



HAL
open science

Effet de l'assurance complémentaire santé sur les consommations médicales, entre risque moral et amélioration de l'accès aux soins

Marc Perronnin

► **To cite this version:**

Marc Perronnin. Effet de l'assurance complémentaire santé sur les consommations médicales, entre risque moral et amélioration de l'accès aux soins. Economies et finances. Université Paris Dauphine - Paris IX, 2013. Français. NNT : 2013PA090068 . tel-01018486

HAL Id: tel-01018486

<https://theses.hal.science/tel-01018486>

Submitted on 4 Jul 2014

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

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

Université Paris-Dauphine (Paris 9)
Ecole Doctorale de Dauphine (ED543),
Laboratoire d'Economie et de gestion des Organisations de Santé (LEDa-LEGOS)

Thèse

Présentée et soutenue par Marc PERRONNIN

Pour obtenir le Grade de DOCTEUR EN SCIENCES ECONOMIQUES

Effet de l'assurance complémentaire santé sur les consommations médicales : entre risque moral et amélioration de l'accès aux soins

Sous la Direction de Jérôme WITTWER

Soutenance publique le 2 décembre 2013

MEMBRES DU JURY :

Rapporteurs

- **BARNAY Thomas**, Université Paris-Est Créteil (UPEC)
- **ETNER Johanna**, Université Paris X Nanterre

Suffragants

- **DORMONT Brigitte**, Université Paris Dauphine
- **GEOFFARD Pierre-Yves**, Paris School of Economics (PSE)
- **LAMIRAUD Karine**, Ecole Supérieure des Sciences Economiques et Commerciales (ESSEC)
- **WITTWER Jérôme**, Université Paris Dauphine

L'UNIVERSITE PARIS-DAUPHINE n'entend donner aucune approbation ni improbation aux opinions émises dans les thèses ; ces opinions doivent être considérées comme propres à leurs auteurs.

Remerciements

Cette thèse est la présentation personnelle d'un travail collectif ayant impliqué de nombreuses personnes directement, comme rédacteurs des trois travaux autour desquels sont articulés la thèse, ou indirectement par leurs conseils, leurs encouragements. Certains appartiennent à ma sphère professionnelle, d'autres à ma sphère familiale ou au cercle de mes amis. Je souhaite leur adresser mes remerciements et leur dédier ce travail. Mes remerciements vont ainsi :

A mon directeur de thèse Jérôme Wittwer, pour sa disponibilité et sa perspicacité. Son humilité et son talent sont pour moi exemplaires.

Aux membres du Jury, Jérôme Wittwer donc mais aussi Thomas Barnay, Brigitte Dormont, Johanna Etner, Pierre-Yves Geoffard, Karine Lamiraud pour avoir accepté de prendre le temps de relire et d'expertiser ma thèse. Puisse ce travail leur apporter en retour des éléments de réflexion.

A Tom Buchmueller, qui fut coauteur du premier article de cette thèse et qui a toujours manifesté beaucoup d'intérêts pour les travaux que nous menons sur l'assurance à l'Irdes.

A Agnès Couffinbal et Michel Grignon eux aussi coauteur du premier article, pour m'avoir mis le pied à l'étrier.

A John Lavis, qui a participé de manière décisive à la rédaction du deuxième article ; sa contribution nous a beaucoup apportée.

A Aurélie Pierre, ma partenaire de travail sur l'Assurance depuis 2005 et coauteur du troisième article. Echanger avec elle fut l'occasion de voir sous un angle nouveau de nombreux aspects économiques et économétriques que je prenais pour acquis.

A Carine Franc, également coauteur du troisième article, pour son dynamisme et ses nombreux conseils. Ses encouragements ont également été stimulants.

A l'Irdes qui m'a permis de mener ce travail dans des conditions favorables.

A Khadidja Ben Larbi et Franck Séverin Clérembault, pour leur soutien de toujours et leur aide lors de la mise en forme du document.

A mes collègues Anne Aligon, Paul Dourgnon, Anne Evans, Pascale Lengagne, Julien Mousquès, Marie-Odile Safon, Aude Sirvain et Véronique Subard, pour leurs encouragements et l'intérêt qu'ils ont manifesté à mon travail.

A la Mutuelle Générale Environnement et Territoires, pour nous avoir fourni les données utilisées pour le troisième article. En particulier à Catherine Hochet et à Jean Sammut pour l'attention portée aux analyses issues de ces données.

A mes parents Pierre et Anne, ma Compagne Odile, mon frère Florent, ma sœur Aude, ma belle-sœur Florence, à Julie et Manuella Sefetius, à Elodie Bordy, à la famille Morgado et à Rémi Hallée pour leur soutien indéfectible leur intérêt et la compréhension dont ils ont fait preuve tout au long de l'avancée du travail de thèse. A Caroline Comby pour son soutien moral au cours de ces deux dernières années.

J'ai enfin une pensée pour les communes du Moule et de Morne-à-l'Eau en Guadeloupe, des Planches en Montagnes dans le Jura et de Nérès-les-Bains dans l'Allier. Outre les attaches familiales que j'y ai, ce sont des lieux où j'ai pu à la fois me ressourcer et rédiger des parties significatives de ma thèse. Que leurs habitants reçoivent mes saluts chaleureux.

SOMMAIRE

INTRODUCTION	11
CHAPITRE 1. CONTEXTE THÉORIQUE	19
1.1. Notations	21
1.1.1. Outils mathématiques	21
1.1.2. Les caractéristiques générales de l'économie considérée.....	22
1.1.3. Demande d'assurance et demande de soins en information symétrique.....	23
1.1.4. Demande d'assurance et demande de soins en information asymétrique.....	24
1.1.5. La production des biens et services médicaux et non médicaux.....	24
1.2. Contexte général de l'analyse	25
1.2.1. De la demande de santé à la demande de soins	25
1.2.2. Caractéristiques des consommateurs.....	27
1.2.3. Les caractéristiques des producteurs	33
1.3. Equilibre dans une économie avec assurance, en information symétrique.....	35
1.3.1. L'offre des assureurs en information symétrique	36
1.3.2. La demande d'assurance et de soins des individus	37
1.3.3. Effet de l'assurance sur les consommations médicales et non médicales	42
1.3.4. Pareto optimalité de l'équilibre.....	43
1.4. De l'asymétrie d'information au risque moral ex-post	44
1.4.1. Position du problème	44
1.4.2. Les apports historiques à la théorie du risque moral	47
1.5. Demande de soins en information asymétrique.....	51
1.5.1. L'offre des assureurs en information asymétrique.....	51
1.5.2. La demande de soins <i>ex-post</i> la survenue du risque santé	52
1.5.3. Effet de l'assurance sur la quantité de biens et services médicaux et non médicaux consommés	54
1.5.4. Travaux empiriques sur l'effet de l'assurance sur les consommations médicales	59
1.6. Inefficiency de l'équilibre en information asymétrique	69
1.6.1. Mise en évidence de l'inefficience.....	69
1.6.2. Mesurer l'effet du niveau de couverture sur le bien-être des individus	72
1.7. Demande d'assurance en présence de risque moral.....	82
1.7.1. Assurance proportionnelle optimale	82
1.7.2. Les mesures empiriques du niveau optimal de couverture.....	85
1.7.3. Effet d'un subventionnement de la couverture sur les niveaux de garanties et les niveaux de consommations de soins	90
1.8. Risque moral et offre de soins.....	96

1.8.1. Contexte économique	97
1.8.2. Quantité et prix des soins prescrits par les professionnels de santé dans le cadre d'une concurrence monopolistique, sans assurance.....	99
1.8.3. Effet de l'assurance sur la quantité et le prix des soins prescrits par les professionnels de santé dans le cadre d'une concurrence monopolistique	105
1.8.4. Assurance santé et qualité des soins administrés	108
1.8.5. Assurance et demande induite.....	122
1.9. Couverture assurantielle et accès aux soins.....	126
1.9.1. Mesurer l'effet revenu induit par une variation du niveau d'assurance	126
1.9.2. L'accès aux soins comme motif de l'assurance	128
1.9.3. Le débat sur l'effet revenu.....	129
1.9.4. Effet à la marge du niveau de couverture : analyses empiriques exploitant les couvertures complémentaires.....	130

CHAPITRE 2. ACCESS TO PHYSICIAN SERVICES: DOES SUPPLEMENTAL INSURANCE MATTER?

EVIDENCE FROM FRANCE.....	133
2.1. Abstract	135
2.2. Introduction	135
2.3. Health insurance in France.....	137
2.4. Previous research on supplemental insurance and utilization	138
2.5. Data and descriptive evidence	140
2.6. Econometric analysis	144
2.6.1. The determinants of supplemental insurance coverage	144
2.6.2. The effect of supplemental insurance on physician visits	147
2.6.3. Insurance and the choice of physician.....	153
2.7. Discussion	157
2.8. Acknowledgements	159

CHAPITRE 3. DOES FREE COMPLEMENTARY HEALTH INSURANCE HELP THE POOR TO ACCESS HEALTH CARE? EVIDENCE FROM FRANCE.....

3.1. Abstract	163
3.2. Introduction	164
3.3. Previous literature.....	164
3.4. The design of a quasi-experiment: removing financial barriers to access in France	166
3.5. Methods and econometric strategy.....	168
3.5.1. Econometric strategy.....	170
3.5.2. Data sources	171
3.6. Results	175
3.6.1. Baseline: before-and-after analysis (Table 3.6 i)	175
3.6.2. Influence of automatic versus voluntary enrolment (Table 3.6-ii)	177

3.6.3. Difference-in-difference analyses.....	179
3.7. Discussion	183
CHAPITRE 4. SUPPLEMENTAL HEALTH INSURANCE AND HEALTH CARE CONSUMPTION: A DYNAMIC APPROACH TO THE MORAL HAZARD PHENOMENON	185
4.1. Abstract	187
4.2. Introduction	187
4.3. Data.....	190
4.3.1. Context of dataset	190
4.3.2. Sample and dataset	191
4.4. Economic model and methods	193
4.4.1. The assumptions of pent-up demand.....	193
4.4.2. Health care expenditures.....	194
4.4.3. The pent-up demand assumption	195
4.4.4. Empirical models.....	195
4.5. Results	198
4.5.1. Effects of ECHI.....	198
4.5.2. Effect of main other covariates	200
4.6. Discussion	203
4.6.1. The effect of ECHI according to the type of care.....	203
4.6.2. The pent-up demand assumption	204
CONCLUSION.....	207
5.1. Les enseignements de la théorie économique	209
5.2. Les principaux résultats des trois articles présentés.....	211
5.3. Mise en perspective des résultats avec l'organisation du système d'assurance maladie...214	
5.4. Quelques remarques additionnelles.....	216
ANNEXE TECHNIQUE.....	217
6.1. Niveau d'assurance préféré en présence de risque moral (résolution par maximisation de l'utilité espérée)	219
BIBLIOGRAPHIE	225

Introduction

Au cours de ces cinquante dernières années, dans les pays de l'OCDE, le taux de croissance des dépenses de santé a été supérieur à celui du produit intérieur brut de sorte que le poids de ces dépenses dans la richesse nationale s'est considérablement alourdi : de 1960 à 2009, la part des dépenses de santé dans le produit intérieur brut est ainsi passé de 5,1% à 17,7% pour les Etats-Unis, de 3,8% à 11,7% pour la France, de 5,4% à 11,4% pour le Canada, de 3,9% à 9,8% pour le Royaume-Unis, de 3,7% à 9,1% pour l'Australie et de 3% à 9,5% pour le Japon¹. En l'état actuel, la pérennité du financement des systèmes de santé est donc source de préoccupations ; par conséquent, les gouvernements tentent d'infléchir le rythme de croissance des dépenses médicales.

Parmi les facteurs expliquant cette évolution, un grand nombre d'économistes ont mis en avant le rôle des couvertures santé : les remboursements offerts par ces couvertures abaissent le coût apparent des biens et services médicaux, incitant les individus à accroître leurs consommations de soins au détriment des biens et services non médicaux. Ce phénomène, qualifié de risque moral, a pu être stimulé par l'extension de la couverture santé dans de nombreux pays, contribuant ainsi à accroître les dépenses médicales. Aussi, l'analyse économique du risque moral a fait l'objet d'une attention particulière en économie de la santé. Elle s'est développée dans les années 60 (K. Arrow, M. Pauly). Néanmoins, le concept de « risque moral » était connu des assureurs bien avant et ceux-ci avaient déjà mis en place de systèmes de coassurance pour lutter contre ce phénomène. Ainsi, Arrow (1963) note que :

« Coinsurance provisions have been introduced into many major medical policies to meet this contingency as well as the risk aversion of the insurance companies. »

En France, on peut noter que dès sa création en 1945, l'Assurance maladie obligatoire de la Sécurité sociale imposait aux assurés sociaux un reste à charge prenant la forme d'une fraction du tarif conventionnel. Depuis 1945, l'assurance complémentaire santé permet de compléter les remboursements de l'Assurance maladie obligatoire, réduisant ainsi les restes à charge auxquels sont soumis les individus. De ce fait, elle est susceptible d'accroître les consommations médicales en solvabilisant la demande de soins ou en générant du risque moral *ex-post*, c'est-à-dire l'achat par les individus de soins dont la valeur est faible au regard de leur coût total. La solvabilisation de la demande de soins est positive puisqu'elle correspond à une amélioration de l'accès aux soins. En revanche le risque moral *ex-post* est négatif car il conduit les individus à dépenser trop pour leurs soins, les privant de ressources pour leurs autres consommations. Cette surconsommation peut être limitée en instaurant une couverture incomplète (via des « coassurance » ou des franchises) incitant les assurés à modérer leurs dépenses de santé. Néanmoins de tels restes à charge peuvent

¹ Source : données de santé de l'OCDE (disponibles sous <http://stats.oecd.org/>)

également limiter l'accès aux soins, notamment chez les populations les plus modestes. Etablir le juste niveau de couverture s'avère complexe car risque moral et solvabilisation de la demande de soins sont susceptibles de coexister et leur ampleur dépend en théorie du niveau d'assurance souscrit, mais aussi des caractéristiques des individus (en particulier leur état de santé et leur revenu) et du type de soin couvert. Un axe de recherche important consiste donc analyser l'effet de l'assurance santé pour différents niveaux de couverture, différents types de soins et différentes sous-population et à déterminer lequel des deux phénomènes (risque moral ou amélioration de l'accès aux soins) prédomine. Cette thèse s'appuie sur trois articles :

Access to physician services: does supplemental insurance matter? Evidence from France (Buchmueller, Couffinhal, Grignon, Perronnin, 2004)

Le premier travail (Buchmueller *et al.*, 2004) analyse l'effet du fait d'être couvert par une complémentaire santé sur les consommations de soins de généraliste et spécialistes. Il s'appuie sur des données de consommations médicales sur un mois recueillies dans le volet « carnet de santé » de l'enquête ESPS 1998. L'analyse porte sur 8161 individus âgés de 25 ans et plus, qui ne sont pas couvertes par l'AMG² et ne bénéficient pas d'une exonération du ticket modérateur. Dans un premier temps, on estime la probabilité de recourir au moins une fois à un médecin (généraliste ou spécialiste) sur un mois. Dans un second temps, on estime la probabilité qu'une consultation ait lieu chez un spécialiste plutôt qu'un généraliste. L'hypothèse sous-jacente est que les individus couverts par une complémentaire santé vont avoir plus tendance à consulter un spécialiste car le surcoût lié au choix de ce type de médecin est partiellement compensé par la complémentaire. Cette deuxième partie de l'étude est menée sur l'ensemble des consultations de médecins, puis uniquement sur celles pouvant être traitées indifféremment par un généraliste ou un spécialiste. L'existence d'un biais d'antisélection est testée de trois manières : en analysant l'effet de l'état de santé (contrôlé du milieu social) sur la probabilité d'être couvert, en analysant conjointement par un probit bivarié, la probabilité d'être couvert et la probabilité de recourir au médecin³ et enfin, en comparant les consommations de soins chez les personnes couvertes par une complémentaire santé

² Dispositif à destination des ménages pauvres et prenant en charge, *a minima*, le ticket modérateur. Il a été remplacé à partir du 1^{er} janvier 2000 par la CMU-C.

³ Plusieurs instruments sont introduits, dont le fait d'être salarié de la fonction publique.

d'entreprises et celles couvertes par une complémentaire individuelle⁴. Ces tests conduisent à rejeter l'hypothèse d'antisélection.

L'étude met en évidence un effet très sensible du fait d'être couvert par une complémentaire santé sur la probabilité de recourir à un médecin, mais pas d'effet significatif sur la probabilité de consulter un spécialiste plutôt qu'un généraliste. Il ne semble donc pas y avoir de risque moral lié au choix du praticien.

Does free complementary health insurance help the poor to access health care? Evidence from France. (Grignon, Perronnin, Lavis, 2007)

Le rôle joué par l'assurance complémentaire dans l'accès aux soins chez les personnes disposant de faibles ressources financières a conduit les pouvoirs publics à mettre en place en 2000 la Couverture maladie universelle complémentaire (CMU-C). Cette couverture gratuite et accessible sous conditions de ressources permet à ses bénéficiaires de disposer d'une couverture complète de leurs soins, sans avance de frais. Ce dispositif remplace l'ancienne AMG qui concernait une population plus restreinte et offrait une couverture moins généreuse. Destinée à lever les barrières financières à l'accès aux soins, l'absence de reste à charge laissé par la CMU-C a suscité chez certains des interrogations quant à la présence potentielle de risque moral. Le travail de Grignon et al. (2007) vise à répondre à ces interrogations en évaluant l'impact de la CMU-C sur les consommations médicales de ses bénéficiaires.

Les données utilisées concernent les années 1998, 1999 et 2000, soient deux ans avant et un an après la mise en place de la CMU-C. Elles sont issues de l'Echantillon permanent des assurés sociaux. Ce fichier administratif contient un historique détaillé des consommations de soins présentées au remboursement de l'Assurance maladie obligatoire. Pour les bénéficiaires de la CMU-C, il renseigne sur la date d'entrée dans le dispositif. Enfin, il permet de disposer de quelques caractéristiques individuelles, en particulier l'âge du patient. Les données sont agrégées par trimestre puis panélistées ; on dispose ainsi au maximum de 12 points dans le temps par individu. Les estimations sont menées en utilisant un modèle en deux parties : la probabilité de recourir au moins une fois aux

⁴ Ce test repose sur le fait que les individus couverts par une complémentaire santé d'entreprise sont a priori peu sujets à l'antisélection. En effet, dans de nombreux cas, la complémentaire santé est obligatoire (ou tout au moins perçue comme telle), l'individu n'a donc pas la possibilité de refuser de se couvrir. D'autre part, pour la très grande majorité des contrats d'entreprise, l'employeur prend en charge tout ou partie de la prime. Suivant le raisonnement déjà adopté par Ettner (1997) cela réduit le coût de la couverture pour le salarié et limite donc l'incitation à ne pas se couvrir.

soins au cours d'un trimestre est estimée au moyen d'un modèle à probabilité linéaire⁵ puis la dépense conditionnellement au fait d'avoir recouru aux soins est estimée au moyen d'un modèle log-linéaire. Les caractéristiques fixes dans le temps sont contrôlées en introduisant un effet fixe dans chacune des régressions.

Dans un premier temps, les analyses sont effectuées sur un échantillon de 2686 individus en distinguant ceux entrés au 1^{er} janvier 2000 et ceux entrés après le 1^{er} janvier 2000. Chez les personnes entrées dans le dispositif au 1er janvier 2000, qui pour la plupart bénéficiaient auparavant de l'Aide médicale gratuite et ont été affiliées de manière automatique à la CMU-C, les dépenses n'augmentent pas plus vite après la mise en place de la CMU-C qu'avant. Ainsi, bien que bénéficiant d'une augmentation de leurs garanties, notamment en optique et en dentaire, ces individus n'ont pas modifié sensiblement leurs consommations de soins. Ce constat semble écarter l'existence d'un aléa moral induit par la gratuité des soins dont bénéficient les CMUistes. Chez les personnes entrées après le 1er janvier 2000, qui pour une part importante ne bénéficiaient d'aucune complémentaire santé avant la mise en place de la CMU-C, les dépenses augmentent beaucoup plus fortement après la mise en place du dispositif qu'avant. Cette hausse des consommations a été interprétée comme la conséquence d'une meilleure accessibilité aux soins, auquel s'ajoute un éventuel phénomène de rattrapage.

Dans un second temps, on apparie l'EPAS avec ESPS 1998 afin de disposer d'informations plus précises sur la couverture complémentaire dont disposaient les CMUistes avant leur entrée dans le dispositif. Le sous-échantillon obtenu comporte alors 475 individus. On mène des analyses séparées selon le type de couverture complémentaire dont bénéficiaient les individus avant leur entrée dans la CMU-C (AMG, couverture complémentaire privée ou pas de couverture complémentaire). On utilise le même type de modélisation que dans la 1^{ère} partie de l'étude (modèles en deux parties avec effets fixes). Ces analyses complémentaires confirment le fait que la CMU-C a un effet plus important chez les personnes qui n'étaient auparavant pas couvertes par une complémentaire santé que chez celles qui étaient couvertes par l'AMG ou par une complémentaire santé privée.

En conclusion, la CMU-C ne paraît pas avoir eu d'effet inflationniste marqué, mais elle semble plutôt avoir favorisé l'accès aux soins de ceux qui n'étaient auparavant pas couverts.

⁵ Autrement dit, on régresse par les MCO l'indicatrice de recours. Pour cette première étape, un modèle logit à effet fixe pourrait paraître plus adapté. Mais ce type de modèle ne permet pas de prédire la probabilité de recours d'un individu dans la mesure où l'effet fixe ne peut pas être estimé. Or cette probabilité est utilisée dans la suite pour estimer la variation de l'espérance de dépense suite à l'entrée dans le dispositif CMU-C.

Supplemental health insurance and health care consumption: A dynamic approach to the moral hazard phenomenon (Franc, Perronnin, Pierre, en cours de soumission à Health Economics).

La troisième étude (Franc *et al.*, en cours de soumission) analyse l'effet dans le temps d'une augmentation « à la marge » des garanties chez des individus déjà couvertes par un même niveau d'assurance complémentaire. Elle se différencie ainsi de la plupart des travaux français qui analysent l'effet du fait d'être couvert plutôt que non couvert par une complémentaire santé. Elle s'appuie sur l'introduction en juillet 2003 d'une surcomplémentaire santé permettant de compléter les remboursements du contrat de base d'une mutuelle de fonctionnaire. Un cinquième des adhérents au contrat de base de la mutuelle ont choisi de souscrire cette option entre juillet 2003 et décembre 2005.

