

Estimation d'un modèle hédoniste de valorisation des caractéristiques des produits alimentaires ¹

Abdelwaheb Chaieb

AgroParisTech et Université de Cergy Pontoise
16, Rue Claude Bernard, 75005, Paris
Equipe ALISS (UR 1303, INRA)
Correspondance auteur : *abdelwaheb-chaieb@hotmail.fr*

Résumé:

Dans cet article, nous estimons la demande du consommateur pour les caractéristiques des produits en utilisant les informations fournies sur l'étiquetage nutritionnel. L'étiquetage sur la composition nutritionnelle des produits est devenu un instrument important pour fournir aux consommateurs l'information au moment de l'achat. Nous considérons le marché de céréales pour le petit déjeuner en France en utilisant une base de données originale et récente qui intègre le panel TNS-Sofres 2006 et des informations nutritionnelles relevées sur 822 boîtes de céréales. Nous utilisons la méthode des prix hédonistes (Rosen, 1974) en proposant d'estimer la demande du consommateur pour les caractéristiques des produits en deux étapes. La première consiste à estimer l'effet de la composition nutritionnelle des produits sur les prix et calculer des prix implicites de caractéristiques. La deuxième permet de calculer des élasticités de la demande et recouvrir les préférences des consommateurs pour les caractéristiques. Nos résultats montrent que la disposition à payer pour les caractéristiques est différente selon le contenu nutritionnel des produits. La demande pour les caractéristiques des produits est faiblement élastique et les consommateurs répondent bien aux incitations du marché pour la diversité de l'offre alimentaire sur des bases nutritionnelles. Les résultats suggèrent qu'une action sur les caractéristiques des produits visant l'amélioration de la qualité nutritionnelle serait un levier efficace de politiques de santé publique.

Code JEL : D12, Q13, I18

¹ L'auteur remercie vivement François Gardes (Université de Paris I) et Jean-François Huneau (AgroParisTech) pour leurs commentaires et suggestions dans l'élaboration de ce projet d'article, ainsi que Christine Boizot (INRA-ALISS) et Raffaella Goglia (OQUALI-INRA) pour leur collaboration dans la collecte des données. Ce travail a été réalisé dans le cadre de mes travaux de thèse financée par AgroParisTech (bourse MESRT, Rectorat de Paris).

Introduction:

Dans le contexte actuel d'une augmentation des préoccupations de santé, on peut imaginer que pour certains produits au moins le consommateur puisse accorder une grande importance aux attributs de la composition nutritionnelle renseignés sur l'étiquetage des produits alimentaires. La demande serait alors une demande de nutriments et non de produits, dont l'élasticité constituerait le véritable outil de définition des politiques alimentaires et nutritionnelles. A ce jour, des élasticités nutriments ont été généralement calculées de deux façons dans la littérature, soit via l'estimation de formes réduites ad hoc ((Fousekis et Lazaridis (2005), Abdulai et al. (2004)), soit en transformant les élasticités classiques de produits en élasticités nutriments à l'aide d'une matrice de conversion produits/nutriments (Allais et al., (2010). En effet, il n'existe pas d'approche structurelle à ce jour prenant en compte la modification des caractéristiques des produits par une modélisation hédonique de la demande de biens.

Une approche alternative consiste à appliquer aux nutriments le modèle hédoniste de Rosen (1974), en modélisant non plus le choix d'une variété ou d'une quantité de produit mais d'une quantité de nutriments. La méthode consiste à régresser les prix sur les caractéristiques des produits dans une première étape, ce qui permet d'obtenir une estimation de la fonction de prix hédonistes et des prix implicites de chaque caractéristique. Cette étape a fait l'objet de quelques travaux qui ont pu estimer des prix implicites de caractéristiques dans le cas de l'alimentation (Donnet et al., 2008 ; Huang et Lin, 2007 ; Kristofferson et Rickertsen (2004, 2007) ; Edmeades (2007) ; Ward et al., 2008 ;). Cependant, les prix marginaux de caractéristiques sont normalement le résultat de l'offre et la demande de caractéristiques. Il conviendrait donc de rajouter à cette étape une dimension offre/demande qui permet d'obtenir une fonction de demande pour chaque caractéristique et recouvrer les préférences des consommateurs. De ce fait, il n'existe pas de consensus dans la communauté scientifique sur le choix du modèle complet en deux étapes pour traiter les modèles hédonistes.

Dans cet article, nous nous intéressons à l'analyse du marché des céréales pour le petit déjeuner en France par la méthode des prix hédonistes (Rosen, 1974). Il s'agit d'un secteur dans lequel les enjeux nutritionnels tiennent une place importante, marquant les stratégies des industriels en matière de segmentation du marché et de communication aux consommateurs. Ces produits peuvent ressembler à une combinaison de caractéristiques nutritionnelles et leurs prix devraient être expliqués, en partie, par la quantité de chaque nutriment contenue dans les produits et par l'utilité qu'elles procurent. Dans ce cadre, les différences de prix peuvent s'expliquer par le nombre et la quantité de caractéristiques présentes dans le produit. Il en découle que l'on peut construire un marché implicite pour chaque caractéristique et déterminer un prix implicite.

1. Modèle hédoniste

Cadre théorique

Soit un bien décrit par le vecteur de ses caractéristiques $z = (z_1, z_2, \dots, z_n)$. Le prix de vente du bien est une fonction hédoniste $p(z)$ telle que $p(z) = p(z_1, z_2, \dots, z_n)$. L'utilité des agents qui découle de la consommation des caractéristiques s'écrit : $U(x, z_1, z_2, \dots, z_n)$ où x est un bien composite de prix unitaire qui regroupe tous les autres biens consommés. L'acheteur cherche à maximiser son utilité sous la contrainte de budget :

$y = x + p(z)$, où y représente son revenu disponible, c'est-à-dire $\frac{\partial p}{\partial z_i} = \frac{\partial U}{\partial z_i} / \frac{\partial U}{\partial x}$ avec

$i = 1, \dots, n$.

