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.*

Les déterminants de la consommation individuelle de vin en France

Mourad K. AYOUZ
Christian CIBENEL
Jean-Pierre LAPORTE

Cahiers d'économie et sociologie rurales, n° 60-61, 2001

Mourad K. AYOUZ*, Christian CIBENEL*, Jean-Pierre LAPORTE*

The determinants of individual consumption of wine in France

Summary – The objective of this article is to study the factors influencing indiduals' consumption of wine in France using cross section data gathered by ONIVINS and INRA for 2000. We first estimate the probatilities of drinking wine and its frequency (number of times per week) using binary and ordered probit models, respectively. The ordered probit model has however some weaknesses. It does not allow to have estimated parameters to take different values to explain the desision to consume wine and the frequencies. If we only want to explain the frequencies of drinking wine, the estimation of a simple ordered probit based on only those who have decided to drink wine (i.e. *those with a positive frequency) produces a selection bias. To overcome this problem, we specify both the probability to drink wine and the probatility to have a given frequency of drinking wine in a common model framework. Such a model framework is called a bivariate ordered probit model because it incorporates two qualitative dependent variables and one of them is ordered. This model enables us to disentangle between the factors that indluence the decision to drink wine from those affecting the frequencies without any selection bias. It also has the advantage of being able to test whether the decision to drink wine and the number of times are correlated or not.*

Estimating this model with French date, we are able to show that the two decisions – drinking wine or not and then how many times per week – are statistically independent. In short, we conclude that the decision process to drink wine is sequential. Variables like age, gender, area of residence, numbers of individuals in the household, number of good meals taken at home, influence both the decision to drink wine and the frequencies. The social status of the household, individuals' family status, individuals' beliefs and individuals' attitudes towards wine (measured by the parents' habits to drink wine and their attitudes regarding their children's alcohol consumption) have an impact only on the decision to drink wine. The individuals' frequencies to consume wine outside home (number of times attending bars etc...), family's habits (presence of alcoholic beverages in the household during their childhood) influence only the frequency per week to drink wine.

Key-words:
qualitative choice models, empirical analysis, wine

Les déterminants de la consommation individuelle de vin en France

Résumé – L'objet de cet article est d'étudier les déterminants de la consommation de vin en France à partir des données de l'enquête réalisée par l'ONIVINS et l'INRA en 2000. Les décisions concernant la consommation de vin et sa fréquence sont prises en compte simultanément à l'aide d'un modèle probit bivarié ordonné. Celui-ci permet de distinguer les déterminants de la décision de consommer de ceux des fréquences, tout en évitant d'éventuels biais de sélection. Les tests effectués montrent que les décisions de consommer et celles relatives aux fréquences sont statistiquement indépendantes. Le processus de décision relatif à la consommation individuelle du vin semble ainsi être séquentiel. Nous montrons aussi que si certaines variables influencent les décisions de consommer, elles n'agissent pas forcément sur les fréquences. Ainsi, l'âge, le sexe, la région géographique de consommation, le nombre d'individus du foyer et la fréquence des bons repas au foyer influencent, à la fois, les décisions de consommation et les décisions de fréquence. Par ailleurs, le statut social du foyer, la situation matrimoniale et les croyances des individus, les habitudes parentales et leurs attitudes par rapport à la consommation d'alcool de leurs enfants ont un impact uniquement sur les décisions de consommer du vin. Enfin, les fréquences de consommation hors du foyer (restauration collective, bars, etc.) et les habitudes familiales (présence de boissons alcoolisées dans le foyer pendant l'enfance) n'influencent que les fréquences de consommation.

Mots-clés:
modèles de choix qualitatifs, analyse empirique, vin

* *INRA-ESR UMR-MOISA, 2 place Viala, 34060 Montpellier cedex 1.*
e-mail : ayouz@ensam.inra.fr, cibenel@ensam.inra.fr, laporte@ensam.inra.fr

Les auteurs remercient P. Combris, M. Fares, Y. Léon, et Y. Surry pour leur aide et leurs commentaires. Les erreurs qui resteraient dans le texte relèvent de la seule responsabilité des auteurs.

L'OBJET de cette recherche est d'analyser les déterminants qui influencent simultanément les décisions et les fréquences de consommation de vin. Ce travail est conduit sur des données individuelles françaises. L'analyse de la consommation de vin est fondamentale du fait de l'importance économique et culturelle de ce bien en France, pays du plus grand marché de consommation (environ 16% de la consommation mondiale en 1999, selon les estimations de l'OIV) et l'un des plus gros producteurs (environ 20% de la production) et exportateur (environ 25% des volumes et 40% en valeur des échanges mondiaux). Or, la consommation française moyenne par adulte (plus de 14 ans) est structurellement décroissante (160 litres dans les années 1960 contre environ 70 litres en 2000). Depuis le milieu des années 1970, cette baisse (concernant seulement les vins de table), qui n'est plus compensée par la croissance démographique, a des répercussions sur l'équilibre global du marché. Ainsi, les mutations structurelles qui affectent l'économie viti-vinicole française depuis un quart de siècle sont, dans une large mesure, commandées par les évolutions de cette demande intérieure.

La recherche de l'explication des évolutions longues de la consommation a d'abord fait l'objet d'études sur des séries longitudinales de la demande agrégée (Milhau, 1935, Badouin, 1953, Fouquet, 1973, Terraza, 1981). Ces divers travaux révèlent que la consommation de vin de table est inélastique par rapport à son propre prix et qu'elle baisse lorsque le revenu agrégé augmente. A l'opposé, les vins de qualité (AOC-VDQS) apparaissent comme des biens de luxe. Au fur et à mesure de l'augmentation de leur niveau de vie, les ménages substituent des vins de qualité au vin de table.

Une deuxième vague d'études examine les comportements des consommateurs sur des données individuelles et pose comme hypothèse que l'évolution observée de la demande renvoie plus fondamentalement à une modification des modes de vie. L'hétérogénéité des comportements est notamment expliquée par les différences constatées dans les caractéristiques socio-démographiques des individus. Ainsi, depuis 1980, plusieurs travaux ont analysé l'évolution de la consommation de vin à partir d'enquêtes réalisées auprès d'échantillons représentatifs et mises en place par l'ONIVINS et l'INRA. Ces recherches montrent que la fréquence de consommation est principalement expliquée par les variables socio-démographiques (sexe, âge dans ses dimensions cycle de vie et génération, type d'activité) et d'environnement économique et culturel (région de consommation, statut social du foyer) (Boulet *et al.*, 1997). Enfin, les études économétriques les plus récentes sur la demande de vin en France estiment des courbes d'Engel[1] (Boizot, 1999 ; Seyte et Terraza, 2000).

[1] Dans d'autres pays, certains auteurs cherchent à tester si la publicité et la fiscalité ont des effets sur la consommation de vin (Saffer et Chaloupka, 1994; McGuiness, 1983 ; Moulton, 1995).

De tous ces travaux récents, on peut retenir les faits empiriques suivants. En utilisant les données SECODIP sur les consommations des ménages[2], Boizot (1999) a montré avec un modèle probit binaire que l'âge du chef du ménage et l'équipement du foyer (avoir une voiture, une cave) augmentent la probabilité d'acheter du vin. Le même auteur montre également avec un modèle *Quadratic Almost Ideal Demand System (QUAIDS)*[3] qu'être cadre, vivre en couple sans enfant, être une femme seule, ou encore habiter certaines régions de France augmentaient la part du vin dans la dépense totale de consommation. Le vin apparaît comme un bien supérieur et l'estimation des élasticités-prix compensées montre que le vin est faiblement substituable aux autres types de boissons. En utilisant les données longitudinales de l'INSEE (1959-1995), Seyte et Terraza (2000) estiment un modèle *AIDS* et montrent, contrairement à Boizot, que le vin de consommation courante est un bien nécessaire, complémentaire de la bière et du vin de qualité. A l'opposé, le vin de qualité apparaît comme un bien supérieur. Les mêmes auteurs montrent que la consommation de vin de table et la consommation de vin de qualité sont inélastiques par rapport à leur propre prix.

Pour l'Italie, principal concurrent européen de la France, Conforti *et al.* (2000) estiment, comme Boizot, une élasticité-revenu positive mais faible. En revanche, ces auteurs observent la même rigidité au prix que Seyte et Terraza. Cette rigidité est aussi observée dans les études américaines (Leung et Phelps, 1993; Nelson, 1997; Gao *et al.*, 1995; Yen, 1995) et australiennes (Crowley et Richardson, 1997) [4].

Notons aussi que des modèles dynamiques de fréquence d'achat peuvent être une voie d'investigation économétrique pour l'analyse de la consommation de vin. Ces modèles ont été développés au cours des dernières années. Ils s'apparentent aux approches des modèles de durée, et prennent en compte l'hétérogénéité non observée des acheteurs, les coûts de stockage des biens et les coûts fixes liés aux fréquences d'achat (Boizot *et al.*, 2001). Certaines versions de ces modèles permettent de prédire le moment où un acheteur régulier devient un acheteur occasionnel (Allenby *et al.*, 1999). Cependant, ces études ne se sont pas intéressées à la consommation de vin.

Notre travail se situe dans la continuité des recherches réalisées antérieurement sur les données ONIVINS-INRA. Ces dernières portent sur les comportements individuels de consommation et non d'achat du foyer

[2] Ces données sont obtenues par empilement trimestriel des relevés hebdomadaires des années 1991 à 1994 et des relevés quotidiens de 1996.

[3] Ce modèle, développé par Banks *et al.* (1997), est une extension du modèle *AIDS* de Deaton et Muellbauer (1980) auquel est rajouté un terme de dépense quadratique.

[4] Des recherches américaine et australienne plus anciennes donnaient des élasticités de la demande de vin par rapport au prix supérieures à 1 en valeur absolue (Ornstein et Levy, 1983; Selvanathan, 1991).

comme c'est le cas dans plusieurs études pré-citées[5]. En effet, les données que nous utilisons mesurent les **fréquences individuelles** de consommation et non les achats des foyers. Sur le plan des méthodes utilisées, les travaux antérieurs modélisent les décisions d'achat du vin indépendamment des volumes achetés: ainsi, un modèle probit binaire simple est généralement utilisé pour estimer la première variable (décision d'achat) alors que la seconde (niveau d'achat) dépend des dépenses de consommation de biens alimentaires dans le cadre d'un système d'équations.

Dans ce qui suit, les **décisions de consommer** et les niveaux de consommation observés en terme de **fréquences** sont supposés se réaliser simultanément. Ainsi, en plus des spécifications classiques (probit binaire et ordonné), un modèle probit bivarié ordonné est présenté dans cet article. Il repose sur l'idée qu'un individu décide simultanément s'il consomme ou non et la fréquence de sa consommation, et permet ainsi de distinguer les déterminants de la décision de consommer des facteurs affectant les fréquences de consommation, tout en évitant l'éventuel biais de sélection. Il permet aussi de savoir si les deux décisions sont corrélées ou indépendantes (processus de décision séquentiel). Les tests statistiques indiquent si les variables qui expliquent les fréquences (respectivement la décision) réagissent avec le même signe et la même significativité sur les décisions de consommation (respectivement sur les fréquences). Enfin, les données utilisées sont riches d'informations généralement peu disponibles sur les croyances des individus et leurs préférences (mode de vie, système de valeur) et leur passé (éducation à la consommation de vin). Nous pouvons donc tester de nouvelles variables qui n'ont pas été prises en compte dans les travaux antérieurs.

Le plan de l'article est le suivant: les données utilisées sont décrites en première partie; dans la deuxième partie, nous présentons les modèles économétriques proposés et enfin, avant de conclure, nous discutons les résultats des modèles estimés.

DESCRIPTION DES DONNÉES UTILISÉES

La plupart des statistiques disponibles pour étudier la consommation alimentaire en France et plus spécifiquement celle du vin (série INSEE, panels SECODIP) enregistrent les achats des ménages. Les enquêtes

[5] Dans les études économétriques de la consommation, ce sont les dépenses des ménages qui sont estimées. Pour un bien comme le vin, les dépenses observées peuvent sous-estimer ou sur-estimer les dépenses de consommation réelles du foyer. Le vin n'est pas consommé uniquement au foyer mais aussi dans la restauration, ou dans d'autres lieux. De plus, l'achat n'implique pas nécessairement consommation, puisque ce bien peut être offert, stocké et consommé plus ou moins longtemps après l'achat.

mises en place par l'INRA et l'ONIVINS s'intéressent en revanche aux comportements individuels de consommation de vin et mesurent des fréquences individuelles de consommation. De plus, la richesse de cette base de données permet de tester un certain nombre de variables qui ne sont généralement pas prises en compte dans les études de la demande. Certaines d'entre elles permettent notamment de contrôler partiellement l'hétérogénéité des individus (système de valeur des individus, croyances, contexte initiatique de consommation). Ces enquêtes, menées depuis 1980, portent sur des échantillons représentatifs de la population française (environ 4000 personnes de 15 ans et plus, choisies selon la méthode des quotas). Nous utilisons ici uniquement la coupe de l'année 2000, la plus riche en variables explicatives, qui comprend 4010 observations (voir annexe 1 pour plus de détails).

Variables à expliquer

Dans ce travail, deux variables endogènes sont prises en compte: une variable discrète ordonnée, la fréquence de consommation de vin (*FREQ*), et une variable de décision de consommation d ($d = 1$ si $FREQ > 0$, sinon $d = 0$). La variable *FREQ* comprend cinq modalités, précisées dans le tableau 1, qui opposent les individus déclarant ne jamais consommer de vin à ceux déclarant en boire quotidiennement. Par regroupement de ces modalités, on construit une typologie en trois catégories: les consommateurs réguliers (*FREQ* = 3 ou *FREQ* = 4), les consommateurs occasionnels (*FREQ* = 1 ou *FREQ* = 2) et les non-consommateurs (*FREQ* = 0). Cette typologie est utilisée pour interpréter les résultats des modèles.

