.0096)
log_m	-0.2494 (0.0608)	-0.1523 (0.0398)	0.2612 (0.0942)	0.1045 (0.0514)	0.1214 (0.0258)	0.0200 (0.0178)	-0.0100 (0.0176)	-0.0953 (0.0568)
log_m_2	0.0143 (0.0084)	0.0224 (0.0057)	-0.0359 (0.0135)	-0.0142 (0.0072)	-0.0126 (0.0037)	0.0000 (0.0025)	0.0051 (0.0025)	0.0208 (0.0083)

TABLE 3: Paramètres du modèle QUAIDS, estimés sans contraintes d'homogénéité ou de symétrie sur le pseudo-panel en dimension transversale. Ecart-types entre parenthèses.

Source : microdonnées Statistique Canada, calculs de l'auteur.

porte à croire que les paramètres de prix directs sont en grande majorité non significatifs. Ceci rend plus ardue la tâche de comparaison de prix virtuels, mais nous contournons le problème en comparant les Φ_i (voir équation 7).

	gamma	(é-t)
Alim. a domicile	-0.061	(0.049)
Habillement	-0.026	(0.072)
Logement	0.281	(0.198)
Transports	-0.082	(0.030)
Loisirs	0.067	(0.040)
Education	-0.024	(0.019)
Restaurants	0.128	(0.077)
Divers	0.037	(0.045)

TABLE 4: Paramètres de prix directs (gamma-ii), estimés en dimension temporelle sur le pseudo-panel (écarts-types entre parenthèses).

Source : microdonnées Statistique Canada, calculs de l'auteur.

Les élasticités-revenu moyennes obtenues pour les différentes estimations sont reportées dans la table 5. Pour chaque poste, les élasticités moyennes sont calculées en pondérant les élasticités individuelles par le poids de la dépense de l'individu dans la dépense totale de l'échantillon pour ce poste (comme dans Banks *et al* [1]), selon la formule :

$$\overline{El}_i = \sum_{h=1}^H \sum_{t=1}^T El_{iht} \cdot \frac{p_{iht} \cdot q_{iht}}{\sum_{k=1}^H \sum_{s=1}^T p_{iks} \cdot q_{iks}}$$

Les élasticités ainsi obtenues ont pour la plupart les valeurs attendues : l'alimentation à domicile est un poste de dépense de nécessité, l'habillement, le logement et les transports⁷ ont une élasticité-revenu proches de l'unité, et les loisirs, l'éducation, l'alimentation à l'extérieur et les dépenses diverses sont des dépenses de luxe (élasticité-revenu largement supérieure à 1).

	alim	habi	loge	tran	lois	educ	rest	dive
dimension temporelle	0.17	1.21	0.85	0.96	1.67	3.07	1.81	1.42
dimension transversale	0.15	1.09	1.03	1.03	1.48	2.58	1.59	1.28

TABLE 5: Elasticités-revenus moyennes, calculées en dimensions temporelle et transversale.

Source : microdonnées Statistique Canada, calculs de l'auteur.

7. L'élasticité-revenu unitaire pour les transports est étonnante, puisque les dépenses de transports sont dominées par les transports privés, et que nous avons montré précédemment (voir Boelaert [6]) que ce poste est l'un des principaux postes de différenciation sociale.

Les différences visibles dans ce tableau entre les élasticités estimées dans les deux dimensions sont moins importantes que ne peuvent le laisser penser les statistiques d’Hausman présentées ci-dessous ; cela s’explique premièrement par le fait que le tableau ne présente que des élasticités moyennes, et deuxièmement par le fait que les élasticités calculées pour un système QUAIDS ne dépendent uniquement des paramètres de dépense totale, mais aussi de la plupart des autres paramètres estimés, par le biais des termes $a(\mathbf{p})$ et $b(\mathbf{p})$ qui entrent dans la formule. Il est intéressant de noter que nous retrouvons dans ces élasticités moyennes certaines des régularités observées par Gardes *et al* ([13] et [8]), notamment la sous-estimation des élasticités-revenu pour l’alimentation à domicile en dimension transversale (mais pas la sur-estimation pour l’alimentation à l’extérieur).