Dans le cadre de l'analyse théorique de Rosen (1974), le prix d'équilibre, *i.e.* la fonction de prix hédonistes, $p(z)$, s'obtient en égalisant, pour chaque panier de caractéristiques z , la demande à l'offre, ou plus exactement la distribution des caractéristiques demandées à celle des caractéristiques offertes. Le prix implicite d'une caractéristique j , $\frac{\partial p}{\partial z_j}$, s'interprète comme la valeur que les consommateurs attribuent à une unité supplémentaire de cette caractéristique. Ces prix implicites sont utilisés pour estimer la demande et l'offre de chaque caractéristiques j en deux étapes :

$$\hat{p}_j(z) = f_j(z_1, \dots, z_n, Y_1, e_{j1}) \quad (\text{demande})$$

$$\hat{p}_j(z) = g_j(z_1, \dots, z_n, Y_2, e_{j2}) \quad (\text{offre})$$

Avec $j=1,\dots,k$, ou Y_1 le vecteur de revenu et de caractéristiques des consommateurs et Y_2 le facteur prix et de caractéristiques des producteurs ; e_{j1} et e_{j2} des vecteurs de termes d'erreurs. Ce système d'équation est résolu par une méthode d'estimation simultanée, *i.e.* doubles moindres carrés ordinaires, en utilisant Y_1 et Y_2 comme instruments. Une première condition pour que cette procédure en deux étapes aboutisse dans le cadre d'un seul marché stipule que les $p(z)$ soient non linéaire en première étape. Lorsque les $p(z)$ sont linéaire en première étape, les prix implicites sont constants et ont des variances nulles sur l'ensemble d'observations. Cependant, dans ce cas, il est possible d'estimer des prix marginaux au niveau individuel qui représentent les consentements à payer pour la caractéristique.

Cette approche repose sur de nombreuses hypothèses (Rosen, 1974) : chaque caractéristique doit pouvoir être mesuré de façon objective. La mesure est de nature quantitative même si elle se réduit aux valeurs 0 et 1, selon que le bien possède ou non la propriété ; tous les agents qui participent au marché connaissent les quantités ou la présence de chaque caractéristique dans les différents produits ; un bien est entièrement caractérisé par les attributs qui le définissent, ce qui oblige à considérer un large nombre de propriétés afin d'intégrer tous les éléments qui déterminent la décision de consommation. Il faut rappeler aussi que la méthode des prix hédonistes ne peut être appliquée que dans le cadre d'un marché parfaitement concurrentiel (Rosen, (1974)).

Si une approche hédonique semble donc justifiée, il est nécessaire qu'elle tienne compte du fonctionnement du marché. Même si le marché des céréales petit déjeuner en France est constitué d'une dizaine de marques nationales concurrencées par les marques de distributeurs, il n'est pas sûr que la concurrence soit pure et parfaite. Il faut donc être prudent s'agissant de l'interprétation en termes de dispositions à payer. Mais les résultats peuvent toujours s'interpréter en termes de corrélation. Cependant, le marché des céréales petit déjeuner n'est ni plus ni moins concurrentiel que les autres marchés auxquels la méthode des prix hédoniste a été appliquée.

Application d'un modèle de demande de caractéristiques

Nous supposons dans le cadre de ce travail que l'offre de céréales pour le petit déjeuner est exogène et élastique. Cela implique que le prix des caractéristiques est déterminé uniquement par la demande, et que l'estimation de la deuxième étape du modèle hédoniste se réduit à un système de demande de caractéristiques. En pratique, la modification de l'offre (prix et caractéristiques) peut être soumise à certaines rigidités. Cette hypothèse est réductrice pour deux raisons : la première est que la variation des caractéristiques peut être différente au niveau des coûts de production. Par exemple, une variation de 10% de la teneur en sucre peut avoir des coûts moins importants qu'une variation de 10% de la teneur en sel ; la deuxième est qu'il peut exister des rigidités liées aux contraintes technologiques qui ne sont pas toujours considérées comme totalement indépendantes des variations de caractéristiques.

En faisant cette hypothèse, il convient d'ajouter à cette dimension « demande » d'éventuelles composantes de « coûts » : la production et la vente d'un produit regroupant différents contenus est par exemple de nature à réaliser des économies d'échelle dans la production et la vente. Toutefois, cela nécessite des calculs trop compliqués et nous n'avons pas assez d'éléments pour le faire dans le cadre de ce travail. Il faut donc être prudent dans l'interprétation des élasticités prix de nutriments car les effets que nous allons identifier ne reflètent pas les effets réels de modification de prix ou de caractéristiques.

Nous retenons le système de demande² AIDS développé par Deaton-Muellbauer (1980). Les systèmes de demande font dépendre, théoriquement, la demande d'un bien *i.e* caractéristiques en fonction de son prix, du prix des autres biens constituant le panier du ménage et du revenu. Cependant, il n'est pas possible en pratique de prendre en considération les effets prix de tous les biens. Deux solutions se présentent dans la théorie : soit on regroupe les biens en constituant des groupes homogènes, soit on pose l'hypothèse de séparabilité des préférences et l'on considère que les dépenses se répartissent en groupes qui peuvent être étudiés indépendamment les uns des autres (Deaton et Muellbauer, 1980a). Nous posons donc l'hypothèse de séparabilité de préférence et supposons que la quantité de caractéristiques contenues dans les céréales est faiblement séparable de celles contenues dans les autres produits alimentaires.

Par ailleurs, les caractéristiques que nous utilisons sont objectives et issues de l'observation d'indicateurs nutritionnels. Ces caractéristiques sont observables et les agents peuvent les identifier. Le prix des différents biens (les différentes combinaisons de caractéristiques) est déterminé sur un marché concurrentiel. L'hypothèse de marché concurrentiel signifie que les agents sont preneurs de prix et leur décision d'achat n'influence pas le prix des produits. Nous considérons ici un ensemble de nutriments contenus dans les céréales petit déjeuner. Cinq nutriments ont été retenus pour l'analyse : sucres, protéines, fibres, acides gras saturés et sodium. Dans la première étape, la fonction de prix hédoniques est estimée par l'équation :

$$\ln(\text{prix}_{ir}) = x_j \beta + z_j \delta + \varepsilon_{jr}$$

Où prix_{ir} est le prix unitaire du produit j dans l'enseigner, x_j et z_j l'ensemble de caractéristiques (variables continues et binaires), β et δ sont des vecteurs de paramètres à estimer représentant les prix implicites des caractéristiques. Dans la deuxième étape, ces prix³ implicites P_j de caractéristiques sont introduits comme variables indépendantes dans les fonctions de demandes AIDS données par :

$$w_i = \alpha_i + \sum_{j=1}^n \gamma_{ij} \log p_j + \chi_i (\ln X - \ln a(p))$$

Où w_i sont les parts de dépenses des cinq nutriments, X la dépense totale en nutriment et $\ln a(p)$ un indice de prix:

$$\log a(p) = \alpha_0 + \sum_{i=1}^n \alpha_i \log(p_i) + \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n \gamma_{ij} \log(p_i) \log(p_j)$$

α_i , χ_i et γ_{ij} sont les paramètres à estimer.

³ On peut constater qu'ils dépendent de la quantité de caractéristiques. Or, ces prix sont ici régressés sur ces mêmes caractéristiques. Il y a donc endogénéité par construction, du fait que le terme d'erreur de la fonction de demande d'une caractéristique est potentiellement corrélé avec son prix. Nous verrons dans la section suivante la méthode d'estimation approprié pour remédier à ce problème.

