

Fumer ou ne pas fumer ...

Un modèle de l'initiation sur données empiriques

Bertrand PIERRARD

CREDES

Les opinions et les résultats présentés dans ce travail n'engagent que l'auteur qui porte l'entière responsabilité d'éventuelles erreurs.

« Smoking is a bad thing...

Especially bad cigars ! »

(adapt. W. Churchill)

Résumé

L'addiction est une dépendance qui est à la fois cause et conséquence d'une consommation : le niveau de consommation actuelle dépend de la consommation passée et implique la consommation future. La théorie de l'addiction rationnelle explique cette dépendance par l'optimisation d'une fonction d'utilité (Becker et Murphy 1988) et recherche les variables de la fonction de demande du bien addictif.

Notre objectif est : d'une part d'étudier les conclusions de cette théorie dans le cas de l'addiction liée au tabac ; et d'autre part de les valider à partir de données empiriques françaises, avec un intérêt particulier pour les déterminants de l'initiation. En effet, les premières études menées sur l'addiction concernent essentiellement la consommation, et ce n'est que plus récemment que certains auteurs ont commencé à s'intéresser aux déterminants de l'initiation tabagique.

Les recherches ont privilégié le rôle du prix sur la consommation, sans que cela ne permette de conclure de manière tranchée. Par contre, l'impact du prix sur l'initiation, lorsqu'il apparaît dans les analyses, semble très faible. Ceci explique que certains auteurs, comme Douglas et Hariharan (1995), se soient intéressés à l'impact de caractéristiques individuelles socio-économiques.

Notre analyse descriptive des données de l'enquête Santé et Protection Sociale 1998 confirme l'importance de variables telles que la catégorie socioprofessionnelle, le sexe et l'âge ; et l'application des modèles économétriques semblent confirmer les conclusions de Douglas et Hariharan, à savoir une influence du prix sur la consommation mais pas sur l'initiation, renforçant ainsi le caractère explicatif des variables socio-économiques sur l'initiation tabagique.

Table des matières

Résumé.....	3
Introduction.....	5
1. Une revue des modèles théoriques de l’addiction rationnelle et des tests empiriques sur jeune population	7
1.1. Le modèle de Stigler et Becker.....	7
1.2. Le modèle de Becker et Murphy	9
1.2.1. Présentation	9
1.2.2. Résolution	9
1.2.3. Analyse sur cycle de vie.....	12
1.2.4. Conclusions du modèle	13
1.3. Le modèle de Orphanides et Zervos	13
1.3.1. Le modèle.....	14
1.3.2. Les conclusions	15
1.4. Le modèle de Douglas et Hariharan	16
1.4.1. Pourquoi commence-t-on à fumer ?	16
1.4.2. Le modèle de théorique	17
1.4.3. Les données	18
1.4.4. Le modèle empirique.....	18
1.4.5. Les résultats.....	20
1.5. Les tests empiriques sur l’initiation tabagique	20
2. Analyse descriptive des données de consommation tabagique.....	22
2.1. Le CREDES.....	22
2.2. L’enquête Santé et Protection Sociale (SPS).....	23
2.2.1. Objectifs et thèmes étudiés.....	23
2.2.2. La base de sondage.....	23
2.2.3. La méthode d’enquête	24
2.2.4. Les questionnaires et données recueillies.....	25
2.3. Les données	26
2.3.1. De la population enquêtée à la population étudiée.....	26
2.3.2. Qui sont les fumeurs ?.....	27
3. Modélisations et résultats.....	31
3.1. Les modèles logistiques.....	31
3.2. Le modèle de survie.....	32
3.3. La vraisemblance	33
Conclusion	37
Bibliographie.....	38

Introduction

Qu'est-ce que l'addiction¹ ? L'« addiction » vient du latin *ad-dictum* et était utilisé en français au Moyen-Age avant de s'exporter Outre-Manche. A cette époque, « [...] l'addiction était un arrêt de justice qui obligeait l'individu incapable de remplir autrement ses obligations envers un créancier, de payer avec son corps et son comportement, il devenait donc une sorte d'esclave. Esclave de son créancier tout comme aujourd'hui le toxicomane est esclave de sa drogue » (Loonis E. 1997). Au sens large, à l'heure actuelle, une addiction est une dépendance qui est à la fois cause et conséquence d'une consommation. L'addiction est, de plus, souvent opposée à un comportement rationnel.

Qu'est-ce que la rationalité ? Il en existe un grand nombre de définitions. Nous ne considérerons ici que la rationalité standard en économie, ou rationalité substantielle, telle que définie par Simon : « *Un comportement est substantivement rationnel lorsqu'il est approprié à l'accomplissement des buts donnés dans les limites imposées par des conditions et des contraintes données* » (Simon 1976, dans MBR II, 1982, p 425). La notion de rationalité renvoie donc aux décisions d'un individu qui se comporte toujours de façon à ce que les conséquences de ses actions maximisent ses préférences. La rationalité doit être entendue comme un postulat méthodologique qui permet de dégager des règles comportementales, en d'autres termes elle permet de postuler que les comportements ne sont pas le fruit du hasard. En économie, cette règle est la maximisation des préférences.

Il est alors possible de développer une théorie de l'addiction rationnelle qui explique la dépendance par l'optimisation d'une fonction d'utilité (Becker et Murphy 1988).

Nous nous intéresserons ici à l'addiction liée au tabac qui a motivé de nombreuses réflexions, au cours des 20 dernières années notamment, sous le couvert d'une volonté des pouvoirs politiques de maîtriser ce problème de santé publique. Les travaux existants portent principalement sur l'analyse des déterminants de la consommation tabagique et des coûts qu'elle engendre pour la société.

Ce travail se situe en amont de ces recherches puisqu'il vise à préciser, s'ils existent, les facteurs déterminants de l'initiation tabagique.

En première partie, nous exposerons les fondements économiques et les modèles économétriques de la théorie de l'addiction rationnelle. Après avoir présenté, dans une deuxième partie, les résultats de l'analyse descriptive réalisée à partir de l'enquête Santé et Protection Sociale 1998, nous utiliserons ces mêmes données dans des modélisations présentées en troisième partie. Les résultats obtenus semblent confirmer la thèse de Douglas et Hariharan : le prix n'aurait pas d'impact sur l'initiation tabagique.

¹ Le terme « addiction » est aujourd'hui un anglicisme et ne figure pas dans le dictionnaire français. Sa traduction renvoie au terme « assuétude » très peu usité. Nous parlerons donc « d'addiction » qui est le terme le plus répandu pour évoquer la dépendance.

1. Une revue des modèles théoriques de l'addiction rationnelle et des tests empiriques sur jeune population

Notons tout d'abord que les modèles que nous analysons ici concernent l'addiction au sens large et non pas réduite à quelques biens néfastes comme l'alcool, les drogues ou encore le tabac, même si ce dernier est le sujet qui nous occupe. L'addiction concerne tout bien dont la consommation provoque une dépendance, il peut donc s'agir aussi bien de la musique et son appréciation, que de la pratique d'une religion, de l'achat d'une marque de lessive particulière ou encore bien évidemment de la consommation de tabac.

Nous présenterons dans cette partie quatre modèles que nous avons jugé représentatifs de l'évolution de la pensée dans le domaine des comportements liés à des consommations dépendantes. Si le modèle de Stigler et Becker (1977) a été l'un des premiers à analyser ce phénomène d'addiction rationnelle en étudiant plus spécifiquement la consommation de musique, l'article « A theory of rational addiction » de Becker et Murphy (1988) est sans doute le texte fondateur d'une nouvelle manière d'analyser les comportements de dépendance. Orphanides et Zervos (1995) ont enrichi l'analyse en introduisant notamment le concept de regret dans la consommation d'un bien dont le consommateur peut-être dépendant. Enfin, Douglas et Hariharan ont testé la théorie de l'addiction rationnelle sur l'initiation tabagique en tentant d'en dégager les déterminants. Dans un dernier paragraphe, nous ferons une brève revue de la littérature portant sur l'initiation. Nous verrons que, très souvent, les études n'ont analysé qu'un effet des prix à travers les élasticités et que seuls Douglas et Hariharan ont cherché d'autres facteurs explicatifs.

1.1. Le modèle de Stigler et Becker

Stigler et Becker (1977) ont associé à l'addiction une hypothèse forte de rationalité, qui constituera la base des modèles d'addiction rationnelle, à savoir la stabilité des préférences des consommateurs² et

² Le titre de cet article : « Degustibus non est disputandum » est assez évocateur de ce postulat de base sur lequel s'appuient les auteurs, à savoir que les goûts des individus sont stables et les mêmes pour tout le monde (de la même manière que les montagnes rocheuses qui sont les mêmes pour tout le monde et tout le temps).

leur détermination par des circonstances exogènes de prix et de revenus³. Le plan du consommateur consiste alors en une maximisation de sa fonction d'utilité.

On peut dès à présent présenter la fonction d'utilité dans sa forme générale qui servira de cadre d'analyse pour l'ensemble des modèles d'addiction rationnelle $U(C; Y)$, où Y représente la consommation d'un bien composite et C est la consommation d'un bien addictif.

Le bien C sera défini comme addictif si sa consommation présente est dépendante de sa consommation antérieure par le biais d'un stock de capital addictif. Ce stock de capital addictif, créé par les consommations antérieures, correspond en quelque sorte à un apprentissage de la consommation du bien en question. Cet apprentissage influe donc inévitablement sur les consommations présentes et futures. En termes d'utilité, ceci implique donc que l'utilité marginale de la consommation présente par rapport à la consommation antérieure est positive.

L'exemple privilégié développé par les auteurs est l'écoute de la musique, ce qui correspond à une addiction « bénéfique ». Selon eux, le temps passé à écouter de la musique améliore l'appréciation que l'on en retire, l'entraînement permet de l'apprécier et augmente la productivité du temps consacré à sa consommation, ce qui implique une liaison croissante entre l'utilité marginale du bien « écoute de la musique » et son stock d'écoute antérieure.

A contrario, on peut observer une utilité marginale décroissante pour le bien « euphorie » lié à la consommation de drogues dures (type héroïne). Plus la consommation d'héroïne a été élevée dans le passé, plus il est difficile de produire la même quantité d'euphorie dans le présent, sauf à augmenter la quantité consommée. L'addiction peut alors être qualifiée de néfaste.