La première question à laquelle nous voulons répondre est celle de savoir si les estimations en dimensions transversale et temporelle sont suffisamment différentes pour que l’on puisse détecter des variations de prix virtuels. Nous effectuons à cette fin un test d’Hausman. La table 6 donne les différences d’estimation entre dimension transversale et temporelle, et les écarts-types correspondants. La table 7 donne pour chaque poste les statistiques d’Hausman pour les termes de dépense totale. Les statistiques sont calculées à partir des différences entre les estimations en dimensions transversale et temporelle des coefficients concernés, et la matrice de variance-covariance de ces différences, estimée directement par bootstrap (voir Cameron et Trivedi [7], chapitre 8). Le test d’Hausman indique que les différences entre estimateurs sont significatives au seuil de 5% pour les postes Alimentation, Habillement, Logement, Education et Dépenses Diverses, au seuil de 10% pour le poste Loisirs, et ne sont pas significatives pour les postes Transports et Restaurants & Hôtels. Là où les différences sont significatives, l’estimation transversale des paramètres de dépense totale est biaisée, et il existe donc des variables latentes corrélées à la dépense totale qui influencent les décisions de dépense. Les tests portant sur l’ensemble des termes de dépense totale, tous postes confondus, rejettent fortement l’égalité des paramètres estimés (p-value 0). Certaines valeurs de la statistique d’Hausman dans la table 7 peuvent sembler étonnantes au regard des différences reportées dans la table 6 (notamment pour le poste Logement, dont les différences individuelles de paramètres sont non-significatives, et pour le poste Restaurants et Hôtels, dont la différence de β est significative). Ceci est dû au fait que la statistique d’Hausman intègre les termes d’interaction entre paramètres (puisque aucun des estimateurs n’est efficace), termes qui n’apparaissent pas dans ces tableaux.

Le but de notre recherche est de déterminer si les variations de prix virtuels selon la dépense totale - c’est-à-dire les contraintes non-monétaires corrélées au revenu - se sont modifiées au cours du temps. Nous étudions à cet effet les figures 1 à 6, qui présentent les Φ_i calculés pour chaque poste. Ces valeurs sont des approximations des prix virtuels (équation 7), dont seules les variations (et non les valeurs) nous intéressent, comme indiqué plus haut. On remarque sur chacune des figures que les Φ_i sont regroupés en faisceaux paraboliques ou linéaires ; chacun de ces faisceaux correspond à un système de prix différent, c’est-à-dire à un couple année-province (nous représentons sur chaque figure les différentes années par des symboles différents). Cette particularité est due aux valeurs prises par les déflateurs $a(\mathbf{p})$ et $b(\mathbf{p})$, qui entrent de façon non linéaire dans la formule 7 des Φ_i . Les \mathbf{p} varient en sauts discrets entre provinces et entre années, d’où les

	beta	(é-t)	lambda	(é-t)
Alim. a domicile	-0.4399	(0.0755)	0.0619	(0.0110)
Habillement	-0.1519	(0.0429)	0.0205	(0.0062)
Logement	0.1315	(0.1160)	-0.0108	(0.0168)
Transports	0.0742	(0.0603)	-0.0090	(0.0086)
Loisirs	0.0698	(0.0277)	-0.0118	(0.0041)
Education	0.0138	(0.0178)	-0.0028	(0.0026)
Restaurants	-0.0631	(0.0336)	0.0076	(0.0049)
Divers	-0.1997	(0.0622)	0.0247	(0.0094)

TABLE 6: Différences (et écarts-types) entre estimations transversale et temporelle des paramètres de dépense totale (beta) et de dépense totale quadratique (lambda). Source : microdonnées Statistique Canada, calculs de l'auteur.

	alim	habi	loge	tran	lois	educ	rest	dive
H	17.58	56.04	32.14	0.03	5.57	57.25	3.56	69.09
p-value	0.00	0.00	0.00	0.99	0.06	0.00	0.17	0.00

TABLE 7: Statistique d'Hausman et p-value correspondante pour les termes de revenu. Source : microdonnées Statistique Canada, calculs de l'auteur.

sauts entre faisceaux.