2. Données et Méthode d'estimation :

2.1 : Données

Nous utilisons les données TNS 2006 ainsi que des informations nutritionnelles collectées sur les boîtes de céréales. Le panel TNS Worldpanel utilise un échantillon de ménages de consommateurs stratifié en 2 strates : par région et par taille de commune. Il est complété par la méthode des quotas pour s'assurer d'une représentativité sociodémographique. Cette base de données contient de l'information sur les caractéristiques sociodémographiques des ménages vivant en France (à l'exception de la Corse) et sur leur comportement d'achats de plusieurs biens de consommation. Les ménages enregistrent leurs achats alimentaires plusieurs fois par semaine tout au long de l'année. Ils indiquent la quantité et la dépense correspondante pour chaque article acheté. Les produits sont enregistrés à un niveau très fin. Afin que l'effort demandé aux ménages demeure acceptable, les ménages sont répartis en 2 sous-panels, également construits sur une base représentative. Chaque sous-panel enregistre les achats sur un ensemble restreint de produits frais (les fruits et les légumes frais dans l'un ; la viande, le poisson frais et le vin dans l'autre). Nous signalons que les céréales et les biscuits sont enregistrées dans le panel général.

Nous disposons de 822 références de céréales. L'information nutritionnelle associée à ces produits est diverse. Trois types d'étiquetages sont associés à ces produits. Le premier est un format simple qui affiche respectivement les valeurs moyennes pour 100g de calories, protéines, lipides et glucides. Le deuxième est un format plus complet affichant en outre les fibres, sucres, sodium, acide gras saturés. Le troisième indique les vitamines et minéraux en plus des autres indicateurs nutritionnels affichés dans les deux derniers formats d'étiquetage. Les lipides sont composés d'acides gras saturés, mono-insaturés et polyinsaturés. Les glucides se décomposent en sucres totaux et amidon. Les vitamines regroupent les B1, B2, B3, B5, B6, B8, B9, B12, C et E. Enfin, Les minéraux sont donnés par le calcium, fer, magnésium, zinc, phosphore et sodium.

Variables à expliquer

-Prix : dans la première étape, on régresse les prix sur les caractéristiques des céréales. Pour le calcul des prix, nous utilisons comme variables approchées, pour chaque produit, les dépenses annuelles divisées par les quantités correspondantes. Pour que ces valeurs soient les plus représentatives possibles des vrais prix, nous en calculons les moyennes par produit et type d'enseigne. Les enseignes d'achat sont nombreuses ici. Pour pouvoir tenir compte de cette diversité de l'offre, nous avons calculé la fréquence d'achat dans chaque type d'enseigne. Nous avons sélectionné 11 enseignes d'achat les plus fréquentées. Ces magasins sont : Auchan, Carrefour, Champion, Cora, Ed, Intermarché, Leaderprice, Leclerc, Lidl, SuperU et Géant. Les autres enseignes, moins fréquentées, sont regroupées dans un groupe que l'on a nommé « autre enseigne ». Une fois les prix calculés, nous ne conservons qu'une seule observation par produit et type d'enseigne. Au total, 822 observations ont été retenues.

-Caractéristiques: dans la deuxième étape, on estime la demande pour les Cinq nutriments en fonction des prix estimés en première étapes et des variables sociodémographiques des ménages. La quantité consommée de caractéristiques est donnée par :

$$Q_j = \sum_i x_{ij} q_i$$

Où x_{ij} est la quantité de caractéristique j contenue dans le produit i , et q_i la quantité achetée du produit i (en kg).

Variables explicatives

Les informations nutritionnelles sont renseignées sur différents types d'étiquetages. Nous avons recensé quatre types d'étiquetage différents relatifs aux céréales petit déjeuner. Le premier contient des informations sur la teneur en calories, protéines, glucides, lipides. Le deuxième type, contenant les mêmes informations que le premier, est complété par la teneur en vitamines. Le troisième vient compléter le premier type d'étiquetage en décomposant les glucides en sucres totaux et amidon, et les lipides en acides gras saturés, mono insaturés, polyinsaturés. S'ajoutent à cela, les fibres et les minéraux qui sont

décomposés en calcium, magnésium, fer, zinc, sodium et phosphore. Enfin, le dernier étiquetage contient les informations nutritionnelles du troisième type ainsi que les vitamines. Face à l'hétérogénéité des affichages d'un côté, et la corrélation entre les différentes informations de l'autre côté, les caractéristiques les plus intéressantes retenues sont les calories, protéines, sucres, acides gras saturés et fibres. Compte tenu d'une certaine hétérogénéité des unités de mesure, ces nutriments sont décrits en termes d'équivalent grammes pour une valeur moyenne d'un kg de céréales (sauf les calories données par Kcal pour un kg de céréale).

Cependant, pour faciliter ce choix de caractéristiques, nous avons regroupé les quatre types d'étiquetage en deux groupes. Un premier type qualifié de simple qui regroupe le premier et le deuxième type d'étiquetage. Un deuxième type qualifié de complet regroupant le troisième et quatrième type d'étiquetage. La segmentation des céréales par ces deux types d'étiquetage distingue donc essentiellement celles qui n'affichent que les macro-nutriments et celles qui affichent également les micro-nutriments.

Outre les nutriments retenus, il est intéressant de tenir compte de la teneur en vitamines et minéraux. La liste de ces teneurs varie d'un produit à l'autre. Il existe en effet une forte hétérogénéité dans l'affichage de ces deux composantes. Les vitamines sont représentées par les valeurs moyennes en B1, B2, B3, B5, B6, B9, B12, C et E. Ces valeurs sont données en microgramme pour une valeur moyenne de 100g de céréales et sont corrélées entre elles. Nous avons donc agrégé toutes les vitamines en une seule variable donnant le nombre de vitamines renseignées sur l'étiquette. Pour les mêmes raisons, nous avons dû exclure tous les minéraux à l'exception du sodium. Ce choix s'explique par le fait que seul le sodium, parmi les autres minéraux, est présent sur toutes les références qui ont un étiquetage complet.

Il existe une grande variété de produits que l'on a regroupés en trois grandes familles. La première famille nommée segment « santé » regroupent les céréales « ligne et forme » et les céréales riches en fibres. La deuxième famille, nommée segment « gourmand » regroupe les céréales chocolatées, au miel et caramel, les céréales fourrées et muesli croustillant. La dernière famille regroupe les céréales au blé complet en galettes, les pétales de maïs, les autres céréales nature, flocons d'avoines et le muesli floconneux.

Nous soulignons que cette classification a été effectuée sur la base des allégations portées sur l'emballage des produits. Les types de céréales sont donnés par cinq catégories : les céréales à base de riz, maïs, blé, avoine ou mélange.

