



AgEcon SEARCH
RESEARCH IN AGRICULTURAL & APPLIED ECONOMICS

The World's Largest Open Access Agricultural & Applied Economics Digital Library

This document is discoverable and free to researchers across the globe due to the work of AgEcon Search.

Help ensure our sustainability.

Give to AgEcon Search

AgEcon Search

<http://ageconsearch.umn.edu>

aesearch@umn.edu

*Papers downloaded from **AgEcon Search** may be used for non-commercial purposes and personal study only. No other use, including posting to another Internet site, is permitted without permission from the copyright owner (not AgEcon Search), or as allowed under the provisions of Fair Use, U.S. Copyright Act, Title 17 U.S.C.*

RECHERCHES EN ECONOMIE ET SOCIOLOGIE RURALES

COMPORTEMENT DU CONSOMMATEUR ET CARACTÉRISTIQUES DES BIENS : RECHERCHES ACTUELLES ET ENJEUX FUTURS

La saturation quantitative de la demande alimentaire dans les pays développés contraint les entreprises à adopter des politiques de différenciation qualitative et à multiplier les variantes des produits alimentaires offerts sur les marchés. Il importe donc de comprendre comment s'effectuent les choix entre des produits dont les caractéristiques sont très proches, la difficulté étant de prévoir les réactions des consommateurs (à une modification d'un prix ou d'une caractéristique, à l'introduction d'un nouveau produit, etc.), rarement capables de connaître et d'identifier toutes les caractéristiques d'un produit, et dont la perception est influencée par l'information disponible. Ce texte présente les approches modélisatrices utilisées pour prendre en compte les caractéristiques des produits dans l'analyse des choix individuels, en accordant une attention particulière aux enjeux qui sous-tendent les développements récents dans ce domaine.

Introduction

L'analyse des choix individuels se donne pour objectif de trouver des moyens d'observation et de mesure des déterminants des choix et de leur variabilité. Pour cela, elle doit pouvoir disposer d'un cadre conceptuel qui permette d'identifier les déterminants des choix individuels. L'économie et la psychologie fournissent ce cadre en supposant que les individus agissent de telle sorte que leur choix est celui qui leur procure soit la plus grande utilité (économie), soit la plus grande satisfaction (psychologie) dans l'ensemble des choix possibles. Pour identifier, à partir des choix observés, les caractéristiques des produits qui influencent les choix des consommateurs et caractériser les attributs sociodémographiques qui expliquent la variabilité des choix, on introduit une variable non observable, ou variable latente, appelée "utilité". Le modélisateur fait alors dépendre celle-ci des caractéristiques observées des biens et des consommateurs. Il s'appuie sur une théorie : la théorie des caractéristiques, qui stipule que les produits, seuls ou combinés, sont désirés non pas pour eux-mêmes, mais pour les caractéristiques qu'ils possèdent (par exemple, un repas est une combinaison particulière d'attributs nutritionnels et gustatifs). Les autres facteurs affectant l'utilité mais inobservés par le modélisateur sont pris en compte via l'introduction de termes d'erreur. Cette partie "structurelle" du modèle est alors complétée en supposant que le choix observé d'un produit par un consommateur signifie que l'utilité qu'il retire de ce choix est supérieure à l'utilité qu'il retirerait du choix de tout autre produit.

Ce cadre résume l'approche modélisatrice sous-jacente aux deux grandes familles de modèles habituellement utilisées pour analyser les choix individuels de produits différenciés : les modèles hédoniques et les modèles de choix discrets. Les deux sections suivantes illustrent ces deux types de modélisations par des travaux réalisés dans le département SAE2. Une dernière section donne un éclairage sur des pistes de recherche actuellement suivies dans le domaine de la modélisation des choix.

Caractéristiques objectives et caractéristiques sensorielles

La théorie des caractéristiques rationalise une pratique empirique plus ancienne, celle de la régression hédonique, permettant de mesurer la contribution des caractéristiques d'un produit à son prix. Cette pratique consiste à régresser le prix du bien sur certains de ses attributs. Le prix implicite, ou prix caché, d'un attribut est alors défini comme la dérivée du prix du produit par rapport à cet attribut. Dans le cadre d'un marché parfaitement concurrentiel, il s'interprète comme la valeur que les consommateurs attribuent à une unité supplémentaire de la caractéristique. Si le prix implicite estimé est non significativement différent de zéro, soit la caractéristique n'est pas perçue, soit elle n'est pas considérée comme une caractéristique importante et n'est pas valorisée par les consommateurs.