Tableau 1. **Évolution de la distribution de la population française selon les fréquences de consommation de vin et typologie des consommateurs en pourcentages (données redressées)**

Déclarent consommer du vin :		1990	1995	1998	2000
Tous les jours	(*FREQ* = 4)	25,4	23,6	22,6	19,5
Presque tous les jours	(*FREQ* = 3)	4,8	4,3	4,7	4,3
1 à 2 fois par semaine	(*FREQ* = 2)	12,1	17,0	18,8	17,5
Plus rarement	(*FREQ* = 1)	24,8	24,2	23,0	25,1
Jamais	(*FREQ* = 0)	32,9	30,9	30,9	33,6
Typologie des consommateurs					
Consommateurs réguliers	(*FREQ* = 3 ou *FREQ* = 4)	30,2	27,9	27,3	23,8
Consommateurs occasionnels	(*FREQ* = 1 ou *FREQ* = 2)	36,9	41,2	41,8	42,6
Non-consommateurs	(*FREQ* = 0)	32,9	30,9	30,9	33,6

Source: enquêtes ONIVINS-INRA

Variables explicatives

Les déterminants de la consommation individuelle ont été recherchés à la fois dans les caractéristiques des individus et dans celles du foyer auquel ils appartiennent. L'ensemble des variables explicatives retenues ainsi que leurs définitions et modalités sont présentés au tableau 2. Celles-ci sont maintenant expliquées plus en détail.

Les variables socio-démographiques

La consommation du vin est soumise à des contraintes biologiques liées à la fois à l'âge et au sexe. Selon Boulet et Laporte (1997), la consommation de vin est influencée par un effet « cycle de vie » et un effet de génération. Il n'est pas possible de prendre en compte cet aspect à partir de données transversales. Nous avons introduit dans le modèle la variable « âge » sous forme quadratique afin de tester l'effet non-linéaire de l'âge sur les décisions des individus, toutes générations confondues.

Les autres variables socio-démographiques explicatives sont le sexe, la situation familiale et le nombre de personnes au foyer.

Les variables de « style de vie »

Le vin est à la fois un aliment et un bien culturel, et sa consommation peut être associée à un mode de vie. L'effet du mode de vie du foyer sur le comportement individuel de consommation a été approché par l'étude des pratiques des repas conviviaux au domicile (variable *SOCF*). Une caractéristique du mode de vie personnel a également été testée par la fréquentation de certains lieux de consommation hors domicile *(THD)*. La variable *SOCF*, indicatrice du niveau de « sociabilité » du foyer, a été construite à partir de la fréquence des bons repas en famille ou avec des invités au domicile. Pour l'estimation, nous distinguons les foyers de « faible sociabilité » des autres.

La variable *THD* a été construite à partir du nombre de jours de fréquentation de divers lieux de consommation hors domicile. Les modalités de construction de cette variable sont précisées en annexe 2. Elle permet de distinguer les « sédentaires » qui consomment peu hors de chez eux, les « nomades » qui sont à l'extérieur presque tous les jours et plus particulièrement au restaurant, les « habitués de bar », les « convives » qui fréquentent les foyers amis, et les « rationnaires » qui mangent régulièrement sur le lieu de travail. Les goûts personnels et l'activité professionnelle peuvent conduire les individus à un certain niveau de consommation hors domicile. Les individus sont regroupés en quatre catégories : les « sédentaires » et « rationnaires », les « convives », les « habitués de bars » et les « nomades ». Si nous pouvons nous attendre à ce que les individus appartenant à des foyers « conviviaux » soient plus consomma-

teurs de vin, nous ne pouvons rien anticiper sur l'effet des fréquentations hors domicile.

Le système de valeurs des individus

Pour capter l'effet des croyances et de la subjectivité des individus sur les décisions de consommation de vin, nous avons introduit une dimension explicative particulière, le « système de valeurs ». Il est utilisé par les spécialistes du marketing dans leurs travaux sur la psychologie du consommateur. La « valeur » peut être considérée comme le principe ultime qui permet aux individus de donner une cohérence à l'ensemble de leurs comportements (d'Hauteville, 2001). Ainsi, les personnes interviewées étaient appelées à se situer en retenant l'une des quatre propositions suivantes : avoir de bonnes relations avec les autres (*VALEUR* = 1), profiter de l'existence et se réaliser pleinement (*VALEUR* = 2), se sentir en sécurité (*VALEUR* = 3), avoir le respect de soi (*VALEUR* = 4). Les réponses à ces questions permettent de révéler les préférences des individus parmi quatre valeurs : l'insertion sociale (*VALEUR* = 1), l'hédonisme (*VALEUR* = 2), la sécurité (*VALEUR* = 3), et la dignité – respectabilité (*VALEUR* = 4). De part la nature particulière du vin, nous pouvons nous attendre à ce que les personnes caractérisées par les valeurs d'insertion sociale ou hédoniques soient significativement plus régulièrement consommatrices de cette boisson.

La perception « vin et santé »

Les préoccupations des individus face au risque alimentaire peuvent influencer les comportements de consommation. Pour tester l'influence possible des croyances quant au risque vin, nous introduisons une variable *RISQUE,* qui permet de classer les individus selon leur perception des risques ou bienfaits liés à une consommation modérée de vin. Cette variable a été construite à partir de leurs réponses aux deux questions suivantes :

a) à votre avis, une consommation modérée de vin peut-elle prévenir l'apparition de certaines maladies ?

b) à votre avis, une consommation même modérée de vin peut-elle favoriser l'apparition de certaines maladies ?

Ainsi, on peut opposer les personnes pour qui le vin a un effet préventif de certaines maladies (*RISQUE* = 1) à celles qui le considèrent comme risqué (*RISQUE* = 5) ; les autres personnes sont réparties entre celles pour qui le vin peut prévenir l'apparition de certaines maladies mais en favoriser d'autres (*RISQUE* = 2), celles qui pensent qu'il n'a pas d'effet particulier (aliment comme un autre, *RISQUE* = 3) et, enfin, celles qui n'ont pas d'opinion sur la question (*RISQUE* = 4). Nous pouvons supposer que les personnes qui considèrent la consommation de vin comme comportant toujours un risque soient les moins consommatrices.

Tableau 2. **Définition et modalités des variables explicatives**

Désignation des variables	Valeurs ou modalités prises par les variables explicatives
Variables socio-démographiques	
- *SEXE*: Sexe	= 1 pour sexe masculin
	= 0 pour sexe féminin
- *AGE*: Age	Valeurs continues positives
- *AGECARRE*: Age x Age	Valeurs continues positives
- *SIT*: Situation familiale	= 0 si l'individu est en situation maritale
	= 1 si l'individu n'est pas en situation maritale
- *NPF*: Nombre de personnes au foyer	= 1 si personne seule
	= 2 deux personnes
	= 3 trois personnes
	= 4 quatre personnes
	= 5 cinq personnes et plus
Variables de « style de vie »	
- *SOCF*: Niveau de « sociabilité » du foyer	= 1 faible sociabilité
	= 0 autres
- *THD*: Fréquentation des divers lieux de consommation	= 1 « sédentaires » ou « rationnaires »
	= 2 « convives »
	= 3 « habitués de bars »
	= 4 « nomades »
Variable « Système de valeur des individus »	
- *VALEUR*: Système de valeurs	= 1 insertion sociale
	= 2 hédonisme
	= 3 sécurité
	= 4 dignité-respectabilité
Variable « Perception vin et santé »	
- *RISQUE*:	= 1 vin préventif de certaines maladies
	= 2 vin préventif de certaines maladies et risqué pour d'autres
	= 3 sans effet
	= 4 pas d'opinion
	= 5 plutôt risqué
Variables de localisation géographique	
- *REC*: Régions françaises	= 1 Ouest
	= 2 Nord
	= 3 Est
	= 4 Nord-Ouest
	= 5 Sud-Est
	= 6 Centre
	= 7 Sud
	= 8 Ile de France
- *VILLEENF*: « Mobilité » géographique	= 1 ville enfance
	= 0 autre
Variables de contexte initiatique	
- *TFBOYA*: Présence de vins et autres boissons alcoolisées dans le foyer de leur enfance	= 0 jamais de boisson alcoolisée dans le foyer
	= 1 vin présent quotidiennement
	= 2 autres boissons alcoolisés présentes quotidiennement
	= 3 vin présent de façon hebdomadaire
	= 4 autres boissons présentes de façon hebdomadaire
	= 5 boissons alcoolisées rarement présentes
- *INTERDIT*: Attitude des parents vis-à-vis de la consommation de boissons alcoolisées interdites au moment de l'enfance	= 1 aucun interdit
	= 2 boissons alcoolisées autres que le vin interdites
	= 3 toutes boissons alcoolisées interdites
Variables « niveau de vie »	
- *STUT*: Niveau de vie du foyer	= 1 très modeste
	= 2 modeste
	= 3 classe moyenne inférieure
	= 4 classe moyenne supérieure
	= 5 foyers aisés
- *PLOGT*: Situation patrimoniale:	= 1 propriétaire de son logement
	= 0 autre

La localisation géographique

Malgré une certaine tendance à l'homogénéisation de la consommation favorisée, en particulier, par la publicité, la mobilité et le développement de la grande distribution, on peut faire l'hypothèse que la demande de vin, qui reste un produit agricole peu transformé et lié à des zones géographiquement délimitées, reste attachée à des pratiques alimentaires régionales encore différenciées. L'existence de boissons locales traditionnelles (bière, cidre), les facteurs climatiques et le degré d'urbanisation sont des éléments qui contribuent à maintenir cette différenciation. Ces facteurs de localisation géographique sont pris en compte par deux variables. La première d'entre elles (dénotée *REC* et dont les modalités sont présentées en annexe 2) définit huit régions de consommation de vin relativement homogènes. La deuxième variable *(VILLEENF)* précise si l'individu habite toujours sa ville d'origine, indicatrice de «(im)mobilité» géographique. On cherche ainsi à capter l'effet de l'ancrage des individus à leur territoire.

Les variables du contexte initiatique

Ces variables caractérisent les comportements de consommation et l'attitude à l'égard des boissons alcoolisées de l'environnement familial lors de l'enfance. Elles nous permettent de tester l'impact des pratiques alimentaires familiales dans la construction de celles des futurs adultes. Deux types de variables sont utilisés pour représenter ce phénomène: i) boissons alcoolisées habituellement présentes dans le foyer d'enfance *(TFOYBA)* et ii) attitude de l'entourage à l'égard de la consommation des individus au moment de l'enfance *(INTERDIT)*. La première variable répartit les individus selon la présence de vin et de boissons alcoolisées dans le foyer de leur enfance, tandis que la seconde classe les individus selon qu'ils ont ou n'ont pas connu d'interdiction de consommation des boissons alcoolisées lors de leur enfance. Intuitivement, nous pouvons supposer que les individus qui n'ont pas connu d'interdiction et qui avaient du vin dans leur foyer d'enfance soient les plus consommateurs.

Le niveau de vie

Dans les travaux économiques sur la consommation, la variable la plus utilisée pour capter le niveau de vie du foyer est le revenu du foyer. Dans notre enquête, cette donnée est manquante pour environ un quart des observations. Dans ce type de situation, on peut estimer le revenu sur le sous-échantillon des individus qui l'ont déclaré, puis prédire les revenus pour les observations manquantes après correction d'un éventuel biais de sélection. Par manque d'information sur les déterminants du revenu, les estimations que nous avons effectuées ne permettent pas de mener à bien cette stratégie d'estimation[(6)]. Nous avons utilisé à la place une échelle de

[(6)] Nous avons appliqué la méthode de Pascal et Robin (1995) qui permet d'utiliser les prédictions d'une variable observée en tranches.

« statut social » (variable *STUT*). L'utilisation de cette échelle va au-delà du revenu et son emploi se justifie même en l'absence de problèmes de données manquantes; en effet, la variable *STUT* intègre simultanément, outre le revenu, des éléments de statut professionnel et de train de vie du foyer (la méthode de construction de *STUT* est exposée en annexe 2); elle permet ainsi de ne pas réduire le « statut social » du ménage au revenu ou à la catégorie socio-professionnelle du chef de famille. Les foyers sont ainsi répartis en cinq classes, depuis les foyers « très modestes » jusqu'aux foyers « aisés ».

Un élément de la situation patrimoniale de l'individu est aussi testé et concerne la propriété du logement (variable *PLOGT*). Selon les travaux antérieurs, nous nous attendons à ce que les foyers aisés soient les plus consommateurs.

MODÈLES EMPIRIQUES PROPOSÉS

Le travail empirique consiste en l'estimation de trois modèles et la réalisation de tests d'hypothèses. Dans une première phase, les décisions de consommation sont représentées par un modèle probit binaire simple (Modèle I). Dans un deuxième temps, les décisions concernant les fréquences sont analysées à partir d'un modèle probit ordonné avec des seuils inconnus (dénoté Modèle II). Enfin, nous modélisons les décisions de consommation et les décisions de fréquences simultanément à l'aide d'un modèle probit bivarié ordonné (Modèle III).