L'effet de la surcomplémentaire santé sur les consommations de soins est estimé à partir de données issues de fichiers administratifs fournis par la mutuelle. Ces fichiers permettent en particulier de disposer d'un historique détaillé des consommations de soins, 2 ans et $\frac{1}{2}$ avant et 2 ans et $\frac{1}{2}$ après l'introduction de la surcomplémentaire. L'effet du fait d'être couvert par la surcomplémentaire sur les consommations de soins est analysé sur un sous-échantillon de 18 160 individus. Les estimations sont menées en utilisant des modèles en trois parties : dans un premier temps, la probabilité de consommer au moins une fois des soins à chaque semestre est estimée par un probit. Dans un second temps, le nombre de recours conditionnellement au fait d'avoir consommé est estimé au moyen d'un modèle sur variable de compte tronquée ; enfin, la dépense par acte est estimée au moyen d'une régression log-linaire. L'antisélection est prise en compte en introduisant dans les régressions une variable constante dans le temps qui indique à quel semestre l'individu a souscrit la surcomplémentaire (s'il l'a souscrite)⁶ ; en effet, cette variable permet d'appréhender les différences de niveaux de consommations de soins conditionnellement au fait que les individus ne bénéficient pas de la surcomplémentaire. Pour tenir compte des corrélations temporelles des résidus, les estimations sont menées avec des modèles GEE (*General Estimating Equation*).

Partie introductive : contexte théorique

Préalablement à la présentation de ces trois articles, une partie présente le contexte théorique dans lequel s'inscrit cette thèse. Cette partie a pour but de faire quelques rappels permettant de comprendre le rôle de l'assurance santé, les choix de niveau de couverture et de consommations de soins. Les notions de choix et d'assurance renvoient aux notions d'utilité en incertain, de risque et de

⁶ L'introduction de ces variables revient à utiliser la spécification de Chamberlain permettant de modéliser, dans le cadre de données longitudinales, la corrélation entre une variable explicative et une hétérogénéité inobservée fixe dans le temps.

sinistre. La santé est affectées de chocs aléatoires négatifs ce qui rend les consommations de soins elles même incertaines. Sous la condition dite d'aversion pour le risque, les individus vont opter pour des transferts de revenu entre état de la nature de manière à diminuer le risque financier associé aux dépenses de santé.

Afin de comprendre les conséquences économiques du risque moral, il est nécessaire au préalable de décrire ce qui se passe en termes d'allocation des ressources dans le cas « idéal » où l'information est symétrique. Cette sous-partie rappelle brièvement les principaux résultats sur la demande d'assurance en information symétrique.

Les mécanismes conduisant au risque moral sont ensuite explicités. Nous présentons les motifs qui, dans le cadre de la santé, rendent difficile l'observation par l'assureur des comportements des assurés conduisant ainsi à l'asymétrie d'information entre ces deux parties. Nous expliquons pourquoi l'assuré est alors incité à accroître son risque financier. Nous présentons les conséquences en termes de perte d'efficacité et nous montrons que, compte tenu des restrictions sur l'offre, la couverture apportant le maximum de bien-être collectif est incomplète. En contrepoint, nous montrons que la hausse des consommations de soins peut résulter en partie voire en totalité d'un transfert de revenu permettant de solvabiliser la demande de soins : les remboursements permettent aux individus de bénéficier d'une augmentation de leur pouvoir d'achat leur permettant d'avoir un meilleur accès aux soins.

Les outils utilisés sont certes issus de la théorie du risque et de l'incertain mais aussi de l'économie publique (l'assurance pouvant être considérée comme une forme de subventionnement des soins) et de l'économie industrielle. Nous tenterons tout au long de la partie théorique de faire le lien avec ces disciplines.

Chapitre 1.

Contexte théorique

1.1. Notations

Dans l'ensemble de cette partie de contexte théorique, les notations ont été uniformisées autant que possible. Cette partie présente les notations générales adoptées, celles-ci sont rappelées au fil du texte. Lorsqu'un changement de notation s'est avéré nécessaire, celui-ci est mentionné dans le texte.

1.1.1. Outils mathématiques

Notation	Signification
\mathbb{R} et \mathbb{R}_+	Ensemble des nombres réels et ensemble des nombres réels positifs.
\mathbb{R}^n et \mathbb{R}_+^n	Espace de dimension n et à valeurs dans \mathbb{R} et espace de dimension n à valeurs dans \mathbb{R}_+ .
$\forall x$	« Pour tout x ».
$x \in X$	« x appartient à l'ensemble X ».
$(x_k)_{k \in \mathcal{K}}$	Forme condensée pour désigner le vecteur $(x_1 \dots x_k \dots x_n)$, \mathcal{K} étant l'ensemble des entiers allant de 1 à n .
(x_k)	Même signification que $(x_k)_{k \in \mathcal{K}}$, lorsque l'ensemble \mathcal{K} des indices est implicite.
$\frac{df}{dx}(\hat{x})$ ou $f'(\hat{x})$	Dérivée première de f par rapport à x , appliquée à \hat{x} dans le cas où f ne dépend que d'une seule variable x .
$\left. \frac{\partial f}{\partial x_k} \right _{(\hat{x}_1 \dots \hat{x}_k \dots \hat{x}_n)}$	Dérivée première de f par rapport à x_k , appliquée $(\hat{x}_1 \dots \hat{x}_k \dots \hat{x}_n)$, dans le cas où f dépend de plusieurs variables $(x_1 \dots x_k \dots x_n)$. « Dérivée partielle de f par rapport à x_k »
$\frac{d^2 f}{dx^2}(\hat{x})$ ou $f''(\hat{x})$	Dérivée seconde de f par rapport à x , appliquée à \hat{x} , dans le cas où f ne dépend que d'une seule variable x .
$\left. \frac{\partial^2 f}{\partial x_k^2} \right _{(\hat{x}_1 \dots \hat{x}_k \dots \hat{x}_n)}$	Dérivée seconde de f par rapport à x_k , appliquée $(\hat{x}_1 \dots \hat{x}_k \dots \hat{x}_n)$, dans le cas où f dépend de plusieurs variables $(x_1 \dots x_k \dots x_n)$.
$\left. \frac{\partial^2 f}{\partial x_k \partial x_l} \right _{(\hat{x}_1 \dots \hat{x}_k \dots \hat{x}_n)}$	Dérivée seconde de f par rapport à x_k et x_l , appliquée $(\hat{x}_1 \dots \hat{x}_k \dots \hat{x}_n)$, dans le cas où f dépend de plusieurs variables $(x_1 \dots x_k \dots x_n)$.
\bar{x} (avec $x = (x_1 \dots x_k \dots x_n)$)	Moyenne d'une variable prenant les valeurs $x_1 \dots x_k \dots x_n$.
$E(\tilde{x})$	Espérance d'une variable aléatoire \tilde{x} continue ou discrète.
$var(\tilde{x})$	Variance d'une variable aléatoire \tilde{x} continue ou discrète.
$cov(\tilde{x}, \tilde{y})$	Covariance entre variable aléatoire \tilde{x} et une variable aléatoire \tilde{y} .

Remarques :

- Lorsque $x = (x_k)_{k \in \mathcal{K}}$ ou $x = (x_k)$, $(x_k)_{k \in \mathcal{K}}$ ou (x_k) sont aussi appelés « profil de la variable x ».

- Par soucis de simplification, lorsque f est une fonction sur un sous ensemble de \mathbb{R}^n (donc $f(x) = f(x_1 \dots x_k \dots x_n)$), alors :

$$f' = \left(\frac{\partial f}{\partial x_1}, \dots, \frac{\partial f}{\partial x_k}, \dots, \frac{\partial f}{\partial x_n} \right)$$

- Par abus, lorsque \tilde{x} est une variable aléatoire discrète prenant les valeurs $x_1 \dots x_k \dots x_n$, nous noterons \bar{x} l'espérance de \tilde{x} .

1.1.2. Les caractéristiques générales de l'économie considérée

Notation	Signification
N_P	Nombre total d'unités de production
N_{P_x}	Nombre total d'unités de production de biens et services non médicaux
N_{P_m}	Nombre total d'unités de production de biens et services médicaux
W^n	Revenu « nominal » des individus
S	Ensemble $\{0, \dots, S\}$ des $S + 1$ niveaux de gravité pathologiques dont les individus peuvent être affectés
h_s	Etat de santé de l'individu affecté d'une pathologie de gravité s
π_s	Probabilité de survenue du niveau de gravité s
u_x	Utilité attachée aux consommations de biens et services non médicaux
u_h	Utilité attachée à l'état de santé
g_h	Fonction de production de santé
$u_m(s, \cdot)$	Utilité attachée aux consommations de soins lorsque l'individu est dans l'état s : $u_m(s, \cdot) = u_h(h_s + g_h(\cdot))$
U_s	Utilité directe des individus dans l'état s
V_s	Utilité indirecte des individus dans l'état s
p_m	Prix de vente d'une unité de biens et services médicaux
q_m	Prix d'achat des biens et services médicaux
$x_s^u(q_m, W)$	Quantité de biens et services non médicaux consommés par les individus dans l'état s , confrontés à un prix unitaire des soins q_m et disposant du niveau de ressources W <i>ex-post</i> la réalisation du risque.
$m_s^u(q_m, W)$	Quantité de biens et services médicaux consommés par les individus dans l'état s , confrontés à un prix unitaire des soins q_m et disposant du niveau de ressources W <i>ex-post</i> la réalisation du risque.
$\psi_s^u(q_m, v)$	Demande Hicksienne en bien et services médicaux des individus dans l'état s , confrontés à un prix unitaire des soins q_m et disposant du niveau d'utilité v <i>ex-post</i> la réalisation du risque.
$\chi_s^u(q_m, v)$	Demande Hicksienne en bien et services non-médicaux des individus dans l'état s , confrontés à un prix unitaire des soins q_m et disposant du niveau d'utilité v <i>ex-post</i> la réalisation du risque.
$\eta_{m_s^u, q_m}$	Elasticité-prix de m_s^u

Notation	Signification
$\eta_{m_s^u, W}$	Elasticité-revenu de m_s^u
$\eta_{\bar{m}^u, q_m}$	Elasticité-prix de la demande de soins moyenne (ou agrégée) \bar{m}^u
$\eta_{\bar{m}^u, W}$	Elasticité-revenu de la demande moyenne (ou agrégée) \bar{m}^u
$\eta_{\psi_s^u, q_m}$	Elasticité-prix de la demande hicksienne ψ_s^u
$\eta_{\bar{\psi}^u, q_m}$	Elasticité-prix de la demande hicksienne moyenne (ou agrégée) $\bar{\psi}^u$
$\eta_{\chi_s^u, q_m}$	Elasticité-prix de la demande hicksienne χ_s^u

1.1.3. Demande d'assurance et demande de soins en information symétrique

Notation	Signification
I	Désigne un contrat contingent à l'état de santé
$(I_s)_{s \in \mathcal{S}}$ ou (I_s)	Profil de remboursements du contrat I
P_I	Niveau de prime associé au contrat I
$(\Delta W_s)_{s \in \mathcal{S}}$ ou (ΔW_s)	Désigne un profil de versements (ou transfert de revenu)
I^*	Contrat préféré par les individus, parmi ceux tarifés au coût actuariel
$(I_s^*)_{s \in \mathcal{S}}$ ou (I_s^*)	Profil de remboursements du contrat I^* préféré par les individus
P_{I^*}	Niveau de prime associé au contrat I^*
$(\Delta W_s^*)_{s \in \mathcal{S}}$ ou (ΔW_s^*)	Transfert de revenu préféré parmi ceux respectant la neutralité actuarielle (ie tels que $\sum_{s \in \mathcal{S}} \pi_s \cdot \Delta W_s^* = 0$)
$x_s^*(p_m, W)$	Niveau des consommations non-médicales des individus dans l'état s lorsqu'ils sont couverts par I^* .
$m_s^*(p_m, W)$	Niveau des consommations médicales des individus dans l'état s lorsqu'ils sont couverts par I^* .

Remarque :

Les individus couverts par le contrat I^* sont confrontés, ex-post la survenue du risque au prix unitaire des soins p_m et au niveau de revenu $W_s^* = W + I^* - P_{I^*}$. Par conséquent x_s^* et m_s^* sont liés à x_s^u et m_s^u par les relations suivantes :

$$x_s^*(p_m, W) = x_s^u(p_m, W + I^* - P_{I^*})$$

$$m_s^*(p_m, W) = m_s^u(p_m, W + I^* - P_{I^*})$$

1.1.4. Demande d'assurance et demande de soins en information asymétrique

Notation	Signification
A	Désigne un contrat contingent aux consommations médicales
$(A_s)_{s \in \mathcal{S}}$ ou (A_s)	Profil de remboursements du contrat A
P_A	Niveau de prime associé au contrat A
a	Taux de couverture du contrat A (lorsque celui-ci est proportionnel)
c	Taux de coassurance du contrat A ($c = 1 - a$)
$x_s^\#(q_m, W)$	Profil de consommations non-médicales des individus assurés par le contrat A laissant le prix résiduel q_m et le revenu résiduel $W - P_A$.
$m_s^\#(q_m, W)$	Profil de consommations médicales des individus assurés par le contrat A laissant le prix résiduel q_m et le revenu résiduel $W - P_A$.
$\eta_{m_s^\#, q_m}$	Elasticité-prix de $m_s^\#$
$\eta_{m_s^\#, W}$	Elasticité-revenu de $m_s^\#$
$\eta_{\bar{m}^\#, q_m}$	Elasticité-prix de la demande de soins moyenne (ou agrégée) $\bar{m}^\#$
$\eta_{\bar{m}^\#, W}$	Elasticité-revenu de la demande moyenne (ou agrégée) $\bar{m}^\#$
$a^\#$	Taux de couverture optimal
$c^\#$	Taux de coassurance optimal

Remarque :

Les individus couverts par le contrat A sont confrontés, *ex-post* la survenue du risque au prix unitaire des soins q_m et au niveau de revenu $W - P_A$. Par conséquent $x_s^\#$ et $m_s^\#$ sont liés à x_s^u et m_s^u par les relations suivantes :

$$x_s^\#(q_m, W) = x_s^u(q_m, W - P_A)$$

$$m_s^\#(q_m, W) = m_s^u(q_m, W - P_A)$$

1.1.5. La production des biens et services médicaux et non médicaux

Notation	Signification
g_x	Fonction de production d'une unité représentative du secteur des biens et services non médicaux (concurrence pure et parfaite)
$g_{m,j}$	Fonction de production de l'unité j du secteur des biens et services médicaux (concurrence monopolistique sur le secteur de la santé)
Z	Vecteur d'inputs dans la production de l'ensemble des biens et services
Z_x	Quantité d'inputs allouée par unité représentative du secteur des biens et services non médicaux (concurrence pure et parfaite)

Notation	Signification
$z_{m,j}$	Quantité d'inputs allouée par unités j du secteur des biens et services médicaux (concurrence monopolistique sur le secteur de la santé)
$C_x(x)$	Coût de production d'une quantité x de biens et services non-médicaux
$C_{m,j}(m_j)$	Coût de production pour l'unité j d'une quantité m_j unités de biens et services médicaux (concurrence monopolistique sur le secteur de la santé)
x^P	Quantité optimale produite par unité de production du secteur des biens et services non médicaux (concurrence pure et parfaite)
m_j^P	Quantité optimale produite par unité de production j du secteur des biens et services médicaux (concurrence monopolistique sur le secteur de la santé)
p_m	Prix de vente d'une unité de biens et services médicaux
q_m	Prix d'achat des biens et services médicaux

1.2. Contexte général de l'analyse

Le but de cette partie est de présenter le contexte économique général dans lequel s'inscrit notre analyse. Nous nous plaçons dans le cadre d'une économie avec production. En effet, une simple économie d'échange ne permet pas de décrire la réaction des producteurs à la couverture assurancielle des consommateurs. Notre économie est ainsi constituée :

- De N_C consommateurs.
- De N_p producteurs répartis sur deux secteurs de production, le secteur des biens et services médicaux (N_{p_m} producteurs) et celui des biens et services non médicaux (N_{p_x} producteurs) et d'un vecteur d'inputs z dont le coût est p_z (qui est lui-même un vecteur).
- De N_I assureurs en santé, lesquels produisent de la diminution du risque financier.

Nous commençons par examiner l'objet sur lequel porte le risque à savoir la santé, puis nous examinons la demande de santé des individus et la demande de soins qui en découle. Nous nous intéressons enfin aux producteurs.

1.2.1. De la demande de santé à la demande de soins

La santé est un facteur fondamental du bien-être des individus. Ainsi, l'Organisation mondiale de la santé (OMS) définit la santé de la manière suivante :

« La santé est un état complet de bien-être physique, mental et social, et ne consiste pas seulement en une absence de maladie ou d'infirmité. »

Cette définition a le mérite de ne pas être normative médicalement parlant. Néanmoins, bien qu'elle se base sur la notion de bien-être, elle reste limitée d'un point de vue économique. Tout d'abord, elle décrit les conséquences en termes d'utilité d'une santé parfaite, en restant vague sur les conditions

nécessaires à l'obtention de cet état. Elle précise juste qu'elles ne doivent pas seulement aboutir à l'absence de maladies et d'infirmités. D'autre part, la notion d'état complet de bien-être est purement théorique et relative, dans la mesure où la santé peut être toujours améliorée par des techniques médicales de plus en plus sophistiquée. Dans bon nombre de pays, l'espérance de vie en générale et l'espérance de vie en bonne santé en particulier n'ont cessé de croître et continueront probablement à s'allonger. Mais même si on peut espérer vivre plus vieux en bonne santé, la durée de vie demeure limitée, quelques soient les ressources dont nous disposons. De plus, il n'est pas fait référence aux autres éléments entrant dans l'utilité individuelle. Or, l'obtention d'un bon état de santé nécessite des ressources financières au détriment de consommations non-médicales et également des contraintes sur le mode de vie, ce qui pose question sur le fait que la santé parfaite soit associée systématiquement à une forte utilité. Un individu peut très bien arbitrer entre sa santé et certains nombre de consommations et juger ces dernières plus utiles pour lui. Enfin, dans la définition de l'OMS, la notion de santé est attachée à un état unique qualifié de parfait. Or la notion de santé est attachée à plusieurs états possibles plus ou moins désirables : on parle de bonne mais aussi de mauvaise santé.

La définition précise de l'objet « santé » bute sur sa très grande complexité, due au fait que c'est un concept multidimensionnel. Il recouvre une multitude d'états physiques et mentaux difficilement hiérarchisables entre eux. Cette hiérarchie est d'autant plus ardue qu'elle peut se faire selon plusieurs échelles. Blaxter (1990) en dénombre trois: la santé subjective, la santé médicale et la santé fonctionnelle. Ces trois dimensions ne se recouvrent pas forcément. Ainsi, un individu atteint d'asthme a une santé médicale imparfaite (puisqu'il est porteur d'une pathologie), mais il peut cependant jouir de très bonnes capacités fonctionnelles (il n'est pas gêné dans sa vie courante) et s'estimer en bonne santé si son asthme est parfaitement contrôlé. De surcroît, il existe une dimension présente et future et une dimension de certitude et d'incertitude.

Un apport considérable sur le lien entre santé et demande de soin est le fruit des travaux de Grossman (1972) sur l'investissement du capital humain « santé ». Selon Grossman la santé est un bien durable immatériel, qui apporte un bien-être direct (elle entre comme argument de la fonction d'utilité des individus) mais également indirect en permettant d'accroître le temps passé à la production d'activités marchandes (donc sources de revenu) et non marchandes. Le stock de santé subit deux formes de dégradations : une dépréciation liée à l'âge (Grossman, 1972) que nous supposons déterministe et des chocs négatifs (accidents ou maladies) qui affectent aléatoirement les individus. Ceux-ci ont la possibilité d'accroître leur santé, non pas en achetant directement (c'est un bien qui ne peut être échangé), mais en la produisant ou en la faisant produire par des

professionnels de santé à partir d'une quantité m de soins. Cette caractéristique se retrouve sur d'autres biens que la santé, comme par exemple l'éducation. La consommation de soins médicaux n'est donc pas source de bien-être en elle-même ; elle n'apporte de l'utilité que dans la mesure où elle permet d'accroître l'état de santé. Grossman considère que ce rôle des soins est un élément clé du modèle sous-jacent à ses travaux :

"A second justification for the model is that most students of medical economics have long realized that what consumers demand when they purchase medical services are not these services per se but rather "good health".

De manière pratique, cela se traduit par le fait que l'utilité apportée par la consommation de soins est relative à l'état de santé de l'individu. La consommation de soins par un individu en bonne santé ne va s'accompagner que d'une faible amélioration du bien-être, elle peut même générer de la désutilité (la consommation de soins pouvant s'accompagner d'effets secondaires douloureux ou handicapant). Mathématiquement cela signifie que les biens et services médicaux ne sont pas des arguments de la fonction d'utilité de l'individu ; celle-ci ne dépend directement que de la santé et des consommations non médicales ; elle s'écrit donc $u(x, h)$ où h est la santé et x l'ensemble des consommations non médicales. Néanmoins, puisque les individus peuvent modifier leur état de santé par la consommation de soins, leur utilité dérive indirectement de ces consommations. Plus exactement, elle peut s'écrire :

$$U_s(x_s, m_s) = u(x, h_s + g_h(m_s))$$

Où h_s est l'état de santé après qu'une pathologie de niveau de sévérité s soit survenue et avant consommation de soins ; g_h est la fonction de production de santé. Dans la suite de la partie 1.2, nous allons préciser les hypothèses concernant les caractéristiques des individus et allons décrire succinctement le contexte d'offre, le but étant de déterminer le ou les programmes qui reflètent la manière dont individus et les professionnels de santé choisissent les quantités de soins consommés.

1.2.2. Caractéristiques des consommateurs

Le comportement des agents, plus particulièrement des consommateurs, est analysé sur une seule période $[0, T]$ au cours de laquelle la santé des individus varie aléatoirement : les agents évoluent dans un environnement incertain et unitemporel, étant donné que nous considérons que l'individu ne fait pas d'arbitrage dans le temps sur ses consommations.

1.2.2.1. Les caractéristiques générales des consommateurs

Au début de la période, tous les individus disposent du revenu W^n et sont supposés ne pas être malades, leur état est noté 0. L'état de santé associé est noté h_0 .

Durant la période $[0, T]$, les individus consommateurs peuvent rester en bonne santé h_0 avec une probabilité π_0 ou tomber au plus une fois malade, avec une probabilité π_s , le niveau de gravité de leur pathologie étant gradué sur une échelle croissante $1, \dots, S$; ils peuvent ainsi être dans l'état de santé $h_s < h_0$; h_s est tel que pour $s_1 > s_0$ $h_{s_1} < h_{s_0}$; nous notons $\mathcal{S} = \{0, \dots, S\}$ les $S + 1$ états dans lesquels l'individu peut se trouver. Les individus peuvent restaurer partiellement ou totalement leur santé en consommant une quantité m_s de soins, qu'ils achètent au prix unitaire q_m , auprès d'un producteur de soins (ou professionnel de santé). La consommation de soins ampute les ressources disponibles pour des consommations non médicales d'un montant $q_m \cdot m_s$. Les individus sont donc soumis à un « risque » concernant d'une part leur santé et d'autre part leurs consommations non médicales x . Kaplan et Garrick (1981) caractérisent le risque de la manière suivante :

- Le risque est associé à un dommage (dégradation ou destruction de biens, pertes financières) consécutif à un évènement survenant aléatoirement ;
- Il représente la connaissance que les individus ont de chaque évènement possible, de la probabilité de survenue de l'évènement et des dommages qu'il engendre. Autrement dit, c'est un triplet (s, π_s, D_s) où e est l'évènement π_s la probabilité de survenue de l'évènement et D_s le dommage qu'il engendre.