L'analyse est donc duale : d'une part, il est nécessaire de constater une addiction par un raisonnement en termes de quantité (ou consommation) ; d'autre part, l'étude de la rationalité s'effectuera à l'aide du concept d'utilité des individus.

³ Sensiblement à la même époque (1970), Pollak avait pris une voie d'analyse rigoureusement différente en introduisant dans son modèle la possibilité de modification des préférences d'un consommateur d'une période à l'autre.

1.2. Le modèle de Becker et Murphy

Becker et Murphy ont repris directement le modèle de Stigler et Becker, en incorporant, entre autres, les développements de Iannaccone (1984 et 1986). Ils rejoignent également des résultats de Pollak (1970, 1976), Ryder et Heal (1973), Boyer (1978, 1983) et Spynnewyn (1981) sur les habitudes.

1.2.1. Présentation

Tout individu possède une fonction d'utilité qui dépend de la consommation d'un bien addictif (c) et d'un bien composite (y) représentant l'ensemble des consommations « normales » de la vie courante.

$$U(t) = U(c(t), S(t), y(t)) \quad (1)$$

$$S \text{ et } c \text{ sont liés par la relation : } S'(t) = c(t) - \delta S(t) \text{ avec } S'(t) = dS/dt \quad (2)$$

où δ est le taux de dépréciation du stock de consommation.

La fonction d'utilité (1) sera considérée dans un premier temps comme quadratique en c , y et S ; nous verrons plus tard que les auteurs modifieront cette hypothèse.

L'équation (2) est une équation classique d'investissement, $c(t)$ joue donc à la fois sur l'utilité directe qu'elle procure et sur l'impact qu'elle aura sur le stock de consommation futur. En choisissant $c(t)$, le consommateur rationnel intègre donc à la fois ces deux effets (satisfaction immédiate et investissement en capital de consommation, donc en satisfaction future) ; c'est ce que les auteurs appellent le « *learning by doing* ». En d'autres termes, on peut assimiler la consommation du bien addictif à celle d'un bien durable.

1.2.2. Résolution

Le consommateur maximise donc ici la somme actualisée de son utilité décrite ci-dessus dans l'équation (1) :

$$U = \int_0^T e^{-\sigma t} \cdot U(t) dt$$

avec σ le taux de préférence pour le présent.

Cela revient à un problème d'optimisation dynamique que les auteurs résolvent par l'équation d'Euler en raison de la difficulté de calcul des dérivées dans ce modèle.

i.e. on maximise une fonction de la forme : $\int F[t, x(t), x'(t)]dt$ ($x'(t) = dx/dt$)

Selon l'équation d'Euler : si $x(t)$ est une courbe extrémale, elle doit vérifier :

$$\partial F/\partial x = d[\partial F/\partial x']/dt$$

De ce programme de maximisation sous contrainte de budget, les auteurs déduisent une solution dynamique unique :

$$S(t) = (S_0 - S^*)e^{\lambda t} + S^* \quad (3)$$

- S^* représente le niveau d'équilibre stable obtenu en partant de la situation initiale S_0 et en suivant le chemin optimal.
- λ est la racine négative solution de l'équation différentielle.

L'ensemble des points d'équilibre est donné par la droite $c = \delta S$ (hypothèse de stabilité temporelle) sur laquelle la consommation présente est égale au stock de consommations antérieures pondéré par un coefficient de dépréciation fixe. A l'aide de cette relation, nous pouvons écrire l'équation donnant la consommation courante en fonction du stock de consommations passées :

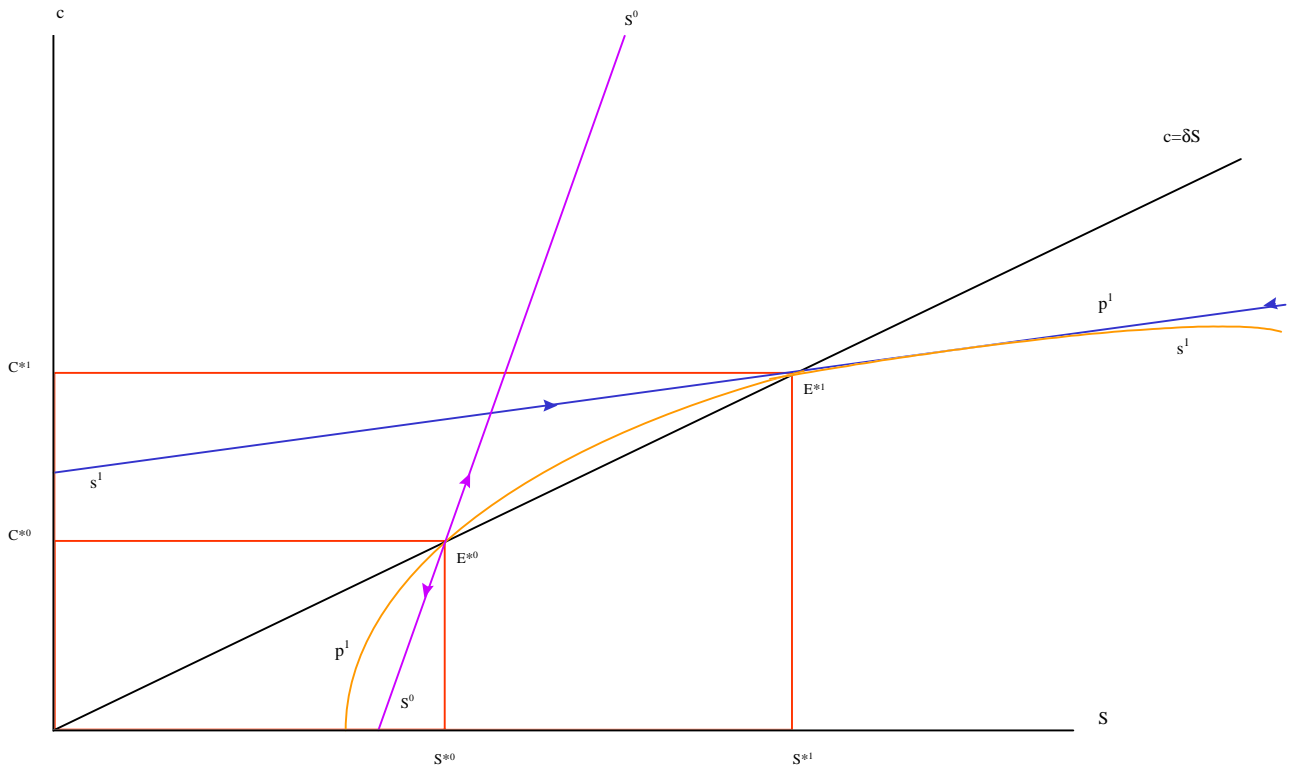
$$c(t) = (\delta + \lambda)S(t) - \lambda S^* \quad (4)$$

La dépendance, qui se traduit par une utilité marginale positive de la consommation présente par rapport à la consommation passée, peut également se décrire comme une complémentarité intertemporelle des consommations d'un bien⁴ (*adjacent complementarity*). Pour un bien classique, une préférence pour le présent importante conduit à privilégier une consommation immédiate des biens au détriment d'une consommation future ; alors que lorsqu'il s'agit d'un bien à dépendance, la préférence pour le présent augmente la consommation future par un effet d'entraînement du S dans la fonction d'utilité.

Le graphique ci-dessous (Becker et Murphy, 1988, p 681) présente les équilibres du modèle et analyse leur stabilité.

⁴ Ainsi lors d'une étude dynamique du modèle, le choix du taux de préférence pour le présent des individus doit être effectué avec beaucoup de précautions sous peine de biaiser l'analyse des comportements qui sera faite par la suite.

Graphique n°1 : Les équilibres dans le modèle d'addiction rationnelle



Le graphique tente d'explicitier la relation existante entre S (stock de consommation d'un bien addictif) et c (sa consommation présente). L'étude s'effectue en observant l'effet d'une légère variation de S sur chacune des situations d'équilibre décrites par les points E^{*0} et E^{*1} positionnés sur la droite d'équilibre $c = \delta S$.

- E^{*0} est un équilibre instable : si la complémentarité adjacente est trop forte nous nous situons dans ce cas de figure sur la droite S^0 dont la pente est supérieure à celle de la droite d'équilibre, une variation à la baisse de S fait alors tendre la consommation présente (c) vers zéro et une variation à la hausse fait tendre la consommation vers l'infini. Cette conclusion est inacceptable car elle met en lumière une possibilité d'état de la nature (la consommation infinie) en dehors de toute réalité. Les auteurs supposent donc une fonction quadratique qui serait une approximation d'une vraie fonction d'utilité cubique en S (avec un terme en S^3) ; il peut donc apparaître un terme négatif qui permet de faire diminuer la complémentarité adjacente quand S augmente. Par conséquent, la courbe $c(S)$ qui passe au-dessus du point d'équilibre se trouve modifiée et passe donc par dessous l'équilibre en E^{*1} pour lequel c^{*1} et S^{*1} sont supérieurs, offrant donc un équilibre stable.

- E^{*1} est un équilibre stable : nous nous situons dans ce cas de figure sur la droite s^1 dont la pente est inférieure à δS , une variation positive ou négative de S fait alors tendre la consommation présente vers le point d'équilibre E^{*1} .

En conclusion, la consommation peut diverger d'un équilibre instable, soit vers zéro, soit vers un équilibre stable. Ces courbes sont notées p_i sur le graphique car la solution S^* dépend du niveau du prix p_c (prix actuel du bien addictif). Nous verrons plus loin que certains auteurs ont remis en cause l'hypothèse de cohérence intertemporelle des préférences en considérant des préférences individuelles hyperboliques, et que cela permet d'expliquer plus naturellement l'impossibilité de consommations infinies.

1.2.3. Analyse sur cycle de vie

Le modèle sous cette forme explique le comportement des fumeurs mais ne permet pas d'en expliquer l'origine, *i.e.* pourquoi certains individus ont un stock initial non nul. L'idée des auteurs a donc consisté à introduire une variable événement $Z(t)$ représentant le début d'une carrière tabagique, cette variable s'ajoute à la formule de constitution du stock :

$$S^{\circ}(t) = c(t) - \delta S(t) + Z(t) \quad (2bis)$$

Becker et Murphy font jouer à ce paramètre un rôle de prix et seulement de prix. Il influe donc directement sur les courbes du graphique n°1 en faisant passer un individu d'une courbe à l'autre. Or, il semblerait plus légitime d'attribuer à Z un rôle dans la détermination des préférences et leur modification d'une période à l'autre. Nous reviendrons sur ce concept d'initiation tabagique avec la présentation du modèle de Douglas et Hariharan.