Le reste de l'hétérogénéité individuelle est contrôlée par l'introduction de variables socio-démographiques (âge, éducation) de la personne qui s'occupe des courses dans le ménage. La proportion des membres de la famille par classe d'âge (proportion 0-13, 14-24, 25-39, 40-64, >=65 ans) permet de capter l'impact de la composition du ménage. Les effets résidentiels sont captés par une variable décrivant la taille de la commune de résidence du ménage (commune rurale ou urbaine). Enfin, nous considérons aussi l'indice de masse corporelle (IMC) du responsable d'achat dans le ménage (normal, surpoids, obèse).

2.2 Méthode d'estimation :

L'estimation d'équations hédoniques pose le problème du choix d'une forme fonctionnelle, lequel déterminera la qualité des résultats. D'habitude, le choix de la forme structurelle de l'équation se limite aux modèles linéaires, semi-logarithmique et bilogarithmique. Dans tous les cas, le but recherché sera d'atteindre la meilleure corrélation possible entre prix et caractéristiques. Halvorsen et Pollakowski (1981) suggèrent d'utiliser la transformation de Box-Cox dans le but de discriminer entre les formes fonctionnelles classiques (log-linéaire, log-log...). Cette méthode permet de transformer les variables dépendantes et/ou indépendantes à l'aide d'un paramètre que l'on nomme ici λ . Si nous voulons transformer les prix, par exemple, la variable

transformée sera $\frac{(prix)^\lambda - 1}{\lambda}$ dans le cas d'un λ non nul,

et $\log (prix)$ dans le cas d'un λ nul⁴. Nous transformons ici les prix et les variables indépendantes continues à l'aide de deux paramètres λ et θ . Cette relation est donnée par :

⁴ L'objectif est d'obtenir entre -1 et 2 les valeurs de ce paramètre λ avec un intervalle de confiance à 95%, qui maximise l'ajustement de la forme générale sur les données. Si cet intervalle comprend la valeur 1, aucune transformation n'est nécessaire et la méthode linéaire offre le meilleur ajustement. En revanche, la spécification idéale pour la variable concernée (dépendante ou indépendante) sera logarithmique dans le cas où cet intervalle comprend la valeur 0. Pour plus de détail voir Marchand, O. et Skhiri, E. (1995).

$$prix_{jr}^{(\lambda)} = x_j^{(\theta)} \beta + z_j \delta + \varepsilon_{jr}$$

Où p_{jr} est le prix unitaire du produit j acheté dans l'enseigne r , x_j est l'ensemble des variables continues représentées ici par les quantités respectives de calories, sucres, fibres, sodium, protéines et vitamines (en nombre), z_j représente les autres caractéristiques qualitatives de type binaire (type d'étiquetage, type de famille, type de grain, marque), β et δ sont des vecteurs de paramètres à estimer. Les valeurs des micronutriments dans les céréales avec un étiquetage simple sont nulles. Or, l'estimation de θ n'est possible qu'en éliminant les observations correspondantes. Nous ne pouvions pas retenir cette option compte tenu de l'effet de l'étiquetage. Le paramètre θ a donc été fixé à 1. Le paramètre λ prend la valeur de -0.03 non significativement différent de 0 avec un intervalle de confiance compris entre -0.26 à 0.20. La spécification optimale pour notre fonction de prix hédoniques est donc la régression du logarithme des prix sur les variables quantitatives et qualitatives sans aucune transformation.

Dans la deuxième étape, les 11 variables du système demande AIDS (5 part de dépense, 5 prix implicites et dépense totale en nutriment) sont potentiellement endogènes. Pour vérifier l'hypothèse d'endogénéité, nous avons réalisé un test de Hausman (1978) qui consiste à comparer l'estimateur des moindres carrés ordinaires (MCO) biaisé sous $H1$ et efficient sous l'hypothèse nulle, avec l'estimateur des doubles moindres carrés (2SLS) non biaisé sous $H1$ et $H0$ et non efficient sous $H0$. L'idée de ce test est que, sous l'hypothèse nulle d'indépendance entre les erreurs et les variables explicatives, l'écart entre les coefficients des deux estimations n'est pas statistiquement différent de zéro. Le test d'Hausman compare la matrice de variance-covariance des deux estimateurs à l'aide de la statistique de Wald suivante :

$$W = (\hat{\beta}_0 - \hat{\beta}_1)'(\hat{V}_1 - \hat{V}_0)^{-1}(\hat{\beta}_0 - \hat{\beta}_1)$$

Où \hat{V}_1 et \hat{V}_0 sont les deux matrices de variance-covariance des estimateurs $\hat{\beta}_0$ (estimateur des MCO) et $\hat{\beta}_1$ (estimateur des 2SLS). Cette statistique suit une loi de Fischer pour 6 (prix implicites et dépense totale) et 4272 degrés de liberté. Les résultats de ce test permettent de rejeter l'hypothèse d'exogénéité des variables en questions (cf. Tableau 10 en annexes).

En conséquent, nous utilisons la méthode des variables instrumentales et on instrumente la dépense globale et les 5 prix implicites par le revenu (en log), des indicatrices de région et de type d'enseigne d'achat. On estime le modèle par triples moindres carré (3SLS) sous Stata version 10. Cette méthode permet d'estimer les 11 équations du système simultanément en commençant par projeter les variables endogènes sur les instruments et les autres variables exogènes, puis utiliser les résidus de cette première étape pour estimer la liaison entre les aléas des différentes équations. Cette méthode utilise les moindres carrés généralisés (MCG) pour estimer globalement l'ensemble du modèle en tenant compte de cette information.

Toutefois, il convient de s'assurer de la validité des instruments utilisés après avoir opté pour la méthode des variables instrumentales. Un instrument valide ne doit pas être corrélé avec le terme d'erreur. Le test de validité des instruments le plus couramment utilisé est celui du test de restriction de sur-identification de Sargan. Il consiste à régresser les résidus de l'estimation des 2SLS sur les variables exogènes et les instruments. La statistique de Sargan est égale au nombre d'observations multiplié par le coefficient de détermination de cette régression (R^2). Cette statistique suit une loi de khi-deux pour $n-K$ degrés de liberté, avec n le nombre de variables exogènes et d'instruments et K le nombre d'exogènes. Par ailleurs, la corrélation avec le terme d'erreur ne peut pas être testé si le nombre d'instruments est égal au nombre de variables endogènes. On dit dans ce cas que le modèle est juste-identifié. Disposant de 12 variables instrumentales et 11 variables endogènes, notre modèle est sur-identifié. Les résultats du test de Sargan montrent que les instruments utilisés sont valides (cf. Tableau 11 en annexes).