La notion de risque nécessite donc que les individus soient non seulement conscients des différents évènements néfastes (ou « sinistres ») pouvant survenir mais également qu'ils aient une idée de la distribution des dommages occasionnés. Dans le domaine de la santé, les évènements vont être le fait d'être ou non malade et le niveau de gravité de la pathologie, π_s la probabilité de survenue de l'évènement et D_s le niveau de dommage. Le niveau de dommage peut être caractérisé par la perte $h_s - h_0$ en termes d'état de santé ou la perte financière $q_m \cdot m_s$ associée coût du traitement. Dans le premier cas, nous parlerons de risque santé et dans le deuxième cas de risque financier. La distinction est importante car l'individu peut choisir de modifier son risque santé et/ou son risque financier. Mathématiquement, le risque peut être résumé par la variable aléatoire \tilde{D} de niveau de sinistre. Ici, $\tilde{D} = \tilde{h}$ pour le risque santé et $\tilde{D} = q_m \cdot \tilde{m}$ pour le risque santé.

Ex-post la survenue du risque, N_{C_s} individus sont dans l'état s . Puisque pour chaque individu, il existe un état pathologique s tel que l'individu soit dans cet état. Nous avons donc :

$$N_C = \sum_{s \in \mathcal{S}} N_{C_s}$$

De plus, le nombre de consommateurs est supposé très grand, nous avons donc par la loi des grands nombres :

$$\frac{N_{C_s}}{N_C} \cong \pi_s$$

Pour décrire la manière dont les individus choisissent leurs consommations médicales et non médicales, nous faisons les hypothèses supplémentaires suivantes :

- **HUQ1** : L'état de santé avant consommation de soins est noté h_s . h_s décroît strictement lorsque le niveau s de sévérité de la pathologie augmente.
- **HUQ2** : La consommation de m_s unités de soins produit $g_h(m_s)$ unités d'état de santé. Ainsi, les consommations médicales font passer l'état de santé du niveau h_s au niveau $h_s + g_h(m_s)$.
- **HUQ3** : Lorsqu'ils sont dans l'état s ($s \in \mathcal{S}$), les individus attribuent une utilité u à leurs consommations non médicales x_s et $u_m(s, m_s) = u_h(h_s + g_h(m_s))$ à la consommation médicale m_s ; leur utilité totale est :

$$U_s(x_s, m_s) = u_x(x_s) + u_m(s, m_s)$$

- **HUQ4** : Les fonctions d'utilité u_x et u_h et la fonction de production de soins g_h respectent les conditions d'Inada : elles sont définies sur \mathbb{R}_+ , nulles en 0, continument dérivable, strictement croissantes et strictement concaves (autrement dit leur dérivée est strictement décroissante) sur \mathbb{R}_+ . Leur dérivée en 0 vaut sur $+\infty$ et leur dérivée en $+\infty$ vaut 0.

Sous les hypothèses précédentes concernant u_h , h_s et g_h , nous obtenons les résultats suivants concernant u_m ⁷:

⁷ Le résultat RUQ1 résulte simplement du fait que v_s est la composée de deux fonction concave positive et croissantes. Le résultat RUQ2 se démontre en dérivant $u_h(h_s + g_h(\cdot)) = u_m(s, \cdot)$.

$$u'_m(s, m) = g'_h(m) \cdot u'_h(h_s + g_h(m)) \quad \forall s \in \mathcal{S}$$

$$h_{s_1} + g'_h(m) < h_{s_0} + g'_h(m)$$

Donc, par stricte décroissance de v' :

$$g'_h(m) \cdot u'_h(h_{s_0} + g(m)) < g'_h(m) \cdot u'_h(h_{s_1} + g'_h(m))$$

Autrement dit :

$$u'_m(s_0, m) < u'_m(s_1, m)$$

- **RUQ1** : $u_m(s, \cdot)$ est une fonction positive, croissante et concave de la quantité de soins consommés.
- **RUQ2** : pour tout $m > 0$, $u'_m(s, m)$ croît strictement avec s . Autrement dit, pour tout $s_0 < s_1$:

$$u'_m(s_0, m) < u'_m(s_1, m)$$

Ainsi, l'utilité apportée par une unité de soins supplémentaire est d'autant plus grande que l'état de santé est dégradé.

Décrivons maintenant le choix des consommations médicales et non médicales opéré par les individus dans l'état s , faisant face au prix des soins q_m et au revenu W (le revenu net des transferts de ressources ayant eu lieu ex-ante la consommation de soins). Sous l'ensemble des hypothèses précédentes, la demande biens et services médicaux et non-médicaux est donnée par le programme suivant :

Programme 1.2-i

$$\max_{(x_s, m_s)} U_s(x_s, m_s)$$

$$sc: x_s + q_m \cdot m_s = W_s$$

La condition de premier ordre s'écrit :

Équation 1.2-i :

$$q_m \cdot \frac{\partial U_s}{\partial x} = \frac{\partial U_s}{\partial m}$$

De l'équation de premier ordre et de la concavité de U_s par rapport à x et m , nous obtenons le respect de la condition de second ordre :

$$-(q_m)^2 \cdot \frac{\partial^2 U_s}{\partial x^2} - \frac{\partial^2 U_s}{\partial m^2} > 0$$

Cette condition garantit l'unicité de la solution au programme de maximisation. Compte tenu des hypothèses de croissance et de concavité de la fonction d'utilité, cette condition est vérifiée. Nous nommons $(x_s^u(q_m, W), m_s^u(q_m, W))$ l'unique solution du programme 1.2-i et V_s la fonction d'utilité indirecte des individus :

$$V_s(q_m, W) = U_s(x_s^u(q_m, W), m_s^u(q_m, W))$$

Les fonctions de demande des individus dans les différentes situations considérées dans ce travail s'écriront à partir de (x_s^u, m_s^u) .

1.2.2.2. La demande de réduction risque financier

Ex-ante la réalisation du risque santé, les individus ont la possibilité de mener des actions qui réduisent ou accroissent leur risque financier, soit en augmentant le risque santé en lui-même, soit en augmentant les dépenses engendrées par ce risque. Considérons le revenu net des dépenses de santé, c'est-à-dire les ressources qui peuvent être allouées aux consommations non médicales. A espérance de revenu net identique, les individus qui choisissent une situation accroissant leur risque sont qualifiés de risquophiles, ceux qui choisissent la situation permettant de réduire leur risque sont qualifiés de risquophobes, ou plus communément d'averse au risque. Enfin, les individus indifférents au niveau de risque sont qualifiés de neutres au risque.

Les préférences vis-à-vis du risque sont communément modélisées au moyen de la fonction dérivant des quatre axiomes de Von-Neumann Morgenstern : complétude, transitivité, continuité et indépendance (*cf* Henriot et Rochet, 1991). Sous la condition que l'ensemble de ces axiomes soient respectés, l'utilité qu'un individu va attribuer à une variable aléatoire va s'écrire sous forme de l'espérance de l'utilité que l'individu peut retirer des réalisations de cette variable. Dans notre économie, la variable aléatoire en question est (\tilde{x}, \tilde{m}) , autrement dit, il s'agit du binôme « consommations non-médicales × consommations médicales ». L'espérance d'utilité s'écrit alors :

$$EU(\tilde{x}, \tilde{m}) = \sum_{s=0}^S \pi_s \cdot u_x(x_s) + \sum_{s=0}^S \pi_s \cdot u_m(s, m_s)$$

Sous l'hypothèse que les comportements des individus vis-à-vis du risque sont régis par la fonction d'utilité de Von-Neumann Morgenstern, il existe un lien fort entre la forme de la fonction d'utilité u_x et l'hypothèse d'aversion pour le risque. Plus précisément, la concavité de u_x implique les individus sont averses au risque puisque par l'inégalité de Jensen :

$$\sum_{s=0}^S \pi_s \cdot u_x(x_s) \leq u_x \left(\sum_{s=0}^S \pi_s \cdot x_s \right)$$

Le degré de risquophobie est alors mesuré par les indices R^A absolus ou relatifs R^R d'aversion pour le risque :

$$R^A = - \frac{u_x''(W^n)}{u_x'(W^n)} > 0 \text{ si l'individu est risquophobe}$$

$$R^R = -W \cdot \frac{u_x''(W^n)}{u_x'(W^n)} > 0 \text{ si l'individu est risquophobe}$$

Nous utiliserons dans l'ensemble de ce travail la fonction de Von-Neumann Morgenstern, malgré ses limites ; Le paradoxe d'Allais en particulier a remis en cause l'axiome d'indépendance sur lequel repose pour partie la validité de cette fonction d'utilité. Nous constatons de plus qu'*a priori*, u_x est concave quel que soit la valeur de W^n et du niveau de perte financière occasionné par le dommage : l'individu devrait donc toujours être averse au risque. Or, la simple observation de la pratique des jeux de hasard nous permet de conclure que l'individu est dans certaines situations risquophile. Des travaux basés sur la « prospect theory » ont abouti à la construction de fonction d'utilité permettant de contourner ces limites. Ces fonctions d'utilité ne seront pas utilisées dans ce travail, excepté dans la partie consacrée à l'accès aux soins.

Plusieurs actions peuvent être entreprises pour réduire le risque financier. La première solution consiste à agir directement sur la distribution des dommages (c'est-à-dire des pertes d'états de santé liés à la pathologie). On énumère ainsi deux types d'actions :

- L'autoprotection qui consiste à agir de manière à réduire la probabilité de survenue du sinistre : dans le domaine de la santé, il s'agit des actes de prévention primaire qui permettent de prévenir l'apparition de pathologie (par exemple la vaccination).
- L'auto-assurance qui consiste à diminuer les dommages occasionnés par le sinistre. Dans le domaine de la santé, il s'agit des actes de prévention secondaire qui permettent de limiter la gravité de la pathologie : par exemple les actes de dépistage des cancers qui permettent de détecter et de traiter la pathologie à un stade où elle est encore précoce.

Les comportements d'autoprotection et d'auto-assurance se traduisent également par la pratique d'activités et la consommation de certains biens non médicaux favorables à la santé (pratique du sport, alimentation équilibrée) et par la limitation de comportements et de consommations à risque (consommations de tabac, d'alcool...). Ainsi, la non-consommation de tabac réduit la probabilité de survenue de crises d'asthme, et leur niveau de sévérité : il s'agit donc d'un acte d'autoprotection et d'auto-assurance.

Un deuxième type d'action consiste à organiser *ex-ante* la survenue de la pathologie un transfert de revenu selon les états de la nature : l'individu recevra en cas de maladie un certain montant I_s , dépendant de sa dépense de santé M_s ; en contrepartie, son revenu est diminué en cas de bonne santé. Ce type de transfert de ressources peut se faire en théorie de manière décentralisée, au moyen d'échanges entre individus opérés en ouvrant un marché de biens contingents. Un bien

contingent (l, s) est un contrat qui garantit la livraison d'une unité du bien l si l'événement s se réalise (s appartenant à l'ensemble d'état de la nature \mathcal{S}). En contrepartie, l'acheteur du contrat (c'est-à-dire l'individu qui se fait livrer le bien) paie le prix unitaire p_{js} quelque soit l'état de la nature qui se réalise. L'individu i choisit la quantité $q_i^l(s)$ de bien contingent à acheter ou à vendre de manière à maximiser son utilité espérée.

Ce type de marché offre une base théorique pour des échanges entre états de la nature permettant de réduire le risque, mais il est impossible à mettre en œuvre de manière pratique, sous sa forme décentralisée. En effet, le nombre d'états de la nature est généralement immense. De plus, le nombre d'individus intervenant sur ce marché doit être très élevé de manière à ce que le comportement des individus soit concurrentiel (Laffont, 1985). Ainsi, avec N_C individus et $S + 1$ états de la nature, le nombre de contrats à mettre en place pour que le marché soit complet est de N_C^{S+1} . Ces produits, en grand nombre donc, sont également complexes puisqu'ils doivent tenir compte de la situation de l'ensemble des individus (Laffont, 1985). Or un contrat portant sur un objet ou une personne est déjà complexe, puisqu'il comporte souvent de nombreuses clauses sur les conditions de remboursement.

Néanmoins, la théorie des marchés contingent contient un cas particulier qui rend ce type de marché opérant (Laffont, 1985, Henriot et Rochet, 1991) : il s'agit du cas où le nombre d'individus échangeant leur risque est très grand et où un des agents est neutre au risque : l'agent neutre au risque acceptera de vendre des contrats à tous les autres agents et ainsi de se faire transférer l'intégralité de leur risque. En effet, si un grand nombre d'individus confient à cet agent leur risque, et que ces risques sont indépendants et identiquement distribués, la somme des risques devient un montant quasi-certain ; l'agent est donc en mesure d'offrir une couverture complète à l'ensemble des individus. Ce cas particulier où les échanges de risques se font de manière centralisée constitue le service d'assurance.

1.2.3. Les caractéristiques des producteurs

1.2.3.1. Présentation des trois types de producteurs

Le secteur des biens et services médicaux est constitué de N_m producteurs (chirurgiens, médecins, dentistes, auxiliaires), offrant différents types de soins (consultations, c'est-à-dire expertises, actes médicaux, pharmacie, lunettes, prothèses dentaires...), en différents lieux (ville, hôpital). Le secteur des biens et services non médicaux est constitué de N_x producteurs, offrant un bien composite qui agrège l'ensemble des biens et services n'appartenant pas au secteur médical (alimentation, logement, loisirs, etc...).

L'ensemble des producteurs est complété par le secteur de l'assurance santé. Nous supposons qu'il existe N_I opérateurs d'assurances (les assureurs) et supposons que ces opérateurs prélèvent auprès des assurés une prime P_I et versent quel que soit l'état pathologique qui se réalise et versent en cas de réalisation de l'état pathologique s , le montant I_s .

A l'instar des hypothèses faites sur les professionnels de santé, les assureurs sont supposés présenter des caractéristiques identiques de sorte que les individus se répartissent uniformément sur chacun d'entre eux. De plus, le nombre $N_C^I = N_C/N_I$ d'assuré par assureur est supposé très grand de sorte que par la Loi des grands nombres :

$$\frac{N_{C_s}^I}{N_C^I} \cong \pi_s$$

$N_{C_s}^I = N_{C_s}/N_I$ étant le nombre d'assurés dans l'état s , au sein de chaque assureur. Ainsi, *ex-post* la réalisation du risque santé, les assureurs ont tous le même pool d'assuré et le risque présenté par ce pool est supposé nul.

1.2.3.2. Hypothèse sur la structuration des marchés

Hypothèse de concurrence pure et parfaite sur l'ensemble des secteurs

Dans la majeure partie de ce travail, nous supposons que la concurrence entre producteurs au sein de chacun des secteurs est pure et parfaite. Dranove et Satterthwaite (2000) rappellent les conditions qui doivent être satisfaites pour obtenir un tel marché : un grand nombre de consommateurs et de producteurs (entraînant l'atomicité de ces agents) ; la liberté d'entrer et de sortir du marché ; l'existence d'un marché pour chaque bien et service, y compris le risque ; la symétrie de l'information entre agents et un accès gratuit à cette information ; l'absence de rendements croissants, d'externalités et de collusions. Nous pouvons ajouter l'homogénéité des produits ce qui nous conduit à supposer que tous les producteurs de biens et services médicaux, ceux de biens et services non-médicaux et les assureurs présentent des caractéristiques identiques. La présentation théorique standard du phénomène de risque moral *ex-post* en santé (par exemple, Pauly, 1968) utilise implicitement l'hypothèse de concurrence pure et parfaite : l'analyse de l'effet d'une diminution du prix des soins sur la quantité de biens et services médicaux consommés est réalisée au moyen d'un graphique où la courbe de demande est horizontale, traduisant une élasticité-prix infinie de l'offre de soins. Dans le contexte qui vient d'être exposé, la concurrence pure et parfaite représente un cas particulier où le coût moyen des unités produites est égal à leur coût marginal.

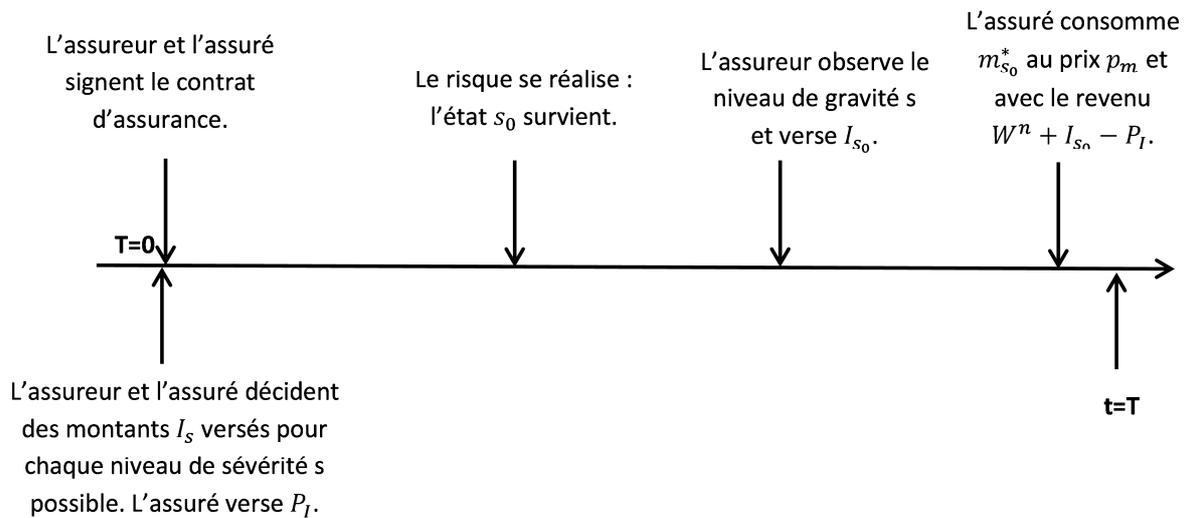
Hypothèse de concurrence monopolistique sur le secteur des biens et services médicaux

L'hypothèse de concurrence pure et parfaite facilite fortement les analyses, mais elle est extrêmement restrictive compte tenu des caractéristiques du marché de la santé. Arrow (1963) présente les aspects qui font que ce marché s'écarte du cadre de la concurrence pure et parfaite insistant entre autre sur les barrières à l'entrée de nouveaux professionnels de santé et sur les discriminations par les prix. Par ailleurs, McGuire (2000) note que du fait des différences de localisation, de spécialisation, les patients ne considèrent pas les professionnels de santé comme parfaitement substituables les uns aux autres. Ces caractéristiques ainsi que l'information imparfaite peuvent générer un marché où les producteurs de soins sont en concurrence monopolistique, cadre moins restrictif où les producteurs ont un pouvoir de marché et peuvent réaliser un profit non nul, au moins à court terme. Ce type d'organisation correspond à la seconde hypothèse sur laquelle nous porterons notre attention, dans la partie 2.8.1.

1.3. Equilibre dans une économie avec assurance, en information symétrique

Nous supposons que l'assureur et l'assuré disposent d'une information symétrique sur le risque de l'assuré, autrement dit que le risque de l'assuré est connu de manière identique par les deux parties. Dans ce cas de figure, l'assureur a la capacité d'observer les comportements entraînant un accroissement du risque santé, *ex-ante* la réalisation de ce risque, il est ainsi en mesure de sanctionner ces comportements par une augmentation de la prime ou une diminution des remboursements ; *ex-post* la réalisation de ce risque, il peut observer l'état de santé de l'individu et opérer un transfert de revenu contingent à cet état, indépendamment de la quantité de soins que l'individu va effectivement consommer, ce transfert intervenant *ex-ante* la consommation de soins. L'individu peut alors allouer la somme transférée à ses consommations médicales ou non médicales. Le montant de revenu transféré est décidé entre les deux parties – assureur et assuré – au moment de la signature du contrat. Le séquençage du comportement des agents sur la période $[0, T]$ est décrit dans le graphique 1.3-i ci-après.

Graphique 1.3-i : séquençage du comportement de l'assureur et de l'assuré sur la période [0,T] sur laquelle court le contrat, en information symétrique



1.3.1. L'offre des assureurs en information symétrique

L'offre des assureurs se présente sous la forme de $S + 1$ contrats élémentaires, notés $(e_s)_{s \in \mathcal{S}}$ qui remboursent un euro lorsque le souscripteur est dans l'état s . Le coût de chaque contrat e_s est noté p_{e_s} . Tout contrat souscrit est une combinaison linéaire de ces contrats élémentaires : le contrat $I = (I_s)_{s \in \mathcal{S}}$, remboursant I_s euros lorsque l'état s survient est obtenu en achetant une quantité I_s de chaque contrats e_s . Le coût P_I du contrat s'écrit :

Équation 1.3-i :

$$P_I = \sum_{s \in \mathcal{S}} p_{e_s} \cdot I_s$$

Le contrat I permet donc à ses souscripteurs de bénéficier, lorsque l'état s se réalise, du transfert net de revenu :

$$\Delta W_s = I_s - P_I$$

Un tel contrat qui opère des remboursements conditionnellement à la nature du sinistre est connu dans la littérature anglophone sous le nom de « contingent claims contract » ou « indemnity contract ». Zeckhauser (1970) note :

The indemnity policy is the simplest health insurance policy. In effect, it operates as a contingent claims market; people get paid specified amount depending on which contingency occurs.

Les contrats offrant des remboursements contingents à l'état de santé servent de situation de référence mais n'ont pas d'existence concrète sur le marché de l'assurance « frais de soins de

santé » du fait de l'impossibilité pour l'assureur d'observer précisément l'état de santé des individus, comme nous le repréciserons dans la partie 1.4. Ils existent cependant sur d'autres segments de l'assurance de personnes tels que l'assurance dépendance, couverture qui se présente souvent sous forme d'une rente que l'individu peut utiliser pour acheter des biens et services permettant de pallier son état ou pour toute autre consommation ; cette rente est déclenchée à partir du moment où l'invalidité dépasse un certain niveau (mesuré souvent sur la grille Aggir). Le fait qu'il soit plus aisé d'objectiver qu'un état de dépendance qu'un état de santé quelconque peut expliquer que les contrats dont les remboursements sont contingents à l'état de santé soient offerts dans le cas de la dépendance et pas dans le cas des frais de soins de santé.