Si les auteurs n'ont pas expliqué l'initiation, ils ont poursuivi leur raisonnement pour expliquer l'arrêt de la consommation et particulièrement l'arrêt brutal. Selon eux, si la complémentarité adjacente est suffisamment forte, l'utilité devient alors convexe en $(c ; S)$. Dans ce cas, les solutions de l'équation différentielle sont complexes et le point d'équilibre instable E^{*0} dans le graphique devient un point de discontinuité qui se traduit par un « décrochement » de la courbe p_i . Ainsi une légère variation de la consommation passée ou présente peut entraîner le passage à un point de la courbe situé en dessous du point d'équilibre et donc provoquer un arrêt de la consommation.

1.2.4. Conclusions du modèle

Les auteurs ont montré que l'équilibre de consommation d'un bien addictif est instable quand le degré d'addiction est élevé et la probabilité de devenir dépendant est d'autant plus forte que la préférence pour le présent est importante. De plus, certains facteurs, comme par exemple le niveau de revenu, les événements ponctuellement stressants de la vie courante et le niveau de l'évolution des prix, sont des déterminants significatifs de la vraisemblance de devenir dépendant. Sur le prix, les auteurs ont montré qu'une modification permanente des prix a un impact relativement faible sur la consommation à court terme alors que son effet est plus important sur la consommation à long terme. Selon Becker et Murphy on peut également supposer que l'anticipation d'une augmentation des prix futurs diminue la consommation courante en raison de la complémentarité, ce qui implique un effet supérieur d'une augmentation permanente à celui d'un accroissement ponctuel. Enfin, le modèle démontre que l'on ne se sépare d'une telle habitude de consommation que par un comportement d'arrêt brutal (*cold turkey*). Si l'addiction rationnelle pose les bases d'une analyse permettant des conclusions intéressantes, deux critiques fondamentales ont été développées : l'une tient au programme de maximisation par la formule d'Euler, et l'autre à la forme et à l'interprétation de la courbe p_i dans le graphique (Ferguson 2000). En effet, cette courbe n'est pas clairement légitimée dans l'article et sa présence pose de sérieux doutes dans l'analyse et les conclusions des auteurs.

1.3. Le modèle de Orphanides et Zervos

Partant des analyses de Stigler et Becker (1977) et de Becker et Murphy (1988), Orphanides et Zervos (1995) modèrent l'analyse de l'addiction rationnelle en avançant des doutes sur la compatibilité de ce concept avec l'idée de regret chez les individus. Ils mettent en avant la critique de Winston (1980) selon lequel Stigler et Becker développent une théorie de " l'addiction-heureuse " d'un consommateur qui a soigneusement choisi son addiction après observation et évaluation de toutes les autres. Cette idée est d'ailleurs reprise par Akerlof (1991) pour qui les individus deviennent addictifs " en toute connaissance de cause ".

Les auteurs présentent un modèle dans lequel les individus maximisent des préférences stables pour des biens au pouvoir addictif — pouvoir qu'ils ne connaissent pas à l'avance. Malgré cette incertitude,

les comportements des individus sont optimaux dans le temps, les personnes dépendantes peuvent regretter leur décision passée d'initiation aux substances addictives et leur évaluation initiale de la potentialité nuisible du bien.

L'apport d'Orphanides et Zervos réside dans les trois postulats suivants :

- la consommation d'un bien addictif n'est pas nuisible à tous les individus de manière équivalente ;
- les individus ont des croyances subjectives concernant ces nuisances ;
- ces croyances sont constamment mises à jour au cours des périodes de consommation des biens en question.

Prenant l'exemple de l'alcool, on peut observer qu'une même quantité donnée consommée provoquera un simple effet euphorique chez les uns et une sérieuse dépendance chez les autres. La prédisposition plus ou moins forte d'une addiction à l'alcool ne peut donc être connue qu'après essai de chacun des individus. A l'inverse, certains biens, comme l'héroïne ou la cocaïne, ont un potentiel de pouvoir addictif suffisamment fort pour décourager tout essai à ces substances⁵. Ainsi les individus ne sont pas égaux devant les différentes addictions tant dans les consommations que devant leur appréhension, et de plus, leur situation personnelle leur est inconnue *ex ante*. L'analyse des auteurs met donc en avant le rôle important des croyances et convictions des individus en face de produits addictifs et de leurs effets, et accorde par conséquent une place prépondérante à l'information et l'éducation de la population.

1.3.1. Le modèle

Il consiste en l'analyse de la consommation de deux biens : un bien ordinaire et un bien potentiellement addictif. Le bien addictif est, comme dans les modèles béckeriens, à la fois cause et conséquence du stock de capital dépendance.

Dans le modèle, tous les individus maximisent la même fonction d'utilité mais il existe deux sous-populations : l'une ayant une tendance à devenir dépendante (les "potentiellement dépendants"), et

⁵ Il s'agit là d'une simple hypothèse des auteurs : en fait, les résultats cliniques montrent que la dépendance à doses faibles est moindre pour la cocaïne ou l'héroïne que pour la cigarette; s'il y a moins d'essais à ces deux substances, cela serait dû à leur prix et à leur caractère prohibé.

l'autre non (les "non dépendants"). Les individus sont supposés adopter un comportement bayésien, c'est-à-dire qu'ils intègrent l'information issue de leur expérience dans leur traitement de l'incertitude.

Des hypothèses de continuité et de dérivabilité des fonctions d'utilité et de complémentarité des consommations du bien addictif entre les différentes périodes sont nécessaires à la détermination d'un optimum.

Pour les auteurs, les regrets concernant leur consommation passée et la sous-estimation initiale de leur capacité à devenir dépendant, sont motivés par une information présente plus importante que celle de la période dans laquelle les individus ont pris la décision de consommer le bien addictif. L'analyser des comportements selon des critères de choix permet à Orphanides et Zervos de conclure leur article en étudiant le problème sous l'angle de la valeur de l'information, de l'éducation et des politiques publiques. Une meilleure éducation et information minimisent la probabilité de regrets *ex post* puisque le choix d'initiation de la consommation est supposé être effectué à l'aide d'une meilleure connaissance du marché et de ses produits, donc du risque de dépendance.

1.3.2. Les conclusions

En résumé, les auteurs mettent en évidence les rôles importants de l'expérimentation de la consommation et des croyances subjectives *a priori*. Théoriquement, en se plaçant dans le cadre de la rationalité limitée, ils expliquent la contradiction entre la volonté de consommer un bien provoquant une addiction et les regrets postérieurs à celle-ci. Concrètement, ils défendent le rôle fondamental des politiques publiques dans leur capacité de réglementation de la publicité et dans leur fonction de réhabilitation des individus dépendants (héroïnomanes...). Enfin, ils mettent également en lumière la raison des nuisances supérieures provoquées par des biens dont l'addiction est moindre (ex : alcool...) par rapport aux drogues illicites par exemple, les croyances subjectives *a priori* conduisant à une moindre appréhension des conséquences de l'initiation. Ces appréhensions semblent être différentes selon les caractéristiques des individus, comme le montrent Douglas et Hariharan dans leur modèle présenté ci-après.

1.4. Le modèle de Douglas et Hariharan

1.4.1. Pourquoi commence-t-on à fumer ?

Dans un article de 1994, Douglas et Hariharan ont présenté un modèle par lequel ils expliquent le phénomène d'initiation à la consommation de tabac. La littérature sur l'addiction étudie principalement les comportements des fumeurs, leur consommation, leur coût pour la société, leur disposition à payer afin d'éviter une morbidité voire une mortalité liée à la consommation de tabac, etc. Or, il semblerait plus logique de s'intéresser d'abord aux causes de ce comportement puis à leurs conséquences. En effet, il paraît plus logique d'empêcher les gens de commencer à fumer plutôt que de tenter de les persuader d'arrêter une fois que la dépendance s'est installée. La littérature sur ce thème de l'initiation au tabac est paradoxalement relativement pauvre. On pourra toutefois citer Lewitt et Coate (1982) et Wasserman et col. (1981) dont nous exposerons les résultats dans un paragraphe à la fin de cette partie.

Le modèle de Douglas et Hariharan (1994) est un modèle de durée en deux sous-populations (Schmidt et Witte 1989) qui prend ses bases dans la théorie de l'addiction rationnelle (Becker et Murphy, 1988). Ils étudient, la probabilité de ne jamais fumer tout en observant la date à laquelle la dépendance est apparue si la personne enquêtée fume⁶.

Alors que Lewitt et Coate (1982) et Wasserman et col. (1981) n'avaient observé le début de consommation qu'à travers des élasticités prix et revenus de la décision d'être fumeur, le modèle de Douglas et Hariharan est novateur puisqu'il recherche d'autres aux déterminants de cette décision.

⁶ En fait, les auteurs ont tout d'abord testé l'effet prix décrit dans le modèle de Becker et Murphy sur l'initiation tabagique, mais n'ont pas trouvé de résultat satisfaisant. Ils se sont alors tournés vers des variables explicatives de type « culturel ».

1.4.2. Le modèle de théorique

On considère tout d'abord une fonction d'utilité concave :

$$U_t = U(C_t, Y_t, S_t, L_t) \quad (1)$$

avec : C_t la consommation du bien addictif dont le prix est P_t ,

Y_t la consommation numéraire non addictif dont le prix est l'unité,

S_t le stock de capital addictif accumulé,

L_t les autres variables du cycle de vie qui affectent l'utilité.

Le stock de capital addictif (nul à la période initiale) de la période suivante se déprécie au taux γ mais s'accroît par l'intermédiaire de la consommation courante du bien dont on est dépendant.

$$S_{t+1} = (1 - \gamma)S_t + C_t \quad (2)$$

Faisant l'hypothèse d'une durée de vie infinie, l'utilité actualisée sur cycle de vie s'écrit :

$$V_t = \sum_{i=t}^{\infty} \frac{1}{(1 + \rho)^{i-t}} U_i \quad (3)$$

avec : ρ le taux de préférence pour le présent .