La décision de consommer du vin

Pour expliquer la décision de consommer ou non du vin, nous pouvons considérer qu'un individu compare les utilités obtenues en consommant ou non du vin. Supposons que, pour chaque individu (repéré avec l'indice *i)* de l'échantillon de taille n, cette différence d'utilité entre chaque alternative est une variable latente U_i^*. Admettons aussi que cette variable latente est une combinaison linéaire d'un certain nombre de variables explicatives :

$$U_i^* = \boldsymbol{\alpha}' \boldsymbol{X_i} + \mu_i \qquad \text{pour } i = 1,\ldots n \tag{1}$$

où $\boldsymbol{\alpha}$ est un vecteur de paramètres, μ_i désigne un terme d'erreur de moyenne nulle et de variance $\sigma^2_{\mu i}$, et $\boldsymbol{X_i}$ est l'ensemble des variables qui influencent les préférences des individus i. Nous pouvons définir pour chaque individu i, une variable binaire d_i qui vaut 1 quand l'individu consomme du vin, sinon $d_i = 0$. La probabilité de consommer du vin et son complémentaire sont alors respectivement égaux à

$$P(d_i = 1 \mid \boldsymbol{X_i}) = E(d_i \mid \boldsymbol{X_i}) = P(U_i^* > 0 \mid \boldsymbol{X_i}) = P(\mu_i > \boldsymbol{\alpha' X_i} \mid \boldsymbol{X_i}) = 1 - F\mu(-\boldsymbol{\alpha' X_i} \mid \sigma_{\mu i}) = \int_{-\infty}^{\boldsymbol{\alpha' X_i}/\sigma_{\mu i}} f_\mu(k)dk \quad (2)$$

et
$$P(d_i = 0 \mid \boldsymbol{X_i}) = F_\mu(-\boldsymbol{\alpha' X_i} \mid \sigma_{\mu i}) \quad (3)$$

où $E(.)$ représente l'opérateur « l'espérance mathématique », $F_\mu(.)$ désigne la fonction de répartition et $f_\mu(.)$ la fonction de densité de la loi normale centrée réduite quand il s'agit d'un probit.

Quand les termes d'erreur μ_i ont le même écart-type σ_μ et qu'ils sont indépendants des variables explicatives, l'estimation correcte des paramètres $\boldsymbol{\alpha}$ est réalisée par la maximisation du logarithme de la fonction de vraisemblance suivante :

$$\max_{\boldsymbol{\alpha}} \sum_{i=1}^{n} LogL(d_i; \boldsymbol{X}, \boldsymbol{\alpha}, \sigma_\mu = 1) = \max_{\boldsymbol{\alpha}} \sum_{i=1}^{n} \left[P(d_i = 1 \mid \boldsymbol{X_i}) \right]^{d_i} \left[P(d_i = 0 \mid \boldsymbol{X_i}) \right]^{1-d_i} \quad (4)$$

Comme les variables d_i et U_i^* sont respectivement ordonnées et non-observables, nous posons $\sigma_\mu = 1$ pour rendre le modèle identifiable. Les erreurs μ_i suivent alors une loi normale centrée réduite. Ainsi, le modèle probit permet l'identification des paramètres en opposant seulement les consommateurs et les non-consommateurs.

Détermination des fréquences de consommation de vin

Pour analyser les fréquences de consommation, la voie naturelle est d'utiliser un modèle probit ordonné. La variable à expliquer *FREQ* décrit les fréquences de consommation (en cinq modalités $j = 0,...,4$). L'individu i consomme du vin si $FREQ_i > 0$, sinon $FREQ_i = 0$. Les probit ordonnés reposent sur l'idée qu'il existe une variable latente Q_i^* continue, déterminée par les caractéristiques observées $\boldsymbol{Z_i}$ et les caractéristiques aléatoires ε_i :

$$Q_i^* = \boldsymbol{\beta' Z_i} + \varepsilon_i \quad \text{pour } i = 1,\ldots,n \quad (5)$$

ε_i est un terme stochastique de moyenne nulle et d'écart-type σ_{ε_i}, $\boldsymbol{\beta}$ un vecteur de paramètres qui contient un terme constant. Quand la variable latente Q_i^* est dans des intervalles délimités par des seuils inconnus s_j, une fréquence particulière j de consommation est observée :

$$FREQ_i = j \quad \text{si} \quad s_{j-1} < Q_i^* \leq s_j \quad \text{pour } j = 0,...,4 \quad (6)$$

Pour ce probit ordonné simple (modèle II), la première modalité $j = 0$ est la fréquence nulle (l'individu ne consomme pas). Les probabilités p_{ij} d'avoir une fréquence particulière $j = 0,...,4$ de consommation pour l'individu i sont :

$$p_{ij} = P\left[FREQ_i = j\right] = P\left[s_{j-1} < Q_i^* \leq s_j \,|\, \boldsymbol{Z_i}\right] = P\left[\frac{s_{j-1}}{\sigma_{\varepsilon i}} < \frac{\boldsymbol{\beta' Z_i} + \varepsilon_i}{\sigma_{\varepsilon i}} \leq \frac{s_j}{\sigma_{\varepsilon i}}\right] \quad (7)$$

$$= F_\varepsilon\left[\frac{s_j - \boldsymbol{\beta' Z_i}}{\sigma_{\varepsilon i}}\right] - F_\varepsilon\left[\frac{s_{j-1} - \boldsymbol{\beta' Z_i}}{\sigma_{\varepsilon i}}\right]$$

avec $F_\varepsilon(.)$ la fonction de répartition de la loi normale, s_j les seuils ordonnés $s_0<s_1<...<s_3$ inconnus estimés avec les paramètres $(\boldsymbol{\beta}, \sigma_\varepsilon)$ du modèle. On pose $s_{-1} = -\infty$,$s_4 = +\infty$ pour que la somme des probabilités sur tous les intervalles soit égale à 1. Comme le vecteur des paramètres $\boldsymbol{\beta}$ contient une constante, un des seuils s_j n'est pas identifiable. Nous procédons donc à la normalisation suivante: $s_0 = 0$. Quand les termes d'erreur sont homoscédastiques et indépendants statistiquement des variables explicatives, l'estimation correcte des paramètres $\{s, \boldsymbol{\beta}, \sigma_\varepsilon\}$ du modèle probit ordonné simple (modèle II) est obtenue par la maximisation du logarithme de la fonction de vraisemblance de l'échantillon:

$$LogL = \sum_{i=1}^{n}\sum_{j=0}^{4} d_{ij} * Log\left[\frac{F_\varepsilon(s_{j+1} - \boldsymbol{\beta' Z_i})}{\sigma_\varepsilon} - \frac{F_\varepsilon(s_j - \boldsymbol{\beta' Z_i})}{\sigma_\varepsilon}\right] \text{ pour } j = 0,\ldots,4 \quad (8)$$

avec d_{ij} une variable pour chaque individu i de l'échantillon, qui vaut 1 lorsque $FREQ_i$ est égale à une fréquence j particulière. Pour identifier le modèle, nous normalisons les paramètres en supposant que $\sigma_\varepsilon = 1$. Les paramètres estimés (s et $\boldsymbol{\beta}$) n'ont alors plus d'interprétation absolue et fournissent seulement une échelle d'intensité de consommation[7].

Avec ce modèle II, la probabilité de consommer du vin et celle de ne pas en consommer sont respectivement égales à

$$p_{i1...4} = \sum_{j=1}^{4} P\left[FREQ_i = j\right]$$

et

$$p_{i0} = P\left[FREQ_i = 0\right] = F_\varepsilon\left[(-\boldsymbol{\beta' Z_i})/\sigma_{\varepsilon i}\right]$$

La seule différence entre les modèles I et II réside dans le fait que l'information disponible sur la distribution des fréquences positives de consommation ($FREQ_i = 1,\ldots,4$) est maintenant prise en compte dans la procédure d'estimation.

Le modèle II possède toutefois une faiblesse notable qui est liée au fait que les paramètres $\boldsymbol{\beta}$ qui expliquent les fréquences positives ($FREQ_i > 0$) sont identiques à ceux qui interviennent dans la décision de

[7] Nous verrons par la suite comment calculer les effets marginaux pour avoir une lecture quantitative des résultats.

consommation ($FREQ = 0$ ou $FREQ_i > 0$). Seules les constantes du modèle changent avec les valeurs des seuils s_j. En d'autres termes, les variables explicatives Z_i agissent sur la variable latente Q_i^* avec la même intensité et avec le même signe quel que soit le niveau de la fréquence. Un individu i passe de la situation de non-consommateur ($FREQ_i = 0$) à celle de consommateur ($FREQ_i > 0$) quand les variations des variables explicatives Z_i entraînent le dépassement du seuil s_0 par la variable latente Q_i^*.

Or il n'est pas sûr que les déterminants de la décision de consommer et des fréquences soient les mêmes. Il n'est pas sûr non plus que ces déterminants agissent sur ces deux types de décisions selon la même intensité et le même signe. Un certain nombre de travaux ont montré que la décision de consommation pouvait se modéliser comme un processus en deux étapes: décider de consommer, puis choisir le volume de consommation. Par exemple, Haines *et al.*, (1988), et Blaylock et Blisard (1992) ont présenté des méthodes pour tester la corrélation entre les décisions de consommation et les volumes consommés, ainsi que l'existence de déterminants communs. Mais ces méthodes ne sont applicables que si la variable à expliquer est continue, ce qui n'est pas le cas ici. Pour analyser les facteurs qui influencent les fréquences de consommation indépendamment de ceux qui affectent la décision de consommation, une stratégie consisterait à estimer le modèle II sur le sous-échantillon des individus qui consomment du vin ($FREQ_i = 1,...,4$). Cependant, une telle approche induit un biais de sélection qui nécessite un traitement différent de celui proposé par Heckman (1979). Pour remédier à ce dernier problème, il est nécessaire de modéliser les décisions et les fréquences de consommation simultanément.

Détermination simultanée des décisions et des fréquences de consommation

Dans ce qui suit, un modèle probit bivarié ordonné (modèle III) est présenté, s'inspirant du travail de Hall *et al.* (2002). Ces économistes ont développé une méthode qui évite le biais de sélection lors de l'estimation d'un modèle probit ordonné sur un échantillon non aléatoire. Appliqué à notre étude, ce cadre de travail permet de modéliser simultanément la décision de consommer et le choix de fréquence de consommation. L'objectif est alors double: i) tester la dépendance des deux types de décision, et ii) tester si celles-ci répondent aux mêmes déterminants. Pour décrire ce dernier modèle, nous présentons en premier lieu le cas général où les décisions de consommation et de fréquence sont corrélées. Puis, nous développons le cas particulier où les décisions sont indépendantes.

Cas où les décisions sont corrélées

Le modèle III repose sur une représentation simultanée de deux variables latentes. En utilisant les deux spécifications précédentes – expressions (1) et (5) – le modèle III est décrit comme suit :

$$FREQ_{ij} = j \text{ si } s_{j-1} < Q_i^* = \boldsymbol{\beta}' \mathbf{Z}_i + \varepsilon_i \leq s_j \text{ et } U_i^* = \boldsymbol{\alpha}' \mathbf{X}_i + \mu_i > 0 \quad \text{pour } j = 1,..,4 \tag{9}$$

et

$$FREQ_{i0} = 0 \text{ si } U_i^* = \boldsymbol{\alpha}' \mathbf{X}_i + \mu_i \leq 0 \tag{10}$$

Selon (9), nous observons une fréquence de consommation positive $j > 0$ si la variable latente Q_i^* est située dans un intervalle particulier $]s_{j-1}, s_j]$ et si simultanément l'individu a intérêt à consommer du vin, $U_i^* > 0$. Dans le cas contraire, l'individu ne consomme pas de vin, $FREQ_i = 0$ (ou $d_i = 0$). Nous supposons que les termes d'erreur (ε_i, μ_i) suivent une distribution normale bivariée de moyenne nulle avec une matrice de variances-covariances $\boldsymbol{\Omega}$:

$$\boldsymbol{\Omega} = \begin{Bmatrix} \sigma_\varepsilon & \rho\sigma_\varepsilon\sigma_\mu \\ \rho\sigma_\varepsilon\sigma_\mu & \sigma_\mu \end{Bmatrix} \tag{11}$$

avec ρ le coefficient de corrélation entre les termes d'erreurs (ε_i, μ_i). La probabilité d'avoir une fréquence positive, $P(FREQ_i > 0)$, de consommation et la probabilité de ne pas consommer ($P[FREQ_i = 0]$) sont respectivement égales à

$$p_{ij} = P\left[FREQ_{ij} = j \,\middle|\, \mathbf{Z}_i, \mathbf{X}_i\right] = P\left[s_{j-1} < Q_i^* \leq s_j \text{ (et) } U^* > 0\right] \tag{12}$$

$$= P(\varepsilon_i \leq s_j - \boldsymbol{\beta}' \mathbf{Z}_i,\ \mu_i > -\boldsymbol{\alpha}' \mathbf{X}_i) - P(\varepsilon_i \leq s_{j-1} - \boldsymbol{\beta}' \mathbf{Z}_i,\ \mu_i > \boldsymbol{\alpha}' \mathbf{X}_i)$$

$$= F\left(\frac{s - \boldsymbol{\beta}' \mathbf{Z}_i}{\sigma_\varepsilon}, \frac{\boldsymbol{\alpha}' \mathbf{X}_i}{\sigma_\mu}, \rho\right) - F\left(\frac{s_{j-1} - \boldsymbol{\beta}' \mathbf{Z}_i}{\sigma_\varepsilon}, \frac{\boldsymbol{\alpha}' \mathbf{X}_i}{\sigma_\mu}, \rho\right)$$

et

$$p_{i0} = P\left[FREQ_{i1} = 0 \,\middle|\, \mathbf{Z}_i, \mathbf{X}_i\right] = P\left[\mu_i \leq -\boldsymbol{\alpha}' \mathbf{X}_i\right] = F\left(\frac{-\boldsymbol{\alpha}' \mathbf{X}_i}{\sigma_\mu}\right) \tag{13}$$

avec $F(x, y, \rho)$ étant la fonction de répartition normale bivariée évaluée aux points de coordonnées $\{s_j - \boldsymbol{\beta}' \mathbf{Z}_i / \sigma_\varepsilon, \boldsymbol{\alpha}_i' / \sigma_\mu\}$, et $F(x)$ désignant la fonction de répartition univariée évaluée au point $\boldsymbol{\alpha}' \mathbf{X}_i / \sigma_\mu$. Nous pouvons ainsi constater que la relation (13) est équivalente à l'expression (3). Par ailleurs, pour permettre l'identification du modèle, nous posons $\sigma_\varepsilon = \sigma_\mu = 1$. Le logarithme de la fonction de vraisemblance du modèle III à maximiser est alors :

$$LogL(\boldsymbol{\beta}, \boldsymbol{\alpha}, s; \rho) = \sum_{i=1}^{4} \left(1_{(d_i=0)} \times LogF(-\boldsymbol{\alpha' X_i}) + \sum_{j=1}^{4} 1_{(d_{ij}=j)} \times Log\left[F(s_j - \boldsymbol{\beta' Z_i}, \boldsymbol{\alpha' X_i}, \rho) - F(s_{j-1} - \boldsymbol{\beta' Z_i}, \rho\right] \right) \quad (14)$$

où $1_{(d_i = 0)}$ et $1_{(d_{ij} = j)}$ sont des fonctions indicatrices qui spécifient respectivement si l'individu i ne consomme pas de vin et si l'individu a une fréquence de consommation particulière j. Nous pouvons alors calculer la probabilité d'avoir une fréquence particulière sachant que l'individu a décidé de consommer du vin :

$$P(FREQ_i = j \mid d_i = 1) = \frac{F(s_j - \boldsymbol{\beta' Z_i}, \boldsymbol{\alpha' X_i}, \rho)}{F(\boldsymbol{\alpha' X_i})}$$

Un simple examen de l'expression (14) permet d'évaluer les éventuels biais qui résulteraient de l'estimation d'un probit ordonné simple (modèle II) sur le sous-échantillon des consommateurs ($FREQ_i = j > 0$).