Si la méthode des prix hédoniques a principalement été appliquée à des biens durables, en raison de la facilité

avec laquelle on peut identifier et mesurer leurs attributs, elle peut aussi être utilisée pour des produits alimentaires. Les caractéristiques pertinentes alors ne se limitent pas aux propriétés physiques apparentes ou informationnelles. Dans le cas du vin, par exemple, si l'étiquette d'une bouteille fournit des informations sur le millésime, le nom du domaine, le classement éventuel du vin, etc., la plupart des caractéristiques importantes restent inconnues. En effet, l'étiquette ne donne en général aucune information sur les caractéristiques sensorielles (apparence, odeur et goût) d'un vin. Or, dans la mesure où elles sont généralement les facteurs essentiels de l'acceptabilité ou du rejet des aliments par les consommateurs, les propriétés sensorielles doivent également être considérées.

Partant de cette idée, Combris, Lecocq et Visser (1997 et 2000) estiment des équations de prix hédoniques pour les vins de Bordeaux et de Bourgogne, en considérant non seulement les caractéristiques objectivées du vin (c'est-à-dire celles qui sont inscrites sur l'étiquette : millésime, appellation, classement), mais aussi les caractéristiques sensorielles, de manière à connaître la valeur que les consommateurs attribuent implicitement à ces deux types de caractéristiques. Ils utilisent des données issues de deux numéros spéciaux de *50 Millions de Consommateurs* présentant l'évaluation circonstanciée d'un échantillon de 519 vins de Bordeaux et de 613 vins de Bourgogne par des jurys d'experts. Les vins ont été sélectionnés de manière aléatoire parmi ceux dont le prix de vente "départ cave" se situait entre 20 et 200 francs la bouteille. En outre, l'échantillon a été tiré de façon à être représentatif des principales appellations d'origine contrôlée.

Les experts ont évalué les vins à l'aveugle. Pour chaque vin, la revue publie un commentaire rendant compte de leurs appréciations olfactives (intensité aromatique, finesse, complexité des arômes), de leurs appréciations gustatives (franchise de l'attaque, souplesse, gras en bouche, concentration, harmonie entre les constituants, longueur de la finale, etc.), et quelques remarques d'ordre général (sur le niveau d'alcool, sur le potentiel de vieillissement, etc.). Dans ce qui suit, ces variables sont appelées variables sensorielles. Une note moyenne (entre 0 et 20) attribuée à chaque vin est également reportée. La revue fournit des informations sur le nom du château ou du domaine, le classement, la couleur, le millésime et l'appellation. Ces variables sont appelées variables objectives. Enfin, la revue fournit pour chaque vin le prix auquel la bouteille a été achetée.

La variable dépendante dans les équations de prix hédoniques est le logarithme du prix d'une bouteille de vin. Les variables explicatives sont tous les attributs de la bouteille, représentés par les variables objectives et sensorielles définies plus haut. Les résultats montrent que si les prix dépendent des attributs objectivés, ils dépendent beaucoup moins nettement des attributs sensoriels, et sont plus faiblement corrélés avec les notes des experts qu'avec le classement par exemple.

Valorisation d'un signe de qualité

Une autre manière de rendre compte de l'influence des caractéristiques des biens sur les choix individuels consiste à utiliser des modèles de choix discrets. Ceux-ci permettent de mesurer l'effet d'une caractéristique sur la probabilité de choisir un produit. Tous sont basés sur

des modèles à utilité aléatoire, où l'utilité est décomposée en une partie "systématique", fonction des caractéristiques des biens et des attributs des individus, et une partie aléatoire représentant non seulement l'aléa "pur", mais aussi l'hétérogénéité inobservée dans les préférences, à savoir l'expérience et l'information des consommateurs sur les caractéristiques des biens. Paramétrer la fonction d'utilité permet alors d'obtenir des expressions pour les probabilités de choix dépendant des caractéristiques des produits et des attributs des individus. Si certains modèles sont connus pour imposer beaucoup de structure sur les préférences (dans le logit multinomial, par exemple, celles-ci sont identiques pour tous les consommateurs), d'autres, comme le logit mélangé, fournissent une représentation flexible des préférences en permettant à celles-ci de varier dans la population étudiée.

Un domaine d'application des modèles de choix discrets est celui des signes de qualité. Les consommateurs sont confrontés, dans le domaine alimentaire en particulier, à un grand nombre de signaux informationnels : AOC, labels, etc. Ces derniers se proposent d'améliorer l'information des consommateurs. Mais encore faut-il que l'information véhiculée soit correctement perçue, d'autant plus que la prolifération des signaux peut conduire les consommateurs à une certaine confusion quant à leur signification. Certains travaux ont cherché à caractériser la demande des consommateurs pour ce type d'attributs et à mesurer la façon dont les signes de qualité étaient valorisés par le marché.