La contrainte de budget actualisée sur cycle de vie est déterminée par la richesse à la date t et se définit comme suit :

$$\sum_{i=t}^{\infty} \frac{1}{(1 + r)^{i-t}} (Y_i + P_i C_i) = \sum_{i=t}^{\infty} \frac{I_i}{(1 + r)^{i-t}} \equiv A_t \quad (4)$$

où : r est le taux d'intérêt,

I_i est le revenu à la date i ,

A_t est la richesse à la date t .

L'équation (4) égalise la somme actualisée des dépenses en bien addictif et en bien non addictif avec la somme actualisée des revenus qui est équivalente à la richesse actuelle des individus.

S'intéressant à la décision de début de consommation de tabac, un individu commence à fumer si et seulement si le bénéfice marginal de la première cigarette est supérieur à son coût marginal. Ceci se traduit mathématiquement par la relation :

$$MB_C(C_t, Y_t, L_t | S_t = 0) > MC_C(C_t, Y_t, L_t | S_t = 0) \quad (5)$$

où : MB_C est le bénéfice marginal décompté restant sur cycle de vie de la consommation de cigarette

MC_C est le coût marginal décompté restant sur cycle de vie de la consommation de cigarette

1.4.3. Les données

Les données utilisées par les auteurs sont issues de la NHIS (*National Health Interview Survey*) de 1978 et 1979. Le modèle a pour but de déterminer une liaison entre la présence d'une dépendance au tabac, l'âge d'occurrence de cette dépendance et quelques variables démographiques comme l'âge, le sexe, le niveau de vie de la famille, etc.

La NHIS comporte des informations sur les aspects sociaux, démographiques et économiques de la santé de la population des Etats-Unis. Un questionnaire supplémentaire a été ajouté en 1978 et 79 comportant des questions sur le tabagisme (fumeur ou non, quantité de cigarette/tabac consommée, âge d'initiation, etc.). Les données de prix proviennent du *Tobacco Institute* qui a calculé une moyenne pondérée du prix des paquets de cigarettes par Etat et par an pour la période 1954-1991 (les auteurs ont ensuite pondéré les prix par l'indice de prix à la consommation).

1.4.4. Le modèle empirique

Le modèle de durée en deux sous-populations pondère la vraisemblance de chaque observation en utilisant la probabilité estimée que l'individu va commencer à fumer ou non (une première analyse des données avait permis d'observer que 49 % de l'échantillon enquêté n'avait jamais fumé).

Soit :

- z_i et x_i les vecteurs des caractéristiques socio-démographiques de l'individu i ,
- θ et β des vecteurs de paramètres. θ : détermine la probabilité qu'un individu commencera à fumer un jour ; β : détermine l'âge auquel l'individu commencera (si jamais il commence). Ces paramètres sont déterminés par le modèle de sous-populations,
- $G(z_i, \theta)$ la probabilité que l'individu i fume un jour,
- T_i l'âge auquel l'individu i a pris l'habitude de fumer,
- $f(T_i; x_i, \beta)$ la valeur de la fonction de densité de l'âge d'initiation évaluée à la date T_i ,
- $S(t; x_i, \beta) \equiv \int_t^{\infty} f(T; x_i, \beta) dT$ la fonction de survie liée à l'interview à la date t . La survie est ici un terme générique économique, elle représente en fait le complémentaire de la fonction de densité

et s'interprète comme la représentation des individus qui n'ont pas encore commencé à fumer mais qui vont le faire,

- δ_i un indicateur égal à 1 si l'individu a été fumeur avant ou à la date t et 0 sinon.

La contribution de l'individu i à la fonction de vraisemblance est alors :

$$L_i(\theta, \beta; T_i, z_i, x_i) = \delta_i \cdot G(z_i, \theta) \cdot f(T_i; x_i, \beta) + (1 - \delta_i) \cdot [(1 - G(z_i, \theta)) + G(z_i, \theta) \cdot S(t; x_i, \beta)] \quad (6)$$

⇒ **Si $\delta_i = 1$ (i.e. si i déclare avoir commencé à fumer en $T_i \leq t$)** : la vraisemblance de l'individu est alors égale au produit de la probabilité d'être fumeur un jour et de la valeur de la fonction de densité ;

⇒ **Si $\delta_i = 0$ (i.e. si i déclare ne jamais avoir fumé en t)** : la vraisemblance de l'individu est alors égale à la somme de la probabilité que l'individu ne sera jamais fumeur et de la probabilité qu'il sera dépendant au tabac un jour.

En raison de la forme de la distribution de la variable endogène de participation tabagique ($f(T_i; x_i, \beta)$), le modèle de vraisemblance sera évalué sous une forme logistique.

Les variables exogènes sont composées de données relatives au revenu, aux prix présents et futurs anticipés, à la démographie (statut marital, race, sexe, âge et âge² pour contrôler la non auto-corrélation du modèle), à la connaissance des coûts et méfaits du tabac.

Dans le modèle, entre également en ligne de compte le prix du tabac. Les prix passés et actuels influencent le fumeur et le non-fumeur. Pour le fumeur sa consommation présente dépend des prix passés à travers leurs effets sur le stock de capital addictif. Un non fumeur, quant à lui — s'il n'est pas un consommateur myope⁷ —, anticipe un prix futur en fonction des prix passés et présents associés à des actions gouvernementales éventuelles sur les taxes ; ce qui constitue un déterminant important de son initiation tabagique potentielle puisqu'il est un facteur direct jouant sur les coûts et bénéfices marginaux.

⁷ Un consommateur myope se définit comme un consommateur qui ne se soucie pas des conséquences futures de ses consommations présentes.

Les données empiriques nous montrent que l'initiation tabagique est la plus importante entre 15 et 25 ans (surtout entre 17 et 20). Cela conduit à introduire des variables de prix adaptées aux besoins de l'analyse et conformes aux intuitions que l'on veut valider empiriquement. Les adolescents étant très sensibles aux variations de prix autour de 17 et 20 ans, les auteurs ont considéré une variable PRICE18 qui représente le log du prix des cigarettes affecté à l'individu l'année de ses 18 ans. Une augmentation de ce prix retarde ou empêche donc en théorie l'initiation tabagique. De plus, pour contrôler la perception de l'augmentation du prix par l'individu, les auteurs ont calculé une variable Δ PRICE18-15 qui devrait avoir le même effet que la variable PRICE18.

1.4.5. Les résultats

Le modèle fait apparaître que les variables : race, âge, sexe et niveau d'éducation ont un effet significatif sur l'initiation et sur la date d'initiation tabagique. Les individus ayant le niveau d'éducation le plus élevé sont les moins susceptibles de commencer à fumer. De la même manière, la variable « divorcé » a un effet très significatif en raison du stress qu'elle engendre chez les individus. Le revenu a également un effet mais cette variable contenant des « bruits », il faut l'interpréter avec prudence. Par contre, le prix n'a pas d'effet significatif. En fait, on observe un coefficient à la limite de la significativité mais une élasticité associée faible, ce qui contredit Wasserman et al. (1991).

1.5. Les tests empiriques sur l'initiation tabagique

L'initiation tabagique est un des points d'achoppement de la théorie de l'addiction rationnelle. En effet, si la consommation présente d'un bien addictif est liée à sa consommation passée, comment expliquer la première consommation pour laquelle il n'existe pas d'antécédent ? Becker et Murphy ont « résolu » cette lacune de leur modèle en considérant qu'un stock initial d'addiction se forme lors du passage d'un individu dans une situation stressante de sa vie ce qui attire celui-ci vers une consommation addictive : le bénéfice marginal devient supérieur au coût marginal. Douglas et Hariharan ont, quant à eux, développé un modèle spécifique permettant de tester l'initiation tabagique.

D'autres auteurs ont réalisé des tests empiriques portant notamment sur l'effet des prix (et leurs variations) sur la consommation des jeunes à travers le calcul d'élasticités.

En premier lieu, notons qu'il n'existe pas de consensus sur les résultats de ces tests : d'une part, certaines études ont montré aux Etats-Unis un rôle important de l'augmentation des taxes fédérales et nationales sur la consommation tabagique des adolescents (Lewitt et al., 1981 ; National Cancer Institute, 1993 ; Chaloupka et Grossman, 1996 ; Evans et Huang, 1998 ; Gruber, 2000) ; mais d'autre part, de nombreux autres articles ont infirmé cette relation (Chaloupka, 1991 ; Wasserman et al., 1991 ; Douglas et Hariharan, 1994 ; DeCicca et al., 1998).

Le premier article traitant de l'initiation (Lewitt et coll. 1981) décompose l'élasticité-prix en deux parties : l'élasticité de la participation (fumer ou non) et l'élasticité de la consommation (conditionnelle à une consommation déjà présente). Leurs résultats montrent que l'élasticité de la participation représenterait 60 % de l'élasticité totale et qu'elle concerne en majorité les 20-25 ans ; c'est-à-dire qu'une augmentation du prix du tabac diminuerait plus la probabilité que les jeunes commencent à fumer qu'elle ne diminuerait la consommation des fumeurs. Ces conclusions sont toutefois infirmées notamment par Wasserman et al. et par Douglas et Hariharan. Les résultats de ces derniers apparaissent comme ayant un attrait particulier puisqu'ils montrent que ce n'est pas le prix qui influence la consommation tabagique mais des variables d'ordre socio-économique. L'étude que nous avons réalisée à partir des données de l'enquête Santé et Protection Sociale et que nous présentons dans la partie suivante décrit les liens entre la consommation tabagique et certaines de ces variables. Elle permet d'apporter certains éléments de réponse sur la validité de cette hypothèse sur la population française.

2. Analyse descriptive des données de consommation tabagique

Les données que nous décrirons dans cette partie nous permettront de tester le modèle de Douglas et Hariharan dans une troisième partie. Il s'agit des données de l'enquête Santé et Protection Sociale (SPS) de 1998. Cette enquête a été réalisée par le CREDES (Centre de Recherche d'Etude et de Documentation en Economie de la Santé) auprès des ménages issus des trois principales caisses d'assurance maladie : CNAMTS (Caisse Nationale d'Assurance Maladie des Travailleurs Salariés), CANAM (Caisse Nationale des Professions Indépendantes) et MSA (Mutualité Sociale Agricole). Elle permet de recueillir des données socio-économiques et socio-démographiques ainsi que la morbidité et le niveau de protection sociale pour chaque personne composant le ménage.