Cas où les processus de décision sont indépendants

Si le coefficient de corrélation ρ est nul, les décisions de consommation sont indépendantes des décisions de fréquence. La relation (12) se simplifie à

$$F(s_j - \boldsymbol{\beta' Z_i}, \boldsymbol{\alpha' X_i}, \rho) - F(s_{j-1} - \boldsymbol{\beta' Z_i}, \boldsymbol{\alpha' X_i}, \rho) = F(\boldsymbol{\alpha' X_i}) \left[F(s_j - \boldsymbol{\beta' Z_i}) - F(s_{j-1} - \boldsymbol{\beta' Z_i}\right] \quad (15)$$

Le logarithme de la fonction de vraisemblance à maximiser est donc :

$$LogL\ (\boldsymbol{\beta}, \boldsymbol{\alpha}, s) = \sum_{i=1}^{n} \left(\left\{1_{(d_i=0)} \times LogF(-\boldsymbol{\alpha' X_i}) + 1_{(d_i \neq 0)} \times LogF(\boldsymbol{\alpha' X_i})\right\} + \sum_{j=1}^{4} 1_{(d_{ij} = j)} \times \left\{Log[F(s_j - \boldsymbol{\beta' Z_i}) - F(s_{j-1} - \boldsymbol{\beta' Z_i})]\right\} \right) \quad (16)$$

$LogL(\boldsymbol{\beta}, \boldsymbol{\alpha}, s)$ est désormais la somme de deux fonctions de vraisemblance, la première correspondant à un probit binaire et la seconde à un probit ordonné (modèle II). Notons également que l'estimation du modèle II (modèle probit ordonné) sur le seul sous-échantillon des consommateurs ($FREQ_i = j > 0$) induit un biais de sélection du fait de l'omission du terme $1_{(d_i = 0)} \times LogF(-\boldsymbol{\alpha' X_i}) + 1_{(d_i \neq 0)} \times LogF(-\boldsymbol{\alpha' X_i})$.

Les travaux économétriques effectués sur des données individuelles peuvent souffrir d'hétéroscédasticité. Yatchew et Griliches (1984) ont montré que lorsque les perturbations sont hétéroscédastiques, les estimateurs du maximum de vraisemblance ne sont pas convergents et les matrices des variances-covariances sont inappropriées pour réaliser les

tests usuels. Pour tenir compte de ce dernier problème, tous les écarts-types des paramètres estimés ont été obtenus en utilisant la procédure de White (1982). Toutes les estimations économétriques des modèles I, II et III ont été réalisées avec le logiciel SAS.

RÉSULTATS ÉCONOMÉTRIQUES: PRÉSENTATION ET DISCUSSION

Les résultats économétriques des probit binaire (modèle I) et ordonné (modèle II) sont présentés dans le tableau 3. Les estimations des paramètres du modèle III (modèle probit bivarié ordonné) sont rassemblées dans le tableau 4. Les trois modèles estimés sont validés à l'aide des critères d'Akaike et de Schwartz, le test du rapport de vraisemblance et le pseudo-R^2 de McFadden[8]. Dans l'ensemble, un examen des valeurs prises par ces critères montre que la représentation de la demande individuelle de vin par des modèles probit simple et ordonné est acceptable sur le plan statistique. Le coefficient de corrélation ρ entre les termes d'erreur des variables latentes (ε_i, μ_i) dans le modèle III (modèle probit bivarié ordonné) est statistiquement non significatif et de signe positif comme on pouvait s'y attendre. La valeur estimée de ρ est 0,104 avec un écart-type de 0,336 et une *p-value* de 0,775. Ainsi, selon le modèle III, les décisions de consommation et de fréquence sont indépendantes. Nous pouvons alors admettre que les consommateurs en France déterminent leur niveau de consommation de vin selon un processus séquentiel à deux niveaux : ils décident d'abord de consommer, puis ils déterminent la fréquence de consommation.

Les paramètres estimés des modèles I, II et III n'ont pas d'interprétation directe. Seuls leurs signes nous fournissent des informations qualitatives sur l'effet (négatif ou positif) de chaque variable explicative sur les probabilités estimées. Pour quantifier les impacts des variables exogènes sur les différentes probabilités d'être consommateur régulier, occasionnel ou non-consommateur, des effets marginaux ont été calculés pour chacun des trois modèles estimés (voir annexe 3). Ces différents effets marginaux sont mesurés par les variations moyennes de probabilités et

[8] Si n est le nombre d'observations, LogL($\boldsymbol{\theta}$) le logarithme de la fonction de vraisemblance obtenue à partir du vecteur de paramètres $\boldsymbol{\theta}$, LogL($\boldsymbol{\theta}^0$) le logarithme de la fonction de vraisemblance obtenue seulement avec un vecteur de constantes $\boldsymbol{\theta}^0$, et p le nombre de paramètres, les statistiques utilisées dans l'article sont alors obtenues à l'aide des expressions suivantes :

critère d'Akaike	AIC=-2LogL+2p
critère de Schwarz	SC=-2LogL+pLog(n)
rapport de vraisemblance	LRT=-2[LogL($\boldsymbol{\theta}$)-LogL($\boldsymbol{\theta}^0$)]
pseudo-R^2	pseudo-R^2=1-[LogL($\boldsymbol{\theta}$)/LogL($\boldsymbol{\theta}^0$)]

les quasi-élasticités moyennes calculées pour l'ensemble de l'échantillon (voir tableau 5).

L'ensemble de ces résultats économétriques est tellement riche en informations qu'il faut mettre en place une méthode pour procéder à leur interprétation et discussion. Pour cela, nous avons décidé de regrouper les influences des variables explicatives selon qu'elles agissent, et ce de manière indépendante, sur la décision de consommer du vin ou sur les niveaux de fréquences, ou selon qu'elles influencent simultanément les deux types de décisions.

Variables influençant uniquement la décision de consommer

Dans ce premier groupe de variables explicatives se trouvent le statut social du foyer (variable *STUT*), la situation familiale (variable *SIT*), l'éducation à la consommation de vin pendant l'enfance (variable *INTERDIT*), les valeurs (variable *VALEUR*) et les croyances (variable *RISQUE*).

Les consommateurs de vin appartiennent plus fréquemment aux ménages « aisés » : un test de Wald appliqué aux paramètres associés à la variable *STUT* dans le modèle III montre que l'hypothèse nulle correspondant à l'absence d'effet du statut social est rejetée quand il s'agit d'expliquer les décisions de consommer; en revanche, elle n'est pas rejetée pour expliquer les fréquences[9]. Ces résultats sont conformes à ceux obtenus par Boizot (1999) avec un modèle probit. En revanche, le modèle II estimé (probit ordonné) révèle que les individus appartenant à un foyer de statut social élevé (*STUT* = 5) ont plus de chance d'avoir une forte fréquence de consommation, alors que le modèle III confirme que cette dimension explicative agit seulement sur la décision de consommer. Ainsi, selon cette dernière spécification et par rapport aux individus appartenant à des foyers de statut social « aisé » (*STUT* = 5), la variation de probabilité associée aux individus de catégorie sociale « moyenne supérieure » (*STUT* = 4) est de -0,09 (respectivement -0,046 et -0,053 pour les modèles I et II). Pour les « catégories sociales », l'écart en terme de variation moyenne de probabilité par rapport au groupe de référence pour le modèle III se distribue de la manière suivante: -0,153 (-0,094 pour le modèle I) pour la classe « moyenne inférieure » (*STUT* = 3), 0,156 (-0,102 pour le modèle I) pour les foyers « modestes » (*STUT* = 2), et -0,236 (-0,182 pour le modèle I) pour les foyers « très modestes » (*STUT* = 1).

[9] La statistique de Wald pour les décisions de consommation est égale à 53,41 pour quatre degrés de liberté avec une *p-value* inférieure à 0,001. Dans le cas des fréquences, la statistique de Wald est de 7,07 pour une *p-value* de 0,132.

Tableau 3. **Résultats économétriques des modèles probit simple et ordonné (Modèles I et II)**

Variables explicatives		Modèle I (Probit simple)		Modèle II (Probit ordonné)	
		Paramètres	Ecart-type	Paramètres	Ecart-type
Constante		-0,369	0,227	-0,973***	0,173
AGE âge		0,077***	0,007	0,065***	0,006
AGECARRE âge au carré		-0,001***	0,000	-0,0004***	0,000
SIT	= 1 Non maritale	-0,190***	0,068	-0,176***	0,055
PLOGT	= 1 Propriétaire logement	0,050	0,056	0,106***	0,041
THD	= 1 Sédentaire/rationnaire	Référence		Référence	
	= 2 Convive	-0,111*	0,061	0,134***	0,045
	= 3 Habitué de bar	0,072	0,089	0,233***	0,065
	= 4 Nomade	0,077	0,222	0,291***	0,135
INTERDIT	= 1 Aucun interdit	Référence		Référence	
	= 2 Interdit hors vin	0,234***	0,066	0,106**	0,045
	= 3 Vin interdit	0,041	0,055	0,018	0,042
REC	= 1 Ouest	Référence		Référence	
	= 2 Nord	-0,354	0,104	-0,281***	0,080
	= 3 Est	-0,301***	0,103	-0,293***	0,073
	= 4 Nord-Ouest	-0,155	0,106	-0,238***	0,074
	= 5 Sud-Est	-0,229***	0,095	-0,213***	0,066
	= 6 Centre	-0,054***	0,109	-0,086	0,075
	= 7 Sud	-0,155***	0,115	-0,095	0,080
	= 8 Ile-de-France	-0,406***	0,092	-0,301***	0,066
STUT	= 1 Très modeste	-0,675***	0,132	-0,539***	0,092
	= 2 Modeste	-0,435***	0,113	-0,270***	0,065
	= 3 Moyen inférieur	-0,416***	0,112	-0,319***	0,064
	= 4 Moyen supérieur	-0,225***	0,107	-0,197***	0,056
	= 5 Aisé	Référence		Référence	
NPF	= 1 Nombre	0,533***	0,103	0,514***	0,081
	= 2 de personnes	0,219***	0,082	0,342***	0,062
	= 3 au foyer	0,221***	0,081	0,161***	0,062
	= 4	0,231***	0,081	0,066	0,060
	= 5	Référence		Référence	
RISQUE	= 1 Vin préventif	Référence		Référence	
	= 2 Préventif et risqué	-0,113	0,084	-0,039	0,056
	= 3 Sans effet	-0,267***	0,064	-0,155***	0,045
	= 4 Pas d'opinion	-0,394***	0,081	-0,329***	0,064
	= 5 Risqué	-0,477***	0,079	-0,497***	0,062
SEXE	= 1 Masculin	0,372***	0,050	0,609***	0,037
SOCF	= 1 Foyer faible « sociabilité »	-0,264***	0,059	-0,400***	0,050
TFOYBA	= 0 Pas de boisson alcoolisée	-0,776***	0,081	-0,775***	0,080
	= 1 Vin quotidien	Référence		Référence	
	= 2 Autre b. a. quotidienne	-0,166*	0,090	-0,160**	0,066
	= 3 Vin hebdomadaire	0,075	0,079	-0,082	0,050
	= 4 Autres b. a. hebdomadaires	-0,077	0,118	-0,305***	0,089
	= 5 b. a. plus rarement	-0,132*	0,074	-0,422***	0,054
VILLEENF	= 1 Ville d'enfance	-0,069	0,055	-0,026	0,041
VALEUR	= 1 Insertion sociale	-0,110	0,064	-0,031	0,045
	= 2 Hédonisme	Référence		Référence	
	= 3 Sécuritaire	-0,326***	0,086	-0,232***	0,069
	= 4 Dignité, respect de soi	-0,159***	0,069	-0,082*	0,049
SEUIL 1 2e seuil du probit ordonné				0,866***	0,024
SEUIL 2 3e seuil du probit ordonné				1,531***	0,031
SEUIL 3 4e seuil du probit ordonné				1,728***	0,032
Nombre d'observations :		4010		4010	
Nombre de paramètres :		40		43	
Pseudo-R^2 :		0,189		0,146	
LogL :		-1763,433		-5041,000	
LRT :		824,213		-1723,346	
Critère d'Akaike :		3606,866		10169,904	
Critère de Schwartz :		3530,555		10087,665	

Notes : b. a. : boisson alcoolisée
*** significatif à 1% ; ** significatif à 5% ; * significatif à 10%
Les écarts-types des paramètres sont obtenus à l'aide de la procédure de White (écarts-types robustes à l'hétéroscédasticité).