Bonnet et Simioni (2001) utilisent un modèle logit mélangé pour modéliser les choix des consommateurs entre plusieurs marques de camembert et identifier leur disposition à payer pour le label "AOC Camembert de Normandie". Ils utilisent des données issues de la base SECODIP pour l'année 1998, qui permettent de suivre les achats d'un panel représentatif de 4267 ménages français. Pour chaque achat, les données fournissent le prix unitaire et certaines caractéristiques du camembert acheté (marque, présence ou non de l'AOC), ainsi que des informations sur les attributs sociodémographiques du ménage acheteur.

Si les données fournissent des informations sur le bien qui a été acheté, elles n'en donnent aucune sur les biens qui étaient présents au moment de l'achat mais qui n'ont pas été achetés. Autrement dit, l'ensemble de choix du consommateur est inconnu. Pour contourner ce problème, les auteurs supposent que l'ensemble des marques proposées à un ménage à une date donnée dans une enseigne donnée correspond à l'ensemble des marques achetées (et enregistrées dans les données) à la même date dans la même enseigne.

L'estimation du modèle permet d'obtenir une courbe de demande pour l'AOC et de caractériser les dispositions à payer pour ce label en fonction du revenu et de l'âge des individus. Les résultats montrent que la très grande majorité des consommateurs ne valorisent pas le signe de qualité fourni par le label AOC : pour un même prix, seulement 16% d'entre eux choisirait un camembert AOC plutôt qu'un camembert non AOC. Sur ce marché, la marque semble être un signal plus pertinent. Les résultats montrent également l'importance du revenu dans le consentement à payer, les consommateurs les plus aisés étant disposés à payer plus que les autres pour le label AOC.

Les pistes en modélisation des choix individuels

La prise en compte des caractéristiques non observées

Les modèles considérés dans les sections précédentes supposent que toutes les caractéristiques des produits sont observées par l'économètre. Or, cette hypothèse est particulièrement forte lorsque les choix sont étudiés sur la base de données d'achats. En effet, dans ce type de données, on n'observe généralement que la marque achetée par le consommateur, le prix qu'il a payé, et quelques-unes des caractéristiques physiques du bien (camembert au lait cru ou pasteurisé, par exemple). D'autres caractéristiques peuvent intervenir dans le choix du consommateur et ne pas être observées, soit parce qu'elles sont difficiles à quantifier (le style, etc.), soit parce qu'elles sont perçues différemment d'un consommateur à l'autre (la qualité, la durabilité, etc.). D'ailleurs, il n'est pas rare d'observer une demande positive pour un produit qui n'est pas meilleur qu'un autre en termes de caractéristiques physiques et ayant un prix plus élevé. Dans un tel cas, il est probable que l'on n'observe pas la variable qui rend le produit attractif pour certains consommateurs.

Les modèles de choix discrets, et plus récemment les modèles hédoniques, permettent de prendre en compte l'existence d'éventuelle caractéristiques non observées par l'économètre. Cette introduction permet alors au modélisateur de proposer un traitement au problème de l'endogénéité du prix. En effet, de nombreux travaux sur les choix du consommateur négligent le fait que certaines variables affectant l'utilité, comme le prix, la présence d'une promotion ou d'une opération publicitaire, sont fixées par l'offreur (producteur ou distributeur) sur la base d'informations relatives à la situation du marché étudié qui sont non observables par l'économètre mais qui, néanmoins, influent sur le choix du consommateur. Ces variables peuvent donc être corrélées avec le terme d'erreur introduit dans l'utilité du consommateur et entraîner des biais sérieux lors de l'estimation des paramètres, laquelle nécessite alors le recours à des techniques plus élaborées et, parfois, à d'autres sources de données. Bonnet, Dubois et Simioni (2005) illustrent cette démarche lors de l'estimation d'un modèle de choix de type logit emboîté pour étudier la demande d'eau embouteillée en France à partir des données de la base SECODIP. Des travaux récents généralisent également

les modèles hédoniques pour la modélisation des cas où des caractéristiques sont inobservées.