Nous supposons que les remboursements des frais de soins de santé sont les seuls coûts auxquels les assureurs font face. Puisque ces derniers sont supposés en concurrence pure et parfaite, les contrats sont tarifés au risque actuariel : leur coût est égal à l'espérance d'utilité de chaque individu. Nous avons donc $p_{e_s} = \pi_s$; le coût P_I de chaque contrat est donné par :

Équation 1.3-ii :

$$P_I = \sum_{s \in \mathcal{S}} \pi_s \cdot I_s$$

Sous cette dernière hypothèse, le transfert de revenu opéré par l'assurance est actuariellement neutre, autrement dit l'espérance des revenus transférés est nulle :

Équation 1.3-iii :

$$\sum_{s \in \mathcal{S}} \pi_s \cdot \Delta W_s = 0$$

Ex-post la réalisation du risque chaque individu dans l'état s se voit transférer la quantité ΔW_s de ressources, ce qui lui permet de bénéficier du revenu $W_s = W^n + \Delta W_s$ au lieu de W^n .

1.3.2. La demande d'assurance et de soins des individus

1.3.2.1. Choix des consommations médicales pour un niveau d'assurance donné

Ex-post la réalisation du risque, les individus qui disposent d'une assurance offrant le profil de transferts de revenu $(\Delta W_s)_{s \in \mathcal{S}}$, choisissent leurs consommations d'après le programme suivant :

Programme 1.3-i

$$\max_{(x_s, m_s)} U_s(x_s, m_s)$$

$$sc: x_s + p_m \cdot m_s = W^n + \Delta W_s$$

Ce programme correspond au problème d'optimisation élémentaire exposé dans la partie 1.2.2. Il admet une unique solution $(x_s^u(p_m, W^n + \Delta W_s), m_s^u(p_m, W^n + \Delta W_s))$. L'utilité retirée par l'individu de ces consommations s'écrit :

$$V_s(p_m, W^n + \Delta W_s) = U_s(x_s^u(p_m, W^n + \Delta W_s), m_s^u(p_m, W^n + \Delta W_s))$$

1.3.2.2. **Choix du niveau d'assurance ex-ante et des consommations ex-post**

Les modèles « simples » de demande d'assurance en information symétrique considèrent comme données (donc totalement exogènes) les consommations occasionnées par le risque auquel est soumis l'individu. Au moyen de tels modèles Mossin (1968) et Ehrlich et Becker (1972) ont montré qu'en information symétrique, lorsque les contrats sont tarifés au risque actuariel, le niveau d'assurance préféré par les individus est une couverture complète. Ce type de modélisation pose problème par rapport à un de nos objectifs qui est de déterminer comment l'assurance santé modifie les choix de consommations de soins et l'allocation des ressources en information symétrique ; cet objectif impose que les individus puissent choisir leurs consommations médicale et donc leur assurance santé au regard de ces consommations. Nous allons relâcher l'hypothèse d'exogénéité stricte des consommations médicales et allons considérer un modèle où l'individu choisit simultanément son niveau d'assurance et son niveau de consommations médicales. Avec un tel modèle, nous obtenons également que les individus demandent une assurance complète, sous réserve toutefois que la fonction d'utilité des individus soit séparable selon les consommations non-médicales et médicales.

Nous supposons donc qu'ex-ante la réalisation du risque, les individus choisissent les transferts de revenu qu'ils vont recevoir dans chaque état s au regard des quantités de biens et services médicaux et non-médicaux qu'ils s'attendent à consommer. Le choix du profil de remboursements optimal est ainsi décrit par le problème suivant :

Programme 1.3-ii

$$\max_{\substack{(x_s, m_s) \in \mathbb{R}_+^{2S+2} \\ (\Delta W_s) \in \mathbb{R}^{S+1}}} \sum_{s \in \mathcal{S}} \pi_s \cdot U_s(x_s, m_s)$$

$$sc: x_s + p_m \cdot m_s = W^n + \Delta W_s \quad \forall s \in \mathcal{S}$$

$$sc: \sum_{s \in \mathcal{S}} \pi_s \cdot \Delta W_s = 0$$

Il est possible de faire disparaître les $S + 1$ premières contraintes en réécrivant le programme sous la forme suivante :

Programme 1.3-iii

$$\begin{aligned} \max_{\substack{m_s \in \mathbb{R}_+^{S+1} \\ (\Delta W_s) \in \mathbb{R}_+^{S+1}}} \sum_{s \in \mathcal{S}} \pi_s \cdot U_s(W^n + \Delta W_s - p_m \cdot m_s, m_s) \\ sc: \sum_{s \in \mathcal{S}} \pi_s \cdot \Delta W_s = 0 \end{aligned}$$

La fonction optimisée étant concave et les contraintes linéaires, ce problème admet une unique solution que nous notons $((m_s^*), (\Delta W_s^*)) \in \mathbb{R}_+^{S+1} \times \mathbb{R}^{S+1}$. Nous allons montrer que :

$$\Delta W_s^* = p_m \cdot m_s^* - \sum_{s \in \mathcal{S}} \pi_s \cdot p_m \cdot m_s^*$$

Soit un profil de transfert (ΔW_s) respectant l'égalité $\sum_{s \in \mathcal{S}} \pi_s \cdot \Delta W_s = 0$. L'inégalité de Jensen permet d'écrire :

$$\begin{aligned} \sum_{s \in \mathcal{S}} \pi_s \cdot [u_x(W^n + \Delta W_s - p_m \cdot m_s^*) + u_m(m_s^*)] \\ \leq u_x \left(W^n + \sum_{s \in \mathcal{S}} \pi_s \cdot (\Delta W_s - p_m \cdot m_s^*) \right) + \sum_{s \in \mathcal{S}} \pi_s \cdot u_m(m_s^*) \end{aligned}$$

Puisque $\sum_{s \in \mathcal{S}} \pi_s \cdot \Delta W_s = 0$, cette égalité devient :

$$\sum_{s \in \mathcal{S}} \pi_s \cdot [u_x(W^n + \Delta W_s - p_m \cdot m_s^*) + u_m(m_s^*)] \leq u_x \left(W^n - \sum_{s \in \mathcal{S}} \pi_s \cdot p_m \cdot m_s^* \right) + \sum_{s \in \mathcal{S}} \pi_s \cdot u_m(m_s^*)$$

Donc :

$$\sum_{s \in \mathcal{S}} \pi_s \cdot U_s(W^n + \Delta W_s - p_m \cdot m_s^*, m_s^*) \leq \sum_{s \in \mathcal{S}} \pi_s \cdot U_s \left(W^n - \sum_{\tau \in \mathcal{S}} \pi_\tau \cdot p_m \cdot m_\tau^*, m_s^* \right)$$

Nous constatons qu'en posant $I_s^* = p_m \cdot m_s^*$ (avec $(I_s^*) \in \mathbb{R}_+^{S+1}$ le profil de remboursement optimal), le niveau de prime vaut $P_{I^*} = \sum_{s \in \mathcal{S}} \pi_s \cdot p_m \cdot m_s^*$ et le transfert de revenu $\Delta W_s^* = I_s^* - \sum_{s \in \mathcal{S}} \pi_s \cdot p_m \cdot m_s^*$:

$$\sum_{s \in \mathcal{S}} \pi_s \cdot U_s(W^n + \Delta W_s^* - p_m \cdot m_s^*, m_s^*) = \sum_{s \in \mathcal{S}} \pi_s \cdot U_s \left(W^n - \sum_{\tau \in \mathcal{S}} \pi_\tau \cdot p_m \cdot m_\tau^*, m_s^* \right)$$

Par conséquent, pour tout profil de transfert (ΔW_s) qui vérifie $\sum_{s \in \mathcal{S}} \pi_s \cdot \Delta W_s = 0$, nous avons :

$$\sum_{s \in \mathcal{S}} \pi_s \cdot U_s(W^n + \Delta W_s - p_m \cdot m_s^*, m_s^*) \leq \sum_{s \in \mathcal{S}} \pi_s \cdot U_s(W^n + \Delta W_s^* - p_m \cdot m_s^*, m_s^*)$$

(ΔW_s^*) est donc le profil de transfert préféré : $I_s^* = p_m \cdot m_s^*$, la couverture est donc complète : le niveau d'assurance préféré par les individus égalise les consommations non médicales selon les états s .

Pour déterminer m_s^* , il suffit d'écrire le lagrangien associé au programme 1.3-ii. Des conditions de premier ordre et des contraintes de budgets, nous déduisons que pour tout s :

$$\frac{\partial U_s}{\partial m} \Big|_{(x_s^*, m_s^*)} = p_m \cdot \frac{\partial U_s}{\partial x} \Big|_{(x_s^*, m_s^*)}$$

$$x_s^* + p_m \cdot m_s^* = W^n + \Delta W_s$$

Ces conditions sont identiques à celles du Programme 1.3-i régissant les consommations de soins *ex-post* la survenue de la pathologie, dans le cas où le profil de transferts est (ΔW_s^*) : les consommations réalisées *ex-post* correspondant aux consommations anticipées *ex-ante* :

$$x_s^* = x_s^u(p_m, W^n + \Delta W_s^*) = x_s^u \left(p_m, W^n + p_m \cdot m_s^* - \sum_{\tau \in \mathcal{S}} \pi_\tau \cdot p_m \cdot m_\tau^*, m_\tau^* \right)$$

$$m_s^* = m_s^u(p_m, W^n + \Delta W_s^*) = m_s^u \left(p_m, W^n + p_m \cdot m_s^* - \sum_{\tau \in \mathcal{S}} \pi_\tau \cdot p_m \cdot m_\tau^*, m_\tau^* \right)$$

Enfin, toujours d'après les conditions de premier ordre du Programme 2.3-ii, nous avons $\forall s_0, s_1 \in \mathcal{S}$:

$$\frac{\partial U_{s_0}}{\partial m} \Big|_{(x_{s_0}^*, m_{s_0}^*)} = \frac{\partial U_{s_1}}{\partial m} \Big|_{(x_{s_1}^*, m_{s_1}^*)}$$

Donc $u'_m(s_0, m_{s_0}) = u'_m(s_1, m_{s_1})$. D'après les conditions énoncées dans partie 1.2.2.1, $u'_m(s_0, m_{s_0}) < u'_m(s_1, m_{s_0})$. Donc $u'_m(s_1, m_{s_1}) < u'_m(s_1, m_{s_0})$. Par stricte décroissance de u'_m , $m_{s_1} > m_{s_0}$.

L'ensemble de ces résultats sont synthétisés dans la proposition suivante :

Proposition 1.3-1 : En situation d'information symétrique, lorsque le contrat est tarifé à son coût actuariel et la fonction d'utilité est séparable selon les consommations médicales et non médicales :

- Le niveau d'assurance préféré égalise les consommations non médicales : la couverture est complète.
- Les remboursements sont égaux à $p_m \cdot m_s^*$.
- Les transferts de revenu sont égaux à :

$$\Delta W_s^* = p_m \cdot m_s^* - \sum_{\tau \in \mathcal{S}} \pi_\tau \cdot p_m \cdot m_\tau^*, m_\tau^*$$

- Les consommations de soins ex-post la survenue de la pathologie vérifient :

$$m_s^*(p_m, W^n) = m_s^u(p_m, W^n + \Delta W_s^*) = m_s^u \left(p_m, W^n + p_m \cdot m_s^* - \sum_{\tau \in \mathcal{S}} \pi_\tau \cdot p_m \cdot m_\tau^*, m_\tau^* \right)$$

- Les consommations de soins m_s^* sont croissantes avec le niveau de sévérité s de la pathologie.

En conclusion, avec ce modèle où l'individu choisit simultanément ses consommations de soins et son niveau d'assurance au regard de ces consommations, nous obtenons des résultats analogues à ceux de Mossin (1968) et Ehrlich et Becker (1972), à savoir que l'individu choisit une assurance complète.

1.3.2.3. Analogie avec le problème de taxation optimale du revenu en information symétrique

Réécrivons le programme 1.3-i en remplaçant x_s et m_s par x_s^* et m_s^* . La maximisation se fait alors par rapport aux seuls transferts de revenu :

Programme 1.3-iv

$$\max_{(\Delta W_s) \in \mathbb{R}^{\mathcal{S}+1}} \sum_{s \in \mathcal{S}} \pi_s \cdot V_s(p_m, W^n + \Delta W_s)$$

$$SC: \sum_{s \in \mathcal{S}} \pi_s \cdot \Delta W_s = 0$$

Ce programme appelle deux remarques :

- *Ex-ante* la réalisation du risque, le choix du contrat optimal est équivalent au choix d'un transfert de revenu parmi les transferts actuariellement neutres.
- *Ex-post* la réalisation du risque, la somme des transferts est nulle, le budget alloué à la couverture du risque santé est donc budgétairement équilibré : l'assurance santé opère une simple redistribution des richesses qui pourrait être obtenue par une intervention publique *ex-post*. La recherche du transfert optimal est donc équivalente à un problème de taxation non linéaire optimale du revenu visant à gommer les disparités de richesses résultant des différences d'état de santé.

Réécrit convenablement, le programme de la partie précédente est analogue à un problème de taxation optimale. Par résolution du programme 1.3-iv, nous obtenons que les transferts de revenu sont tels que les individus obtiennent la même utilité marginale du revenu quel que soit l'état de la nature s :

$$\left. \frac{\partial V_0}{\partial W} \right|_{(p_m, W^n + \Delta W_0)} = \left. \frac{\partial V_1}{\partial W} \right|_{(p_m, W^n + \Delta W_1)} = \dots = \left. \frac{\partial V_s}{\partial W} \right|_{(p_m, W^n + \Delta W_s)}$$

Cette équation traduit simplement le fait que les individus transfèrent du revenu des états associés à des niveaux de sévérité faibles vers les états associés à des niveaux de sévérité élevés jusqu'à égaliser les utilités marginales du revenu pour tous les états de la nature.

1.3.3. Effet de l'assurance sur les consommations médicales et non médicales

Par rapport à un individu non couvert, les consommations médicales et non médicales d'un assuré sont :

- Moins élevées lorsqu'elles sont associées à un état peu sévère ;
- Plus élevées lorsqu'elles sont associées à un état sévère.

En effet, dans le premier cas, les transferts nets de revenu (c'est-à-dire les remboursements nets de la prime) sont négatifs, faisant décroître les consommations médicales et non-médicales. Dans le deuxième cas, les transferts nets de revenus sont positifs, faisant croître les deux types de consommations.

Deux points sont à noter. Le premier, et sans doute le plus important, est qu'en information symétrique, l'assurance modifie la consommation de soins par un effet revenu. Le transfert de

revenu qui s'opère des individus en bonne santé vers ceux en mauvaise santé permet à ces derniers de solvabiliser leur demande de soins. Ainsi, les transferts de revenu opérés par l'assurance permettent non seulement de réduire la variabilité du risque financier à laquelle font face les individus mais également d'améliorer l'accès aux soins des individus en cas de maladie. Nous approfondirons cet effet dans la partie 1.9). Nous pouvons également noter que la variabilité des dépenses est plus importante pour les individus couverts que pour les individus non couverts ; en effet, la couverture réduit les consommations médicales en cas de bonne santé et les augmente en cas de mauvaise santé. Cet accroissement du risque financier n'entraîne aucune désutilité chez l'assuré puisque celui-ci est totalement couvert.

1.3.4. Pareto optimalité de l'équilibre

Pour tout état s , la condition de premier du programme permet d'établir que :

$$q_m = TMS_s(x_s^u, m_s^u) = \frac{\frac{\partial U_s}{\partial m} \Big|_{(x_s^u, m_s^u)}}{\frac{\partial U_s}{\partial x} \Big|_{(x_s^u, m_s^u)}} = p_m = TMT_z(x^P, m^P)$$

Ainsi, les consommations médicales et non médicales des individus sont telles que :

- Les taux marginaux de substitution TMS_s sont égaux pour tout s : à niveau de ressources inchangé, il n'est donc pas possible d'accroître strictement l'utilité d'un agent sans réduire l'utilité d'un autre agent.
- (x_s^*, m_s^*) se situe sur $\partial \mathcal{P}_z$, la frontière de l'ensemble \mathcal{P}_z des productions possibles : il n'est donc pas possible d'accroître strictement l'utilité de l'ensemble des agents en accroissant le niveau de ressources utilisées.

L'équilibre est donc Pareto optima. Après avoir étudié de manière théorique la demande de couverture et de soins en information symétrique, Zeckhauser (2000) conclue ainsi :

"The amount paid equals the cost of the appropriate treatment for the person's disease; if there is more than one disease, the payments vary. Since each disease requires a fixed amount of care --there is no more nor less that a person can consume -- there are no wasted resources in the policy"

Nous allons voir qu'en information asymétrique, l'offre de couverture et les choix de consommations de soins des individus sont tels que le caractère optimal de l'équilibre n'est plus assuré.

1.4. De l'asymétrie d'information au risque moral ex-post

1.4.1. Position du problème

L'hypothèse d'information symétrique suppose que l'individu transmette à l'assureur toute l'information dont il dispose sur son état de santé. Nous constatons aisément qu'un système qui reposerait uniquement sur la bonne foi de l'assuré n'est pas viable, puisque celui-ci a toujours intérêt à se déclarer dans l'état de santé le plus mauvais afin de bénéficier du plus haut niveau de remboursement, autrement dit, du transfert de revenu le plus avantageux. Ainsi, si l'assuré est dans l'état $s_0 < S$, il va juger plus profitable de se déclarer dans l'état S (le pire) afin de bénéficier du transfert de revenu $p_m \cdot m_S^*$ au lieu de $p_m \cdot m_{s_0}^*$, le montant qui lui aurait été versé s'il n'avait pas majoré son niveau de sévérité. Le supplément de ressource $p_m \cdot (m_S^* - m_{s_0}^*)$ lui permettra de consommer plus de biens et services médicaux et non médicaux. Il se fera donc systématiquement verser la somme $p_m \cdot m_S^*$. Or, cette somme est strictement supérieure à la prime, l'assureur va donc subir un déficit.

Une manière de solutionner ce problème consiste à supposer qu'il existe un expert médical indépendant qui estime l'état de santé de l'individu et le transmet à chacune des deux parties. Pour que ce mode de transmission de l'information fonctionne de manière symétrique, il semble important que la fonction d'expertise soit dissociée de celle de producteur de soins. En pratique ce sont les mêmes agents qui réalisent l'expertise, sous la forme d'un diagnostic, et qui exécutent la prestation découlant de l'expertise, sous la forme d'une offre de soins⁸. Dans ce cas, le professionnel de santé peut avoir intérêt à favoriser une des deux parties, *a priori* celle dont dépend sa rémunération. Dans le cas d'une collusion avec l'assureur, il a tout intérêt à minorer la sévérité de la pathologie de l'individu de manière à ce que l'assureur verse moins de prestations. Dans le cas d'une collusion avec l'assuré, il a au contraire intérêt à majorer le niveau de sévérité de manière à ce que l'assuré bénéficie d'un supplément de prestations.

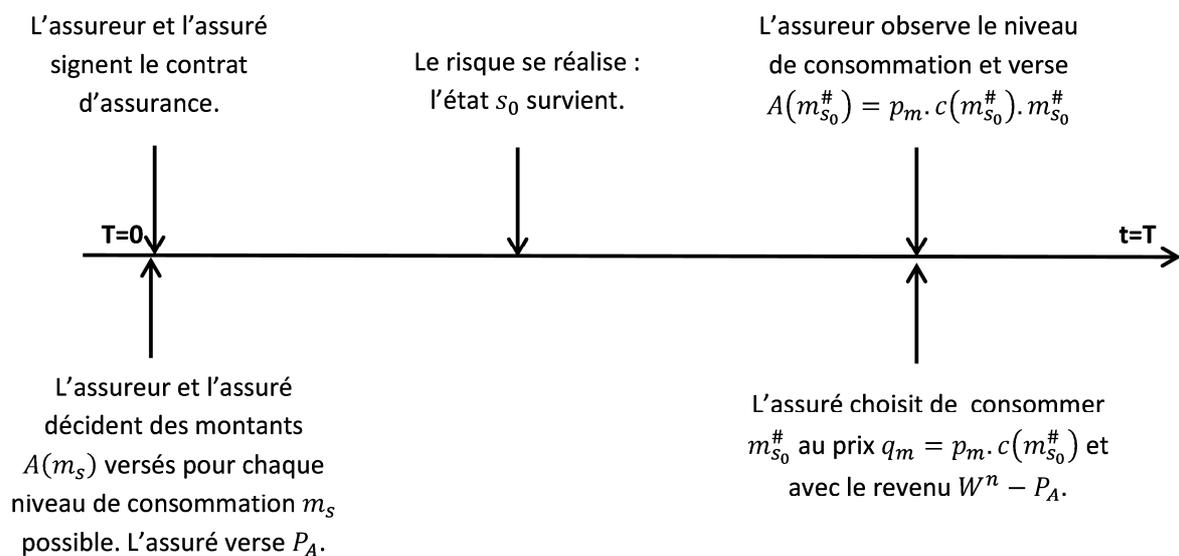
Ne pouvant pas se baser directement sur l'état de santé pour verser les prestations, l'assureur va offrir des remboursements conditionnels aux consommations médicales. Puisque par nature ces consommations ne sont connues qu'une fois qu'elles sont réalisées, l'assureur effectue ses remboursements après que l'individu ait décidé de la quantité de soins m_s qu'il va consommer. A la différence de ce qui se passe en information symétrique, les remboursements ne peuvent par nature qu'être reliés aux consommations médicales et non à l'état de santé. Ces remboursements viennent

⁸ Ce problème est commun à d'autres professions, telles que celle de conseiller patrimonial qui conseil les individus et place leurs avoirs.

donc diminuer le coût des soins au moment de leur achat, modifiant les consommations de soins par effet prix.

Le séquençage du comportement de l'assuré en information asymétrique est donné dans le graphique 1.4-i ci-dessous. Dans ce schéma, A représente la fonction de remboursement en fonction de la quantité de soins consommés P_A la prime associée à A et $m_s^\#$ la demande de soins des individus en information asymétrique. Cette demande sera étudiée précisément dans la partie 1.5.

Graphique 1.4-i : séquençage du comportement de l'assureur et de l'assuré sur la période $[0,T]$ sur laquelle court le contrat, en information asymétrique



La baisse du coût apparent des soins induite par l'assurance santé en situation d'information asymétrique génère un comportement dit d' « aléa moral *ex-post* ». De manière générale, l'adoption, de comportements d'aléa moral résulte de deux ingrédients :

- d'une part, une incitation pour l'assuré à accroître son risque financier ; l'individu se retrouve face à des choix qui lui apportent de l'utilité mais qui accroissent le risque financier.
- D'autre part, une impossibilité pour l'assureur de sanctionner ces comportements d'accroissement du risque en augmentant de manière individuelle les primes ; en effet, en situation d'information asymétrique, l'assureur est incapable d'observer les comportements des individus *ex-ante* la réalisation du risque et leur état *ex-post*. Il est donc dans l'incapacité de juger si une consommation fréquente ou une dépense élevée est justifiée ou non.