L'analyse de la représentation des Φ_i permet trois types d'observations : premièrement la magnitude des différences de prix virtuels d'une année sur l'autre (les contraintes non monétaires - toutes cohortes confondues - se modifient-elles de plus en plus vite au fil des ans?), deuxièmement l'évolution de la dispersion des prix virtuels selon la dépense totale (les contraintes sont-elles de plus en plus inégales entre pauvres et riches?), et troisièmement le sens de variation des prix virtuels selon la dépense totale, année par année (telle année, pour quelle classe de revenus la charge des contraintes non monétaires est-elle la plus forte, pour un poste de dépenses donné?).

On remarque tout d'abord que, pour chaque poste, les Φ_i - considérés toutes cohortes confondues - évoluent de façon monotone au cours du temps, et de moins en moins vite. Ceci semble indiquer *l'existence, pour chaque poste, d'une tendance historique continue à la modification des contraintes non monétaires liées aux revenus, tendance qui s'affaiblit au fil du temps.*⁸ Ce résultat est à rapprocher des constats établis dans Boelaert [6], tant sur données internationales qu'intranationales : les structures de consommation des ménages se transforment au cours temps selon une tendance continue mais de moins en moins forte, les structures budgétaires tendant à se stabiliser.

8. Il est important de noter que le sens de cette tendance historique (faisceaux de Φ_i de plus en plus hauts ou de plus en plus bas sur le graphique) ne peut pas être interprété; en effet, les Φ_i ne renseignent pas sur le signe des prix virtuels correspondant, et subséquemment une augmentation des Φ_i peut tout aussi bien correspondre à une augmentation qu'à une diminution de la valeur absolue des prix virtuels.

L'analyse poste par poste des Φ_i donne des résultats plus détaillés :

- Pour l'alimentation à domicile, les prix virtuels sont une fonction décroissante de la dépense totale en début de période, une fonction quasi-constante au cours des années 1980, et une fonction croissante de la dépense totale à partir de 1990. Cela signifie que les contraintes non monétaires affectant ce poste étaient différenciées en début et en fin de période, mais dans un sens différent. *Il semble que dans les années 1970 il coûtait plus cher à un ménage pauvre qu'à un ménage riche de se restaurer à domicile, que cette différence s'est estompée au cours des années 1980, et qu'à partir de 1990 on retrouve le résultat classique selon lequel l'alimentation à domicile est plus coûteuse pour un ménage riche que pour un ménage pauvre. Les contraintes non monétaires affectant les dépenses d'habillement et les dépenses diverses ont connu la même évolution.*
- Sur la figure 3 on observe que *sur toute la période, les contraintes non monétaires relatives au logement affectent plus sévèrement les ménages pauvres que les ménages riches, et que cette inégalité s'est accentuée au fil des ans.*
- Au contraire, *les dépenses de loisirs et d'éducation coûtent moins cher aux pauvres qu'aux riches en début de période, mais cette inégalité s'estompe au fil du temps jusqu'à quasiment disparaître en fin de période.*

Nous ne reproduisons pas les graphiques de Φ_i pour les postes Transports et Restaurants & Hôtels, puisque les tests d'Hausman indiquent que pour ces postes les différences de paramètres de revenus ne sont pas significatives, et que par conséquent ces postes ne sont pas affectés par des variations de prix virtuels selon le revenu. Le résultat de non-significativité pour le poste Restaurants & Hôtels est néanmoins étonnant puisque les travaux de F. Gardes cités précédemment établissent sur la plupart des données étudiées que l'alimentation à l'extérieur coûte moins cher aux ménages riches qu'aux ménages pauvres.