3. Résultats:

3.1 Effets marginaux

Les résultats d'estimation des prix sont donnés dans le tableau 1 : le prix augmente si les céréales sont caloriques, si elles sont sucrées, si elles sont riches en fibres, si elles sont plus salées, si elles contiennent un nombre important de vitamines. Enfin, le prix baisse si les céréales sont riches en acides gras saturés. Nous soulignons un effet important de l'étiquetage nutritionnel sur le prix des céréales. Plus l'étiquetage est détaillé, plus le prix est élevé. Le besoin d'information nutritionnelle est donc une caractéristique essentielle

des céréales petit déjeuner. La différenciation des produits par l'étiquetage nutritionnel qui rend l'information nutritionnelle plus accessible répond à ce besoin.

La segmentation du marché sur des bases nutritionnelles affecte le prix des céréales. Les céréales se trouvant sur le marché « santé » ont un prix plus élevé. Inversement, les céréales se trouvant sur le marché « gourmand » ont un prix plus faible. Nous rappelons que la segmentation du marché « santé » dans notre étude repose sur les allégations associées aux céréales de type « riche en fibre » et « ligne et forme ». Ce résultat indique que, toutes choses égales par ailleurs, le prix hédonique de ces allégations est positif. Ainsi, la disposition à payer pour des céréales affichant ces allégations est plus élevée que celles ne les affichant pas.

Le type de grains utilisé pour la fabrication des céréales petit déjeuner jouent un rôle moins important dans la détermination des prix. Les céréales à base de riz ou maïs ont un prix plus faible, à l'inverse des céréales à base d'avoine ou mélange qui ont un prix plus élevé. Le consentement à payer pour des céréales constituées d'un mélange de grains ou d'avoine est plus élevé que pour des céréales constituées de maïs ou de riz. En introduisant les marques comme variables de contrôle, les résultats sont identiques en termes de coefficients estimés, ce qui confirme leur robustesse. Le R^2 est de 0.72 et traduit un fort pouvoir explicatif des marques. Le coefficient des céréales premier prix est négatif et de -0.193. Ce résultat indique que la disposition à payer pour des céréales haut de gamme, représentées ici par les marques nationales, est plus importante que pour celles se situant en bas de gamme. La caractéristique « marque de céréales » est donc bien valorisée par les consommateurs.

Tableau 1
 Estimation de la fonction des prix hédoniques par les MCO
 Variable dépendante : logarithme du prix unitaire (euro/kg)
 (822 observations)

Variables indépendantes	Coefficients
Kilocalories	0,002 ***
Nutriments (g/100g)	
Protéines	-0,001
Sucres	0,004 ***
Acides gras saturés	-0,012 ***
Fibres	-0,005 **
Sodium	0,207 ***
Vitamines (en nombre)	0,013 ***
Type d'étiquetage	
Etiquetage détaillé	Réf
Etiquetage simple	0,317 ***
Famille produit	
Céréales « santé »	0,195 ***
Céréales « gourmand »	Réf
Autres céréales	-0,017
Type de céréales	
Riz	-0,145 ***
Maïs	-0,168 ***
Blé	Réf
Avoine	0,082
Mélange	0,090 ***
Marque	
Premier prix	-0,249 ***
MN1	0,297 ***
MDD	Réf
MN2	0,344 ***
MN3	0,204 ***
MN4	0,382 ***
Autres marque	0,441 ***
R ²	0,720

*, **, *** : significatif au taux de 10, 5, 1%

Source : TNS-2006

En termes d'interprétation, une unité supplémentaire du sucre conduit à une augmentation statistiquement significative du prix moyen du produit, de l'ordre de 0,4%. Pour un prix moyen du produit de 1 euro, cette unité supplémentaire a un prix implicite de 0,004 euro. Cependant, une diminution d'un gramme de protéine dans les céréales conduit à une augmentation du prix moyen de l'ordre de 0,1%. Il faut souligner que la valeur élevée du

sodium par rapport aux autres nutriments s'explique par l'effet de la transformation des quantités en grammes. La quantité de cette composante affichée sur l'étiquetage nutritionnel est donnée en milligrammes. A titre de comparaison, les prix implicites des nutriments dans d'autres études sont proches de nos résultats. Le tableau suivant résume les résultats de quelques travaux :

Tableau 2 : Résultats de quelques travaux sur les céréales

	Notre étude, Céréales, Valeur pour 1 euro	Beatty (2007), ensemble de l'alimentation	Eastwood et al.,(1986), ensemble de l'alimentation	Morgan et al., (1979), Céréales (valeur pour un dollar)	Shi et Price (1998), Céréales (valeur pour un Pound)
Sucres	0,004(€/g)	0,003 (\$/g) (glucides)	0,00021(\$/mg) (glucides)	--	--
Fibres	-0,005(€/g)	0,02(\$/g)	--	-0,003	-0,008
Sodium	0,207(€/g)	0,00025(\$/mg)	--	--	--
Acides gras saturés	-0,012(€/g)	0,005(\$/g) (total lipide)	0,002(\$/mg) (total lipide)	--	-0,04
Protéines	-0,001(€/g)	0,006(\$/g)	0,004(\$/mg)	0,018	0,07

3.2 Demande de caractéristiques :

Elasticités revenu des nutriments

Le tableau 3 présente les élasticités dépense des nutriments. D'après ce tableau, les élasticités revenu sont globalement faibles. Comme on peut le constater, si le revenu d'un ménage moyen augmente de 1% par exemple, la consommation de sucres ou de protéines augmente dans une moindre mesure que celle de fibres. A titre de comparaison, les élasticités revenu vont de 0,14 pour les glucides à 0,39 pour les matières grasses dans Huang (1996) et nos résultats sont très proches. Des résultats similaires ont été trouvés par LaFrance (1999) : les élasticités revenus pour l'énergie, protéines, matières grasses, glucides et cholestérol aux USA varient autour de 0,3 entre 1945 et 1994. Huang et Lin (2000) trouvent des valeurs plus élevées qui varient autour de 0,7 pour le cholestérol et de 0,98 pour les vitamines. En revanche, Adelaya et al., (1997) montrent que les élasticité revenu des matières grasses, cholestérol, glucides et calcium sont nulles, à l'exception des vitamines C avec une élasticité de 0,11.

Tableau 3
Élasticités dépense des nutriments
(4272 ménages)

Groupe de nutriments	Elasticités revenu
Sucre (kg)	0,133 ***
Fibres (kg)	0,154 ***
Sodium (kg)	0,149 ***
Acides gras saturés (kg)	0,171 ***
Protéines (kg)	0,128 ***

*, **, *** : significatif au taux de 10, 5, 1%.