2.1. Le CREDES

Le Centre de Recherche d'Etudes et de Documentation en Economie de la Santé a été créé le 30 janvier 1985 ; il est issu de l'ancienne division d'économie médicale du CREDOC.

Producteur de données et d'analyses en économie de la santé, le CREDES observe et analyse l'évolution des comportements des consommateurs et des producteurs de soins à la fois sous l'angle médical, économique et géographique.

Le CREDES conçoit et réalise des enquêtes périodiques ou ponctuelles sur des échantillons de population, d'institutions et de professionnels de santé, pour collecter des données sur la consommation et la production de soins. Des accords de partenariat lui permettent également d'exploiter des enquêtes sur des panels ou des échantillons réalisés par d'autres organismes (INSEE et IMS France par exemple). Les deux principaux axes de recherche du CREDES sont l'offre et la demande de soins, au niveau macro et/ou micro-économique. Parmi les thèmes abordés on trouvera par exemple : l'analyse de l'évolution de la morbidité, l'accès aux soins, l'étude des pratiques de prescription des médecins, la recherche et l'analyse des déterminants de la consommation médicale, etc.

Différents moyens et supports ont été développés par le CREDES pour faciliter l'accès à l'information. Il s'agit notamment du service de documentation et du logiciel Eco-Santé.

2.2. L'enquête Santé et Protection Sociale (SPS)

2.2.1. Objectifs et thèmes étudiés

Le CREDES réalise depuis 1988 l'enquête SPS. La finalité de l'enquête est de mieux connaître l'état de santé, le niveau de protection sociale et les consommations médicales des personnes résidant en France. Les différentes données recueillies auprès de chaque personnes sont à la fois :

- des renseignements socio-démographiques et socio-économiques : âge, sexe, profession, occupation principale, département de résidence, revenu, niveau d'étude, nationalité... ;
- des renseignements sur la protection sociale : régime de sécurité sociale, exonération éventuelle du ticket modérateur et ses motifs, possession d'une couverture complémentaire maladie... ;
- des données de morbidité : morbidité prévalente, facteurs de risque (dont le tabac)... ;
- des données de consommation médicale : pharmacie, séances de médecin...

L'ensemble de ces renseignements permettent d'analyser, entre autres, les thèmes suivants :

- les modes de protection sociale (obligatoires et facultatives),
- les maladies et les troubles de santé déclarés,
- les recours au médecin,
- la consommation de biens médicaux,
- l'hospitalisation,
- le renoncement aux soins et les opinions sur la santé...

2.2.2. La base de sondage

Depuis une vingtaine d'années, la CNAMTS a constitué un dispositif de suivi des dépenses de santé du Régime Général. Il s'agit d'un échantillon aléatoire de bénéficiaires du régime appelé : l'EPAS (Echantillon Permanent d'Assurés Sociaux). Cet échantillon, dont la composition a évolué en 1990 pour couvrir l'ensemble du territoire, permet d'obtenir des informations sur un échantillon au 1/1200^{ème} des assurés du régime général et de leurs ayants droit.

C'est en 1994 et en 1996 que deux échantillons du même type ont été créés respectivement pour les assurés sociaux du régime des professions indépendantes (la CANAM), et du régime des professions agricoles (la MSA). Nous disposons, depuis lors, d'un échantillon global représentatif de 95 % de la population française.

Font partie de l'EPAS, les assurés pour lesquels le numéro national d'identité vérifie simultanément les trois conditions suivantes :

- l'année de naissance est paire,
- le mois de naissance est octobre,
- les deux derniers chiffres du numéro de commune de naissance sont identiques aux deux derniers chiffres du numéro d'ordre, dans le même ordre ou dans l'ordre inverse.

Le tirage du sous-échantillon à enquêter est effectué par le Centre National de Traitement Informatique (CENTI) de la CNAMTS, et par les services informatiques de la CNAM, de la CANAM et de la MSA.

Il convient ici de noter un biais de sélection car nous travaillons sur des ménages ordinaires, c'est-à-dire les ménages ayant un logement. Les personnes sans domicile fixe ou vivant en institution ne sont pas interrogées.

Notons également que si la plupart des assurés constituant l'EPAS sont les mêmes depuis 1988, la composition de l'échantillon évolue. Par exemple, les étudiants correspondant aux critères entrent dans l'EPAS dès qu'ils passent du statut d'ayant droit à celui d'assuré.

2.2.3. La méthode d'enquête

On considère la population enquêtée sous deux unités différentes :

- ⇒ l'assuré principal : personne répondant aux critères de choix énoncés plus haut,
- ⇒ le ménage : comprend l'assuré principal, les ayants droit cohabitants et les autres personnes vivant sous le même toit mais ne dépendant pas de l'assuré principal pour l'assurance maladie.

On interroge donc l'ensemble des personnes citées ci-dessus selon deux modalités différentes :

- ① les assurés dont on a trouvé les coordonnées téléphoniques sont enquêtée par téléphone,
- ☞ les assurés pour lesquels les coordonnées téléphoniques n'ont pu être trouvées sont enquêtés en face-à-face (il s'agit ici des personnes sur liste rouge, habitant en concubinage avec une personne portant un nom différent ou sans ligne téléphonique).

Afin de tenir compte des variations saisonnières de la morbidité et des consommations de soins, chaque enquête se décompose en deux vagues successives d'effectif égal. La première vague a lieu au printemps entre mars et juin, et la seconde à l'automne entre octobre et décembre ; la répartition des adresses étant bien entendu effectuée de manière à maintenir la représentativité géographique des deux sous-échantillons.

2.2.4. Les questionnaires et données recueillies

Deux types de questionnaires forment l'ensemble de l'enquête : le questionnaire principal et les questionnaires auto-administrés.

⇒ Le questionnaire principal.

Il porte sur les caractéristiques de l'ensemble du ménage auquel l'assuré principal tiré au sort appartient. On y trouve notamment :

- ☞ des questions d'opinion (posées à une seule personne dans le ménage) relatives au système de santé et de protection sociale,
- ☞ les caractéristiques socio-démographiques et socio-économiques des membres du ménage,
- ☞ le niveau de revenu.

⇒ Les questionnaires auto-administrés.

☞ Le questionnaire « santé ». Chacune des personnes du ménage remplit son propre questionnaire. On y recueille la morbidité déclarée, la consommation de médicaments dans les dernières 24 heures, et une note de 0 à 10 que l'individu se donne concernant son état de santé et d'autres questions par exemple portant sur la consommation tabagique.

↳ Le questionnaire « couverture complémentaire ». On remplit autant de questionnaires qu'il y a de contrats de couvertures complémentaires signés auprès d'organismes dans le foyer. Sont recueillis ici le nom des organismes mutualistes, les prestations offertes par chaque contrat ainsi que les cotisations versées.

↳ Le carnet de soins. Il permet de relever pendant 30 jours les consommations médicales de tout le ménage, c'est-à-dire tous les soins ou examens médicaux reçus, tous les biens acquis et les hospitalisations.

↳ Le questionnaire « hospitalisation ». On recherche ici les hospitalisations survenues lors des trois derniers mois : dans quel établissement, quel service et pour quel motif.

↳ Le questionnaire « établissement spécialisé ». On recueille également ici l'établissement, le motif et le nombre de fréquentation dans des structures pour handicapés ou invalides.

↳ Le questionnaire « accident de la vie courante ». Il s'agit ici de répertorier les accidents de la vie courante survenus au cours des trois derniers mois et ayant fait l'objet d'une intervention par un professionnel de santé.

↳ Le questionnaire « contrat de suivi médical ». Pour les personnes de 16 ans et plus ayant signé un contrat de ce type, on souhaite obtenir leur opinion sur le système d'abonnement à un généraliste.

2.3. Les données

La base de l'enquête SPS 1998 comporte 23036 individus, mais seulement 12753 ont des caractéristiques exploitables pour l'étude qui nous intéresse. Il s'agit des personnes ayant accepté de participer à l'intégralité de l'enquête, ceci se traduit par le fait que le questionnaire santé et le carnet de soins sont remplis correctement et donc exploitables.

2.3.1. De la population enquêtée à la population étudiée

Les conditions que nous avons introduites afin de permettre une analyse fiable des résultats sont les suivantes :

- comme nous l'avons dit, le questionnaire santé et le carnet de soins doivent être remplis et exploitables (ce filtre élimine 7853 individus),

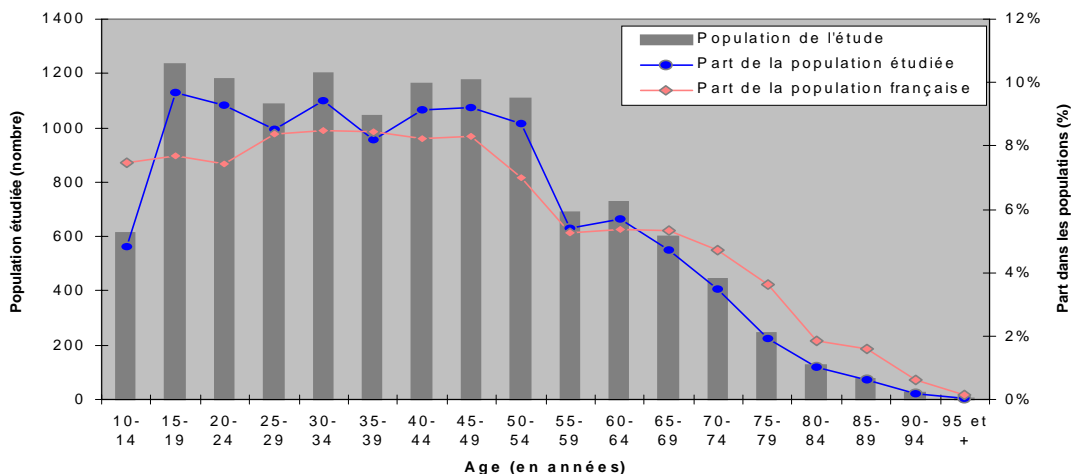
- la réponse à la question « Fumez-vous de façon habituelle ? » ne doit pas être inconnue (ce filtre élimine 1093 individus),
- les personnes enquêtées doivent être âgées de 12 ans et plus, âge en-dessous duquel les réponses ne sont certainement pas exploitables (ce filtre élimine 3360 individus).