Le risque moral désigne donc le fait que face à la possibilité d'actions dont il retire de l'utilité mais qui accroît son risque financier, l'individu choisit d'adopter ces actions car leurs conséquences financières directes (les dépenses dont l'individu doit s'acquitter au moment de l'achat des soins)

sont partiellement réduites ou totalement supprimées par la couverture assurantielle. Dionne (1981) rappelle qu'il existe deux catégories de risque moral dans la littérature sur l'assurance :

- Le risque moral *ex-ante* qui est défini selon Dionne comme la réduction de la part de l'assuré des actions d'autoprotection. En effet, ces actions sont difficiles à observer par l'assureur, celui-ci ne peut donc pas faire varier la prime en fonction des efforts de l'assuré : l'assuré est donc moins incité à pratiquer des efforts d'autoprotection.
- Le risque moral *ex-post* est défini comme l'accroissement des consommations des biens et services couverts par l'assurance due à la diminution du prix de ces biens et services au moment de leur achat. L'assurance subventionne les soins des personnes qu'elle couvre ; par rapport au cas où ils ne sont pas couverts, les assurés consomment un supplément de soins qu'elles valorisent en deçà de leur coût réel, mais au-delà du prix payé au moment de leur achat. L'assureur est dans l'impossibilité de déterminer l'état de l'individu, il peut juste observer les consommations générées par cet état. Il s'agit exactement du contexte que nous considérons et que nous avons décrit dans le graphique 1.4-i.

Dans les deux cas, le coût de l'action entraînant un surcroît de consommation est sous-estimé par les individus, du fait des remboursements opérés par l'assurance. Or, si l'assureur est incapable de déterminer précisément quels individus surconsomment, sa connaissance statistique lui permet de déterminer quel sera le montant de la dépense associée à ces surconsommations sur l'ensemble de son pool. Il peut donc se prémunir contre un risque de perte en accroissant la prime de l'ensemble des assurés. Via cette augmentation de prime, les assurés payent donc finalement l'intégralité du coût de leur surcroît de consommations, alors même qu'ils valorisent ce surcroît de consommation en dessous de ce coût. Les individus allouent ainsi trop de ressources à leurs soins et pas assez aux autres biens et services, ce qui correspond à un phénomène d'inefficience allocative (voir partie 2.6.1). Il y a perte de surplus collectif, perte que nous étudierons dans la partie 2.6.2.

Si cette situation est problématique, pourquoi les individus adoptent-ils ce comportement de risque moral ? L'explication tient à un phénomène de dilemme du prisonnier : Pauly (1968) souligne que tous les individus auraient intérêt à coopérer en maintenant leurs consommations de soins au niveau atteint avant la souscription de la couverture. Cependant, ils reçoivent un supplément d'utilité en consommant plus, et ils ne supportent individuellement qu'une part très faible du coût de cette surconsommation, le reste étant transféré aux autres assurés (*risk spreading* entraînant une externalisation du coût du supplément de consommations de soins sur l'ensemble de la population). Par conséquent, l'individu n'a pas intérêt à effectuer un effort de modération de ses consommations. Dowd (1982) détaille le « jeu » auquel sont confrontés les individus et l'équilibre de Nash qui en

découle. Il insiste sur le fait que les individus considèrent la prime comme un coût fixe déterminé de manière exogène, indépendant de leur niveau de consommation personnel.

Le risque moral *ex-ante*, qui se traduit par une augmentation de la probabilité de développer une pathologie, est à la source de nombreux modèles dans les domaines du risque moral et de la théorie de l'agence (cf par exemple Arnott et Stiglitz 1988, Henriot et Rochet 1991, Laffont et Martimort, 2009) mais il est considéré comme marginal dans le domaine de l'assurance santé du fait de l'importance des coûts non monétaires (temps de trajet, d'attente chez le médecin, augmentation du risque de décès, des incapacités, douleur) non couverts par l'assurance santé (Bardey et al, 2002). Le risque moral *ex-post* en revanche est une problématique prégnante du fait de la multitude des sources d'asymétrie d'information concernant le coût des soins (nature et niveau de gravité de la pathologie, mais également efficacité du traitement) ; il s'agit d'un enjeu important en termes d'efficacité du système de santé, enjeu qui se traduit par de nombreux travaux sur ce sujet (cf Zeckhauser, 2000).

1.4.2. Les apports historiques à la théorie du risque moral

Les notions d'agence et de risque moral apparaissent dans des domaines aussi variés que l'économie du travail, l'économie industrielle, l'économie environnementale, l'économie des marchés financiers. Si la théorie du risque moral en santé a été développée à partir des années 1960 par Arrow et Pauly, le concept de risque moral a été appliqué depuis bien plus longtemps à d'autres branches de l'économie. Zeckhauser (2000), fait remonter les premières réflexions sur ce sujet à Adam Smith (1776). Suivant l'historique que fait Laffont et Martimort (2009), nous pouvons aller plus loin encore, en nous basant sur l'analyse de Hume (1740). Celui-ci a été le premier à faire référence à la notion de passager clandestin (Laffont et Martimort, 2009) : il montre ainsi au travers de l'exemple d'individus qui veulent creuser un canal de drainage que lorsque le nombre d'agents est important, il y a difficulté à planifier et à mettre en œuvre un projet en commun. Lorsque les agents sont deux, chacun sait qu'il est observé par l'autre et que la réussite du projet dépend de son propre effort. Par contre, lorsque les individus sont mille, les responsabilités sont diluées si bien que personne n'a intérêt à faire sa part du travail. Hume conclue ainsi :

“Each seeks a pretext to free himself of the trouble and expense, and wou'd lay the whole burden to others”.

Laffont et Martimort (2009) citent, toujours sur le comportement de passager clandestin, Wicksell (1896) qui généralise le problème induit par ce comportement au financement des biens publics. Il énonce ainsi qu'un individu est enclin à n'effectuer aucune dépense pour un bien public, quel que soit la valeur qu'il accorde à ce bien. En effet, le caractère atomistique de chaque agent pris

isolément entraîne une dilution de la responsabilité individuelle. Chaque agent pense que son seul propre comportement ne va pas influencer les finances publiques, mais cette croyance étant partagée par l'ensemble des acteurs, personne ne va être prêt à financer les biens publics.

Les premières analyses du risque moral en santé

La première analyse significative des problèmes économiques posés par les phénomènes de risque moral a été réalisée par Arrow dans son article « *uncertainty and the welfare economics of medical care* ». Dans cet article, Arrow examine dans quelle mesure le marché soins médicaux s'écarte du modèle compétitif, aboutissant à des équilibres qui ne sont pas Pareto-optimaux. Il s'interroge notamment sur le fait qu'« un nombre important de risques ne sont pas couverts, et de ce fait, les marchés des services de couverture des risques sont peu ou pas développés » alors même que « la maladie est dans une très large mesure un phénomène imprédictible ». Les individus devraient retirer un gain en bien être en transférant le risque lié au coût de la maladie aux assureurs. Pourtant, à l'époque où Arrow écrit son article, seule une faible proportion des dépenses de santé sont couvertes⁹.

Pour expliquer le caractère très incomplet de la couverture santé, Arrow avance en premier lieu l'existence de phénomènes de risque moral. Contrairement à l'hypothèse faite dans le cadre « idéal » de l'information symétrique, les individus ont la capacité de modifier leur risque. Arrow note ainsi que dans le cadre de la santé, le coût des soins n'est pas seulement déterminée par la maladie (en principe hors de contrôle des individus) mais également par le choix du médecin et par la disposition des individus à recourir aux soins. Puisque l'assurance couvre les conséquences financières liées à une augmentation du risque santé, les individus sont désincités à limiter leur risque. A l'appui de cette hypothèse, Arrow souligne qu'on observe une hausse des consommations de soins suite au développement des couvertures santé. Les assureurs n'étant pas en mesure de distinguer entre risques évitables et inévitables, le rôle d'agent de contrôle est en principe délégué aux médecins qui garantissent la nécessité ou non du traitement administré au regard de l'état de santé du patient. Néanmoins, comme le note Arrow, ce contrôle est loin d'être parfait dans la mesure où le médecin peut avoir intérêt pour lui-même ou pour satisfaire son patient à prescrire plus de soins ou des soins plus chers. Par ailleurs, les individus peuvent être désincités à rechercher les producteurs de soins offrant les prix les plus avantageux. Arrow conclut ainsi :

⁹ Arrow note qu'à l'époque où il écrit cet article, entre un cinquième et un quart des dépenses de santé sont couvertes.

“It is impossible to draw up insurance policies which will sufficiently distinguish among risks, particularly since observation of the results will be incapable of distinguishing between avoidable and unavoidable risks, so that incentives to avoid losses are diluted.”

Le phénomène de risque moral n'est pas le seul facteur expliquant le caractère incomplet du marché de l'assurance santé. L'existence de chargements techniques, de frais de gestion et de commercialisation ainsi que le phénomène d'antisélection y contribuent également. Une préconisation importante d'Arrow est que l'Etat doit intervenir pour couvrir les risques qui ne sont pas assurés par le biais du marché. Ainsi, les frais de gestion et de commercialisation pourraient être limités en instaurant une couverture obligatoire : en couvrant un groupe important de personnes elle permettrait de réaliser d'importantes économies d'échelles et réduirait ainsi ces frais. En revanche, le risque moral constitue une limite potentielle, mais pas un obstacle majeur à l'intervention de l'Etat pour la couverture de nouvelles classes de risques.

“The welfare case for insurance policies of all sorts is overwhelming. It follows that the government should undertake insurance in those cases where this market, for whatever reason has failed to emerge. Nevertheless, there are a number of significant practical limitations on the use of insurance. It is important to understand them, though I do not believe that they alter the case for the creation of a much wider class of insurance policies than now exists”.

L'article d'Arrow présente plusieurs intérêts. Il identifie le risque moral comme un facteur expliquant l'incomplétude du marché de l'assurance santé. Il situe le risque moral comme un problème d'asymétrie d'information entre assureur et assuré. Enfin, il a le mérite de mettre en exergue le rôle de la relation médecin-patient dans ce phénomène. Mais il ne fait pas de réelle description des mécanismes économiques sous-jacents à ce phénomène, et ne donne que peu d'éléments sur ses conséquences pour l'assureur ou pour l'assuré. Il n'explique en particulier pas vraiment pourquoi ce phénomène amène les assureurs à offrir une couverture incomplète ou à ne pas offrir de couverture.

C'est à Pauly (1968) qu'on doit une première description des mécanismes du risque moral *ex-post*. A partir de cette description, Pauly montre que le risque moral génère une perte de bien-être collectif, perte qui explique la sous-couverture ou la non couverture de certains risques. Le raisonnement de Pauly s'appuie sur la théorie du consommateur. L'hypothèse clé est que la demande de soins possède une élasticité prix non-nulle : la quantité de soins consommée m^u est décroissante lorsque le prix unitaire q_m des soins augmente.

$$\frac{dm^u}{dq_m} < 0$$

Or, les remboursements effectués par l'assurance réduisent le coût des consommations médicales pour les patients assurés, incitant ceux-ci à accroître leurs consommations médicales. La valeur que l'individu accorde aux unités de soins supplémentaires consommées est supérieure à leur coût apparent mais inférieur à leur coût réel. Or, l'assureur constatant cette augmentation des consommations de soins doit accroître la prime du contrat en conséquence. L'individu paie donc la totalité du coût lié à ce supplément de soins : ce qu'il ne finance pas directement via l'augmentation du reste à charge (le coût direct, visible pour le patient), il le finance indirectement via l'augmentation de la prime. Ainsi, *in fine*, l'individu alloue à ses soins une dépense supérieure à la valeur qu'il leur attribue, générant une perte d'utilité.

Un exemple simple

Le modèle de Pauly explicite ce mécanisme en supposant que les individus peuvent subir trois niveaux d'état de santé possibles H (indexés, 1, 2, 3), avec pour chaque état de santé H : Les individus ont une parfaite santé (état 1 = pas de maladie) avec une probabilité $\pi_1 = 1/2$, une santé moyenne (état 2 = maladie peu grave) avec une probabilité $\pi_2 = 1/4$ et une mauvaise santé (état 3 = maladie grave) avec une probabilité $\pi_3 = 1/4$.

A chaque état de santé correspond une demande de soins propre, exprimée en fonction du taux de coassurance auquel font face les individus :

- Les individus en parfaite santé (pas de maladie) ont une demande de soins $m_1(q_m) = 0$ quelque soit p_m . Autrement dit, les individus ne consomment alors aucun soin, quelque soit le prix des soins.
- Les individus ont une santé moyenne (maladie peu grave) avec une probabilité $\pi_2 = 1/4$; leur demande de soins est donnée par l'équation $m_2(q_m) = 16 - 12 \cdot \frac{q_m}{p_m}$.
- Les individus sont en mauvaise santé (maladie grave) avec une probabilité $\pi_3 = 1/4$; leur demande de soins est donnée par l'équation $m_3(q_m) = 32 - 12 \cdot \frac{q_m}{p_m}$.

Les demandes associées aux états de santé 1, 2 et 3 sont représentées graphiquement respectivement par les courbes D'_1 (confondue avec l'axe des ordonnées), D'_2 et D'_3 . Soit p_m le coût unitaire total des soins (ou « coût de production »). Pauly considère l'effet passage d'une couverture nulle à une couverture complète :

- Lorsqu'ils ne sont pas couverts les individus font face au coût unitaire total des soins ($p_m = q_m$). Ceux qui sont en parfaite santé consomment $m_1^u(p_m) = 0$ UM (unités de soins

médicaux), ceux ayant une santé moyenne consomment $m_2^u(p_m) = 4$ UM et ceux qui sont en mauvaise santé consomment $m_3^u(p_m) = 20$ UM.

- Lorsque la couverture est complète, le prix apparent des soins est réduit à 0, incitant les malades (individus ayant une santé moyenne ou mauvaise) à augmenter leurs consommations de soins. Les individus qui sont en parfaite santé continuent à consommer $m_1(0) = m_1^l = 0$ UM (unités de soins médicaux), ceux ayant une santé moyenne consomment $m_2(0) = m_2^l = 16$ UM et ceux qui sont en mauvaise santé consomment $m_3(0) = m_3^l = 32$ UM.

Par ailleurs, pour être en mesure de servir les prestations promises, l'assureur doit tarifier le contrat sur la base des consommations des assurés. Ainsi, le prix actuariellement juste vaut $P_I = E(p_m \cdot m_s^l) = p_m \cdot E(m_s^l) = p_m \cdot \left(0 \times \frac{1}{2} + 16 \times \frac{1}{4} + 32 \times \frac{1}{4}\right) = 12 \cdot p_m$.

Or, un prix actuariellement juste pour les individus non couverts correspond au montant de la prime que les assurés paieraient si l'élasticité prix de la demande de soins était nulle, donc si le niveau de consommation ne variait pas avec le niveau de couverture : $P_I = E(p_m \cdot m_s^{NI}) = p_m \cdot \left(0 \times \frac{1}{2} + 4 \times \frac{1}{4} + 20 \times \frac{1}{4}\right) = 6 \cdot p_m$. Ainsi, le surcoût lié à l'accroissement des consommations médicales est répercuté indirectement sur les assurés au travers d'une augmentation de la prime ($= 12 \cdot p_m - 6 \cdot p_m = 6 \cdot p_m$). Le coût effectif du contrat excède donc très largement le niveau de la prime actuarielle pour les non-assurés. Ces derniers peuvent avoir intérêt à ne pas souscrire le contrat. D'autre part, chez les assurés, la valeur attribuée au surcroît de consommation de soins est plus faible que le coût total occasionné par le financement de ces soins. La perte de bien être résultant du risque moral résulte ainsi d'un problème d'inefficience : les individus allouent trop de ressources à leurs soins et pas assez aux consommations non-médicales.

1.5. Demande de soins en information asymétrique

1.5.1. L'offre des assureurs en information asymétrique

L'information privée dont les individus disposent sur leur état de santé leur offre donc la possibilité de comportements stratégiques de majoration du niveau de sévérité. Nous supposons désormais que l'assureur est dans l'impossibilité d'observer directement l'état s dans lequel se trouve l'assuré. Le seul signal qu'il reçoit sur cet état est la consommation de soins que l'individu effectue pour se soigner, ainsi que son prix. Ce signal ne fournit qu'une information imparfaite sur l'état de santé dans la mesure où une même quantité de soins peut être associée à différents états de santé. Cette imperfection a deux conséquences :

- L'assureur propose désormais des remboursements conditionnels à la consommation de soins effectuée, l'information la plus fine dont il dispose : pour une quantité de soins m , il rembourse la somme $A(m)$. Il doit donc attendre que l'assuré ait choisi son traitement médical (représenté ici par la quantité de soins m) avant de connaître la somme qu'il aura à rembourser. Dès lors, pour un état s donné, les niveaux de remboursements deviennent endogènes à la consommation de soins : plus celle-ci est importante, plus le remboursement sera élevé.
- L'assureur ne peut pas observer l'état s_0 dans lequel se trouve l'individu, il est donc incapable de déterminer sur quelle courbe de demande il se situe.

Nous supposons que $A(\cdot)$ est une fonction dérivable par morceaux. Soit P_A le niveau de prime du contrat. Le contrat est donc noté (A, P_A) . Dans la majeure partie de ce chapitre, nous allons supposer que $A(\cdot)$ est une fonction linéaire de la dépense. Cette hypothèse présente en effet l'avantage d'offrir un contexte simple qui contribue à clarifier la présentation des mécanismes de risque moral ; la méthode d'analyse est analogue à celle utilisée pour étudier les effets et le niveau optimal de la taxation linéaire d'un bien. Dans la partie 1.7.2, nous nous placerons dans le cadre de remboursements linéaires par morceaux.

1.5.2. La demande de soins *ex-post* la survenue du risque santé

Nous supposons donc qu'il existe une constante $a \in [0,1]$ telle que $A(m_s) = a \cdot p_m \cdot m_s$. Dans le contexte d'une offre de contrats offrant des remboursements proportionnels à la dépense, le coefficient multiplicatif a , appelé « taux de couverture », représente la quantité d'assurance achetée. *Ex-post* la réalisation du risque santé, les individus choisissent leurs consommations en résolvant le problème suivant :

Programme 1.5-i

$$\max_{(x_s, m_s) \in \mathbb{R}_+^2} U_s(x_s, m_s)$$

$$sc: x_s + p_m \cdot m_s = W^n + a \cdot p_m \cdot m_s - P_A$$

Nous introduisons la quantité $c = 1 - a$, dénommée « taux de coassurance » (ou « copayment » en anglais) de manière à faire apparaître, dans la contrainte de budget, le coût résiduel auquel fait face l'assuré. Nous posons $q_m = c \cdot p_m$ la valeur de ce prix résiduel.

Programme 1.5-ii

$$\max_{(x_s, m_s) \in \mathbb{R}_+^2} U_s(x_s, m_s)$$

$$sc: x_s + q_m \cdot m_s = W^n - P_A$$

La prime P_A dépend des consommations de soins des individus de sorte que nous pourrions imaginer que les choix des individus concernant x_s et m_s tiennent compte de l'influence de ces quantités sur P_A . Une hypothèse importante est que les individus ne tiennent pas compte de cette influence.

La fonction à maximiser est concave et la contrainte de budget linéaire, il existe donc une unique solution, notée $(x_s^\#, m_s^\#)$, au programme 1.5-i:

$$m_s^\#(q_m, W^n) = m_s^u(q_m, W^n - P_A)$$

$$x_s^\#(q_m, W^n) = x_s^u(q_m, W^n - P_A)$$

Nous définissons les fonctions de demande agrégées :

$$\bar{m}^\#(q_m, W^n) = \sum_{\tau=0}^S \pi_\tau \cdot m_\tau^\#(q_m, W^n) = \sum_{\tau=0}^S \pi_\tau \cdot m_\tau^u(q_m, W^n - P_A)$$

$$\bar{x}^\#(q_m, W^n) = \sum_{\tau=0}^S \pi_\tau \cdot x_\tau^\#(q_m, W^n) = \sum_{\tau=0}^S \pi_\tau \cdot x_\tau^u(q_m, W^n - P_A)$$

Le contrat étant tarifé à son coût actuariel, la prime P_A est égale à l'espérance des remboursements. Puisque l'individu consomme $m_s^\#(q_m, W^n) = m_s^u(q_m, W^n - P_A)$ unités de soins lorsqu'il est dans l'état s , l'assureur rembourse avec une probabilité π_s la quantité $(p_m - q_m) \cdot m_\tau^u(q_m, W^n - P_A)$. Ainsi, P_A vérifie l'équation :

Équation 1.5-i :

$$P_A = \sum_{\tau=0}^S \pi_\tau \cdot (p_m - q_m) \cdot m_\tau^u(q_m, W^n - P_A)$$

Au moyen de la consommation de soins agrégée \bar{m}^u , cette égalité se réécrit¹⁰ :

Équation 1.5-ii :

$$P_A = (p_m - q_m) \cdot \bar{m}^u(q_m, W^n - P_A)$$

¹⁰ Autrement dit, $P_A = (p_m - q_m) \cdot E_{\tau \in \mathcal{S}}(m_\tau^u(q_m, W^n - P_A))$.

Cette équation admet une unique solution¹¹, l'assureur peut donc trouver un niveau de prime tel que le contrat soit tarifé à son coût actuariel, c'est-à-dire tel que P_A couvre exactement l'espérance des remboursements. Concrètement, l'assureur peut déterminer P_A en graphant la courbe $K \rightarrow (p_m - q_m) \cdot \bar{m}^u(q_m, W^n - K)$ dans le quart de plan \mathbb{R}_+^2 et cherchant l'intersection de cette courbe avec la première bissectrice.

- Dans le cas d'une offre de couverture proportionnelle, $\bar{m}^u(q_m, W^n - P_A)$ représente le coût unitaire du contrat. A la différence de ce qui se passe en situation d'information symétrique, ce coût unitaire dépend de q_m , donc de la quantité d'assurance achetée.
- La recherche du juste niveau de prime P_A nécessite une connaissance précise de la fonction de demande agrégée \bar{m}^u puisque l'assureur doit savoir comment cette demande varie en fonction du niveau de revenu mais également du prix des soins.