Tableau 4. **Résultats économétriques du modèle probit bivarié ordonné (Modèle III)**

Variables explicatives		Décisions de consommer		Décisions des fréquences	
		Paramètres	Ecart-type	Paramètres	Ecart-type
Constante		-0,750***	0,219	-1,426***	0,230
AGE âge		0,079***	0,007	0,045***	0,008
AGECARRE âge au carré		-0,001***	0,000	-0,00017**	0,000
SIT	= 1 Non maritale	-0,187***	0,067	-0,090	0,075
PLOGT	= 1 Propriétaire logement	0,056	0,052	0,133***	0,051
THD	= 1 Sédentaire/rationnaire	Référence		Référence	
	= 2 Convive	0,058	0,060	0,216***	0,057
	= 3 Habitué de bar	0,153***	0,097	0,307***	0,079
	= 4 Nomade	0,198	0,210	0,405***	0,143
INTERDIT	= 1 Aucun interdit	Référence		Référence	
	= 2 Interdit hors vin	0,193***	0,061	-0,003	0,057
	= 3 Vin interdit	0,024	0,053	0,030	0,054
REC	= 1 Ouest	Référence		Référence	
	= 2 Nord	-0,312***	0,097	-0,189***	0,100
	= 3 Est	-0,213***	0,095	-0,398***	0,092
	= 4 Nord-Ouest	-0,137	0,097	-0,365***	0,094
	= 5 Sud-Est	-0,171*	0,088	-0,225***	0,079
	= 6 Centre	-0,091	0,097	-0,123	0,090
	= 7 Sud	-0,030	0,106	-0,192**	0,097
	= 8 Ile-de-France	-0,289***	0,086	-0,276***	0,081
STUT	= 1 Très modeste	-0,806***	0,123	-0,189	0,118
	= 2 Modeste	-0,587***	0,104	-0,005	0,081
	= 3 Moyen inférieur	-0,590***	0,103	-0,098	0,080
	= 4 Moyen supérieur	-0,363***	0,098	-0,118*	0,071
	= 5 Aisé	Référence		Référence	
NPF	= 1 Nombre	0,608***	0,102	0,307***	0,111
	= 2 de personnes	0,269***	0,079	0,337***	0,084
	= 3 au foyer	0,175***	0,079	0,117	0,087
	= 4	0,095	0,077	0,016	0,085
	= 5	Référence		Référence	
RISQUE	= 1 Vin préventif	Référence		Référence	
	= 2 Préventif et risqué	-0,098	0,077	-0,023	0,071
	= 3 Sans effet	-0,227***	0,060	-0,039	0,057
	= 4 Pas d'opinion	-0,422***	0,077	-0,083	0,083
	= 5 Risqué	-0,606***	0,074	-0,220***	0,084
SEXE	= 1 Masculin	0,502***	0,047	0,591***	0,047
SOCF	= 1 Foyer faible « sociabilité »	-0,420***	0,057	-0,266***	0,067
TFOYBA	= 0 Pas de boisson alcoolisée	-0,834***	0,083	-0,243***	0,109
	= 1 Vin quotidien	Référence		Référence	
	= 2 Autre b. a. quotidienne	-0,072	0,087	-0,276***	0,084
	= 3 Vin hebdomadaire	0,117	0,075	-0,245***	0,064
	= 4 Autres b. a. hebdomadaires	-0,244**	0,114	-0,303***	0,116
	= 5 b. a. plus rarement	-0,256***	0,069	-0,622***	0,082
VILLEENF	= 1 Ville d'enfance	-0,079	0,052	0,026	0,052
VALEUR	= 1 Insertion sociale	-0,103*	0,060	0,029	0,058
	= 2 Hédonisme	Référence		Référence	
	= 3 Sécuritaire	-0,293***	0,083	-0,141	0,094
	= 4 Dignité, respect de soi	-0,144***	0,065	0,003	0,062
SEUIL 1 2e seuil du probit ordonné				0,853***	0,028
SEUIL 2 3e seuil du probit ordonné				1,091***	0,031
Nombre d'observations :			4010		
Nombre de paramètres :			82		
Pseudo-R^2 :			0,173		
LogL :			-4883,883		
LRT :			2039,487		
Critère d'Akaike :			9931,766		
Critère de Schwartz :			9772,172		

Notes : b. a. : boisson alcoolisée

*** significatif à 1% ; ** significatif à 5% ; * significatif à 10%

Les écarts-types des paramètres sont obtenus à l'aide de la procédure de White (écarts-types robustes à l'hétéroscédasticité).

Tableau 5. Effets marginaux des variables explicatives sur les probabilités de consommation (effets calculés à partir des modèles I, II, et III)

Noms symboliques	Modalités	Probit binaire (Modèle I)	Probit ordonné simple (Modèle II)			Probit bivarié ordonné (Modèle III)		
		Consommer	Réguliers	Occasionnels	Pas de consommation	Réguliers	Occasionnels	Pas de consommation
Probabilités moyennes calculées (fréquences observées)		0,7670(0,766)	0,231 (0,23)	0,43(0,44)	0,32(0,33)	0,2332(0,23)	0,4364(0,44)	0,3303(0,23)
AGE	Age (quasi-élasticité)	0,112***	0,211***	0,059***	-0,271***	0,256***	-0,101***	-0,154***
SIT =1	Non maritale	-0,052***	-0,037***	-0,013***	0,050***	-0,028***	-0,027***	0,055***
PLOGT=1	Propriétaire logement	0,011***	0,029***	-0,003***	-0,026***	0,034***	-0,020***	-0,014***
THD =2	Convive	-0,028***	0,030***	0,007***	-0,038***	0,043***	-0,027***	-0,016***
THD =3	Habitué de bar	0,018***	0,053***	0,014***	-0,067***	0,068***	-0,025***	-0,043***
THD =4	Nomade	0,012***	0,093***	-0,030***	-0,063***	0,127***	-0,087***	-0,040***
INTERDIT=2	Interdit sauf vin	0,053***	0,027***	0,002**	-0,028***	0,012***	0,040***	-0,052***
INTERDIT=3	Vin interdit	0,011***	0,004***	0,001***	-0,005***	0,007***	-0,00010	-0,007***
REC=2	Nord	-0,091***	-0,066***	-0,014***	0,080***	-0,062***	-0,031***	0,093***
REC=3	Est	-0,07***	-0,068***	-0,013***	0,081***	-0,092***	0,031***	0,060***
REC=4	Nord-Ouest	-0,037***	-0,058***	-0,010***	0,068***	-0,083***	0,043***	0,040***
REC=5	Sud-Est	-0,051***	-0,054***	-0,002	0,056***	-0,060***	0,015***	0,045***
REC=6	Centre	-0,011***	-0,023***	0,001	0,022***	-0,034***	0,010***	0,024***
REC=7	Sud	-0,034***	-0,02***	0,002*	0,024***	-0,045***	0,037***	0,008***
REC=8	Ile-de-France	-0,098***	-0,072***	-0,010***	0,081***	-0,076***	-0,003	0,078***
STUT=1	Foyer très modeste	-0,182***	-0,127***	-0,021***	0,148***	-0,108***	-0,128***	0,236***
STUT=2	Foyer modeste	-0,102***	-0,068***	-0,001	0,069***	-0,044***	-0,112***	0,156***
STUT=3	Foyer moyen inférieur	-0,094***	-0,077***	-0,008***	0,085***	-0,053***	-0,100***	0,153***
STUT=4	Foyer moyen supérieur	-0,046***	-0,051***	-0,003**	0,053***	-0,043***	-0,047***	0,090***
NPF =1	Nombre de personnes au foyer	0,152***	0,121***	0,031***	-0,152***	0,117***	0,070***	-0,187***
NPF =2		0,053***	0,094***	-0,010***	-0,085***	0,101***	-0,028***	-0,074***
NPF =3		0,059***	0,033***	0,015***	-0,048***	0,029***	0,023***	-0,052***
NPF =4		0,061***	0,013***	0,008***	-0,021***	0,007***	0,022***	-0,029***
RISQUE=2	Vin préventif et risqué	-0,024***	-0,010***	0,00004	0,010***	-0,001***	-0,025***	0,026***
RISQUE=3	Sans effet	-0,066***	-0,038***	-0,004***	0,042***	-0,024***	-0,040***	0,064***
RISQUE=4	Pas d'opinion	-0,105***	-0,070***	-0,022***	0,092***	-0,047***	-0,077***	0,124***
RISQUE=5	Risqué	-0,127***	-0,101***	-0,048***	0,149***	-0,076***	-0,109***	0,184***
SEXE=1	Homme	0,088***	0,157***	0,004	-0,162***	0,166***	-0,028***	-0,138***
SOCF=1	Faible «sociabilité» du foyer	-0,071***	-0,095***	-0,017***	0,111***	-0,091***	-0,036***	0,127***
TFOYBA=0	Pas de boisson alcoolisée	-0,244***	-0,129***	-0,104***	0,233***	-0,104***	-0,155***	0,258***
TFOYBA=2	Autre b. alcoolisée quotidienne	-0,041***	-0,043***	0,001	0,043***	-0,068***	0,047***	0,021***
TFOYBA=3	Vin hebdomadaire	0,018***	-0,020***	-0,005***	0,024***	-0,042***	0,075***	-0,033***
TFOYBA=4	Autre b. alcoolisée hebdomadaire	-0,022***	-0,060***	-0,035***	0,095***	-0,065***	-0,012***	0,077***
TFOYBA=5	b. alcoolisée plus rarement	-0,035***	-0,085***	-0,043***	0,128***	-0,113***	0,035***	0,078***
VILLEENF=1	Ville d'enfance	-0,018***	-0,006***	-0,002***	0,007***	-0,001***	-0,022***	0,023***
VALEUR=1	Insertion sociale	-0,025***	-0,008***	-0,0002	0,008***	-0,001***	-0,026***	0,027***
VALEUR=3	Sécuritaire	-0,089***	-0,051***	-0,015***	0,065***	-0,048***	-0,040***	0,088***
VALEUR=4	Dignité, respect de soi	-0,039***	-0,020***	-0,003***	0,023***	-0,010***	-0,031***	0,041***

Notes: « b.» signifie « boisson» . ***, ** et * veulent respectivement dire « significativement différent de 0» pour des risques de première espèce de 1 %, 5 % et 10 %. La somme des quasi-élasticités de l'âge et la somme des variations des probabilités sont nulles pour chaque individu *i* de l'échantillon (voir annexe 3).

Les personnes vivant non maritalement sont moins consommatrices de vin. Le fait d'être célibataire, veuf ou divorcé diminue les chances d'être consommateur. La probabilité de consommer du vin diminue de 0,05 en moyenne pour les modèles I et III.

L'éducation à la consommation des boissons pendant l'enfance a aussi des effets significatifs sur la décision de consommer du vin. Ainsi, en prenant comme référentiel les individus auxquels aucune boisson alcoolisée n'a été interdite durant l'enfance (*INTERDIT* = 1), ceux qui se sont vu interdire de consommer des boissons alcoolisées autres que le vin (*INTERDIT* = 2) ont plus de chance d'être consommateurs de vin (variation moyenne de probabilité de 0,05 pour les modèles I et III). Un test de Wald de nullité de l'ensemble des paramètres liés à cette variable, effectué à partir du modèle III, révèle que l'éducation à la consommation des boissons lors de l'enfance agit sur les décisions de consommation[(10)] et non sur les fréquences[(11)]. Nous pouvons faire l'hypothèse, d'après ce résultat, qu'il y a concurrence entre les autres boissons alcoolisées et le vin dans la formation des habitudes. D'autre part, le fait d'interdire toute boisson alcoolisée (*INTERDIT* = 3) durant l'enfance n'a pas d'influence significative sur la décision de consommer et sur le choix des fréquences.

Pour ce qui est des croyances concernant le risque pour la santé entraîné par une consommation « modérée » de vin, les personnes qui considèrent celle-ci comme risquée (*RISQUE* = 5) ou qui n'ont pas d'opinion (*RISQUE* = 4), voire même qui considèrent que le vin n'a pas d'effet (*RISQUE* = 3) ont moins de chance d'être consommatrices de vin que celles qui le considèrent exclusivement comme bénéfique (groupe de référence *RISQUE* = 1); les variations moyennes de probabilité sont respectivement de -0,127, -0,105, -0,066 pour le modèle I et -0,184, -0,124, -0,064 pour le modèle III. Si nous retenons un seuil de confiance de 5%, cette dernière spécification (Modèle III) montre que cette dimension n'a aucun impact sur les fréquences de consommation, mais au contraire elle a une influence sur les décisions de consommer (voir tableau 4)[(12)].

[(10)] La statistique de Wald est égale à 10,39 pour deux degrés de liberté et la *p-value* est égale à 0,006.

[(11)] Wald = 0,401 pour deux degrés de liberté et une *p-value* égale à 0,819.

[(12)] Le test de Wald de nullité des paramètres associé à la variable *RISQUE* génère une statistique de 8,367 et une *p-value* de 0,079 pour quatre degrés de liberté dans le cas des fréquences de consommer alors qu'elles sont respectivement égales à 80,472 et 0,001 pour les décisions de consommer du vin. Dans un autre travail, un test économétrique effectué à l'aide de la méthode de Heckman (1979), portant sur des valeurs de consommation individuelle à l'année, indique en revanche que les personnes considérant que la consommation modérée de vin protège de certaines maladies consomment plus (Laporte, 2001).