Une approche en termes de modèle généralisé de choix

Les principaux enjeux dans le domaine de la modélisation des choix du consommateur peuvent être illustrés à partir du graphique 1. Les modèles usuels de choix, tels que ceux décrits plus haut, y sont représentés dans la partie centrale. L'ouverture de la "boîte noire" que constitue la formation des préférences des individus est alors proposée en suivant les deux entrées complémentaires suivantes :

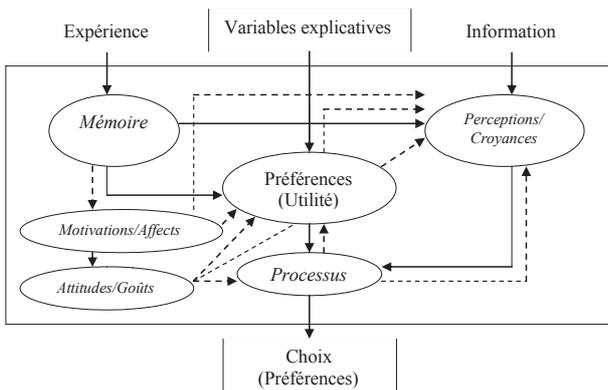
(1) La première est décrite par les flèches pleines. Elle consiste à replacer les choix des consommateurs dans une perspective temporelle et à introduire dans la modélisation économique des concepts empruntés à la psychologie : la mémoire qui dépend des expériences de consommation passée, les perceptions et croyances qui dépendent de l'information disponible, etc. Les choix sont alors le fruit d'un processus dans lequel les consommateurs collectent d'abord de l'information sur les produits. Cette information est ensuite convertie en "attributs perçus" par les individus, la mémoire jouant alors un rôle dans la révision de leurs perceptions et de leurs croyances. Le processus de choix consiste en l'agrégation de ces perceptions en un indice d'utilité qui est alors maximisé.

(2) La deuxième entrée, résumée par les flèches sous forme de tirets, complexifie le schéma précédent en y introduisant des "construits" (motivations, attitudes, etc.), eux aussi empruntés à la psychologie.

Dans le but d'incorporer certains de ces éléments d'un point de vue opérationnel, il a été proposé dans la littérature de généraliser le cadre des modèles de choix discrets en leur adjoignant deux types de modélisations déjà utilisées en psychométrie et en marketing : une modélisation à variables latentes et une modélisation à classes latentes. Les interactions entre ces modèles sont présentées dans le graphique 2. La lecture de ce graphique appelle les commentaires suivants :

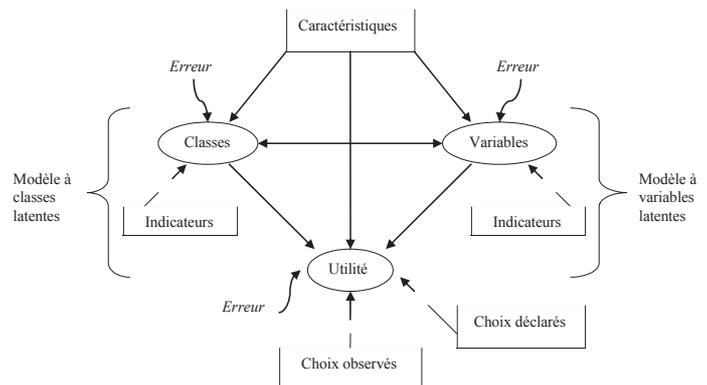
(1) La partie centrale d'un modèle de choix généralisé est constituée de la structure usuelle des modèles de choix discrets. Il s'agit là du "noyau" du modèle. Ce cadre peut ensuite être enrichi par la prise en compte de l'hétérogénéité qui peut exister dans les préférences des individus en utilisant, par exemple, le modèle logit mélangé.

Graphique 1 - Une vision générale du processus de choix



Source : McFadden (2001). *Economic choices*, Nobel Lecture. *American Economic Review*, vol. 91, pp. 351-378

Graphique 2 - Structure d'un modèle généralisé de choix



Source : Walker et Ben-Akiva (2002). *Generalized random utility model*. *Mathematical Social Sciences*, vol. 43, pp. 303-343

(2) À ce cadre central, certains auteurs ont envisagé d'ajouter un modèle à variables latentes. Son objectif est d'expliquer certains choix irrationnels du point de vue de l'économie, en incorporant dans la modélisation de l'utilité des facteurs psychologiques inobservés par l'économètre (donc latents) tels que des attitudes ou perceptions. Cette modélisation propose donc de capter l'hétérogénéité observée dans les choix des individus en combinant une information "dure" sur ces derniers, mesurée par des caractéristiques sociodémographiques, et une information plus "molle" sur certaines de leurs caractéristiques psychologiques comme la perception du risque ou l'impatience, qui, elles, sont difficilement mesurables. Le modélisateur doit alors avoir recours à des indicateurs construits, par exemple, à partir de réponses à des questions subjectives.