On remarque enfin qu'à certaines exceptions près (dépenses de logement sur toute la période, et décennie 2000 pour les dépenses d'alimentation à domicile, d'habillement et diverses), les courbes de Φ_i par année ont tendance à s'aplanir. Ceci signifie que, de manière générale, les contraintes non monétaires qui pèsent sur les choix des ménages sont moins différenciés selon la position dans l'échelle des revenus. De sorte qu'*au fil du temps, les ménages canadiens sont de moins en moins contraints par leur catégorie sociale dans leurs choix de consommation.* Ce résultat rejoint les conclusions des travaux précédents cités en introduction, et il contribue à expliquer à la fois pourquoi les structures de consommation moyennes par classe se rapprochent entre elles, et pourquoi au sein de chaque classe les comportements de consommation éclatent. En outre, on peut interpréter les nouvelles différences de prix virtuels observées pour l'alimentation à domicile, l'habillement et les dépenses diverses au cours des années 2000 comme une cause de la fin de la convergence entre classes.

FIGURE 1: Représentation des Φ_i (approximation des prix virtuels) pour les dépenses d'alimentation à domicile.
 Source : Microdonnées Statistique Canada, calculs des auteurs.

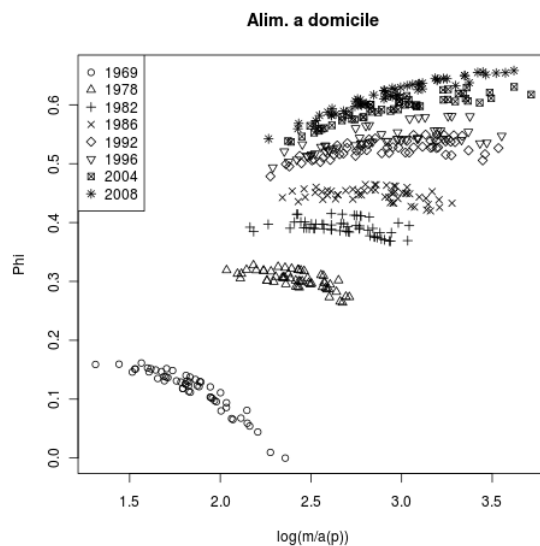


FIGURE 2: Représentation des Φ_i (approximation des prix virtuels) pour les dépenses d'habillement.
 Source : Microdonnées Statistique Canada, calculs des auteurs.

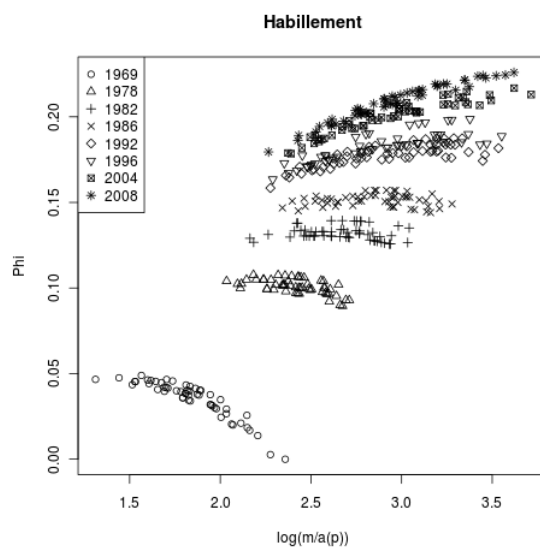


FIGURE 3: Représentation des Φ_i (approximation des prix virtuels) pour les dépenses de logement.
 Source : Microdonnées Statistique Canada, calculs des auteurs.

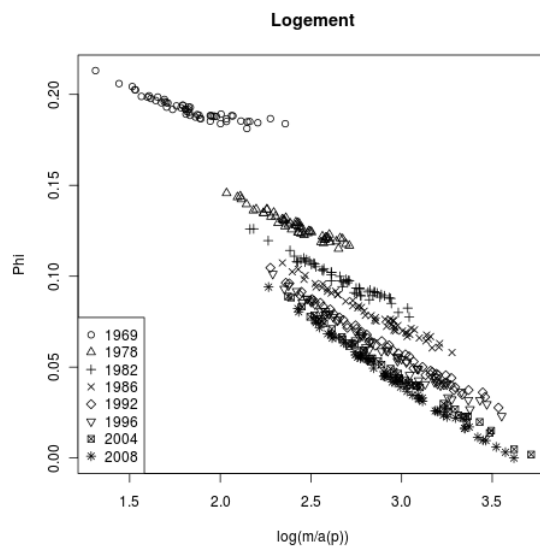


FIGURE 4: Représentation des Φ_i (approximation des prix virtuels) pour les dépenses de loisirs.
 Source : Microdonnées Statistique Canada, calculs des auteurs.