Ainsi l'étude sera réalisée sur 12753 personnes de 12 ans et plus.

La population de l'étude est celle du panel 1998 de l'enquête SPS telle qu'elle a été décrite plus haut.

La représentativité de cet échantillon est illustrée dans le graphique suivant. Le diagramme représente le nombre d'individus contenu dans chaque tranche d'âge, puis les deux courbes qui représentent le pourcentage d'individus de chaque tranche d'âge dans la population enquêtée et dans la population générale de la France. On observe une certaine cohérence entre les deux populations ce qui prouve la représentativité de notre échantillon, du moins concernant la répartition de la population. La sous-représentation des 10-14 ans s'explique par le fait que nous n'avons considéré que des individus de 12 ans et plus dans notre échantillon.

Graphique n°2 : La représentativité de l'échantillon



2.3.2. Qui sont les fumeurs ?

On observe dans notre échantillon qu'un quart de la population fume et que 44 % fume ou a fumé de façon habituelle. Le graphique 3 donnera une répartition des fumeurs en fonction de leur âge.

La proportion de fumeurs par sexe et par âge n'est pas toujours significativement différente

En effet, même s'il existe une différence significative de la proportion de fumeurs hommes et femmes sur la population globale, une observation plus minutieuse par classe d'âge montre que l'écart significatif apparaît à partir de 35 ans seulement. En fait, les données nous montrent que les jeunes femmes commencent à fumer plus tôt que leurs camarades masculins mais également qu'elles « décrochent » plus volontiers peut être à l'occasion d'une grossesse ou bien parce que la gent féminine possède une plus grande sensibilité pour sa santé. Enfin, on peut signaler que la probabilité maximale d'être fumeur, quelque soit le sexe se situe entre 30 et 35 ans.

Le mieux on se sent ... le plus on se détruit

Comme nous l'avons déjà présenté, il existe dans l'enquête une question sur l'état de santé ressenti par les individus (une note de 0 à 10). Il semble que l'on fume d'autant plus que cet état ressenti est bon, du moins au niveau de l'ensemble de la population puisque les résultats par sexe ne sont pas significativement différents. Afin de compléter l'analyse sur cette question, il serait nécessaire de croiser ces résultats avec les variables recueillies dans le questionnaire santé à l'aide duquel nous pourrions étudier la morbidité de la population, mais ce point constitue à lui seul le sujet d'un autre mémoire !

Tabac, revenu et PCS⁸...

Les « plus gros » fumeurs sont les chefs d'entreprise, les professions libérales et les ouvriers alors que les commerçants et les agriculteurs fument en moindre proportion. De plus, ces « gros » fumeurs sont plus concentrés dans le deuxième quartile des tranches de revenu par unité de consommation, sans doute en raison de la classe des ouvriers dont la représentation est importante.

⁸ Profession et Catégorie Sociale

Mais qu'est-ce qu'un « gros fumeur » ? ... Les fumeurs sont-ils gros ?

Nous avons évalué un indice de masse corporelle ($IMC = \text{Poids} / \text{Taille}^2$). A partir de cet indice, il est possible d'affirmer que les fumeurs de cigarettes sont des gens présentant un IMC « normal » pour plus de 60 % d'entre eux (au sens des différentes classifications d'IMC de la littérature : $19 \text{ kg/m}^2 < IMC \text{ « normal »} < 25 \text{ kg/m}^2$). Les fumeurs de cigares quant à eux présentent plus volontiers une surcharge pondérale voire une obésité dans environ 55 % des cas. Cette notion est confirmée sur les fumeurs de cigarettes par la relation croissante qui existe entre IMC et nombre de cigarettes consommées quotidiennement.

En fait, on définira un « gros fumeur » comme un individu fumant plus d'un paquet par jour. A partir des déclarations de consommations quotidiennes, on observe que plus de 80 % des fumeurs consomment moins d'un paquet de cigarettes par jour.

Il est également extrêmement intéressant d'étudier l'âge des fumeurs selon leur consommation. Les individus déclarant fumer moins de 10 cigarettes par jour ont entre 15 et 35 ans (56,3 %), les individus fumant entre la moitié et un paquet par jour ont entre 25 et 45 ans (56,8 %) et les « gros » fumeurs ont entre 30 et 50 ans (64,6 %).

Les jeunes commencent à fumer de plus en plus jeune

Les réponses aux questions étudiées ici sont de type déclaratif, ainsi, il est nécessaire d'en appréhender l'analyse avec certaines précautions. En effet, lorsque l'on demande à un individu depuis combien d'années il fume, la réponse peut être approximative et le sera d'autant plus que la personne est âgée.

Toutefois, l'analyse de ces réponses permet d'observer les fréquences suivantes :

80 % des fumeurs de moins de 26 ans ont commencé avant 18 ans,

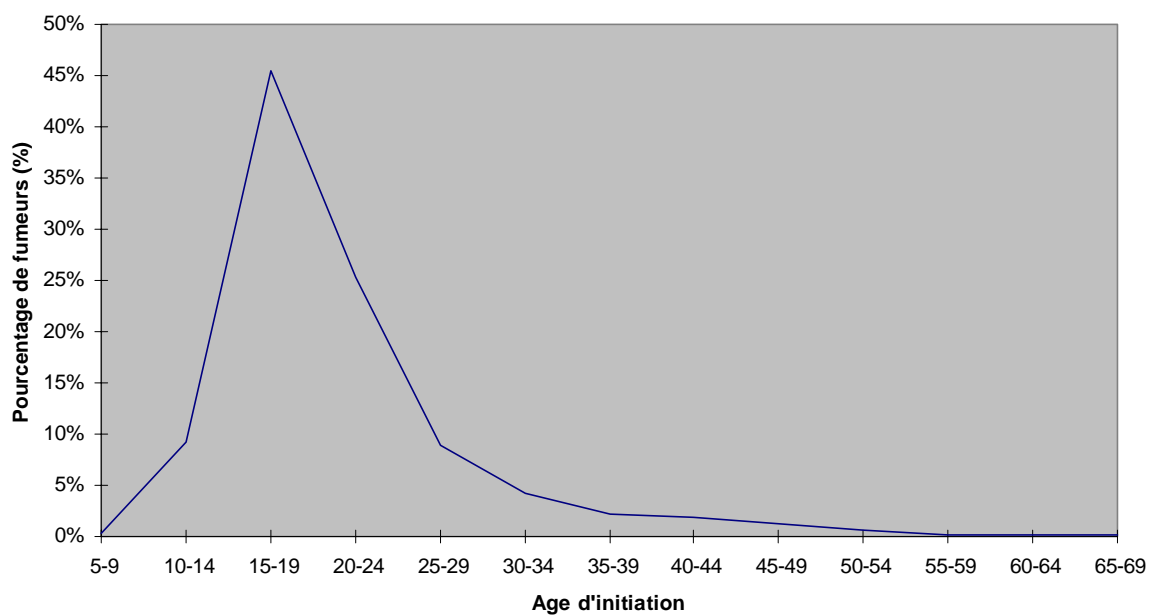
50 % des fumeurs âgés de 26 à 35 ans ont commencé avant 18 ans,

29 % des fumeurs âgés de 36 à 45 ans ont commencé avant 18 ans,

23 % des fumeurs âgés de 46 à 60 ans ont commencé avant 18 ans,

Le graphique ci-dessous permet de visualiser la répartition des fumeurs en fonction de leur âge d'initiation.

Graphique n°3 : Fréquence d'initiation tabagique par classe d'âge



La fonction de répartition est très explicite : 80 % des fumeurs interrogés avaient déjà effectué leur initiation à 25 ans. Cette proportion s'élève même à plus de 90 % à 30 ans. Malheureusement, ce calcul d'âge d'initiation n'est pas possible avec les personnes ayant fumé mais s'étant arrêtées, il aurait pourtant été intéressant de savoir combien, parmi ces 80 ou 90 % s'étaient arrêtées avant 35 ans par exemple, nous reviendrons sur ce problème de censure dans la partie suivante.

3. Modélisations et résultats

Outre les variables qui viennent d'être présentées, la modélisation de Douglas et Hariharan introduit le prix du tabac (prix du paquet de cigarettes lorsque l'individu était âgé de 18 ans et variation de ce prix sur les deux années avant ses 18 ans). Or, l'INSEE ne nous a fourni la série de prix que pour la période 1968-2000, ce qui réduit donc notre échantillon des deux côtés de la distribution : d'une part, en enlevant les individus de moins de 16 ans ; et d'autre part, les individus de plus de 46 ans en 1998. Nous présenterons dans cette partie tout d'abord des modèles logistiques expliquant différentes parties de notre échantillon (les fumeurs, les ex-fumeurs, et l'ensemble des gens ayant fumé ou fumant actuellement) ; puis un modèle dit « de survie » expliquant l'âge d'initiation tabagique ; et enfin un modèle de vraisemblance intégrant les deux modèles sus-cités.

3.1. Les modèles logistiques

Nous prendrons ici comme référence un homme marié. Les résultats de modélisations logistiques (*proc logistic* dans SAS) expliquant : premièrement le fait de fumer ; deuxièmement d'avoir fumé et troisièmement, le fait de fumer ou d'avoir fumé sont donnés dans le tableau ci-après.

Tableau n°1 : Les modèles logistiques.

Variables explicatives:	Fumeurs		Anciens Fumeurs		Fumeurs et anciens	
	Coefficients	Pr > ChiSq	Coefficients	Pr > ChiSq	Coefficients	Pr > ChiSq
Constante	3.7737	<.0001	2.2147	0.0197	3.2924	<.0001
Divorcé, séparé ou veuf	0.7012	0.0002	-0.3012	0.2589	0.4088	0.0234
Célibataire	-0.2824	0.0041	-0.7283	<.0001	-0.4004	<.0001
Femme	-0.5631	<.0001	-0.2412	0.0081	-0.4469	<.0001
Revenu	-0.00003	0.0009	-6.16E-6	0.4549	-0.00002	0.0049
Age	-0.0329	0.0089	-0.0376	0.0398	-0.0228	0.0459
Prix du tabac	-0.0860	<.0001	-0.1908	<.0001	-0.0905	<.0001
Indice Masse Corporelle	-0.0301	0.0019	0.0239	0.0493	-0.0116	0.1769
Note (état de santé)	-0.0986	<.0001	-0.0727	0.0219	-0.0910	<.0001
Nb de pers. dans le ménage	-0.0961	0.0003	-0.0942	0.0136	-0.0902	0.0002
Nationalité étrangère	-0.5517	0.0008	-0.9595	0.0004	-0.6571	<.0001

Il peut être étonnant de ne pas voir apparaître certains critères explicatifs dans ces modélisations mais ceci est dû à leur non significativité dans tous les modèles (ex. : le niveau d'étude) ou bien au fait

qu'ils apparaissent comme une combinaison linéaire des autres paramètres des modèles (ex : la catégorie socio-professionnelle).