1.5.3. Effet de l'assurance sur la quantité de biens et services médicaux et non médicaux consommés

L'effet du niveau de couverture sur le prix de la couverture P_A , la demande de soins $m_s^\#$ et de biens et services non médicaux $x_s^\#$ est déterminé au moyen du signe des dérivées de P_A , de $m_s^\#$ et $x_s^\#$ par rapport à q_m . Le calcul de ces dérivés conduit aux formules suivantes :

Équations 1.5-iii :

$$(a) \quad \frac{\partial P_A}{\partial q_m} = \frac{-\bar{m}^u(q_m, W^n - P_A) + (p_m - q_m) \frac{\partial \bar{m}^u}{\partial q_m} \Big|_{(q_m, W^n - P_A)}}{1 + (p_m - q_m) \cdot \frac{\partial \bar{m}^u}{\partial W} \Big|_{(q_m, W^n - P_A)}} < 0$$

$$(b) \quad \frac{\partial m_s^\#}{\partial q_m} \Big|_{(q_m, W^n)} \underset{< > 0?}{=} \frac{\partial m_s^u}{\partial q_m} \Big|_{(q_m, W^n - P_A)} - \frac{\partial P_A}{\partial q_m} \cdot \frac{\partial m_s^u}{\partial W} \Big|_{(q_m, W^n - P_A)} \underset{< 0}{\quad} \underset{> 0}{\quad}$$

$$(c) \quad \frac{\partial x_s^\#}{\partial q_m} \Big|_{(q_m, W^n)} \underset{< > 0?}{=} \frac{\partial x_s^u}{\partial q_m} \Big|_{(q_m, W^n - P_A)} - \frac{\partial P_A}{\partial q_m} \cdot \frac{\partial x_s^u}{\partial W} \Big|_{(q_m, W^n - P_A)} \underset{< > 0?}{\quad} \underset{> 0}{\quad}$$

¹¹ Posons $\varphi(K) = (p_m - q_m) \cdot \bar{m}^u(q_m, W^n - K)$, $K \in [0, W^n]$. Puisque $0 < \varphi(K) < p_m \cdot \bar{m}^u(q_m, W^n - K) < W^n$, $\varphi(K) \in [0, W^n]$ est une fonction du compact $[0, W^n]$ dans $[0, W^n]$. La fonction φ étant de plus continue, par le lemme de Brouwer, l'équation $\varphi(K) = K$ admet au moins une solution.

Par ailleurs, $K \rightarrow \varphi(K) - K$ est strictement décroissante φ , donc injective. Par conséquent la solution de $\varphi(K) = K$ est unique.

Concernant la prime P_A , plus le coût résiduel q_m est faible et plus le niveau de prime est élevé : une augmentation du niveau de couverture se traduit par un accroissement de la prime. Concernant les biens médicaux et non médicaux, l'effet de q_m peut se décomposer en deux parties :

- Un effet prix direct : une augmentation du niveau de couverture diminue le prix résiduel des soins et a donc un effet positif sur les consommations médicales et indéterminé sur les consommations non médicales.
- Un effet revenu indirect : une augmentation du niveau de couverture accroît comme on l'a vu la prime, donc réduit le revenu disponible pour les consommations médicales et non médicales. Elle a donc un effet négatif sur les consommations médicales et non médicales.

L'effet du niveau de couverture sur les consommations médicales et non médicales est *a priori* indéterminé. Néanmoins, nous montrons que lorsque la prime P_A est considérée comme très petite au regard du revenu W , un accroissement du niveau de couverture entraîne un accroissement des consommations de soins. Pour démontrer ce résultat, nous écrivons les équations 1.5-iii sous forme d'élasticités¹² :

Équation 1.5-iv :

(a)

$$\eta_{m_s^\#, q_m} = \eta_{m_s^u, q_m} - \frac{\theta_{P_A}}{1 - \theta_{P_A}} \cdot \eta_{P_A, q_m} \cdot \eta_{m_s^u, W}$$

(b)

$$\eta_{x_s^\#, q_m} = \eta_{x_s^u, q_m} - \frac{\theta_{P_A}}{1 - \theta_{P_A}} \cdot \eta_{P_A, q_m} \cdot \eta_{x_s^u, W}$$

Où :

¹² Nous pouvons réécrire l'équation 1.5-iii (b) sous la forme suivante :

$$\left. \frac{\partial m_s^\#}{\partial q_m} \right| = \frac{m_s^u(q_m, W^n - P_A)}{q_m} \cdot \eta_{m_s^u, q_m} - \frac{P_A}{q_m} \cdot \eta_{P_A, q_m} \cdot \frac{m_s^u(q_m, W^n - P_A)}{W - P_A} \cdot \eta_{m_s^u, W}$$

Sachant que $m_s^u(q_m, W^n - P_A) = m_s^\#$, nous obtenons :

$$\left. \frac{\partial m_s^\#}{\partial q_m} \right| = \frac{m_s^\#}{q_m} \cdot \left[\eta_{m_s^u, q_m} - \frac{\theta_{P_A}}{1 - \theta_{P_A}} \cdot \eta_{P_A, q_m} \cdot \eta_{m_s^u, W} \right]$$

Autrement dit :

$$\eta_{m_s^\#, q_m} = \eta_{m_s^u, q_m} - \frac{\theta_{P_A}}{1 - \theta_{P_A}} \cdot \eta_{P_A, q_m} \cdot \eta_{m_s^u, W}$$

Nous procédons de même pour calculer $\eta_{x_s^\#, q_m}$.

- $\eta_{m_s^{\#}, q_m}$ et $\eta_{q_m, x_s^{\#}}$ sont les élasticités prix de la demande de soins et de la demande de biens et services non médicaux ;
- $\eta_{m_s^u, q_m}$ et $\eta_{m_s^u, W}$ sont les élasticités prix et revenu de la demande de soins ;
- $\eta_{x_s^u, q_m}$ et $\eta_{x_s^u, W}$ sont les élasticités prix et revenu de la demande de biens et services non médicaux ;
- θ_{P_A} est le rapport entre le coût du contrat et le revenu d'un individu, autrement dit le poids de la prime dans le budget des individus ;

Or, la formule suivante lie l'élasticité de la prime au coût résiduel q_m avec les élasticités prix et revenus de la demande agrégée m^u de soins¹³ :

Équation 1.5-v :

$$\eta_{P_A, q_m} = \frac{1 - \theta_{P_A}}{1 - c} \cdot \frac{(1 - c) \cdot \eta_{\bar{m}^u, q_m} - c}{1 - \theta_{P_A} + \theta_{P_A} \cdot \eta_{\bar{m}^u, W}}$$

Par conséquent :

Équation 1.5-vi :

(a)

$$\eta_{m_s^{\#}, q_m} = \eta_{m_s^u, q_m} - \frac{\theta_{P_A}}{1 - \theta_{P_A}} \cdot \eta_{P_A, q_m} \cdot \eta_{m_s^u, W} = \eta_{m_s^u, q_m} + \eta_{m_s^u, W} \cdot \theta_M \cdot \frac{c - (1 - c) \cdot \eta_{\bar{m}^u, q_m}}{1 - \theta_{P_A} + \theta_{P_A} \cdot \eta_{\bar{m}^u, W}}$$

(b)

$$\eta_{x_s^{\#}, q_m} = \eta_{x_s^u, q_m} - \frac{\theta_{P_A}}{1 - \theta_{P_A}} \cdot \eta_{P_A, q_m} \cdot \eta_{x_s^u, W} = \eta_{x_s^u, q_m} + \eta_{x_s^u, W} \cdot \theta_M \cdot \frac{c - (1 - c) \cdot \eta_{\bar{m}^u, q_m}}{1 - \theta_{P_A} + \theta_{P_A} \cdot \eta_{\bar{m}^u, W}}$$

¹³ Nous avons les égalités suivantes :

$$\begin{aligned} (p_m - q_m) \cdot \left. \frac{\partial \bar{m}^u}{\partial q_m} \right|_{(q_m, W - P_A)} &= \frac{(p_m - q_m) \cdot \bar{m}^u(q_m, W^n - P_A)}{q_m} \cdot \eta_{\bar{m}^u, q_m} = \frac{P_A}{q_m} \cdot \eta_{q_m, \bar{m}^u} \\ \bar{m}^u(q_m, W^n - P_A) &= \frac{P_A}{(p_m - q_m)} = \frac{P_A}{q_m \cdot \left(\frac{p_m}{q_m} - 1\right)} = \frac{c}{1 - c} \cdot \frac{P_A}{q_m} \\ (p_m - q_m) \cdot \left. \frac{\partial \bar{m}^u}{\partial W} \right|_{(q_m, W^n - P_A)} &= \frac{P_A}{W^n - P_A} \cdot \eta_{\bar{m}^u, W} = \frac{\theta_{P_A}}{1 - \theta_{P_A}} \cdot \eta_{\bar{m}^u, q_m} \end{aligned}$$

En appliquant ces égalités dans l'équation 1.5-iii (b), nous obtenons la formule suivante pour l'élasticité de la prime au prix des soins :

$$\eta_{P_A, q_m} = \frac{q_m}{P_A} \cdot \frac{\partial P_A}{\partial q_m} = \frac{-\frac{c}{1 - c} + \eta_{\bar{m}^u, q_m}}{1 + \frac{\theta_{P_A}}{1 - \theta_{P_A}} \cdot \eta_{\bar{m}^u, W}}$$

Où $c = q_m/p_m$ est le taux de coassurance du contrat et $\theta_M = p_m \cdot \bar{m}^u/W^n$ le montant moyen des dépenses de santé ramené au budget des individus.

D'après l'identité de Slutsky, les élasticité-prix $\eta_{m_s^u, q_m}$ de la demande de soins et $\eta_{x_s^u, q_m}$ de la demande de biens et services non-médicaux, sans effet de la prime, se décompose de la manière suivante¹⁴ :

Équation 1.5-vii :

(a)

$$\eta_{m_s^u, q_m} = \eta_{\psi_s^u, q_m} - \frac{c \cdot \theta_{M_s}}{1 - \theta_{P_A}} \cdot \eta_{m_s^u, W}$$

(b)

$$\eta_{x_s^u, q_m} = \eta_{\chi_s^u, q_m} - \frac{c \cdot \theta_{M_s}}{1 - \theta_{P_A}} \cdot \eta_{x_s^u, W}$$

Où $\eta_{\psi_s^u, q_m}$ et $\eta_{\chi_s^u, q_m}$ sont les élasticité-prix de la demande hicksienne en biens et services médicaux et non-médicaux. De plus, $\theta_{P_A} = \theta_M \cdot (1 - c)$. Ainsi les élasticité-prix totales $\eta_{q_m, m_s^\#}$ de la demande de soins et $\eta_{q_m, x_s^\#}$ de la demande de biens et services non-médicaux peuvent se réécrire :

Équation 1.5-viii :

(a)

$$\eta_{m_s^\#, q_m} = \eta_{\psi_s^u, q_m} + \eta_{W, m_s^u} \left[\theta_M \cdot \frac{c - (1 - c) \cdot \eta_{\bar{m}^u, q_m}}{1 - \theta_M \cdot (1 - c) \cdot (1 - \theta_M \cdot \eta_{\bar{m}^u, W})} - \frac{c \cdot \theta_{M_s}}{1 - \theta_M \cdot (1 - c)} \right]$$

¹⁴ Soit $u_{q,s} = V_s(q_m, W^n - P_A)$ l'utilité que retirent les individus dans l'état s , couverts par le contrat laissant un coût résiduel q_m et de prime P_A . L'identité de Slutsky s'écrit :

$$\left. \frac{\partial \psi_s^u}{\partial q_m} \right|_{(q_m, u_{q,s})} = \left. \frac{d m_s^u}{d q_m} \right|_{(q_m, W - P_A)} - \left. \frac{d m_s^u}{d W} \right|_{(q_m, W - P_A)} \cdot m_s^u(q_m, W^n - P_A)$$

C'est-à-dire :

$$\left. \frac{\partial \psi_s^u}{\partial q_m} \right|_{(q_m, u_{q,s})} = \left. \frac{d m_s^u}{d q_m} \right|_{(q_m, W - P_A)} - \frac{q_m \cdot m_s^u(q_m, W^n - P_A)}{W^n - P_A} \cdot \eta_{m_s^u, W}$$

C'est-à-dire :

$$\eta_{\psi_s^u, q_m} = \eta_{m_s^u, q_m} - \frac{c \cdot p_m \cdot m_s^u(q_m, W^n - P_A)}{W^n \cdot \left(1 - \frac{P_A}{W^n}\right)} \cdot \eta_{m_s^u, W}$$

Or, nous avons $c \cdot p_m \cdot m_s^u(q_m, W^n - P_A) = c \cdot M_s$ et $M_s = p_m \cdot m_s^u(q_m, W^n - P_A)$ et $W^n \cdot \theta_{M_s} = M_s$. Ainsi :

$$\eta_{\psi_s^u, q_m} = \eta_{m_s^u, q_m} - \frac{c \cdot \theta_{M_s}}{1 - \theta_{P_A}} \cdot \eta_{m_s^u, W}$$

(b)

$$\eta_{x_s^\#, q_m} = \eta_{x_s^u, q_m} + \eta_{W, x_s^u} \left[\theta_M \cdot \frac{c - (1-c) \cdot \eta_{\bar{m}^u, q_m}}{1 - \theta_M \cdot (1-c) \cdot (1 - \eta_{\bar{m}^u, W})} - \frac{c \cdot \theta_{M_s}}{1 - \theta_M \cdot (1-c)} \right]$$

L'effet de l'assurance sur les consommations médicales et non-médicales est donc décomposable en :

- **Un effet substitution** qui vaut $\eta_{\psi_s^u, q_m} < 0$ pour les consommations médicales et $\eta_{x_s^u, q_m} > 0$ pour les consommations non-médicales : la baisse du coût résiduel des soins q_m occasionné par une hausse du niveau de couverture entraîne une réallocation des ressources des biens et services non médicaux vers les biens et services médicaux.
- **Un effet revenu** qui vaut :
 - Pour les consommations médicales :

$$\eta_{m_s^u, W} \left[\theta_M \cdot \frac{c - (1-c) \cdot \eta_{\bar{m}^u, q_m}}{1 - \theta_M \cdot (1-c) \cdot (1 - \eta_{\bar{m}^u, W})} - \frac{c \cdot \theta_{M_s}}{1 - \theta_M \cdot (1-c)} \right]$$

- Pour les consommations non-médicales :

$$\eta_{x_s^u, W} \left[\theta_M \cdot \frac{c - (1-c) \cdot \eta_{\bar{m}^u, q_m}}{1 - \theta_M \cdot (1-c) \cdot (1 - \eta_{\bar{m}^u, W})} - \frac{c \cdot \theta_{M_s}}{1 - \theta_M \cdot (1-c)} \right]$$

Effet du niveau d'assurance lorsque la part des dépenses de santé est négligeable

Lorsque les dépenses de santé sont faibles au regard du budget des individus, les élasticités $\eta_{m_s^\#, q_m}$ et $\eta_{x_s^\#, q_m}$ sont à peu près égales aux élasticités $\eta_{\psi_s^u, q_m}$ et $\eta_{x_s^u, q_m}$ de la demande hickienne. Nous sommes donc en mesure d'énoncer la proposition suivante :

Proposition 1.5-1 : Lorsque le niveau de couverture augmente :

- La prime P_A du contrat d'assurance s'accroît ;
- La consommation de biens et services médicaux $m_s^\#$ s'accroît pour tout état s , pourvu que la part des dépenses de santé dans le budget des individus soit faible.
- La consommation de biens et services non-médicaux $x_s^\#$ décroît pour état s , pourvu que la part des dépenses de santé dans le budget des individus soit faible.

Effet du niveau d'assurance lorsque la part des dépenses de santé est élevée

L'effet revenu tient compte du fait qu'une augmentation du niveau de couverture accroît la prime, entraînant une baisse des ressources disponibles pour acheter des soins, mais aussi du fait que la baisse du coût résiduel des soins accroît le pouvoir d'achat des individus, leur permettant d'acheter plus de soins et plus de biens et services non médicaux. Plus précisément, la hausse du pouvoir d'achat induite par la baisse du coût résiduel des biens et services médicaux au moment de l'achat des soins a pour contrepartie une baisse du pouvoir d'achat induite par l'augmentation de la prime. Lequel de ces deux effets domine ? Pour répondre à cette question, considérons l'inégalité suivante :

$$\frac{c \cdot \theta_{M_s}}{1 - \theta_M \cdot (1 - c)} - \theta_M \cdot \frac{c - (1 - c) \cdot \eta_{\bar{m}^u, q_m}}{1 - \theta_M \cdot (1 - c) \cdot (1 - \eta_{\bar{m}^u, W})} > \frac{c \cdot (\theta_{M_s} - \theta_M) - \theta_M \cdot (1 - c) \cdot \eta_{\bar{m}^u, q_m}}{1 - \theta_M \cdot (1 - c)}$$

Donc :

$$c > \frac{\theta_M \cdot \eta_{\bar{m}^u, q_m}}{\theta_{M_s} - \theta_M \cdot (1 - \eta_{\bar{m}^u, q_m})} \Rightarrow \frac{c \cdot \theta_{M_s}}{1 - \theta_M \cdot (1 - c)} - \theta_M \cdot \frac{c - (1 - c) \cdot \eta_{\bar{m}^u, q_m}}{1 - \theta_M \cdot (1 - c) \cdot (1 - \eta_{\bar{m}^u, W})} > 0$$

Ainsi lorsque θ_M , la part de la dépense de santé moyenne dans le revenu des ménages, est faible devant θ_{M_s} , la part de la dépense de santé des individus dans l'état s , l'élasticité revenu est positive. Dans ce cas, l'effet substitution et l'effet revenu contribuent tous deux à accroître les consommations médicales et non-médicales. Reprenant les équations 1.5-viii (a) et (b), nous observons que lorsque $\eta_{m_s^u, W}$, θ_{M_s} et c sont élevés, proches de 1, l'effet revenu l'emporte sur l'effet substitution η_{ψ, m_s^u} . A l'inverse, quand les quantités précédentes sont faibles, l'effet revenu devient négligeable et l'effet substitution domine. Ainsi, un impact positif de la couverture sur les consommations de soins peut traduire des effets différents, notamment selon le niveau de coassurance, l'état de santé des individus (qui joue sur la valeur de θ_{M_s}) et leur niveau de revenu (qui joue sur les valeurs de θ_{M_s}). Comprendre quel effet domine dans ces différents contextes représente un enjeu important en termes d'efficacité et d'équité. Cette problématique est à l'origine d'un débat entre Nyman pour qui l'effet revenu est élevé et surpasse l'effet substitution dans un certain nombre de cas de figure et les économistes de la RAND pour qui cet effet est plus anecdotique. Nous reviendrons sur ce débat et ses enjeux dans la partie 1.9.

1.5.4. Travaux empiriques sur l'effet de l'assurance sur les consommations médicales

Au regard du poids croissant des dépenses de santé dans le budget des nations, le risque moral, supposé expliquer cette hausse représente un enjeu important en termes de politique publique ; il donc fait l'objet de nombreux travaux. Ces travaux visaient tout d'abord à vérifier l'existence du risque moral *ex-post* en testant l'effet de l'assurance sur les consommations de soins. Certains

travaux ont également cherché mesurer l'ampleur de la perte monétaire induite par ce phénomène et / ou à estimer le taux de couverture optimal au regard de l'intensité de ce phénomène. Une partie importante d'entre eux a été réalisée aux Etats-Unis. Cutler et Zeckhauser (2000) en font une revue de littérature assez riche.

Les premières analyses empiriques

Selon cette revue de littérature, les premières mesures du risque moral sont réalisées dès le milieu des années 60, dans la foulée de l'article d'Arrow (1963) et avant même celui de Pauly (1968)¹⁵ :

Concernant les soins ambulatoires :

- Scitovsky et Snyder (1972) s'appuient sur une expérience naturelle résultant de l'introduction en avril 1967, sur un contrat de l'Université de Stanford, d'un taux de coassurance de 25% pour les soins de médecins et d'auxiliaire, les analyses médicales (biologie, radiologie...) et les actes de chirurgie fournis par la « Palo Alto Medical Clinic ». Ce contrat, le « Group Health Plan » (GHP) offrait auparavant une couverture complète pour l'ensemble de ces soins. Les auteurs étudient l'évolution entre 1966 et 1968 des consommations des soins offerts par la « Palo Alto Medical Clinic », suite à cette variation du niveau de couverture. L'analyse est effectuée sur les individus couverts par la GHP en 1966 et en 1968. Ce cylindrage permet de travailler sur les mêmes individus en 1966 et 1968, donc de travailler à structure sociodémographique identique, hormis celle d'âge qui évolue du fait du vieillissement de l'échantillon. L'échantillon comporte 2567 individus dont 859 sont souscripteurs du contrat. Les auteurs comparent le nombre de visites de médecins d'une part et d'actes d'analyse médicales et d'auxiliaire d'autre part ainsi que le coût moyen de ces soins par individu en calculant l'évolution de ces quantités brutes puis à âge comparable. Le nombre de visites de médecin décroît de 24,1% sans ajustement par l'âge et de 24,8% avec ajustement soit dans les deux cas une élasticité prix d'environ 0,14. Le nombre d'actes d'auxiliaire et d'analyses médicales décroît de 11,5% sans ajustement et de 19,2% avec ajustement, soient des élasticités-prix égales respectivement à -0,06 et -0,11.
- Phelps et Newhouse (1972) prolongent l'analyse de Scitovsky et Snyder (1972), en menant une étude économétrique du nombre de consultations chez les spécialistes, la dépense en soins de spécialistes, le nombre de recours aux analyses médicales et aux soins d'auxiliaire et

¹⁵ Le premier article cité dans la revue de littérature de Cutler et Zeckhauser (2000) est celui de Feldstein et Severson, publié en 1964 et qui analyse l'effet de différents facteurs (dont le coût à la charge des patients) sur la demande de visites de médecins. L'analyse décrite dans cet article s'appuie sur des données de l'enquête du National opinion research center (NORC) de Chicago. Les estimations sont menées au moyen des moindres carrés ordinaire et conduisent à une élasticité prix de -0,19 (Cutler et Zeckhauser, 2000).

la dépense pour ces soins. Les données de 1966 et de 1968 sont groupées de sorte que les auteurs travaillent sur une base de 5134 unités (des individus – années). La variable explicative d'intérêt est une dichotomique 1_{1968} qui indique que l'année d'observation est 1968. Les coefficients associés à cette variable sont interprétés comme les effets propres de l'introduction du niveau de coassurance sur les différents indicateurs de consommations de soins régressés. Cet effet est contrôlé de plusieurs variables : l'âge, le genre, la relation de l'individu avec le souscripteur du contrat (souscripteur homme / femme, ayant-droit homme / femme / enfant), la distance entre le domicile et la « Palo Alto Medical Clinic », le type d'emploi occupé à l'Université de Stanford et la taille de la famille. Les résultats font apparaître un effet important de la dichotomique 1_{1968} sur le nombre de consultations et la dépense de médecins et un effet moindre pour les analyses médicales et les soins d'auxiliaire. Cette différence d'effet s'explique par le fait que la décision de consommer des analyses médicales et des soins d'auxiliaire n'incombe pas totalement au patient mais dépend en partie du médecin, alors que les consultations chez le médecin sont initiées par les patients. Traduit en termes d'élasticité-prix, les résultats trouvés sont proches de ceux de Scitovsky et Snyder (1972). Afin de tester l'uniformité de l'effet de l'augmentation du niveau de coassurance sur les différentes sous population, Phelps et Newhouse régressent la différence entre le nombre de recours en 1967 et celui en 1968, et la différence de dépenses entre ces deux années. La constante est alors interprétée comme l'effet « global » de l'introduction du taux de coassurance et les coefficients associés à chaque variable comme des différences d'effets du niveau de coassurance par rapport à l'effet global. Aucune variable ne ressort significativement hormis le fait d'être ayant-droit femme. De plus, le test de Fisher tend à rejeter l'hypothèse d'une significativité global des variables : l'effet du niveau de la variation du taux de coassurance tend donc à être uniforme sur l'ensemble des catégories de population. Un apport essentiel de ce travail concerne la manière dont l'effet de l'assurance santé est affecté par les coûts non assurables (en particulier les coûts d'opportunité liés au temps passé à se soigner) : Phelps et Newhouse pointent le fait que les ayants-droit femmes ont, toutes choses égales par ailleurs, un niveau de consommation plus élevé que les autres catégories d'assurés. De plus, cette catégorie est la seule à être affectée différemment par la baisse du taux de coassurance : une baisse significativement plus forte que sur le reste de la population est observée, l'élasticité de la demande est donc plus importante. Présument qu'un grand nombre d'ayants-droit femmes ne travaillent pas, ils y voient un indice d'une plus forte élasticité-prix liée à des coûts d'opportunité moindres.