Source : TNS-2006

Elasticités-prix des nutriments

Le tableau 4 regroupe les élasticités prix compensées qui mesurent l'effet de substitution. Sur la diagonale sont données les élasticités-prix directes. On remarque que les sucres et les fibres sont les moins sensibles aux variations de prix avec une élasticité de -0.49 et -0.50 respectivement. Quand le prix du sucre augmente de 1%, la demande pour ce nutriment diminue de 0.49%. On peut constater que les protéines sont les plus sensibles aux variations du prix avec une élasticité de -1.18. Dans la littérature, Park et Davis (2001) montrent que les élasticités-prix nutriments sont plus élevées lorsque la méthode des variables instrumentales est utilisée. Les élasticités prix des fibres et des graisses sont de l'ordre de -0,54 et -1,25 respectivement en utilisant la méthode des variables instrumentales et de 0,11 et -0,05 respectivement en utilisant la méthode des moindres carrés ordinaires. D'autres études montrent que la demande de nutriments est inélastique (Eastwood et al. (1986), Huang et Lin (2000), Beatty et LaFrance (2005), Allais et al. (2010)).

Il ressort également à la lecture de ce tableau que les relations mises en évidence sont des relations de substitution pures. Les calculs ne font l'objet d'aucune relation de complémentarité significative. Ainsi, les sucres sont substitués des acides gras saturés dans les deux sens. Si le prix du sucre augmente de 1% par exemple, sa consommation diminue mais celle des acides gras saturés augmente de 0.10%. Les relations mises en évidence montrent que les ménages consomment davantage de céréales grasses en cas de hausse du prix des sucres. D'autre part, les acides gras saturés sont substitués du sodium

(sel) dans les deux sens également. Cette dernière relation est plus significative avec des élasticités de substitution de 0.46 et 0.40 respectivement. Ces relations entre nutriments semblent confirmer l'hypothèse sous-jacente qui considère les produits alimentaires comme une combinaison de caractéristiques nutritionnelles. Du point de vue nutritionnel, le goût d'un produit peut être associé à un mélange de nutriments. Nos résultats indiquent que si les recommandations vont à l'encontre de la consommation de certains nutriments moins favorables sur le plan nutritionnel, il est difficile de produire des céréales qui soient à la fois moins sucrés, moins grasses et moins salés. De ce fait, c'est l'une des contraintes technologiques qu'avancent les industries agroalimentaires.

Tableau 4
Élasticités prix compensées (effet substitution) directes et croisées (4272 ménages)

	Sucre	Fibres	Sodium	Acide gras saturé	Protéines
Sucre (kg)	-0,495 (0,176) ***	-0,081 (0,166)	-0,057 (0,106)	0,196 (0,060) ***	0,436 (0,266)
Fibres (kg)	-0,071 (0,146)	-0,503 (0,231) **	0,037 (0,134)	0,084 (0,076)	0,453 (0,293)
Sodium (kg)	-0,036 (0,066)	0,027 (0,095)	-0,664 (0,119) ***	0,467 (0,037) ***	0,207 (0,123)
Acide gras saturé (kg)	0,107 (0,034) ***	0,052 (0,050)	0,400 (0,036) ***	-0,540 (0,026) ***	-0,025 (0,057)
Protéines (kg)	0,421 (0,266)	0,497 (0,332)	0,312 (0,198)	-0,040 (0,093)	-1,186 (0,569) **

*, **, *** : significatif au taux de 10, 5, 1% respectivement

Ecart types entre parenthèse

Source : TNS-2006

Effet des variables sociodémographiques

Le tableau 5 donne les coefficients estimés des variables socio-démographiques utilisées dans les équations de demande de nutriments. On constate que la demande pour le sucre est plus importante dans les ménages comportant des jeunes membres, tandis que la demande pour des fibres est plus élevée chez les adultes. Par ailleurs, on relève une influence significative du niveau de diplôme sur la demande des fibres. Le coefficient budgétaire augmente avec un diplôme élevé mais diminue lorsqu'il s'agit de la demande de sucre. On remarque aussi un effet important du statut pondéral sur la demande pour des acides gras saturés. Le fait que le responsable des achats soit obèse accroît la part budgétaire de cette caractéristique. S'agissant de la demande des protéines, la présence d'enfants et de jeunes dans le ménage favorise leur consommation. En revanche, un état d'obésité ou de surpoids du chef du ménage joue contre la demande de ce nutriment.

Tableau 5
Effet des variables socio-démographiques sur la demande de nutriments
(4272 ménages)

	Sucres (kg)	Fibres (kg)	Sodium (kg)	Acides gras saturés (kg)	Protéines (kg)
Age du panéliste (ans)	ε	ε	ε	ε	ε
Enfants (0_13 ans)	0,036 *	-0,044 **	-0,004	-0,008	0,020 **
Jeunes (14_24 ans)	0,039 **	-0,041 **	-0,021	-0,009	0,032 **
Adultes (40_65 ans)	0,003	0,009 *	-0,011	-0,003	0,002
Personnes âgées (>65ans)	0,000	0,012 *	-0,008	-0,012	0,008
Sexe masculin	0,004	0,001	-0,004	0,002	-0,003
Commune urbaine	-0,002	0,002	ε	-0,003	0,003
IMC obésité	-0,001	-0,005	-0,003	0,014 **	-0,005 *
IMC surpoids	0,000	0,001	-0,005	0,005	-0,001 **
IMC non déclaré	0,001	0,001	-0,011	0,016 *	-0,007
< à bac	0,003 *	0,002	-0,001	-0,007	0,003
> à bac	ε	0,005 *	-0,003	-0,005	0,003
Constante	0,420 ***	0,289 ***	0,031	0,049	0,211 *
R ² (moyen)	0,041	0,054	0,111	0,126	0,09

*, **, *** : significatif au taux de 10, 5, 1%.

Source : TNS-2006

Note : jeunes adultes est la référence pour la composition du foyer ; sexe féminin est la référence pour la variable sexe ; commune rurale est la référence pour la catégorie de commune ; IMC normal est la référence pour la catégorie d'IMC ; niveau bac est la référence pour la variable diplôme ;

Conclusion:

Dans un contexte de prise de conscience croissante des enjeux de santé publique, l'objectif de cet article était d'estimer la demande de caractéristiques et de montrer comment les quantités de nutriments consommées peuvent être affectées par les prix et les caractéristiques sociodémographiques des ménages. La méthode utilisée pour estimer la demande de nutriments est celle de la méthode de prix hédoniques inspirée de la théorie des caractéristiques (Rosen, 1974, Lancaster, 1966). Depuis le développement des travaux empiriques sur ces méthodologies, très peu d'applications se sont intéressées à l'estimation de la demande de caractéristiques dans le cadre de l'alimentation. C'est ce qui fait l'apport de ce travail que nous avons consacré aux céréales pour le petit déjeuner en France. Nous avons estimé la demande de nutriments en deux étapes. Dans la première, nous avons régressé les prix sur les caractéristiques des produits, ce qui a permis d'obtenir une estimation de la fonction de prix hédoniques. Dans la deuxième étape, les prix implicites de chaque caractéristique ont été régressés sur la quantité consommée de la caractéristique correspondante, ainsi que sur des variables sociodémographiques décrivant les individus.