Les cases grisées mettent en évidence les paramètres non-significatifs au seuil de 5 %. La plupart des coefficients des modèles montrent un effet négatif des paramètres sur la variable endogène. Ainsi, le fait d'être une femme, d'être âgé, de bien gagner sa vie, de se sentir en bonne santé, de faire partie d'une famille nombreuse ou encore d'être de nationalité étrangère sont des facteurs diminuant la probabilité de fumer. De plus, nous observons un effet négatif du prix sur la consommation tabagique (ce qui est tout à fait logique), mais il est plus important de noter le coefficient plus élevé chez les anciens fumeurs, ce qui signifie que leur plus grande sensibilité au prix est peut-être un facteur déterminant de l'arrêt de la consommation. Enfin, si être célibataire ne favorise pas la consommation tabagique, le fait d'avoir subi un événement stressant et donc être séparé, divorcé ou veuf sont des éléments fortement explicatifs de la consommation tabagique.

C'est, à peu de différences près (en dehors des deux variables de statut matrimonial), le premier modèle présenté dans le tableau 1 qui sera réutilisé dans la fonction de vraisemblance que nous présenterons plus bas.

3.2. Le modèle de survie

Le modèle dit « de survie » est estimé dans le logiciel SAS à l'aide d'une « *proc lifereg* ». Il teste l'âge d'initiation tabagique en fonction de plusieurs critères explicatifs. Ce modèle est donc estimé sur une population plus restreinte car elle ne comprend que les fumeurs actuels pour lesquels nous disposons de l'information : « depuis combien d'année fumez-vous ? ». Le résultat de cette procédure est donné dans le tableau 2.

Tableau n°2 : Le modèle de survie.

Variables explicatives:	Age d'initiation	
	Coefficients	Pr > ChiSq
Constante	2.92755	<.0001
Divorcé, séparé ou veuf	0.09662	0.0004
Célibataire	0.08400	<.0001
Revenu	2.98501E-6	0.0027
Prix du tabac	-0.02563	<.0001
Note (état de santé)	0.01109	0.0033
Nb de pers. dans le ménage	0.01462	0.0001

La quasi totalité des variables explicatives ont un effet significativement positif sur l'âge d'initiation tabagique. Deux éléments sont cependant particulièrement intéressants à noter dans ces résultats. Tout d'abord, l'effet du niveau de revenu : même si le coefficient est significatif, il est extrêmement faible ($2.98E-6$) ; l'influence de cette variable est donc très limitée. Et deuxièmement l'effet négatif du prix sur l'initiation conforte les résultats de l'étude de Emery (Emery 2000) par exemple en montrant une insensibilité au prix des individus commençant à fumer ; en effet, les jeunes qui commencent à fumer n'achètent pas leurs cigarettes eux-mêmes, ils sont donc insensibles au prix. Ce n'est qu'une fois la dépendance acquise qu'ils s'intéressent à l'achat de tabac et deviennent donc sensibles au prix comme nous l'avons montré dans les modèles logistiques.

3.3. La vraisemblance

Afin de compléter l'analyse sur les fumeurs et particulièrement sur l'initiation tabagique, nous avons adapté le modèle de vraisemblance de Douglas et Hariharan décrit en première partie. En effet, c'est à l'aide d'un modèle de ce type qu'il est possible de prendre en compte la censure à gauche⁹ de notre exemple.

La modélisation, telle que les auteurs la présentent dans leur article, ne peut être transposée directement. En effet, ils ont utilisé une fonction de vraisemblance maximisée à l'aide du logiciel LIMDEP dont nous ne disposons que d'une version d'évaluation, limitée à des échantillons de 1000 individus. Il a donc fallu adapter leurs équations afin d'effectuer une résolution sous SAS qui n'est pas programmé pour la maximisation de vraisemblance. Ceci signifie que la réalisation de certaines étapes de cette maximisation s'est effectuée à la main et à l'aide d'une programmation spécifiquement adaptée dans le logiciel SAS.

La fonction de vraisemblance pour un individu, dans notre cas, s'écrit comme suit :

$$L_i(\theta, \beta; T_i, z_i, x_i) = [G(z_i; \theta) \cdot f(T_i; x_i, \beta)]^{\delta_i} \times [1 - G(z_i; \theta) + G(z_i; \theta) \cdot S(T_i; x_i, \beta)]^{(1-\delta_i)}$$

⁹ La censure à gauche est présente en raison des anciens fumeurs pour lesquels il nous est impossible de déterminer l'âge d'initiation ni l'âge d'arrêt de la consommation addictive.

Soit, sous sa forme logarithmique :

$$\text{Log}L_i = \delta_i \cdot [\log G(z_i; \theta) + \log f(T_i; x_i, \beta)] + (1 - \delta_i) \cdot [\log(1 + G(z_i; \theta) \cdot (S(T_i; x_i, \beta) - 1))]$$

avec :

- z_i et x_i les vecteurs des caractéristiques de l'individu i ,
- θ et β des vecteurs de paramètres. θ : détermine la probabilité qu'un individu commencera à fumer un jour ; β : détermine l'âge auquel l'individu commencera (si jamais il commence). Ces paramètres sont déterminés par le modèle de sous populations,

- $G(z_i, \theta)$ la probabilité que l'individu i fume un jour : $G(z_i, \theta) = \frac{1}{1 + \exp^{(-z_i \theta)}}$

- T_i l'âge auquel l'individu i a pris l'habitude de fumer,

- $f(T_i; x_i, \beta)$ la valeur de la fonction de densité de l'âge d'initiation évaluée à la date T_i considérée ici à l'aide d'un modèle log-logistique :

$$\text{soit } \begin{cases} T_i = \exp^{(x_i \beta + \sigma \varepsilon)} \\ \varepsilon \approx \frac{\exp^\varepsilon}{(1 + \exp^\varepsilon)^2} \end{cases}$$

$$\text{donc : } f(T_i; x_i, \beta) = \exp^{(x_i \beta)} \cdot \left[\frac{\exp^{\frac{1}{\sigma} (\log T_i - x_i \beta)}}{(1 + \exp^{\frac{1}{\sigma} (\log T_i - x_i \beta)})^2} \right]$$

- $S(t; x_i, \beta) \equiv \int_t^\infty f(T; x_i, \beta) dT$ la fonction de survie liée à l'interview à la date t . La survie est ici un terme générique économique, elle représente en fait ici, en quelque sorte, le complémentaire de la fonction de densité et s'interprète donc comme la représentation des individus qui n'ont pas encore commencé à fumer mais qui vont le faire. Dans notre exemple, celle-ci s'écrit :

$$S(T_i; x_i, \beta) = \frac{1}{1 + [t \cdot \exp^{-x_i \beta}]^{\frac{1}{\sigma}}}$$

- δ_i un indicateur égal à 1 si l'individu a été fumeur avant ou à la date t et 0 sinon.

Au final, la log-vraisemblance à maximiser s'écrit comme suit :

$$\log L_i = \delta_i \cdot \left[\log \left(\frac{1}{1 + e^{-z_i \theta}} \right) + x_i \beta \cdot \frac{e^{-\frac{\log T_i - x_i \beta}{\sigma}}}{\left(1 + e^{-\frac{\log T_i - x_i \beta}{\sigma}} \right)^2} \right] + (1 - \delta_i) \cdot \left[\log \left[1 + \frac{1}{1 + e^{-z_i \theta}} \cdot \left(\frac{1}{1 + (t \cdot e^{-x_i \beta})^{1/\sigma}} - 1 \right) \right] \right]$$

Les vecteurs paramètres (β et θ) liés aux caractéristiques des individus (x_i et z_i) ainsi que le paramètre σ (qui détermine la forme initiale de la distribution de l'initiation tabagique) ont été estimés respectivement par le modèle de survie et le modèle logistique présentés dans les deux paragraphes précédents. Le faible nombre de paramètres provient de critères de sélection sévères en raison du fait qu'il sera nécessaire de calculer chacune des dérivées partielles de la log vraisemblance à la main ; mais également, et surtout, car un trop grand nombre de paramètres constitue un critère favorisant la non-convergence du modèle de vraisemblance.

La méthode de résolution de cette log-vraisemblance s'est effectuée à l'aide d'une « *proc nlin* » dans le logiciel SAS qui minimise l'inverse de la fonction de vraisemblance et des dérivées partielles de premier ordre par rapport à chacun des coefficients des vecteurs β et θ . Les coefficients de départ ainsi que les résultats sont présentés dans le tableau 3 ci-dessous.

Tableau n°3 : Les résultats de la fonction de vraisemblance.

Coefficients de θ	Valeurs initiales		Coefficients de la vraisemblance		
	Coefficients	Pr > ChiSq	Coefficients	Intervalles de confiance (95%)	
Constante	3,6106	<,0001	3,6159	1,9068	5,325
Femme	-0,5205	<,0001	-0,4373	-0,6003	-0,2744
Revenu	-0,00003	0,0004	-0,00003	-0,00006	-1,17E-07
Age	-0,0273	0,0274	-0,0273	-0,0609	0,00624
Prix du tabac	-0,0986	<,0001	-0,0947	-0,1609	-0,0284
Indice Masse Corporelle	-0,0287	0,0029	-0,0285	-0,0506	-0,00629
Note (état de santé)	-0,0967	<,0001	-0,0952	-0,1348	-0,0556
Nb de pers. dans le ménage	-0,1072	<,0001	-0,1024	-0,1646	-0,0403
Nationalité étrangère	-0,5445	0,0009	-0,3936	-0,8706	0,0833
Coefficients de β	Coefficients	Pr > ChiSq	Coefficients	Intervalles de confiance (95%)	
Constante	2,92755	<,0001	2,9275	2,9271	2,9279
Divorcé, séparé ou veuf	0,09662	0,0004	0,0965	0,0962	0,0969
Célibataire	0,084	<,0001	0,0837	0,0832	0,0842
Revenu	2,99E-06	0,0027	0,00003	0,00003	0,00003
Prix du tabac	-0,02563	<,0001	-0,0257	-0,0258	-0,0256
Note (état de santé)	0,01109	0,0033	0,0111	0,0111	0,0111
Nb de pers. dans le ménage	0,01462	0,0001	0,0146	0,0146	0,0147
σ	0,11466	-	0,1087	0,1019	0,1155

Ces coefficients sont obtenus au terme de 46 itérations de la procédure *nlin* dans SAS. Pour le logiciel, la convergence n'est pas atteinte ce qui signifie que ces résultats ne sont pas les optima que l'on désirait obtenir. Toutefois, les valeurs finales ne sont pas très éloignées des valeurs initiales, donc plusieurs hypothèses sont possibles :

- soit nous étions déjà proche de l'optimum avec les valeurs initiales et donc le modèle reste « bloqué » autour de celui-ci ;
- soit il existe plusieurs optima et les valeurs présentées ici n'en déterminent qu'un seul.