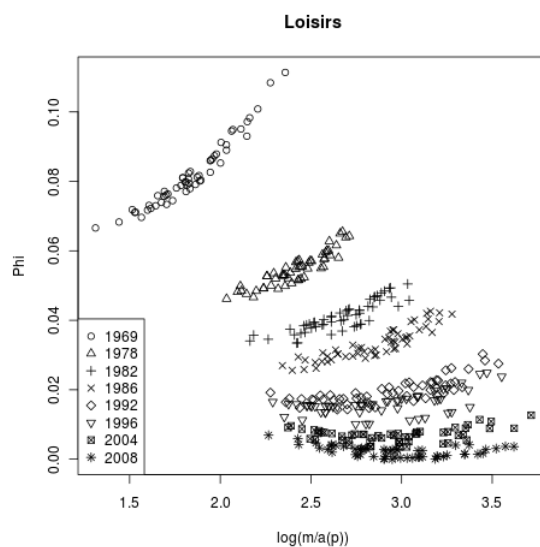


FIGURE 5: Représentation des Φ_i (approximation des prix virtuels) pour les dépenses d'éducation.

Source : Microdonnées Statistique Canada, calculs des auteurs.

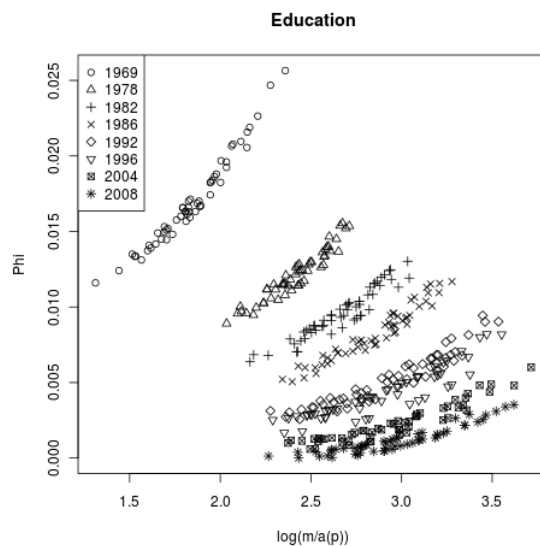
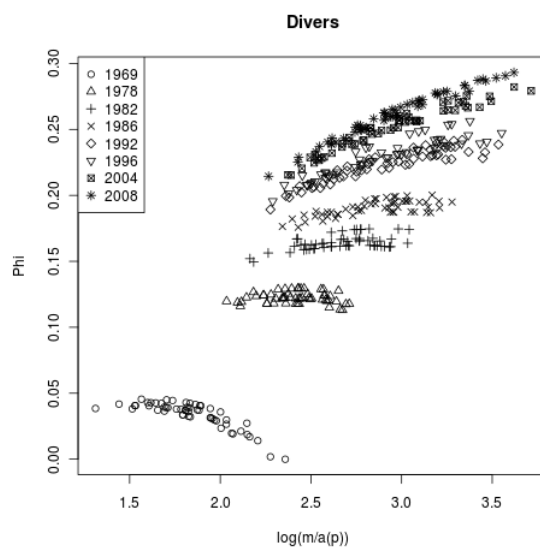


FIGURE 6: Représentation des Φ_i (approximation des prix virtuels) pour les dépenses diverses.

Source : Microdonnées Statistique Canada, calculs des auteurs.