Dernière variable explicative appartenant à ce groupe et influençant uniquement la décision de consommer: le système de valeurs (*VALEUR*). Selon les résultats obtenus à partir du modèle III, un test de Wald sur le pouvoir explicatif de la variable *VALEUR* génère une statistique de 3,58 et une *p-value* de 0,31 pour les décisions de fréquence. Ces deux données sont respectivement égales à 13,30 et 0,004 pour la décision de consommation. Ceux qui ont déclaré une valeur « sécuritaire » (*VALEUR* = 3) ou de « respectabilité » (*VALEUR* = 4) ont moins de chance d'être consommateurs de vin que ceux qui se sont affichés dans une valeur « hédoniste » (groupe de référence *VALEUR* = 2). Le paramètre associé à la variable *VALEUR* = 1, qui distingue les individus ayant déclaré leur attachement « aux bonnes relations avec les autres », n'est pas significatif (sauf à considérer le seuil de signification de 10% dans le modèle III). Globalement, ce sont les « hédonistes » qui ont le plus de chance d'être des consommateurs de vin. Les variations moyennes de probabilité calculées à partir du modèle I sont similaires à celles estimées à partir du modèle III. Notons enfin que la probabilité d'être consommateur pour ceux qui ont déclaré la *VALEUR* = 4 (« respectabilité ») est plus faible que la probabilité de ceux qui ont déclaré la *VALEUR* = 3 (« sécuritaire »). Ces résultats semblent confirmer l'idée intuitive selon laquelle le vin fait partie du répertoire alimentaire des « bons vivants ».

Variables explicatives affectant plus particulièrement les fréquences de consommation : le rythme de fréquentation des divers lieux de consommation hors domicile et les habitudes familiales

Le rythme de fréquentation des divers lieux de consommation hors domicile (variable *THD*) semble n'influencer que les fréquences. Par rapport aux groupes des individus « sédentaires » ou « rationnaires » (*THD* = 1), le fait que les personnes soient « convives », « habituées de bar » ou « nomades » ne semble pas jouer sur la probabilité de consommer du vin (voir tableau 4). Le modèle I (probit binaire) confirme ce résultat (tableau 3). Cette dimension explicative a sûrement une influence sur le volume ou l'intensité de la consommation, mais pas sur la décision de consommer. Un test de Wald de nullité de l'ensemble des paramètres associés à la variable *THD* effectué à l'aide des estimateurs du modèle III donne pour les décisions de consommer une statistique de Wald qui vaut 4,20 avec une p-value de 0,241 alors que pour les choix des fréquences, nous obtenons une valeur de 28,80 pour une p-value inférieure à 0,001. Quand nous comparons les résultats économétriques des modèles II (probit ordonné simple) et III (probit bivarié ordonné), nous observons que, par rapport aux individus classés comme « sédentaires » ou « rationnaires » (*THD* = 1), les individus qui pratiquent le « hors domicile » ont

plus de chance d'avoir des fréquences de consommation élevées. Ce sont les « nomades » qui ont la probabilité la plus élevée d'être des consommateurs de vin réguliers, ils sont suivis par les « habitués des bars ». Entre un individu qui fréquente peu les lieux de consommation hors domicile et un « nomade », on peut prédire une variation moyenne de probabilité d'être un consommateur régulier égale à 0,127 (Tableau 5). Ces résultats confirment le caractère convivial de cette consommation.

Les habitudes familiales (variable *TFOYBA*) pèsent sur les comportements actuels : les résultats économétriques des modèles II et III montrent l'importance de la présence de vin dans le foyer d'enfance. Selon le modèle III, par rapport au groupe de référence (avec vin quotidien dans le foyer d'enfance *TFOYBA* = 1), tous les autres groupes d'individus (caractérisés par des valeurs de *TFOYBA* = 2, 3, 4, 5) ont tous des chances d'avoir de plus faibles fréquences de consommation de vin. En effet, dans le modèle III, la probabilité d'être consommateur régulier diminue respectivement de -0,104, -0,068, -0,042, -0,065, -0,113 (Tableau 5). Nous pouvons donc conclure que les fréquences élevées de consommation de vin se retrouvent principalement chez les individus pour lesquels le vin était un bien de consommation courante dans la tradition familiale.

Variables explicatives influençant à la fois les décisions et les fréquences de consommation : âge, sexe, localisation géographique, nombre de personnes au foyer et sociabilité du foyer

La consommation de vin est une fonction concave de l'âge. La probabilité d'être consommateur de vin et celle d'avoir une fréquence de consommation élevée dépendent en effet de l'âge de manière significative. Ainsi les probabilités de consommer et d'être consommateur quotidien (*FREQ* = 4) augmentent pour les jeunes individus et diminuent pour les plus âgés (voir figure 1). La quasi-élasticité moyenne est égale à 0,112 pour le modèle I et proche de la valeur moyenne calculée à partir du modèle III, qui est de 0,154. L'effet négatif de l'âge sur la probabilité de consommer du vin se situe vers 60 ans ; à partir de cet âge, l'élasticité-âge devient fortement négative (voir figure 2). Ces résultats sont conformes à ceux obtenus par Boulet *et al.* (1997). Dans les trois modèles estimés, le test de Wald portant sur l'hypothèse nulle (égalités à zéro des paramètres associés aux variables *AGE* et *AGECARRE*) révèle que cette dernière est rejetée.

Les hommes sont plus consommateurs de vin que les femmes : pour une personne du sexe masculin, la probabilité d'être consommateur de vin augmente de 0,088 pour le modèle I contre 0,138 pour le modèle III. Les hommes ont aussi plus de chance d'être des consommateurs

réguliers; la variation moyenne de probabilité d'être un consommateur régulier (0,166) est plus élevée, alors que celle d'être un consommateur occasionnel est plus faible (-0,028). Nous remarquons ainsi la forte significativité du paramètre associé au sexe dans tous les modèles, ce qui confirme le rôle important de cette variable explicative.

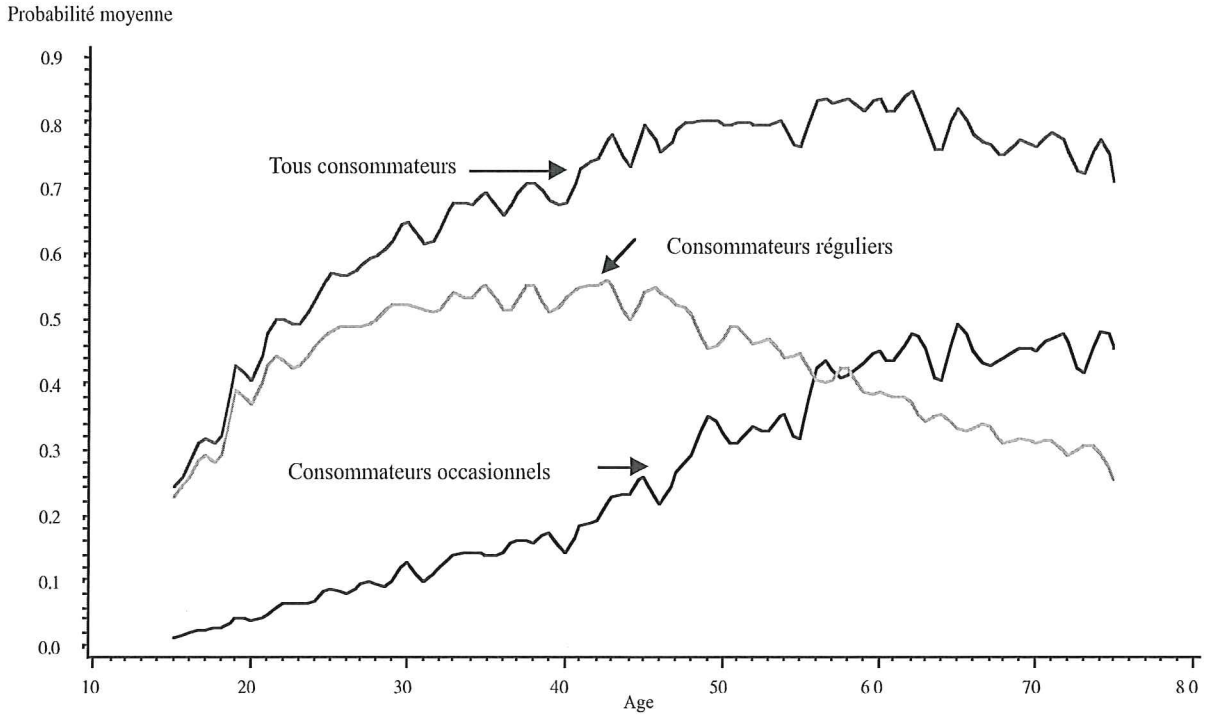
C'est dans la région « Ouest » que l'on rencontre le plus de consommateurs de vin. Par rapport aux personnes qui vivent dans cette région, les autres Français ont une plus faible probabilité d'être des consommateurs de vin et d'avoir de fortes fréquences de consommation (toutes autres choses étant égales par ailleurs). C'est en Ile-de-France, dans le Nord et l'Est que ces probabilités sont les plus faibles. Selon le modèle III, la variation moyenne de probabilité quand on passe de la région Ouest à la région Ile-de-France est de -0,078 (-0,098 pour le probit binaire), proche de la variation obtenue quand on passe de l'Ouest au Nord -0,093 (-0,091 pour le probit binaire). D'après les résultats empiriques des spécifications II et III, c'est dans l'Ouest du pays que la probabilité d'être consommateur quotidien de vin est la plus élevée (Tableau 5). Le test de Wald sur la significativité de l'effet de la localisation géographique sur les fréquences et sur les décisions de consommer du vin montre que l'hypothèse nulle est rejetée pour les deux types de décisions[13]. Enfin, rester dans la ville de son enfance (*VILLEENF*) n'apparaît jamais significatif quelle que soit la spécification retenue.

Un examen des résultats présentés aux tableaux 3, 4 et 5 indique que l'on consomme moins souvent de vin dans les familles nombreuses (*NPF* = 5); d'après les estimations du modèle III, les variations moyennes de probabilité d'être consommateur passent de 0,187 pour les personnes seules à 0,029 pour celles vivant dans des foyers de quatre individus. Les estimations des modèles II et III montrent également que les individus, appartenant à des foyers dont le nombre de personnes est faible, ont plus de chance d'être des consommateurs réguliers. La variation de probabilité pour un individu seul d'être un consommateur régulier est de 0,117 dans le cas du modèle III (0,121 pour le modèle II) et de 0,070 (0,031 pour le modèle II) d'être un consommateur occasionnel. Le test de Wald montre que l'influence de cette variable explicative est significative à la fois sur les décisions et les fréquences de consommation du vin[14].

[13] La statistique de Wald est égale à 27,71 pour les choix de fréquences pour sept degrés de liberté, ce qui équivaut à une *p-value* inférieure à 0,001 alors que pour les décisions, la même donnée vaut 20,29 pour une *p-value* de 0,005.

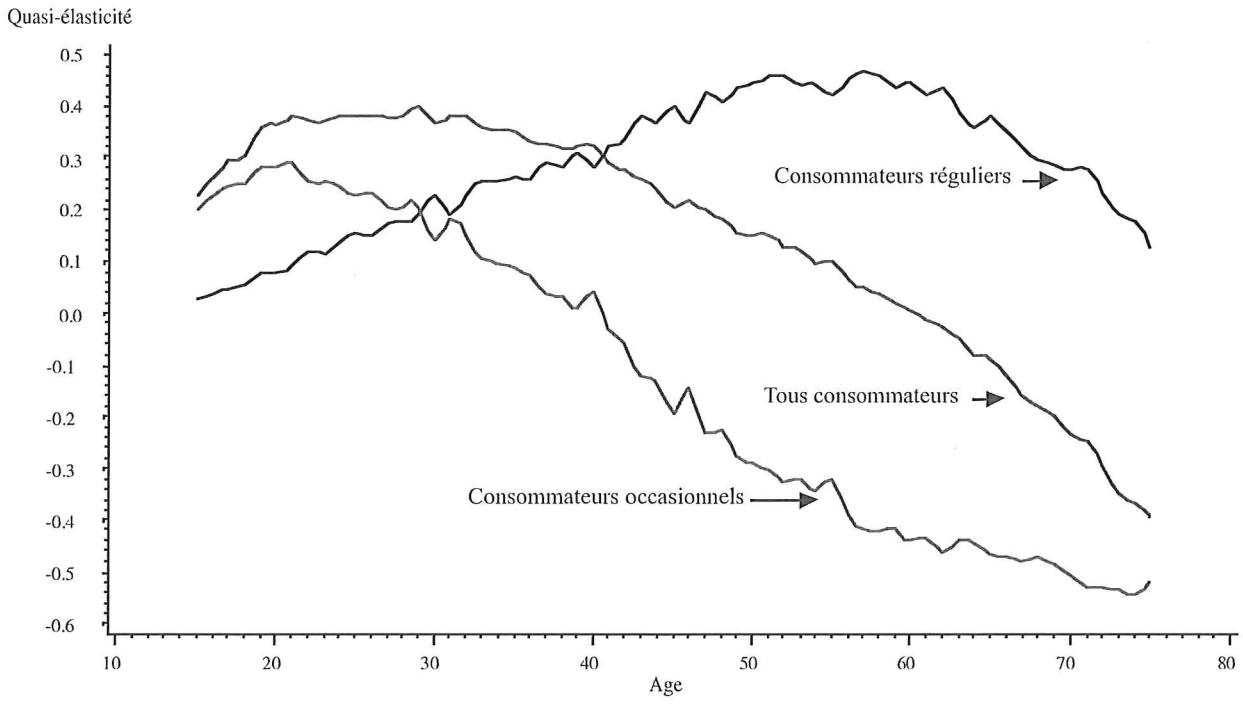
[14] Les statistiques de Wald sont respectivement égales à 27,71 pour les décisions de fréquences et 38,61 pour les décisions de consommation, ce qui donne pour les deux valeurs un *p-value* inférieure à 0,001. Le nombre de degrés de liberté pour ce test est de quatre.