Les résultats de l'estimation de l'équation de prix hédoniques à partir d'un échantillon de 822 produits de céréales sont cohérents avec la segmentation du marché sur des bases nutritionnelles, et montrent que l'influence positive des vitamines dans les céréales, par exemple, correspond à une valorisation de l'adéquation de l'offre alimentaire aux besoins des consommateurs. Les prix implicites négatifs de certaines caractéristiques indiquent un consentement à payer des produits moins conformes aux exigences nutritionnelles. Les résultats montrent également que la demande de nutriments est faiblement élastique, et les ménages réagissent bien aux variations des caractéristiques des produits.

Plusieurs perspectives de recherches s'ouvrent à l'issue de ce travail. La prise en compte des composantes de « coût » de production dans l'estimation de la demande ou l'offre de caractéristiques pourrait être une piste de recherche intéressante. Aussi, il serait intéressant d'appliquer cette approche sur des segments de marché différents et de vérifier la relation de l'hétérogénéité des consommateurs à la diversité de l'offre alimentaire. Ce type d'études pourrait intéresser les acteurs publics et privés dans la définition des politiques alimentaires et nutritionnelles.

REFERENCES

- Abdulai, A., Aubert, D., (2004) "A cross-section analysis of household demand for food and nutrients in Tanzania". *Agricultural Economics*, vol. 31, pp. 67-79
- Allais, O., Bertail, P., et Nichèle, V. (2010). "The Effects of a Fat Tax on French Households' Purchases: A Nutritional Approach." *American Journal of Agricultural Economics*, Vol. 92, pp. 228-245.
- Beatty, K.M., (2007), "Recovering the shadow value of nutrients". *American Journal of Agricultural Economics*, vol.89, pp. 52-62.
- Chaieb, A., (2011), « Deux essais sur la valorisation par les consommateurs des caractéristiques des produits alimentaires en France » Thèse de doctorat, AgroParisTech.
- Deaton, A. et Muellbauer, J. (1980). 'An Almost Ideal Demand System.' *The American Economic Review*, Vol. 70, pp. 312-326
- Donnet M. L., Weatherspoon, D. D., et Hoehn, J. J., (2008). "Price determinants in top-quality e-auctioned specialty coffees". *Agricultural Economics*, vol. 38, pp. 267-276.
- Eastwood et al., (1986). "Household nutrient demand: use of characteristics theory and a common attribute model". *Southern Journal of Agricultural Economics*, pp. 235-246.
- Edmeades, S. (2007). "A hedonic approach to estimating the supply of variety attributes of a subsistence crop". *Agricultural Economics*, vol. 37, pp.19-28.
- Fousekis, P. et Lazaridis, P. (2005), "The demand for selected nutrients by Greek households: an empirical analysis with quantile regressions." *Agricultural Economics*, Vol. 3, pp. 267-279.
- Griffith, R. et Nesheim, L., (2008). 'Household willingness to pay for organic products' *The Institute for Fiscal Studies Department of Economics*, UCL cemmap working paper CWP18/08.
- Halvorsen, R. et Pollakowski, H. O. (1981). "Choice of functional form for hedonic price equations." *Journal of Urban Economics*, vol. 10, pp. 37-49.
- Huang CL, et BH Lin, (2007). "A Hedonic Analysis of Fresh Tomato Pricing among Regional Markets". *Review of Agricultural Economics*, vol. 29, pp. 783-800.
- Huang, K.S. et Lin, B.H. (2000). 'Estimation of Food Demand and Nutrient Elasticities from Household Survey Data'. Food and Rural Economics Division, Economic Research Service, U.S. Department of Agriculture. Technical Bulletin No. 1887.
- Jeffrey T. et LaFrance, J.T., (1999). "Inferring the Nutrient Content of Food with Prior Information". *American Journal of Agricultural Economics*, vol. 81, pp. 728-734.
- Kristofersson, D. et Rickertsen, K. (2007). "Hedonic Price Models for Dynamic Markets." *Oxford Bulletin of Economics and Statistics*, vol. 69, pp. 387-412.

Lancaster, Kelvin J. (1966), "A New Approach to Consumer Theory." *Journal of political Economy*, 74, 132-57.

Marchand, O. et Skhiri, E. (1995). 'Prix hédoniques et estimation d'un modèle structurel d'offre et de demande de caractéristiques.' *Economie & Prévision*.

Morgan, K. J., Metzen, E. J. et Johnson, S. R. (1979). "An Hedonic Index for Breakfast Cereals." *The Journal of Consumer Research*, Vol. 6, pp. 67-75.

Park, J. et Davis, G. C. (2001). "The Theory and Econometrics of Health Information in Cross-Sectional Nutrient Demand Analysis". *American Journal of Agricultural Economics*, vol. 83, pp. 840-851.

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

Stanley, Linda R. and John Tschirhart. 1991. "Hedonic Prices for a Nondurable Good: The Case of Breakfast Cereals," *Review of Economics and Statistics*, 73: 537-541

Teuber, R., (2009). "Estimating the Demand for Characteristics via Hedonic-Theoretical Considerations and an Empirical Application on Specialty Coffee." *Contributed Paper for presentation at the 2009 EAAE PhD Workshop*, Giessen, Germany.

Annexes

Tableau 6 : Statistiques descriptives sur les caractéristiques des céréales pour le petit déjeuner : nombre d'observations 822.

Variabes	Modalité	Moyenne	Ecart type
Prix (€/kg)	Continue (euro/kg)	5,850	1,762
Kilocalories	Continue (en Kcal)	386,831	36,590
Nutriments (g/100g)			
Protéines	Continue (en g)	8,515	2,400
Sucres	Continue (en g)	25,054	9,782
Acides gras saturés	Continue (en g)	2,627	2,839
Fibres	Continue (en g)	5,459	4,366
Sodium	Continue (en g)	0,366	0,236
Vitamine	Continue (en nombre)	5,156	3,189
Famille de produit			
Céréales « santé »	Oui=1 ; Non=0	0,300	0,459
Céréales « gourmand »	Oui=1 ; Non=0	0,583	0,493
Céréale « autre »	Oui=1 ; Non=0	0,105	0,306
Marque			
Premier prix	Oui=1 ; Non=0	0,080	0,272
MDD	Oui=1 ; Non=0	0,225	0,418
MN1	Oui=1 ; Non=0	0,060	0,237
MN2	Oui=1 ; Non=0	0,328	0,470
MN3	Oui=1 ; Non=0	0,190	0,392
MN4	Oui=1 ; Non=0	0,055	0,228
Autre marque	Oui=1 ; Non=0	0,062	0,241
Type d'étiquetage			
Etiquetage détaillé	Oui=1 ; Non=0	0,951	0,215
Etiquetage simple	Oui=1 ; Non=0	0,048	0,215
Type de céréales			
Riz	Oui=1 ; Non=0	0,047	0,213
Maïs	Oui=1 ; Non=0	0,108	0,311
Blé	Oui=1 ; Non=0	0,305	0,461
Avoine	Oui=1 ; Non=0	0,024	0,154
Mélange	Oui=1 ; Non=0	0,500	0,500