Références

- [1] J. BANKS, R. BLUNDELL et A. LEWBEL : Quadratic engel curves and consumer demand. *Review of Economics and Statistics*, 79(4):527–539, 1997.
- [2] A. BARTEN : Family composition, prices and expenditure patterns. In P. HART, G. MILL et J. WHITTAKER, éditeurs : *Economic Analysis for National Income Planning*. Butterworth, 1964.
- [3] G. BECKER : A theory of the allocation of time. *The Economic Journal*, 75(299):493–517, 1965.
- [4] R. BLUNDELL, P. PASHARDES et G. WEBER : What do we learn about consumer demand patterns from micro data ? *The American Economic Review*, 83(3):570–597, 1993.
- [5] R. BLUNDELL et J.-M. ROBIN : Estimation in large and disaggregated demand systems : An estimator for conditionally linear systems. *Journal of applied econometrics*, 14(3):209–232, 1999.
- [6] J. BOELAERT : *La convergence internationale et intranationale des structures de consommation des ménages*. Thèse de doctorat, Université Paris 1 Panthéon-Sorbonne, Centre d'économie de la Sorbonne, 2012.
- [7] A. C. CAMERON et P. K. TRIVEDI : *Microeconometrics, methods and applications*. Cambridge University Press, 2005.
- [8] N. CARDOSO et F. GARDES : Estimation de lois de consommation sur un pseudo-panel d'enquêtes de l'insee (1979, 1984, 1989). *Economie et Prévision*, 5:111–125, 1996.
- [9] L. DAVESIES et X. D'HAULTFOEUILLE : Faut-il pondérer ?... ou l'éternelle question de l'économètre confronté à des données d'enquête. INSEE, *Série des documents de travail de la Direction des Études et Synthèses Économiques*, G 2009/06, 2009.
- [10] A. DEATON : Panel data from time series of cross-sections. *Journal of Econometrics*, 30:109–126, 1985.
- [11] A. DEATON et J. MUELLBAUER : An almost ideal demand system. *The American Economic Review*, 70(3):312–326, 1980.
- [12] F. GARDES : Différenciation sociale et dynamique des prix dans une économie walrassienne. In R. BARANZINI, A. LEGRIS et L. RAGNI, éditeurs : *Léon Walras et l'équilibre économique général, Recherches récentes*, pages 291–307. Economica, 2011.
- [13] F. GARDES, G. J. DUNCAN, P. GAUBERT, M. GURGAND et C. STARZEC : Panel and pseudo-panel estimation of cross-sectional and time series elasticities of food consumption : The case of u.s. and polish data. *Journal of Business & Economic Statistics*, 23(2):242–253, 2005.
- [14] H. S. HOUTHAKKER : New evidence on demand elasticities. *Econometrica*, 33(2): 277–288, 1965.

- [15] A. LEWBEL et K. PENDAKUR : Tricks with Hicks : the EASI demand system. *The American Economic Review*, 99(3):827–863, 2009.
- [16] C. LLUCH, A. A. POWELL et R. A. WILLIAMS : *Patterns in Household Demand and Saving*. Oxford University Press, 1977.
- [17] Y. MUNDLAK : On the pooling of time series and cross-section data. *Econometrica*, 46(1):483–509, 1978.
- [18] J. P. NEARY et K. W. S. ROBERTS : The theory of household behaviour under rationing. *European Economic Review*, 13(1):25–42, 1980.
- [19] P. PASHARDES : Bias in estimating the almost ideal demand system with the stone index approximation. *The Economic Journal*, 103(419):908–915, 1993.
- [20] S. SELVANATHAN : *A System-Wide Analysis of International Consumption Patterns*. Kluwer Academic Publishers, 1993.
- [21] H. THEIL et K. W. CLEMENTS : *Applied Demand Analysis : Results from System-Wide Approaches*. Ballinger, Cambridge, 1987.
- [22] H. THEIL, F. E. SUHM et J. F. MEISNER : *International Consumption Comparisons : a System-Wide Approach*. Elsevier, 1981.
- [23] UNITED NATIONS STATISTICS DIVISION : *System of National Accounts 1993*. United Nations, 1993.
- [24] M. VERBEEK et Nijman T. : Can cohort data be treated as genuine panel data? *Empirical Economics*, 17:9–23, 1992.
- [25] M. VERBEEK et Nijman T. : Minimum MSE estimation of a regression model with fixed effects from a series of cross-sections. *Journal of Econometrics*, 59:125–136, 1993.
- [26] A. ZELLNER : An efficient method of estimating seemingly unrelated regression equations and tests for aggregation bias. *Journal of the American Statistical Association*, 57:348–368, 1962.