(3) Parallèlement aux modèles mélangés ou à variables psychologiques latentes, une modélisation en termes de classes latentes peut aussi être utilisée pour capter une partie de l'hétérogénéité non observée entre individus. L'idée est que la population des consommateurs peut être décomposée en classes qui ne peuvent pas être identifiées directement à partir des données. À chacune de ces classes correspond alors une fonction d'utilité et donc des paramètres propres. Ce type de modélisation a été utilisée par Etilé (2005) pour étudier les déterminants de certaines consommations risquées (cannabis et alcool). Il met en évidence trois classes de consommateurs et étudie la probabilité d'appartenance à chacune d'entre elles en fonction de variables latentes comme l'hyperactivité ou l'aversion au risque pour lesquelles il dispose de différents indicateurs dans ses données.

Choix observés et choix déclarés

De nombreux travaux portant sur l'analyse des choix individuels utilisent des données issues de l'observation de situations de choix réelles. Par exemple, Bonnet et Simioni (2001) utilisent des données extraites du panel de consommateurs SECODIP. Ces données retracent tous les achats de ces consommateurs. Il s'agit là de données qui entrent dans la catégorie "choix observés". Ces don-

nées nous informent sur les choix réalisés dans des situations concrètes, c'est-à-dire intégrant des contraintes réelles sur les décisions des individus, en particulier la contrainte budgétaire. En revanche, elles soulèvent plusieurs problèmes. Par exemple, l'ensemble de choix d'un consommateur qui achète un bien n'est pas observé par le modélisateur. Celui-ci doit alors faire des hypothèses sur cet ensemble ou avoir recours à des techniques d'estimation assez complexes. De plus, peu de caractéristiques des produits sont observées dans ce type de données.

Pour pallier les défauts des données de choix observés, on peut avoir recours à des expérimentations où l'on demande à des individus de choisir entre différentes alternatives (Combris et Ruffieux, 2005). Ces données appartiennent à la catégorie "choix déclarés". L'idée est de placer les individus dans une situation d'arbitrage entre différents niveaux de prix et combinaisons de caractéristiques. Un des avantages de ces données est qu'elles permettent de s'intéresser à des caractéristiques qui ne sont pas forcément présentes dans les produits proposés aux consommateurs sur le marché (présence d'OGM ou pas, par exemple). Leur défaut majeur est que la situation de choix proposée à l'individu est totalement hypothétique, celui-ci ne subissant aucune des contraintes auxquelles il est confronté dans une situation de choix réelle.

La littérature sur les choix individuels propose d'utiliser les deux catégories de données afin que les forces des uns pallient les défauts des autres. Par exemple, les données de choix observés peuvent être utilisées avec l'objectif de décrire les choix des consommateurs dans le cadre de l'équilibre sur le marché considéré. Peu d'attention est alors donnée aux arbitrages opérés par le consommateur, d'une part entre les niveaux de caractéristiques des produits présents sur le marché, ceux-ci étant le plus souvent peu différenciés dans ce type de données, et d'autre part lorsqu'une nouvelle caractéristique ou un nouveau produit est introduit sur le marché. Les données de choix déclarés peuvent en revanche être utilisées pour apporter un éclairage sur ces arbitrages.

Sébastien Lecocq, INRA - SAE2/CORELA Ivry
Michel Simioni, INRA - SAE2/MAIA Toulouse
simioni@toulouse.inra.fr - lecocq@ivry.inra.fr

Pour en savoir plus

Bonnet, C. ; Dubois, P. ; Simioni, M. (2005). *Two-part tariffs versus linear pricing in manufacturers and retailers strategic relationships: empirical tests on differentiated products markets*. Document de travail, INRA-ESR, Toulouse, 40 p.

Bonnet, C. ; Simioni, M. (2001). Assessing consumer response to protected designation of origin labelling: a mixed multinomial logit approach. *European Review of Agricultural Economics*, vol. 28, pp. 433-449.

Combris, P. ; Lecocq, S. ; Visser, M. (1997). Estimation of a hedonic price equation for Bordeaux wine: does quality matter? *Economic Journal*, vol. 107, pp. 390-402.

Combris, P. ; Lecocq, S. ; Visser, M. (2000). Estimation of a hedonic price equation for Burgundy wine. *Applied Economics*, vol. 32, pp. 961-967.

Combris, P. ; Ruffieux, B. (2005). La révélation expérimentale des préférences des consommateurs. *INRA Sciences Sociales*, n° 3-4, 4 p.

Etilé, F. (2005). Who the hats fit? Drugs, public policy and teenager heterogeneity. *Health Economics*, à paraître.