- Dans le cadre d'un modèle d'investissement dans le capital santé des enfants de 5 ans et moins, Colle et Grossman (1978) analysent entre autre l'effet du niveau d'assurance santé sur les consommations de soins pédiatriques à partir de l'échantillon du NORC¹⁶, une enquête représentative de la population américaine. Ces consommations sont appréhendées par la probabilité sur un an la probabilité d'avoir un examen à titre préventif, la probabilité de recourir à des soins de médecin, le nombre de consultations en cas de recours et la « qualité » des médecins consultés. Cet indicateur de qualité est construit en rapportant le coût des médecins consultés à un indicateur construit à partir des honoraires moyens des médecins selon leur type (généraliste, pédiatre, spécialiste) et leur niveau de certification. La probabilité de recours est étudiée au moyen d'un modèle logit, le nombre de consultations et le niveau de qualité des consultations par des régressions par les moindres carrés ordinaires. Les variables explicatives introduites sont le fait de bénéficier d'une couverture ou de l'aide sociale (pour la probabilité de recourir), la dépense à la charge du patient (pour le nombre de visites et la qualité des soins), le niveau d'éducation de la mère, les connaissances des parents concernant la santé, leur goût pour la santé, des variables indiquant les coûts non monétaires liés aux recours aux soins (temps d'attente avant un rendez-vous, temps de trajet jusqu'au médecin, mode de déplacement, temps d'attente chez le médecin) ainsi qu'un grand nombre de variables socio-économiques et d'état de santé. Le traitement du biais engendré par la sélection adverse n'est pas clairement explicité, mais il se fait implicitement par l'introduction des variables d'état de santé et de goût pour la santé. Les régressions sont menées sur 839 enfants, moitié moins pour celles portant sur la qualité des soins. Les auteurs trouvent une élasticité prix de 0,11 pour le nombre de visites et de 0,04 pour la qualité du médecin.

Concernant les soins hospitaliers :

- Feldstein (1971) estime l'effet de de l'assurance sur les admissions et les durées de séjour dans les hôpitaux à but non lucratif. Son modèle théorique s'appuie sur un arbitrage des managers d'hôpitaux entre le nombre de jours d'hospitalisation et la qualité des soins administrés aux patients. Les managers ont une fonction d'utilité qui dépend de ces deux arguments. Une diminution du volume de journées d'hospitalisation disponibles (ce qui réduit l'utilité) augmente le coût de ces journées ce qui permet d'améliorer la qualité des soins administrés en augmentant le personnel par patient et le plateau technique (ce qui augmente l'utilité). Feldstein suppose de plus qu'il existe un prix d'équilibre équilibrant offre

¹⁶ Le NORC est le National Opinion Research Center.

et demande de journées d'hospitalisation et que le prix effectif des soins suit un processus dynamique : le prix effectif à la date t diminue si le prix effectif à la date $t-1$ excède le prix d'équilibre. Il en finit que la consommation de soins hospitaliers au temps t dépend de la consommation au temps $t-1$: les normes au sein de la communauté et le comportement des médecins empêchent un ajustement immédiat des consommations aux nouvelles caractéristiques, en particulier aux nouveaux prix. L'auteur s'appuie donc sur un modèle économétrique dynamique. Il utilise des données agrégées par état (47 états sont pris en compte) sur la période allant de 1958 à 1967. Les variables expliquées sont le logarithme du nombre d'admissions et de la durée moyenne d'hospitalisation. La principale variable explicative d'intérêt est le prix des soins auquel le patient fait face, c'est-à-dire net des remboursements de l'assurance (lorsque l'individu en possède une). Deux spécifications sont testées : une avec le prix des soins nets des remboursements et une avec le prix nominal des soins plus une variable d'assurance¹⁷. Les autres variables explicatives introduites sont le logarithme du revenu, le temps, le logarithme de la densité de la population, le fait d'être couvert par Medicaid, le logarithme de la densité de médecin, le logarithme de la proportion de généralistes parmi les médecins et la variable dépendante retardée. L'estimation est menée selon la méthode des variables instrumentales¹⁸. Feldstein obtient une élasticité-prix égale à -0,67 pour l'ensemble des journées d'hospitalisations

- Davis et Russel (1972) analysent l'effet de l'assurance sur la substitution entre soins de ville et soins hospitaliers. Les analyses sont menées à partir de données agrégées au niveau de 48 états américains. Les variables expliquées sont le logarithme du nombre de visites chez le médecin et le logarithme du nombre d'admission à l'hôpital par habitant. Seuls les hôpitaux à but non lucratif sont considérés. Le prix des soins de médecins et des soins hospitaliers à la charge des patients est un facteur explicatif fondamental de ces variables. Ces prix agissent en terme absolu et en terme relatif : plus le prix des séances médecin (respectivement des hospitalisations) est faible et plus le patient va consommer des soins de médecins (respectivement des soins hospitaliers). D'autre part, plus le prix des séances de médecins est élevé relativement à celui des hospitalisations et plus l'individu va être enclin à substituer des soins de médecin par des soins hospitaliers. Ces deux effets sont pris en compte en introduisant le prix total de ces deux catégories de soins ainsi que la proportion de personnes qui sont couvertes par le programme d'assurance « Blue Cross » qui, par rapport aux autres

¹⁷ La variable d'assurance est construite au moyen de la proportion de personnes couvertes par état et du niveau de coassurance moyen sur les Etats-Unis.

¹⁸ Les variables instrumentales introduites sont les variables exogènes introduites dans le modèle et des variables qui auraient été introduites si une analyse complète du système de santé avait été effectuée.

couvertures, couvre mieux les soins ambulatoires. Les autres variables explicatives introduites sont le revenu, le taux d'occupation des lits d'hôpitaux, la densité de médecins libéraux, la proportion de spécialistes parmi les médecins libéraux, la proportion de la population âgée de moins de 65 ans et le taux d'immigration dans l'état. Les régressions sont menées par les moindres carrés ordinaires. Les résultats font apparaître une forte élasticité-prix de la demande de soins de médecin (comprise entre -0,85 et -1,46) et une élasticité-prix de la demande de soins hospitaliers.

- Rosett et Huang (1973) mesurent l'effet de l'assurance et du revenu sur les dépenses de soins totales (ambulatoires et hospitalières) en utilisant l'enquête Survey of Consumer Expenditures du Bureau of Labor Statistics. Cette base, représentative de la population américaine, leur permet de disposer pour chaque individu enquêté, des dépenses de santé à sa charge ainsi que du niveau de prime de son contrat, s'il est assuré. Les auteurs reconstruisent à partir de ces deux données la dépense totale et le niveau de couverture (sous la forme d'un niveau de coassurance de chaque assuré) pour 11 000 individus environ. Ils estiment ensuite une fonction de demande de soins en régressant au moyen d'un Tobit généralisé, la dépense totale sur le niveau de couverture, sur le revenu, sur le carré de ces variables, sur ces variables croisées entre elles et sur des variables de contrôle sociodémographiques et socio-économiques. Aucune correction de la sélection adverse n'est spécifiée¹⁹. Les auteurs obtiennent des élasticités-prix de la dépense de santé comprises entre -0,35 pour un niveau de coassurance de -0,2 à -1,5 pour un niveau de coassurance de 0,3.

L'expérience de la RAND

Newhouse (1978) a mené une analyse critique des méthodologies des différents travaux menés dans les années 1960 et 1970 sur le risque moral. Il insiste tout d'abord sur la disparité des mesures d'élasticité-prix fournies par ces premiers travaux est très importante puisque les mesures les plus élevées (de l'ordre de -2,1) sont plus de dix fois supérieures aux mesures les plus faibles (de l'ordre de -0,1). Les principales raisons pouvant expliquer ces différences sont :

- L'utilisation fréquente de données agrégées qui entraîne deux phénomènes :
 - Une surestimation de l'ampleur des élasticités-prix liée à une corrélation entre la variable d'assurance et le terme d'erreur.

¹⁹ Les auteurs contrôlent l'état de santé uniquement au travers d'une variable de recours au dentiste qui indiquant la tolérance à la douleur de l'individu.

- Une Instabilité des estimations lorsque les effets des prix sont très variables d'un individu à l'autre.
- L'omission dans pratiquement tous les travaux des variables de prix des soins qui peuvent se substituer aux biens et services médicaux étudiés (absence d'effet prix croisé).

L'expérience de la Rand a été menée pour pallier à ces problèmes ainsi que celui lié à l'endogénéité. Elle a fourni une des sources de données les plus reconnues pour l'analyse des effets de l'assurance sur les consommations de soins : les données qu'elle a fournies ont servi de base à de nombreux travaux sur le risque moral. 2000 familles appartenant à six sites géographiques américains²⁰ ont été échantillonnées et se sont vues attribuer aléatoirement un contrat de couverture santé parmi quatorze formules de couverture possibles. Ces quatorze formules se différencient par leur niveau de coassurance et le niveau de dépense au-delà duquel les individus sont totalement pris en charge :

- un contrat offrait une couverture complète.
- Trois catégories de contrat offraient une couverture partielle, uniforme sur l'ensemble des soins, 25%, 50%, ou 95% de la dépense et une catégorie de contrat offre un taux de coassurance de 25% pour tous les soins ambulatoires, excepté les soins dentaires et les soins psychiatriques de ville pour lesquels le taux de coassurance était de 50%. Chacune de ces formules offrant une couverture partielle comprenait un *stop-loss*, c'est-à-dire un reste à charge annuel maximum au-delà duquel tout dollars supplémentaire de dépense est intégralement pris en charge ; les différents niveaux de *stop-loss* possibles étaient calculés sous la forme d'un pourcentage du revenu du ménage - 5%, 10%, ou 15% du revenu - avec une limite supérieure de 1000\$.
- Une formule particulière, dite « à franchise individuelle », appliquait un taux de coassurance nul sur les soins hospitaliers et un taux de coassurance de 95% sur les soins ambulatoires, tant que le reste à charge individuel n'atteignait pas 150€ ou que le reste à charge par ménage n'atteignait pas 450€ ; au-delà de cette limite, la couverture était complète.

Sont exclus les individus qui, au début de l'expérience, sont âgés de 62 ans et plus, ceux dont le revenu annuel est supérieur à 25 000 €, ceux éligibles à Medicare au titre d'une invalidité, les personnes en institution pour une durée indéterminée les militaires et leurs ayants-droit, les vétérans ayant des invalidités liées à leur service. Par ailleurs, certains individus ont refusé de

²⁰ Il s'agissait des sites de Dayton dans l'Ohio ; Seattle dans l'Etat de Washington ; Fitchburg dans le Massachusetts ; le Comté de Franklin dans le Massachusetts ; Charleston en Caroline du Sud ; et le Comté de Georgetown en Caroline du Sud. Ces sites, dont certains sont ruraux et d'autres urbains, ont été choisis de manière à tenir compte de la diversité de l'offre de soins sur le territoire américain.

participer à l'expérience, d'autre l'ont quitté avant sa fin. Pour limiter le refus, les individus se sont vu proposer une compensation, versée à l'issue de l'expérience. Le taux de refus s'échelonne de 6% pour le contrat de couverture complète à 23% pour les contrats de niveau de coassurance de 95%. Les seules différences significatives entre ceux qui ont accepté de participer et ceux qui ont refusé sont un niveau de revenu et d'éducation plus bas chez ceux qui ont refusé, caractéristiques qui sont recueillies par l'expérience de la RAND et donc contrôlées dans les travaux économétriques issus de cette expérience. Finalement, l'échantillon comporte environ 6000 individus.

Cette expérience a été menée sur une période de trois ans pour 70% des individus de l'échantillon et de cinq ans pour les 30% restant. Durant cette période, l'exhaustivité des consommations de soins a été recueillie pour chaque individu. La nature de chaque consommation, le niveau de dépense qu'elle occasionne ainsi que le diagnostic de la pathologie associée sont collectés. Des données socio-économiques et d'état de santé sont également recueillies : l'âge, le genre, l'appartenance ethnique, le revenu et la taille du ménage, la zone de résidence, la perception générale de l'état de santé²¹, la présence de limitations dans les activités quotidiennes résultant de problèmes de santé, le nombre de maladies chroniques et l'état de santé mental²².

Une première analyse a été menée sur les dépenses annuelles (Manning et al, 1987). Cette analyse porte sur des individus/années ce qui correspond à environ 20 000 unités. Pour les analyses descriptives et économétriques, les familles de contrats ont été regroupées par taux de coassurance : l'analyse économétrique de l'effet du niveau d'assurance sur les consommations de soins s'appuie sur un modèle à quatre équations non corrélées :

- La première équation, estimée au moyen d'un modèle probit, décrit la probabilité de recourir au moins une fois aux soins plutôt que de ne pas y recourir. L'estimation de cette probabilité vaut $\Phi(X_{it} \cdot \hat{\beta}_1)$ pour chaque individu i et chaque année t , où Φ est la fonction de répartition de la loi normale et X_{it} les variables explicatives constituées du niveau de couverture et de l'ensemble des variables de contrôle.
- La deuxième équation estime, au moyen d'un modèle probit, décrit la probabilité d'avoir au moins une consommation hospitalière plutôt que de n'avoir que des consommations ambulatoires, sachant que l'individu a recouru au moins une fois aux soins. La probabilité estimée vaut $\Phi(X_{it} \cdot \hat{\beta}_2)$ pour chaque individu et chaque année t .

²¹ La perception générale de l'état de santé est un score de 0 à 100 basé sur 22 questions pour les individus âgés de 14 ans et plus et 7 questions pour les individus âgés de moins de 14 ans.

²² L'état de santé mental est un score de 0 à 100 basé sur un questionnaire de 38 questions pour les individus âgés de 14 ans et plus et 12 questions pour les individus âgés de moins de 14 ans.

- La troisième équation, estimée par les moindres carrés ordinaires, décrit le logarithme de la dépense médicale chez les individus qui n'ont reçu que des soins ambulatoires. La dépense estimée vaut $\hat{\varphi}_3 \cdot \exp(X_{it} \cdot \hat{\beta}_3)$ pour chaque individu i à l'année t , où $\hat{\varphi}_3$ correspond à l'estimation de l'espérance de l'exponentiel du résidu par le *smearing estimate*²³.
- La quatrième équation, estimée par les moindres carrés ordinaires, décrit le logarithme de la dépense médicale chez les individus qui ont reçu au moins un soin hospitalier. La dépense estimée vaut $\hat{\varphi}_4 \cdot \exp(X_{it} \cdot \hat{\beta}_4)$ pour chaque individu i et chaque année t , où $\hat{\varphi}_4$ correspond à l'estimation de l'espérance de l'exponentiel du résidu par le *smearing estimate*.

Pour chacune des estimations, un certain nombre de caractéristiques non observées (aversion pour le risque, préférence pour la santé, paramètres d'état de santé inobservés....) peuvent entraîner pour un même individu une corrélation des résidus due à la répétition de données dans le temps. Ces corrélations peuvent biaiser les matrices variance des coefficients estimés, donc les tests sur la significativité des coefficients. Les auteurs corrigent ce biais en utilisant la matrice de White.

La dépense totale estimée est calculée de la manière suivante :

$$\frac{1}{N} \sum_{i,t} \Phi(X_{it} \cdot \hat{\beta}_1) \cdot \left\{ \Phi(X_{it} \cdot \hat{\beta}_2) \cdot \hat{\varphi}_4 \cdot \exp(X_{it} \cdot \hat{\beta}_4) + (1 - \Phi(X_{it} \cdot \hat{\beta}_2)) \cdot \hat{\varphi}_3 \cdot \exp(X_{it} \cdot \hat{\beta}_3) \right\}$$

Où N est le nombre d'années / individus.

Les taux annuels de recours estimés au moyen du probit de la première étape et les niveaux de dépense totale estimés par la formule ci-dessus sont donnés dans le tableau ci-dessous.

Tableau 1.5-i : probabilités de recours et dépenses totales annuelles simulées au moyen des données de la RAND

Formule de couverture	Probabilité estimée de recourir aux soins	Dépense par individu
Franchise individuelle	73%	817\$
Coassurance de 95%	68%	700\$
Coassurance de 50%	74%	764\$
Coassurance de 25%	79%	826\$
0% (couverture complète)	87%	1019\$

Note : les dépenses sont exprimées en dollars 1991

Source : Newhouse, 1993

²³ Le *smearing estimate* est calculé en faisant la moyenne empirique de l'exponentielle des résidus :

$$\hat{\varphi} = \frac{1}{N} \sum_{i,t} \exp(\hat{\varepsilon}_{it})$$

En parallèle, Keeler *et al* (1988) étudient l'effet du taux de coassurance sur des épisodes de soins. Un épisode de soins est défini comme un ensemble de soins reliés à un même diagnostique. Deux grandes catégories de soins sont définies : les soins hospitaliers, les soins dentaires et les soins ambulatoires non-dentaires. Les soins ambulatoires non-dentaires sont divisées en trois sous-catégories : soins aigües, soins chroniques (ou de routine), soins de confort²⁴. Comme unité d'analyse, les épisodes de soins sont considérés comme étant plus pertinents que les consommations annuelles, en particulier parce qu'ils sont associés à une maladie et donc sont naturellement liés à la décision de consommer des biens et services médicaux. Les auteurs décomposent la dépense annuelle de la manière suivante :

$$\text{dép. annuelle} = \text{nb épisodes} \times \text{prix par épisode}$$

L'analyse est donc menée en deux étapes :

- Tout d'abord, pour chaque catégorie d'épisode, une estimation du nombre d'épisodes en fonction du niveau de coassurance et des variables de contrôle au moyen d'un modèle négatif binomial afin de tenir compte de la forte présence de zéros et de valeurs extrêmes.
- Dans un deuxième temps, une régression du logarithme de la dépense par épisode en fonction du niveau de coassurance et des variables de contrôle au moyen d'une estimation par les moindres carrés ordinaires.

L'analyse montre que le niveau de coassurance a une forte influence sur le nombre d'épisodes et une faible influence sur la dépense par épisode. Le revenu a un effet significatif sur le nombre d'épisodes de soins chroniques, de soins de confort et de soins dentaires.

A partir de l'analyse des dépenses annuelles menée par Manning *et al* (1987) et par Keeler *et al* (1988) des élasticités-prix de la demande de soins ont été estimées. Ces élasticités sont calculées sous la forme d'élasticités d'arc²⁵ et sont présentées dans le rapport de Manning *et al*. La présence de *stop-loss* sur les dépenses à la charge de l'individu pour chaque grande catégorie de formule pose problème pour le calcul d'effets prix purs. En effet, le niveau de coassurance devient égal à zéro lorsque les individus dépassent cette limite. Pour les individus dépassant cette limite, l'effet de l'assurance est donc un mixe du taux de coassurance en-dessous du stop-loss et du taux de

²⁴ Les soins de confort correspondent aux traitements administrés lorsque le patient n'est pas considéré comme malade et qui peuvent être reportés dans le temps sans perte substantielle de bien-être.

²⁵ Les élasticités d'arc sont calculées en faisant le rapport :

$$\frac{d(c_2) - d(c_1)}{c_2 - c_1} \times \frac{(c_2 + c_1)/2}{(d(c_2) + d(c_1))/2}$$

Où c_1 et c_2 sont deux niveaux de coassurance et $d(c_1)$ et $d(c_2)$ les niveaux de dépense correspondant à ces niveaux de coassurance.

coassurance au-dessus du *stop-loss*. Pour contrer cette difficulté, Keeler *et al* considèrent uniquement les individus qui ont des dépenses à leur charge largement en-deçà du *stop-loss*. Les résultats sont présentés dans le tableau 1.5-ii ci-dessous.

Tableau 1.5-ii : élasticités-prix calculées à partir du modèle sur les épisodes de soins

Variation de niveau de coassurance	Soins hospitaliers	Soins ambulatoires	Ensemble des soins
0% - 25%	-0,17	-0,17	-0,17
25% - 95%	-0,14	-0,31	-0,22

Source : Manning *et al*, 2008

Pour les soins ambulatoires, les élasticités diffèrent assez sensiblement suivant la variation de coassurance considérée. Néanmoins, concernant la demande totale de soins, les élasticités sont assez proches, de l'ordre de -0,2.

Manning *et al* quant à eux calculent les taux de coassurance moyen par grande catégorie de formule et estiment les élasticités à partir de ces taux de coassurance. Les élasticités sur la demande totale sont assez proches, de l'ordre de -0,12.