Figure 1. **Probabilités moyennes par âge (calculées à partir du modèle III sur 4010 observations)**



Notes : Les consommateurs réguliers sont ceux qui déclarent consommer du vin tous les jours ou presque (*FREQ* = 3 ou *FREQ* = 4).
Les consommateurs occasionnels sont ceux qui en consomment une ou deux fois par semaine ou plus rarement (*FREQ* = 1 ou *FREQ* = 2).

Figure 2. **Quasi-élasticités moyennes par âge (calculées à partir du modèle III sur 4010 observations)**



Notes : Identiques à celles de la figure 1

Les individus appartenant à des foyers de faible « sociabilité » (*SOCF* = 1) sont aussi moins concernés par la consommation de vin. Ils ont plus de chance d'être des non-demandeurs de vin et d'avoir de faibles fréquences de consommation que les foyers de « forte sociabilité » (voir tableau 4). En prenant à titre illustratif les résultats économétriques du modèle III, la probabilité d'être consommateur régulier diminue de -0,091 et celle d'être occasionnel de -0,036. Ainsi, la fréquence de bons repas au domicile agit comme un facteur favorable à la consommation du vin par les membres du foyer.

CONCLUSION

Dans ce travail, la décision de consommation du vin et celle qui concerne le niveau de cette consommation (évalué par sa fréquence) sont supposées dépendre des caractéristiques socio-démographiques et socio-économiques des individus. L'originalité de cette recherche réside dans le fait que le modèle empirique utilisé (probit bivarié ordonné) permet de prendre en compte les **décisions de consommation** du vin et les décisions concernant le **niveau des fréquences** de consommation dans un cadre de travail unifié. Le modèle probit bivarié ordonné permet de distinguer les déterminants de la décision de consommer, de ceux qui influencent les fréquences de consommation, tout en évitant d'éventuels biais de sélection. Il permet ainsi de savoir si les deux décisions étudiées sont corrélées ou indépendantes.

Selon nos estimations, décision de consommation et décision concernant la fréquence apparaissent non corrélées. Il est donc fort possible que le processus de décision de consommation du vin soit séquentiel : l'individu décide dans une premier temps de consommer du vin puis dans une seconde phase, il choisit un niveau de fréquence de consommation. Nous montrons que les variables explicatives retenues n'agissent pas avec la même intensité sur les deux types de décisions. Sans changer de signe, l'âge, le sexe, la localisation géographique, la composition et la « sociabilité » du foyer influencent à la fois la décision et les fréquences de consommation. Le statut social, le statut matrimonial, le « système de valeurs », l'éducation à la consommation et la perception par les individus des risques liés à la consommation de vin pour la santé agiraient uniquement sur les décisions de consommation. Enfin, la fréquentation des lieux de consommation hors domicile (restauration, bars, etc.) et certaines habitudes familiales du foyer d'enfance (présence de boissons alcoolisées, en particulier) influenceraient uniquement les fréquences.

Cette étude présente cependant plusieurs limites. La plus importante d'entre elles est que ce travail empirique peut souffrir d'un biais d'hétérogénéité. En l'absence de données de panel, il aurait été judicieux d'utiliser des pseudo-panels. Avec les données de l'enquête ONIVINS-INRA,

cette stratégie d'estimation n'est pas envisageable faute d'un nombre de degrés de liberté suffisamment élevé. D'autre part, le cadre de travail est statique. Or, c'est par une analyse dynamique qu'il est possible de mettre en évidence le rôle des habitudes et du temps historique sur la consommation de vin. Par ailleurs, des recherches antérieures ont montré le rôle essentiel des effets d'âge et plus particulièrement des effets de génération sur l'évolution des fréquences de consommation (Boulet et Laporte, 1997 ; Laporte, 2001). Sur des pseudo-panels construits à partir des coupes successives ONIVINS-INRA, il est possible de contrôler ces effets par des modèles ACP (âge-cohorte-période), avec un degré de liberté élevé. Enfin, les effets des autres biens concurrents ou complémentaires du vin n'ont été pris en compte que par des variables auxiliaires appropriées.

BIBLIOGRAPHIE

ALLENBY (G. M.), LEONE (R. P.) et JEN (L.), 1999 — A dynamic model of purchase timing with application to direct marketing, *Journal of the American Statistical Association*, 94, n° 446, pp. 365-374.

BADOUIN (R.), 1953 — Analyse économique et détermination du prix du vin, Centre régional de productivité et des études économiques, *Bulletin trimestriel*, 3, pp. 115-157.

BANKS (J. R.), BLUNDELL (R. W.) et LEWBEL (A.), 1997 — Quadratic Engel curves and consumer demand, *The Review of Economics and Statistics*, 79, pp. 527-539.

BLAYLOCK (J. R.), BLISARD (W. N.), 1992 — US cigarette consumption: the case of low-income women, *American Journal of Agricultural Economics*, 74, pp. 698-705.

BOIZOT (C.), 1999 — La demande de boissons des ménages en France: une estimation de la consommation à domicile, *Économie et statistique*, n° 324-235, 4/5, pp. 143-156.

BOIZOT (C.), ROBIN (J.-M.) et VISSER (M.), 2001 — The demand for food products. An analysis of interpurchase times and purchased quantities, *Economic Journal*, vol. 111, pp. 391-419

BOULET (D.), LAPORTE (J.-P.), 1997 — Les comportements de consommation de vin en France. Recherche en économie et sociologie rurales, *INRA Sciences Sociales*, n° 3, juin, 4 p.

BOULET (D.), LAPORTE (J.-P.), AIGRAIN (P.) et MELANI (C.), 1997 — La transformation des comportements alimentaires, cycles de vie et effet de génération : le cas du vin, *Économie et sociétés, Développement agro-alimentaire*, Série A.G., 9, n° 23, pp. 47-67.

CONFORTI (P.), PERANI (P.) et RIZZI (P. L.), 2000 — Food and nutriment demands in Italy. Actual behaviour and forecast through a multistage quadratic system with heterogeneous preferences, Working paper 303, QUADERNI, Dipertimento di economia politica, Università degli studi di Siena, octobre, 22 p.

CROWLEY (S.), RICHARDSON (J.), 1997 — Alcohol taxation to reduce the cost of alcohol induced ill health, Working paper 4, Centre for Health Program Evaluation, Melbourne, February, 62 p.

DEATON (A.), MUELLBAUER (J.), 1980 — An almost ideal demand system, *The American Economic Review*, 70, pp. 312-326.

DIDAY (E.), 1971 — Une nouvelle méthode ou classification et reconnaissance des formes, *Revue d'économie appliquée*, vol. 19, n° 2, pp. 19-34.

FOUQUET (A.), 1973 — Modèle de projection de la demande des ménages, *Collection de l'INSEE*, série M22, 228 p.

GAO (X. M.), WAILES (E. J.) et CRAMER (G. L.), 1995 — A microeconometric model analysis of US consumer demand for alcoholic beverages, *Applied Economics*, 27, pp. 59-69.

HAINES (P. S.), GUILKEY (D. K.) et POPKIN (B. M.), 1998 — Modeling food consumption decisions as a two-step process, *American Journal of Agricultural Economics*, August, pp. 543-552.

HALL (B. H.), LINK (A. N.) et SCOTT (J. T.) 2002 — Universities as research partners, *Review of Economics and Statistics* (à paraître).

HAUTEVILLE (F.) (d'), 2001 — Valeurs, attitude et statut du foyer : un éclairage nouveau sur le comportement de consommation de vin, *in :* LAPORTE (J.-P.) (ed.), La consommation de vin en France. Comportements, attitudes et représentations. Résultats d'enquête ONIVINS-INRA 2000. Evolution 1980-2000 et projections 2010, Série études, n° 1, INRA-ESR-MOISA, Montpellier, ONIVINS Paris, pp. 37-44.

HECKMAN (J. J.), 1979 — Sample selection bias as a specification error, *Econometrica*, 47, pp. 153-161.

LAPORTE (J.-P.) (ed.), 2001 — La consommation de vin en France. Comportements, attitudes et représentations. Résultats d'enquête ONIVINS-INRA 2000. Évolution 1980-2000 et projections 2010, Série études, n° 1, INRA-ESR-MOISA, Montpellier, ONIVINS Paris, 78 p. + annexes.

LEBART (L.), MORINEAU (J.-P.) et TABARD (N.), 1977 — *Techniques de la description statistique: méthodes et logiciels*, Paris, Dunod, 351 p.

LEUNG (S. F.), PHELPS (C. E.), 1993 — My kingdom for a drink? A review of estimates of the price sensitivity of demand for alcoholic beverages, *in:* HILTON (M. E.), BOLSS (G.) (eds), *Economics and the prevention of alcohol-related problems,* Research Monograph n° 25, Rockville, MD, National Institute on Alcohol Abuse and Alcoholism, US Department of Health and Human Services, pp. 1-31.

MCGUINESS (T.), 1983 — The demand for beer, spirits and wine in the UK, 1956-1979, *in:* GRANT (M.), PFANT (M.) and WILLIAMS (A.) (eds), *Economics and Alcohol: Consumption and Controls*, New York, Gardner Press, pp. 238-242.

MILHAU (J.), 1935 — Étude économétrique du prix du vin en France, Montpellier, Causse, Graille et Castelnau, Thèse de Doctorat, 88 p.

MOULTON (K. S.), 1995 — The impact on wine demand of balanced messages about wine and health, Working paper 753, Dept. of Agr. and Resource Econ., University of California, Berkeley, 9 p.

NELSON (J. P.), 1997 — Economic and demographic factors in US alcohol demand: a growth accounting analysis, *Empirical Economics*, 22, pp. 83-102.

ONIVINS, 1999 — Enquête sur la consommation de vin en France en 1998, *Bulletin d'information*, janvier, Paris.

ORNSTEIN (S. J.), LEVY (D.), 1983 — Price and income elasticities and the demand for alcoholic Beverages, *in:* GALANTER (M.), (ed.), *Recent Developments in Alcoholism*, vol. 1, New York, Plenum, pp. 303-345.

PASCAL (A.), ROBIN (J.-M.), 1995 — Le coût de l'activité féminine: estimation à l'aide d'un modèle de demande conditionnel, *Économie et Prévision,* 121, pp. 141-148.

SAFFER (H.), CHALOUPKA (F.), 1994 — Alcohol tax equalization and social costs, *Eastern Economic Review*, 20, pp 33-43.

SELVANATHAN (E. A.), 1991 — Cross-country alcohol consumption comparison: an application of the Rotterdam demand system, *Applied Economics*, 23, pp. 1613-22.

SEYTE (F.), TERRAZA (M.), 2000 — Analyse des substitutions dans le complexe des boissons en France au moyen du modèle AIDS, *Les Cahiers de l'OCVE*, 3, 17 p.

TERRAZA (M.), 1981 — Économie des processus aléatoires: le cas du marché français des vins de consommation courante, Thèse de Doctorat, UER de sciences économiques, Université de Montpellier, 3 vol. et annexes 1000 p.

WHITE (H.), 1982 — Maximum likelihood estimation of miss-specified models, *Econometrica*, 50, pp. 1-26.

YATCHEW (A.), GRILICHES (Z.), 1984 — Specification error in probit models, *Review of Economics and Statistics*, 66, pp. 134-139.

YEN (S.T.), 1995 — Cross-section estimation of US demand for alcoholic beverage, *Applied Economics*, 27, pp. 381-392.

ANNEXE 1

Présentation des données

	Nombre d'interviews	Pourcentage
FRÉQUENCES DE CONSOMMATION DU VIN (= *FREQ*)		
4 - Tous les jours	750	8,7
3 - Presque tous les jours	171	4,3
2 - 1 ou 2 fois par semaine	715	17,8
1 - plus rarement	1049	26,2
0 - Jamais	1325	33,0
VARIABLES EXPLICATIVES		
SEXE		
1-Homme	1896	47,3
0 - Femme	2114	52,7
AGE		
<15 - 17 ans	178	4,4
18 - 24 ans	463	11,5
25 - 34 ans	825	20,6
35 - 49 ans	1135	28,3
50 - 64 ans	703	17,5
65 ans et plus	706	17,6
SOCIABILITÉ FOYER (= *SOCF*)		
0 - Autre	3042	75,9
1 - Faible	968	24,1
SYSTÈME DE VALEURS (= *VALEUR*)		
1 - Insertion sociale	1469	36,6
2 - Hédonisme	1024	25,5
3 - Sécuritaire	428	10,7
4 - Respectabilité	1089	27,2
RÉGION. CONSO. (= *REC*)		
1 - Ouest	534	13,3
2 - Nord	402	10,0
3 - Est	441	11,0
4 - N.-Ouest	423	10,5
5 - S.-Est	674	16,8
6 - Centre	440	11,0
7 - Sud	337	8,4
8 - Ile-de-France	759	18,9
BOISSONS FOYER D'ENFANCE (= *TF0YBA*)		
0 - Pas de b.a.	383	9,6
1 - Vin quotidien	2117	52,8
2 - Autre b.a. quot.	360	9,0
3 - Vin hebdomadaire	482	2,0
4 - Autre b.a. hebdo.	159	4,0
5 - b.a. plus rare	509	12,7
STATUT SOCIAL FOYER (= *STUT*)		
1 - Très Modeste	399	10,0
2 - Modeste	1195	29,8
3 - Moy. inf.	817	20,4
4 - Moy. sup.	1196	29,8
5 - Aisé	403	10,0
SITUATION FAMILIALE (= *SIT*)		
0 - Maritale	2490	62,1
1 - Non maritale	1520	37,9
NBRE PERSONNES FOYER (= *NPF*)		
1 - Personne seule	748	18,7
2 - Deux personnes	1145	28,6
3 - Trois personnes	762	19,0
4 - Quatre personnes	789	19,7
5 - Cinq personnes et plus	566	14,1
FRÉQUENTATION HORS DOMICILE (*THD*)		
1 - Sédentaires	1973	49,2
2 - Conviviaux	770	19,2
3 - Habitués de bar	366	9,1
4 - Nomades	83	2,1
5 - Rationnaires	818	20,4
OPINION VIN/SANTÉ (= *RISQUE*)		
1 - Vin préventif	1225	30,5
2 - Prévent. & risqué	537	13,4
3 - Sans effet	1234	30,8
4 - Sans opinion	478	11,9
5 - Plutôt risqué	536	13,4
HABITE VILLE ENFANCE (= *VILLEENF*)		
1 - Ville enfance	1169	29,2
0 - Autre	2841	70,8
ATTITUDE PARENTS/B. ALCOOL (= *INTERDIT*)		
1 - Aucun interdit	1835	45,8
2 - Interdit hors vin	891	22,2
3 - Vin interdit	1284	32,0
OCCUPATION LOGEMENT (= *PL0GT*)		
Non indiqué	19	0,5
1 - Propriétaire	1917	47,8
0 - Autre	2074	51,7