Source : TNS-2006

Tableau 7 : Statistiques descriptives sur l'échantillon des ménages (4272 ménages) TNS-2006

Variabiles	Moyenne	Ecart type
Variabiles dépendantes		
Part de dépense en sucre (en %)	0,152	0,045
Part de dépense en fibres (en %)	0,173	0,054
Part de dépense en sodium (en %)	0,238	0,115
Part de dépense en acides gras saturés (en %)	0,278	0,136
Part de dépense en protéines (en %)	0,158	0,051
Variabiles indépendantes		
Prix marginal sucre (euro/g)	0,024	0,016
Prix marginal fibres (euro/g)	0,185	0,161
Prix marginal sodium (euro/g)	4,410	8,423
Prix marginal acides gras saturés (euro/g)	0,908	1,084
Prix marginal protéines (euro/g)	0,076	0,035
Age du panéliste (ans)	47,324	13,791
<i>Composition du ménage (en %)</i>		
Enfants (0_13) ans	0,170	0,223
Jeunes (14_24) ans	0,121	0,196
Jeunes Adultes (25_39) ans	0,214	0,307
Adultes (40_65) ans	0,346	0,346
Personnes âgées (>65) ans	0,148	0,336
<i>Sexe du panéliste</i>		
Sexe féminin	0,955	0,207
Sexe masculin	0,045	0,207
<i>Commune d'habitation</i>		
Commune rurale	0,527	0,499
Commune urbaine	0,473	0,499
<i>IMC du panéliste</i>		
IMC normal	0,587	0,492
IMC obésité	0,107	0,309
IMC en surpoids	0,234	0,424
IMC non déclaré	0,057	0,232
<i>Niveau de diplôme du panéliste</i>		
<à bac	0,405	0,491
Niveau bac	0,255	0,436
> à bac	0,311	0,463
<i>Instruments identifiants</i>		
Revenu (en log)	7,124	0,451
Région Paris	0,164	0,370
Région Est	0,103	0,304
Région Nord	0,098	0,297
Région Ouest	0,192	0,394
Région Centre ouest	0,081	0,273
Région Centre est	0,152	0,359
Région Sud est	0,113	0,317
Région Sud ouest	0,097	0,296
Client hypermarché (plus de 80% de fréquentation)	0,330	0,470
Client supermarché (plus de 80% de fréquentation)	0,436	0,496
Client autre enseigne	0,234	0,423

Source : TNS-2006

Tableau 8
Tests de colinéarité entre les régresseurs

Variable	VIF ⁵	1/VIF
Kilocalories	4,850	0,206
Protéines (g)	2,900	0,345
Sucres (g)	4,960	0,202
Acides gras saturés (g)	4,210	0,237
Fibres (g)	2,730	0,366
Sodium (g)	2,980	0,336
Vitamines (nombre)	1,920	0,521
Etiquetage simple	6,220	0,161
Céréales « santé »	2,820	0,355
Autre Céréales	1,730	0,578
Riz	1,670	0,598
Mais	2,590	0,386
Avoine	1,660	0,603
Mélange	2,240	0,447
Premier prix	1,170	0,852
MN1	1,270	0,787
MN2	1,280	0,783
MN3	1,400	0,717
MN4	2,390	0,418
Autres marques	2,820	0,354
Moyenne VIF	2,690	

Source : TNS-2006

⁵ Variation inflation factors : Dans la pratique, on mettra en évidence une multicolinéarité dans la régression si la VIF la plus élevée est plus grande que 10 ou si la moyenne de tous les VIF est beaucoup plus forte que 1. Ici, les deux critères semblent respectés, ce qui nous permet d'être confiants sur l'absence de multicolinéarité.

Tableau 9
Régression instrumentale
Variable dépendante : dépense totale en nutriments (euro/kg)
(4272 ménages)

	Coefficient	Ecart type
Age du panéliste (ans)	-0,002	0,004
Enfants (0_13) ans	1,333***	0,131
Jeunes (14_24) ans	2,024***	0,121
Adultes (40_64) ans	-0,006	0,112
Personnes âgées (>65) ans	-0,057	0,168
Sexe homme	-0,036	0,099
Commune urbaine	0,066	0,043
IMC obésité	-0,166**	0,067
IMC surpoids	-0,007	0,050
IMC non déclaré	0,031	0,086
< à bac	-0,021	0,049
>à bac	0,126**	0,053
Instruments identifiants		
Revenu par uc (en log)	0,157***	0,031
Paris	-0,019	0,077
Est	-0,191**	0,086
Nord	-0,109	0,086
Ouest	-0,075	0,075
Centre ouest	-0,231**	0,091
Sud est	-0,030	0,078
Sud ouest	-0,066	0,087
Client supermarché	-0,637***	0,054
Client hypermarché	-0,390***	0,050
Constante	-1,526***	0,337
R ²	0,149	

*, **, *** : significatif au taux de 10, 5, 1%.

Source : TNS-2006

Note : jeunes adultes est la référence pour la composition du foyer ; sexe féminin est la référence pour la variable sexe ; commune rurale est la référence pour la catégorie de commune ; IMC normal est la référence pour la catégorie d'IMC ; diplôme moyen (niveau bac) est la référence pour la variable diplôme ; centre-est est la référence pour la variable région ; client autre enseigne est la référence de la catégorie enseigne.

Tableau 10
Tests de sur identification des instruments

	Statistique de Sargan**	Probabilité*
Equation sucre (kg)	0,897	0,925
Equation fibres (kg)	0,509	0,973
Equation sodium (kg)	1,265	0,867
Equation acide gras saturé (kg)	0,212	0,995
Equation protéines (kg)	0,850	0,932

Source : TNS-2006

** Statistique du khi-deux pour 10 degrés de liberté

Probabilité de ne pas rejeter l'hypothèse nulle d'orthogonalité des instruments

Tableau 11
Tests d'exogénéité sur l'échantillon complet
4272 observations

	Statistique de Hausman**	Probabilité*
Equation sucre (kg)	2,519	0,020
Equation fibres (kg)	2,923	0,008
Equation sodium (kg)	0,650	0,690
Equation acide gras saturé (kg)	2,398	0,026
Equation protéines (kg)	2,275	0,034

Source : TNS-2006

** Statistique de F sur 6 variables lp1, lp2, lp3, lp4, lp5 et lx (prix implicites et dépense totale)

*Probabilité de ne pas rejeter l'hypothèse nulle d'exogénéité, pour 6 et 4272 degrés de liberté.