Or, des tests de robustesse ont montré que le modèle était extrêmement sensible aux modifications d'une ou plusieurs valeurs initiales, ce qui appuie la deuxième hypothèse. Mais, aucune certitude ne peut être avancée. Il serait donc nécessaire de poursuivre et d'approfondir encore ce travail et surtout d'utiliser un logiciel adéquat afin de parfaire cette modélisation.

Conclusion

Les premières études menées sur l'addiction, et plus particulièrement sur le tabac, ont porté leur intérêt sur l'analyse de l'impact des prix sur la consommation. Le calcul d'élasticité-prix est alors un moyen de nourrir les réflexions sur les politiques de taxations présentes ou à venir. Cependant, on ne considère alors le « problème » que lorsqu'il est trop tard ! Ce n'est que plus récemment que certains auteurs ont commencé à s'intéresser aux déterminants de l'initiation tabagique.

Concernant l'impact du prix sur la consommation, l'analyse de la littérature ne permet pas conclure de manière tranchée. Si on admet l'impact positif du prix sur la consommation, il semble qu'il serait plus important en cas de modifications permanentes (thèse de Becker, Grossman et Murphy 1994). Il reste que d'autres auteurs ont infirmé l'existence d'un effet prix (thèse de DeCicca, 1998).

Concernant l'impact du prix sur l'initiation, il a été analysé en distinguant dans l'élasticité-prix un effet sur la consommation et un effet sur la participation. L'influence du prix sur l'initiation, lorsqu'elle apparaît dans les analyses, semble très faible. Dès lors, certains auteurs, comme Douglas et Hariharan (1995) ou Emery (2000), ont étudié l'impact de caractéristiques individuelles socio-économiques sur la consommation et/ou l'initiation tabagique. L'analyse descriptive des données de l'enquête Santé et Protection Sociale confirme l'importance de variables telles que la catégorie socio-professionnelle, le sexe et l'âge. En particulier, nous avons mis en évidence que l'on consomme sa première cigarette de plus en plus tôt.

L'application des modèles économétriques aux données de l'enquête SPS semblent confirmer les conclusions de Douglas et Hariharan, à savoir une influence du prix sur la consommation mais pas sur l'initiation. Toutefois, ce travail devra faire l'objet d'approfondissements tant théoriques que méthodologiques.

Bibliographie

Ainslie G. *Picoeconomics, The strategic Interaction of Successive Motivational States within the person*, 1992.

Allison P.D. *Survival Analysis using the SAS system : A practical guide*, Cary, NC, USA:1995.

Allison P.D. *Logistic regression using the SAS system : Theory and Application*, Cary, NC, USA:1999.

Anguis M. and Dubeaux D. Les fumeurs face aux récentes hausses du prix du tabac. *INSEE Première* 551 (Oct.), 1997.

Atkinson A.B. and Townsend J.L. Economic aspect of reduced smoking. *Lancet*, ii:492-5, 1977.

Becker Gary S. and Murphy Kevin M. A Theory of Rational Addiction. *Journal of Political Economy* 96, 1988.

Becker Gary S., Grossman M., and Murphy Kevin M. Rational addiction and the effect of price on consumption. *American Economic Review* 81:237-241, 1991.

Becker Gary S., Grossman M., and Murphy Kevin M. An empirical analysis of cigarette addiction. *American Economic Review* 84 (3):396-418, 1994.

Becker Gary S. and Mulligan C.B. The endogeneous determination of Time Preference. *Quarterly Journal of Economics* 112 (3):729-758, 1997.

Bocognano A., Dumesnil S., Frérot (Auvray) L., Grandfils N., Le Fur P., and Sermet C. *Santé, soins et protection sociale en 1998*, Paris:1999. 204 pages.

Broïdo M. and Ordonneau C. *Enquête sur la santé et la protection sociale (1997) - Descriptif de la base de donnée*, Paris:1999. 350 pages.

Chaloupka F.J. and Wechsler H. Price, Tobacco control policies and smoking among young adults. *Journal of Health Economics* 16:359-373, 1997.

Chaloupka F.J. and Warner K.E. The Economics of Smoking. In: *Handbook of Health Economics*, edited by Culyer A.J. and Newhouse J.P., North-Holland: 2000, p. 1539-1627.

DeCicca P. and et al. Putting out the fires : will higher cigarette taxes reduce youth smoking ? New York Annual meeting of AEA. Anonymous. 1998.

Douglas S. and Hariharan G. The hazard of starting smoking : estimates from a split population duration model. *Journal of Health Economics* 13:213-230, 1994.

Dumesnil S., Grandfils N., and Le Fur P. *Méthode et déroulement de l'enquête sur la Santé et Protection Sociale - Mise à jour*, Paris:1998. 175 pages.

Emery S., Gilpin E.A., Ake C., Farkas A.J., and Pierce J.P. Characterizing and Identifying "Hard-Core" smokers : Implications for further Reducing Smoking Prevalence. *AJPH* 90 (3):387-394, 2000.

Emery S., White M.M., and Pierce J.P. Does cigarette price influence adolescent experimentation ? *Journal of Health Economics* 20:261-270, 2001.

Evans W.N. and Farrelly M.C. The compensating behaviour of smokers : Tax, tar and nicotine. *RAND Journal of Economics* 29 (3):578-595, 1998.

- Ferguson Brian S. Interpreting the rational addiction model. *Health Economics* 9:587-598, 2000.
- Genier P. and Jacobzone S. Peut-on parler d'une gestion globale de son capital santé ? *Economie et Statistique* 311:37-49, 1998.
- Gruber J. Youth smoking in the US : prices and policies. 2000.
- Gruber J. and Zinman J. Youth smoking in the US : evidence and implications. 2000.
- Hu Teh-wei, Sung Hai-Yen, and Keeler Theodore E. The state antismoking campaign and the industry response : The effects of advertising on cigarette consumption in California. *American Economic Review* 85 (2):85-90, 1995.
- Iannaccone Laurence R. Consumption capital and habit formation with an application to religious participation. Anonymous. Anonymous. University of Chicago. 1984.
- Jones A.M. Health, Addiction, social interaction and the decision to quit smoking. *Journal of Health Economics* 13:93-110, 1994.
- Jones A.M. Adjustment costs, withdrawal effects and cigarette addiction. *Journal of Health Economics* 18:125-137, 1999.
- Keeler Theodore E., Hu Teh-wei, Barnett Paul G., and Manning W.G. Taxation, regulation, and addiction : a demand function for cigarettes based on time-series evidence. *Journal of Health Economics* 12:1-18, 1993.
- Keeler Theodore E., Hu Teh-wei, Barnett Paul G., Manning W.G., and Sung Hai-Yen. Do cigarette producers price-discriminate by state ? An empirical analysis of local cigarette pricing and taxation. *Journal of Health Economics* 15:499-512, 1996.
- Lewit M. and Coate D. The potential for using excise taxes to reduce smoking. *Unknown*:121-145, 1982.
- Loonis E. *Notre cerveau est un drogué - Vers une théorie générale des addictions*, Toulouse:Presses universitaires du Mirail, 1997. 275 pages.
- Manning W.G., Keeler E.B., Newhouse J.P., Sloss E.M., and Wasserman J. The taxes of sin : do smokers and drinkers pay their way ? *JAMA* 261:1604-1609, 19 A.D.
- Orphanides A. and Zervos D. Rational Addiction with Learning and Regret. *Journal of Political Economy* 103, no. 4, 1995.
- Pinget C. and Cornuz J. Taxation du tabac : enjeux économiques et de santé publique. *Rev.Epidem.et Santé Publ.* 47:465-472, 1999.
- Pollak Robert A. Habit formation and dynamic demand function. *Journal of Political Economy* 78, no. 4:745-763, 1970.
- SAS. *SAS/STAT User's guide*, Cary, NC, USA:1990.
- Schelling T.C. Economics and cigarettes. *Preventive Medicine* 15:549-560, 1986.
- Simon H.A. *Economics, bounded rationality and the cognitive revolution*, England:1992.
- Stigler George J. and Becker Gary S. De Gustibus non est Disputandum. *American Economic Review* 67, no.2:76-90, 1977.

Suranovic Steven M., Goldfarb Robert S., and Leonard Thomas C. An economic theory of cigarette addiction. *Journal of Health Economics* 18:1-29, 1999.

Thaler Richard H. and Shefrin Hersh M. An Economic Theory of self-control. *Journal of Political Economy* 89:392-406, 1981.

Townsend J.L., Roderick P., and Cooper J. Cigarette Smoking by socioeconomic group, sex and age : effects of price, income and health policy. *British Medical Journal* 309:923-927, 1994.

Warner K.E., Chaloupka F.J., Cook P.J., Manning W.G., Newhouse J.P., Novotny T.E., Schelling T.C., and Townsend J.L. Criteria of determining an optimal cigarette tax : the economist's perspective. *Tobacco Control* 4:380-386, 1995.

Wasserman J., Manning W.G., Newhouse J.P., and Winkler J.D. The effects of excise taxes and regulations on cigarette smoking. *Journal of Health Economics* 10:43-64, 1991.

Yen S.T. and Jones A.M. Individual cigarette consumption and addiction : a flexible limited dependent variable approach. *Health Economics* 5:105-117, 1996.