ANNEXE 2

Construction de variables explicatives synthétiques

Variable « comportement hors domicile » (THD)

Il s'agit d'une typologie des individus réalisée, par la méthode des nuées dynamiques (Diday, 1971), à partir de la fréquence de présence dans neuf catégories de situations de consommation en dehors du domicile, estimée en nombre de jours: 1) invitation chez des amis ou de la famille, 2) fréquentation des cafés ou bars, 3) fréquentation des restaurants d'entreprise et cantines, 4) fréquentation des fast-foods, 5) fréquentation des cafétérias et self, 6) repas pris sur le lieu de travail, 7) fréquentation des restaurants en famille, entre amis, 8) fréquentation des restaurants liée à d'autres circonstances (travail, déplacements), 9) fréquentations des boîtes de nuit, dancing. Nous avons choisi une structuration en cinq classes, chacune d'elles étant caractérisée par un lieu de fréquentation dominant. Le tableau A.2.1 permet de caractériser ces groupes (moyenne et coefficient de variation).

Variable « statut social des ménages » (STUT)

Une échelle de « statut social » est établie, selon la méthode de Lebart *et al.* (1977), par l'utilisation simultanée de plusieurs critères. Cette variable synthétique a été créée par une analyse factorielle des correspondances multiples (AFCM) sur cinq variables: a) le revenu du ménage, b) le revenu par unité de consommation dans le ménage, c) le niveau d'équipement électro-ménager, d) la catégorie socio-professionnelle (CSP) du chef de famille, et e) le niveau scolaire du chef de famille. La projection de chaque ménage sur le premier axe factoriel a permis d'élaborer une échelle ordinale de « statut social », qui a été arbitrairement divisée en cinq classes (*STUT* = 1,...,5). Il est évident, à cet égard, que les limites entre les différentes catégories sociales retenues sont subjectives et que notre principal souci de partition a été de fixer des effectifs de classes relativement équilibrés (les deux classes extrêmes: « très modeste » et « aisée » étant par définition des « queues de distribution », qui ont été limitées en effectifs).

Variable géographique des régions (REC)

Nord (= 1) :	Régions Picardie et Nord – Pas-de-Calais
Nord-Ouest (= 2) :	Régions Haute-Normandie, Basse-Normandie et Bretagne
Ouest (= 3) :	Régions Pays-de-la-Loire, Poitou-Charente et Aquitaine
Centre (= 4) :	Régions Centre, Limousin, Bourgogne et Auvergne
Est (= 5) :	Régions Champagne, Lorraine, Alsace et Franche Comté
Sud (= 6) :	Régions Midi - Pyrénées et Languedoc - Roussillon
Sud-Est (= 7) :	Régions Rhône - Alpes et Provence - Côte-d'Azur
Ile de France (= 8) :	Région parisienne

Tableau A.2.1. Typologie des comportements de fréquentation du hors domicile

	« Sédentaires »	« Rationnaires »	« Conviviaux »	« Habitués de bar »	« Nomades »	Ensemble
% de l'effectif total	50.6	19.8	18.8	8.6	2.1	100
Estimation du nombre moyen de jours par an de présence hors du foyer						
Invitations tiers	8 (73)	26 (97)	53 (43)	32 (98)	33 (93)	23 (110)
Café, bar	6 (256)	13 (157)	11 (172)	116 (13)	51 (99)	20 (186)
Fast-food	5 (268)	13 (156)	11 (185)	21 (154)	11 (186)	6 (208)
Cafétéria, self	4 (306)	9 (250)	7 (256)	11 (210)	10 (240)	9 (274)
Cantine, lieu de travail	2 (486)	119 (6)	4 (327)	30 (129)	30 (163)	31 (164)
Restaurant loisir	5 (203)	9 (150)	10 (151)	17 (132)	31 (131)	9 (180)
Restaurant autre	1 (484)	3 (313)	2 (345)	5 (244)	117 (11)	5 (405)
Boîte de nuit	2 (477)	5 (270)	5 (268)	15 (175)	11 (207)	4 (312)
Ensemble	33 (73)	197 (33)	103 (57)	257 (40)	295 (41)	107 (97)

Note: Les données entre parenthèses se rapportent au coefficient de variation qui est égal à (écart-type) x 100/moyenne
Source: ONIVINS-INRA

ANNEXE 3

Méthodes de calcul des effets marginaux des variables explicatives

Cas des variables explicatives continues

Pour l'âge, qui est une variable continue X dont le sous-indice est m, les effets marginaux dépendent de la valeur des paramètres et des fonctions de densité :

Modèle I $$\frac{\partial P[d_i = 1]}{\partial X_{im}} = \alpha_m f_u(\boldsymbol{\alpha' X}/\sigma_\mu) \qquad \text{(A.3.1)}$$

Modèle II $$\frac{\partial P[FREQ_i = j]}{\partial Z_{im}} = \beta_m \left[f_\varepsilon(s_j - \boldsymbol{\beta' Z_i})/\sigma_{\varepsilon i}) - f_\varepsilon(s_{j-1} - \boldsymbol{\beta' Z_i}/\sigma_{\varepsilon i})\right] \qquad \text{(A.3.2)}$$

Modèle III (A.3.3)

$$\frac{\partial P[FREQ_i = j, d_i = 1]}{\partial X_{im}} = [\alpha_m - \beta_m] * \left[\frac{\partial F(s_j - \boldsymbol{\beta' Z_i}, \boldsymbol{\alpha' X_i})}{\partial X_{im}} - \frac{\partial F(s_{j-1} - \boldsymbol{\beta' Z_i}, \boldsymbol{\alpha' X_i})}{\partial X_{im}}\right]$$

où $f_\varepsilon(.)$ et $f_\mu(.)$ sont les fonctions de densité. La relation (A.3.3) donne, pour le modèle III, l'effet marginal de la variable X_m commune aux décisions de consommation et de fréquences. Une simple observation des relations (A.3.1) à (A.3.3) permet de constater les erreurs possibles dans la mesure des effets marginaux quand on ne choisit pas la bonne spécification.

Les quasi-élasticités calculées pour la variable « âge » au tableau 5 sont obtenues à l'aide des expressions suivantes :

Modèle I $$(1/n)\sum_{i=1}^{n} \frac{\partial p_i}{\partial Log X_{im}}$$

Modèle II $$(1/n)\sum_{i=1}^{n} \frac{\partial P\left[FREQ_i = j\right]}{\partial Log Z_{im}}$$

Modèle III $$(1/n)\sum_{i=1}^{n} \partial P\left[FREQ_i = j, d_i = 1\right] / \partial Log X_{im}$$

Celles-ci permettent de mesurer l'impact d'une variation de 1 % de la variable continue considérée (âge) sur les différentes probabilités.

Cas des variables explicatives discrètes

Pour les variables discrètes, les effets marginaux se résument à l'analyse de la déviation de la probabilité calculée par rapport à la probabilité correspondant à la situation de référence. Si on note respectivement α_0 et β_0 les constantes dans chaque modèle, les variations de probabilités dans chaque modèle dues à une variable dichotomique X_m incluse dans Z (par exemple le sexe) sont :

Modèle I

$$dp_{i1} = dP[d_i = 1] = F_\psi(\alpha_0 + \alpha_m + \overline{\boldsymbol{\alpha' X_i}}/\sigma_\mu) - F_\psi(\alpha_0 + \overline{\boldsymbol{\alpha' X_i}}/\sigma_\mu) \qquad \text{(A.3.4)}$$

Modèle II

$$dp_{ij} = dP\left[FREQ_i = j\right] = F_\varepsilon\left(\left(s_j - \beta_0 - \beta_m - \overline{\boldsymbol{\beta}'\boldsymbol{Z}_i}\right)/\sigma_\varepsilon\right) \qquad \text{(A.3.5)}$$
$$-F_\varepsilon\left(\left(s_{j-1} - \beta_0 - \beta_m - \overline{\boldsymbol{\beta}'\boldsymbol{Z}_i}\right)/\sigma_\varepsilon\right)$$
$$-\left[F_\varepsilon\left(\left(s_j - \beta_0 - \overline{\boldsymbol{\beta}\boldsymbol{Z}}\right)/\sigma_\varepsilon\right) - F_\varepsilon\left(\left(s_{j-1} - \beta_0 - \overline{\boldsymbol{\beta}'\boldsymbol{Z}}\right)/\sigma_\varepsilon\right)\right]$$

Modèle III

$$dp_{ij} = dP[FREQ_i = j, d_i = 1] = F(s_j - \beta_0 - \beta_m - \overline{\boldsymbol{\beta}'\boldsymbol{Z}_i}, \alpha_0 + \alpha_m + \overline{\boldsymbol{\alpha}'\boldsymbol{X}_i}, \rho) \qquad \text{(A.3.6)}$$
$$- F(s_j - \beta_0 - \overline{\boldsymbol{\beta}'\boldsymbol{Z}_i}, \alpha_0 + \overline{\boldsymbol{\alpha}'\boldsymbol{X}_i}, \rho)$$
$$- [F(s_j - \beta_0 - \overline{\boldsymbol{\beta}'\boldsymbol{Z}_i}, \alpha_0 + \alpha_m + \overline{\boldsymbol{\alpha}'\boldsymbol{X}_i}, \rho)$$
$$- F(s_j - \beta_0 - \overline{\boldsymbol{\beta}'\boldsymbol{Z}_i}, \alpha_0 + \overline{\boldsymbol{\alpha}'\boldsymbol{X}_i}, \rho)]$$

où $\overline{\boldsymbol{\alpha}'\boldsymbol{X}_i} = \overline{\boldsymbol{\alpha}'\boldsymbol{X}_i} - \alpha_m X_m$ et $\overline{\boldsymbol{\beta}'\boldsymbol{Z}_i} = \overline{\boldsymbol{\beta}'\boldsymbol{Z}_i} - \beta_m Z_m - \beta_0$ sont des combinaisons linéaires – relations (1) et (5) - diminuées de l'effet de la variable X_m. Le premier terme de chacune des relations (A.3.4) à (A.3.6) représente le «taux de pratique» d'un individu *i*.

Effets marginaux par type de consommateur

A partir de ces variations de probabilités calculées pour chaque individu de l'échantillon, nous avons calculé l'effet des variables explicatives sur les probabilités d'être consommateur régulier p_i^R = $[FREQ_i = 4$ (ou) $FREQ_i = 3]$, occasionnel $p_i^0 = P[FREQ_i = 2$ (ou) $FREQ_i = 1]$ et non-consommateur $p_i^{NC} = P[d_i = 0] = P[FREQ_i = 0]$. Ces effets sont calculés à partir des expressions (A.3.4) à (A.3.6) et se présentent comme suit :

Consommateur régulier $p_i^R = [FREQ_i = 4$ (ou) $FREQ_i = 3] = dp_{i4} + dp_{i3}$ (A.3.7)

Consommateur occasionnel $p_i^0 = P[FREQ_i = 2$ (ou) $FREQ_i = 1] = dp_{i2} + dp_{i1}$ (A.3.8)

Non-consommateur $p_i^{NC} = P[d_i = 0] = P[FREQ_i = 0] = dp_{i0}$ (A.3.9)

Enfin, puisque la somme des probabilités doit être égale à 1, $\sum_{j=0}^{4} p_{ij} = 1$, la somme des quasi-élasticités de l'âge et la somme des variations de probabilités doivent être nulles pour chaque individu *i* de l'échantillon :

$$\forall i, \sum_{j=0}^{4} \frac{dp_{ij}}{dLog(\hat{a}ge)} = 0 \quad \text{et} \quad \sum_{j=0}^{4} dp_{ij} = 0$$

Calcul des effets marginaux moyens

Dans la littérature, les effets marginaux sont en général calculés au point moyen. Pour notre part, nous avons calculé pour chaque individu *i* de l'échantillon les variations de probabilités – relations (A.3.4) à (A.3.9) – pour toutes les variables explicatives $A = \{X,Z\}$ et la quasi-élasticité pour l'âge, puis nous avons calculé la moyenne de ces variations et quasi-élasticités.

Si l'ensemble des variables explicatives dichotomiques du modèle est indicé par *m*,

c'est-à-dire $A_m = \{X_m, Z_m\}$, la moyenne des variations de probabilités a été calculée uniquement pour les individus i qui ont $A_m = 1$.

Nous présentons dans le tableau 5 les variations de probabilités moyennes de l'échantillon calculées à partir des trois modèles :

modèle I : $(1/k)\sum_{i=1}^{k} dp_i$

modèle II : $(1/k)\sum_{i=1}^{k} dp_i^R$

modèle III : $(1/k)\sum_{i=1}^{k} dp_i^0$, $(1/k)\sum_{i=1}^{k} dp_i^{NC}$,

avec k l'effectif des individus pour lesquels $A_m = 1$.