Tableau 1.5-iii : élasticités-prix calculées à partir du modèle sur les dépenses annuelles

Variation nominale du niveau de coassurance	Variation réelle du niveau de coassurance	Soins ambulatoires	Ensemble des soins
0% - 25%	0% - 16%	-0,13	-0,10
25% - 95%	16% - 31%	-0,21	-0,14

Source : Manning *et al*, 2008

1.6. Inefficiency de l'équilibre en information asymétrique

1.6.1. Mise en évidence de l'inefficience

Lorsque l'information assureur-assurée est asymétrique, les quantités de biens et services médicaux et non-médicaux consommées aboutissent à un équilibre qui n'est pas Pareto-optimal. En effet, si l'information assureur-assurés était symétrique, l'utilité d'au moins une catégorie d'individu pourrait être augmentée sans diminuer celle des autres, avec autant de ressources. Cette inefficiency est mise en évidence ici en s'inspirant du raisonnement de Nyman (1999a), adapté au cas où l'individu peut être dans plusieurs états s . Supposons donc que l'assureur soit en mesure d'observer l'état s dans lequel se trouve l'individu et qu'au lieu du contrat (A, P_A) contingent à la quantité de soins

consommée, il lui offre le contrat (I, P_I) contingent à l'état de santé, et offrant le profil de remboursements $(I_s)_{s \in \mathcal{S}}$, avec :

$$I_s = a \cdot p_m \cdot m_s^\# = (p_m - q_m) \cdot m_s^\#$$

Autrement dit, le contrat (I, P_I) rembourse exactement la même somme que le contrat (A, P_A) lorsque l'individu est dans l'état s , mais cette somme est versée *ex-ante* la consommation de soins et peut être allouée aux consommations médicales ou non médicales. Le contrat (I, P_I) étant tarifé au coût actuariel, nous avons :

Équation 1.6-i :

$$P_I = \sum_{\tau=0}^S \pi_\tau \cdot I_\tau = \sum_{\tau=0}^S \pi_\tau \cdot (p_m - q_m) \cdot m_s^\# = P_A$$

Du point de vue des ressources consommées, les deux contrats sont équivalents. Par contre, le programme décrivant les choix de consommation des individus diffère dans la mesure où ceux-ci reçoivent le remboursement I_s sous forme de revenu, *ex-ante* la consommation des soins. Ils peuvent donc allouer cette somme à leurs consommations médicales, qu'ils paient au coût p_m , mais aussi à leurs consommations non-médicales. Pour chaque état $s \in \mathcal{S}$, le choix de consommation de l'individu dans l'état s est décrit par le programme s'écrit :

Programme 1.6-i

$$\max_{(x_s, m_s) \in \mathbb{R}_+^2} U_s(x_s, m_s)$$

$$sc: x_s + p_m \cdot m_s = W - P_A + I_s = W^n - P_A + (p_m - q_m) \cdot m_s^\#$$

Le graphique 1.6-i ci-dessous permet de montrer que les choix issus du programme 1.6-i (remboursements contingents à l'état de santé) aboutissent à un plus haut niveau d'utilité que ceux issus du programme 1.5-ii (remboursements contingents aux niveaux de consommations). Les droites \mathcal{D}_{p_m} et \mathcal{D}_{q_m} , de pentes p_m et q_m , représentent respectivement les contraintes de budgets des individus sans assurance et des individus assurés. Les courbes iso-utilité des individus sont tangentes à \mathcal{D}_{p_m} en $(E_s)_{s \in \mathcal{S}}$, et à \mathcal{D}_{q_m} en $(F_s)_{s \in \mathcal{S}}$. Le point $H = (\bar{x}^u, \bar{m}^u)$ correspond aux coordonnées moyennes des individus lorsqu'ils ne sont pas assurés. Les coordonnées du point $K = (\bar{x}^\#, \bar{m}^\#)$ correspondent aux consommations moyennes des individus assurés²⁶. Les coordonnées $(\bar{x}^\#, \bar{m}^\#)$ vérifient :

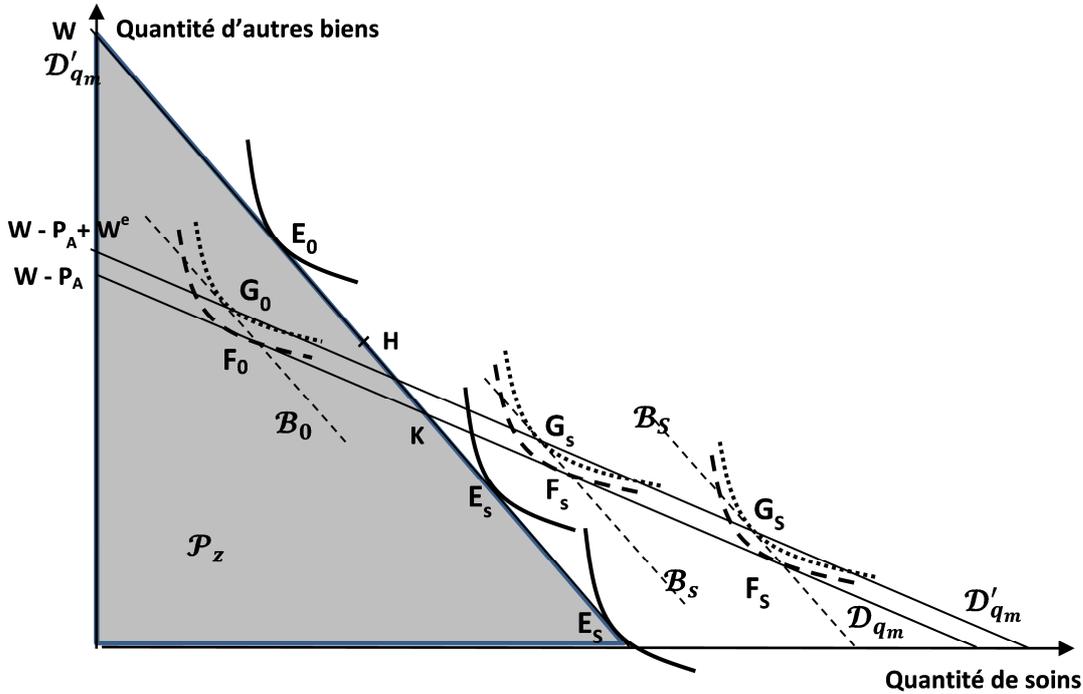
²⁶ Nous pouvons noter que $\bar{x}^\# < \bar{x}^u$ et $\bar{m}^\# < \bar{m}^u$: les consommations non-médicales des non-assurés sont en moyenne plus élevées que celles des assurés et, à l'inverse, les consommations médicales des assurés non assurés sont plus faibles que celles des assurés. De même, nous notons que l'utilité des individus en bonne santé non-assurés est plus élevée que celle des individus en bonne santé assurés. Par contre, l'utilité des individus en mauvaise santé non-assurés est plus faible que celle des individus en mauvaise santé assurés.

$$\bar{x}^\# + q_m \cdot \bar{m}^\# = W^n - P_A$$

$$\bar{x}^\# + p_m \cdot \bar{m}^\# = \bar{x}^\# + q_m \cdot \bar{m}^\# + (p_m - q_m) \cdot \bar{m}^\# = W - P_A + P_A = W^n$$

Le point K est donc l'unique lieu de rencontre de \mathcal{D}_{p_m} et \mathcal{D}_{q_m} .

Graphique 1.6-i : représentation de l'inefficacité liée au risque moral



Soit la droite \mathcal{B}_s correspondant au budget :

$$x_s + p_m \cdot m_s = W^n - P_A + I_s = W^n - P_A + (p_m - q_m) \cdot m_s^\#$$

Nous constatons que le point F_s de coordonnées $(x_s^\#, m_s^\#)$ appartient à cette droite de budget.²⁷ Dans le graphique 1.6-i, cette droite est celle en tirets, parallèle à \mathcal{D}_{p_m} (elles ont même pente) et qui passent par F_s . La solution du programme 1.6-i correspond au point de tangence entre les courbes d'utilité et \mathcal{B}_s ; il s'agit donc sur le graphique du point G_s . Nous remarquons que pour tout $s \in \mathcal{S}$, la

²⁷ Nous avons :

$$x_s^\# + q_m \cdot m_s^\# = W^n - P_A$$

Donc :

$$x_s^\# - (p_m - q_m) \cdot m_s^\# + p_m \cdot m_s^\# = W^n - P_A$$

Et finalement :

$$x_s^\# + p_m \cdot m_s^\# = W^n - P_A + (p_m - q_m) \cdot m_s^\#$$

Donc $(x_s^\#, m_s^\#)$ appartient à la droite de budget :

$$x_s + p_m \cdot m_s = W^n - P_A + I_s = W^n - P_A + (p_m - q_m) \cdot m_s^\#$$

courbe d'utilité passant par G_s est au-dessus de celle passant par F_s . L'utilité retirée des choix issus du programme 1.6-i apportent donc plus d'utilité pour tous les individus que les choix issus du programme 1.5-ii. Si l'information était symétrique, il serait donc possible d'augmenter l'utilité de tous les individus avec le même niveau de ressources qu'en information asymétrique. Cela montre que l'équilibre n'est pas Pareto-optimal.

Il est important de noter qu'en présence du contrat contingent à l'état de santé, les consommations médicales sont moins importantes et les consommations non-médicales plus importantes qu'en présence du contrat contingent aux consommations de soins. Ce résultat permet d'interpréter l'inefficience produite par le second type de contrat : les individus consomment trop de biens et services médicaux et pas assez de biens et services non-médicaux. Le contrat contingent à l'état de santé permet donc de restituer l'optimalité au sens de Pareto en permettant de substituer des consommations médicales par des consommations non-médicales.

Soit W_s^e la valeur monétaire du supplément d'utilité apporté par le contrat dont les remboursements sont contingents à l'état de santé, par rapport au contrat dont les remboursements sont contingents aux niveaux de consommations. Calculer la valeur de W_s^e représente un enjeu important dans la perspective d'une estimation de l'ampleur du phénomène de risque moral. C'est à ce calcul qu'est en partie consacrée la partie 1.6.2 qui suit.

1.6.2. Mesurer l'effet du niveau de couverture sur le bien-être des individus

1.6.2.1. La perte de surplus du consommateur liée au risque moral

Les travaux portant sur la perte de bien-être liée au risque moral s'appuient sur des hypothèses différentes concernant la fonction d'utilité des individus. Convertir les variations d'utilité en équivalents monétaires rend possible la comparaison des résultats de ces travaux sur une même base. De plus, cette valeur monétaire peut être comparée au coût total de la santé, ce qui permet de mesurer l'ampleur du phénomène de risque moral. La conversion en équivalent monétaire de la variation d'utilité générée par la baisse du coût des soins et le paiement de la prime est donnée par la variation équivalente du revenu ΔW_s^e (Laffont, 1985), c'est-à-dire le niveau de revenu qu'il faudrait donner, *ex-post* la réalisation du risque, aux individus non-assurés et dans l'état s pour qu'ils soient indifférents entre d'une part bénéficier du revenu $W - P_A$ et payer les soins au prix $q_m < p_m$ et d'autre part bénéficier du revenu $W^n + \Delta W_s^e$ et payer les soins au prix p_m . La variation équivalente du revenu ΔW_s^e est calculée en s'appuyant sur l'utilité indirecte des individus :

$$V_s(q_m, W_{q_m, s}^e) = V_s(p_m, W^n + \Delta W_s^e) = V_s(p_m, W_{p_m, s}^e)$$

V_s étant l'utilité indirecte des individus dans l'état s , p_m le prix auquel font face les individus non assurés, $W_{p_m,s}^e = W + \Delta W_s^e$, $q_m < p_m$, le prix des soins auquel font face les assurés et $W_{q_m,s}^e = W - P_A$. Soit $u_{q,s}$ le niveau d'utilité des assurés (faisant face à un coût résiduel q) et dans l'état s , *ex-post* la réalisation du risque :

$$V_s(q_m, W_{q_m,s}^e) = V_s(p_m, W_{p_m,s}^e) = u_{q_m,s}$$

W_{s,q_m}^e et W_{s,p_m}^e peuvent donc être exprimés en fonction de $u_{q_m,s}$ et, en fonction de q_m et p_m : $W_{q_m,s}^e = W_s^e(q_m, u_{q_m,s})$ et $W_{p_m,s}^e = W_s^e(p_m, u_{q_m,s})$. En notant que $W_s^e(q, u_{q_m,s})$ est le plus faible niveau de revenu tel que l'utilité $u_{q_m,s}$ soit atteinte au prix q , nous avons :

Équation 1.6-ii :

$$-\Delta W_s^e - P_A = W_{q_m,s}^e - W_{p_m,s}^e = \int_{p_m}^{q_m} \frac{\partial W_s^e}{\partial q} \Big|_{(q, u_{q_m,s})} dq = \int_{p_m}^{q_m} \psi_s(q, u_{q_m,s}) dq$$

ψ_s ayant été précédemment défini comme la demande hicksienne de l'individu dans l'état s . La prime P_A s'écrit :

Équation 1.6-iii :

$$P_A = (p_m - q_m) \cdot \sum_{s \in \mathcal{S}} \pi_s \cdot m_s^u(q_m, W^n - P_A)$$

Or, pour tout $s \in \mathcal{S}$, les formules issues de l'approche duale du consommateur permettent de lier la demande marshalienne m_s et la demande hicksienne ψ_s :

$$m_s^u(q_m, W^n - P_A) = \psi_s(q_m, V_s(q_m, W^n - P_A)) = \psi_s(q_m, u_{q_m,s})$$

Nous avons donc :

$$P_A = (p_m - q_m) \cdot \sum_{s \in \mathcal{S}} \pi_s \cdot \psi_s(q_m, u_{q_m,s})$$

Et ainsi :

$$\begin{aligned} \Delta W_s^e(p_m, q_m, u_{q_m,s}) &= \int_{q_m}^{p_m} \psi_s(q, u_{q_m,s}) dq - P_A \\ &= \int_{q_m}^{p_m} \psi_s(q, u_{q_m,s}) dq - (p_m - q_m) \cdot \sum_{s \in \mathcal{S}} \pi_s \cdot \psi_s(q_m, u_{q_m,s}) \end{aligned}$$

Pris au niveau collectif, la variation de revenu compensé est égal à :

$$\sum_{s \in \mathcal{S}} N_{C_s} \cdot \Delta W_s^e(p_m, q_m, u_{q_m, s}) = \int_{q_m}^{p_m} \sum_{s \in \mathcal{S}} N_{C_s} \cdot \psi_s(q, u_{q_m, s}) dq - N_c \cdot (p_m - q_m) \cdot \sum_{s \in \mathcal{S}} \pi_s \cdot \psi_s(q_m, u_{q_m, s})$$

N_{C_s} étant le nombre d'individus dans l'état s et N_c le nombre total d'individus. A partir de cette formule, nous pouvons nous ramener à un gain moyen par individu, en divisant les égalités de gauche et de droite par N_c :

$$\Delta \bar{W}^e(p_m, q_m, u_{q_m}) = \int_{q_m}^{p_m} \sum_{s \in \mathcal{S}} \pi_s \cdot \psi_s(q, u_{q_m, s}) dq - (p_m - q_m) \cdot \sum_{s \in \mathcal{S}} \pi_s \cdot \psi_s(q_m, u_{q_m, s})$$

Autrement dit :

Équation 1.6-iv :

$$\Delta \bar{W}^e(p_m, q_m, u_{q_m}) = \int_{q_m}^{p_m} \bar{\psi}(q, u_{q_m}) dq - (p_m - q_m) \cdot \bar{\psi}(q_m, u_{q_m})$$

Où u_{q_m} , $\Delta \bar{W}^e$ et $\bar{\psi}$ sont définis de la manière suivante :

- $u_{q_m} = (u_{q_m, s})_{s \in \mathcal{S}}$.
- $\Delta \bar{W}^e(p, q)$ est la variation équivalente agrégée (sous forme de moyenne) du revenu :

$$\Delta \bar{W}^e(p, q, u) = \sum_{s \in \mathcal{S}} \pi_s \cdot \Delta W_s^e(p, q, u)$$

- $\bar{\psi}(q, u)$ est la demande hicksienne agrégée (sous forme de moyenne) des individus :

$$\bar{\psi}(q, u) = \sum_{s \in \mathcal{S}} \pi_s \cdot \psi_s(q, u)$$

Faisant l'hypothèse que pour tout s , $\Delta W_s^e(p_m, q_m, u)$ est indépendant de u ²⁸, nous notons désormais $\Delta W_s^e(p_m, q_m)$ cette quantité et nous avons :

Équation 1.6-v :

$$\Delta \bar{W}^e(p_m, q_m) = \Delta \bar{W}^e(p_m, q_m, u_{p_m}) = \int_{q_m}^{p_m} \bar{\psi}(q, u_{p_m}) dq - (p_m - q_m) \cdot \bar{\psi}(q_m, u_{p_m})$$

²⁸ Cette propriété est équivalente à :

$$\frac{\partial^2 \psi}{\partial q \partial u} = 0$$

La fonction $q \rightarrow \bar{\psi}(q, u_{p_m})$ est strictement décroissante comme moyenne pondérée de fonctions strictement décroissantes selon q . Ainsi, $\bar{\psi}(q, u_{p_m}) \leq \bar{\psi}(p_m, u_{p_m})$ pour tout $q \in [q_m, p_m]$, avec inégalité stricte lorsque $q \in]q_m, p_m[$ et nous avons :

$$\int_{q_m}^{p_m} \bar{\psi}^u(q, u_{p_m}) dq < \int_{q_m}^{p_m} \bar{\psi}^u(q_m, u_{p_m}) dq = (p_m - q_m) \cdot \bar{\psi}(q_m, u_{p_m})$$

La variation $\Delta \bar{W}^e(p_m, q_m)$ est donc négative : **il y a ainsi perte de surplus**. Nous noterons $MHZ(q_m) = -\Delta \bar{W}^e(p_m, q_m) > 0$ la valeur de cette perte :

Équation 1.6-vi :

$$MHZ(q_m) = (p_m - q_m) \cdot \bar{\psi}(q_m, u_{p_m}) - \int_{q_m}^{p_m} \bar{\psi}(q, u_{p_m}) dq$$

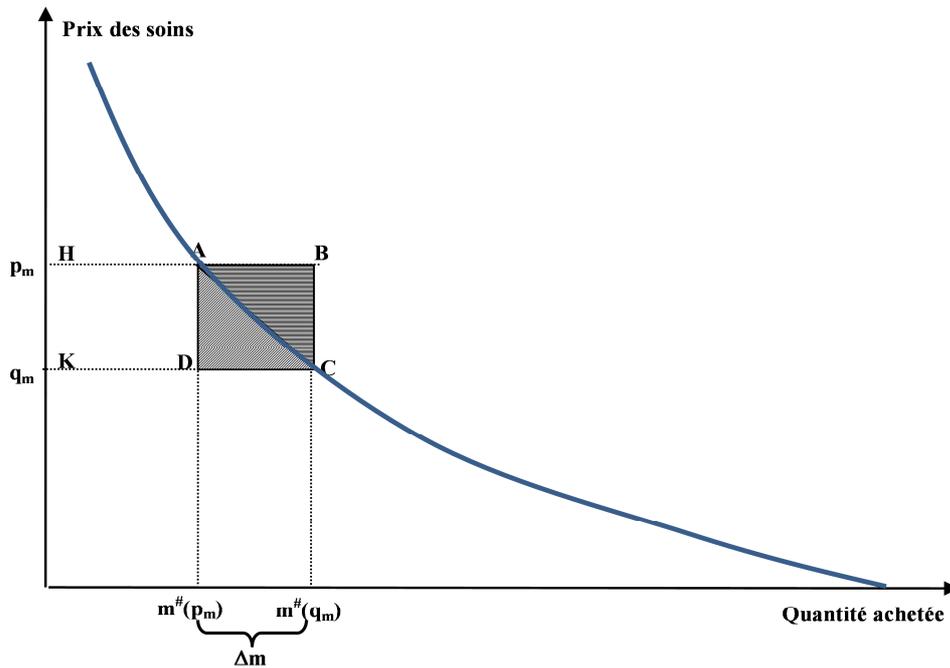
Interprétation économique et graphique de MHZ

L'équation 1.6-vi donne lieu à la fois une interprétation économique et une interprétation graphique :

- le terme intégral de cette équation quantifie la variation du surplus du consommateur lorsque le coût résiduel des soins passe de p_m à q_m euros ; il s'agit donc du gain en bien-être (net du reste à charge) lié à cette baisse de coût. Ce surplus est représenté dans le graphique 1.6-ii ci-dessous par la surface HACK.
- Le terme $(p_m - q_m) \cdot \bar{\psi}(q_m, u_{p_m})$ quantifie l'augmentation de prime liée au surplus Δm de consommations de soins, donc le surcoût (net du reste à charge) lié à Δm ; ce surcoût est représenté dans le graphique 1.6-ii par la surface HACK.

L'équation 1.6-vi traduit donc le fait que la valeur (nette du reste à charge) que les individus attribuent au surplus de consommation Δm est inférieure à son coût réel (net du reste à charge), générant la perte de surplus $MHZ(q_m)$. Cette perte est représentée dans le graphique 1.6-ii par le « triangle d'Harberger » ABC.

Graphique 1.6-ii : représentation de la perte de surplus collectif liée au risque moral



En dérivant ΔMHZ par rapport à q_m , nous obtenons :

Équation 1.6-vii :

$$\frac{dMHZ}{dq_m} = (p_m - q_m) \cdot \frac{d\bar{\psi}}{dq_m}(q_m, u_{p_m}) < 0$$

Ainsi, plus le coût à la charge des patients est faible (donc plus le niveau de couverture est élevé) et plus la perte de surplus est élevée. L'interprétation marginaliste de ce résultat est que chaque unité consommée au-delà de $\bar{\psi}^u(p_m) = \bar{m}^u(p_m)$ est valorisée par l'individu en deçà de son coût effectif, autrement dit, de son coût de production p_m ; ces unités contribuent donc à accroître la perte de surplus : plus le niveau de couverture est élevé, plus q_m est faible, plus l'individu consomme des unités de soins en plus de $\bar{m}^u(p_m)$ et plus la perte de surplus s'accroît.

L'approximation par la demande marshalienne

Dans l'article fondateur écrit en 1968 par Pauly sur la perte de bien-être liée au risque moral en santé et dans nombreuses autres analyses consacrées à l'étude de ce phénomène, le triangle d'Harberger est calculé à partir de la demande marshalienne. *Stricto-sensu*, l'utilisation de la demande marshalienne pour calculer la variation équivalente du revenu n'est valide que si on suppose l'utilité des individus quasi-linéaire (Laffont, 1985), à savoir :

$$U_s(x_s, m_s) = x_s + u_m(s, m_s)$$

Or, une utilité quasi-linaire implique que l'élasticité-revenu de la demande de soins est nulle. Ainsi, en travaillant sur la demande marshalienne, les chercheurs adoptent l'hypothèse selon laquelle l'élasticité-revenu de la demande de soins est négligeable. Dans ce cas :

Équation 1.6-viii :

$$MHZ(q_m) = (p_m - q_m) \cdot \bar{m}^u(q_m) - \int_{q_m}^{p_m} \bar{m}^u(q) dq$$

Pour analyser la conséquence du phénomène de risque moral sur la demande d'assurance santé, nous allons nous affranchir de cette approximation ; nous allons nous appuyer sur l'équation 1.6-vi utilisant la demande Hicksienne.

1.6.2.2. Le gain en bien-être lié à la réduction du risque financier

Le gain en bien-être associé à la réduction du risque financier que procure un service d'assurance est mesuré par la prime de risque, c'est-à-dire le revenu qu'il faudrait retirer à un individu bénéficiant de l'assurance pour qu'il soit indifférent avec le fait de ne pas en bénéficier. Dans le contexte de l'assurance santé, le risque financier est associé à la consommation de frais de soins de santé. La prime de risque W^r est donnée par la formule suivante :

Équation 1.6-ix :

$$\sum_{s \in \mathcal{S}} \pi_s \cdot V_s(q_m, W^n - P_A - W^r) = \sum_{s \in \mathcal{S}} \pi_s \cdot V_s(p_m, W^n)$$

En dérivant l'équation 1.6-ix par rapport à q_m , nous obtenons :

Équation 1.6-x :

$$\sum_{s \in \mathcal{S}} \pi_s \cdot \left[\frac{\partial V_s}{\partial q} \Big|_{(q_m, W - P_A - W^r)} - \left(\frac{dP_A}{dq_m} + \frac{dW^r}{dq_m} \right) \frac{\partial V_s}{\partial W} \Big|_{(q_m, W - P_A - W^r)} \right] = 0$$

Or, d'après l'identité de Roy, nous avons pour tout s :

$$\frac{\partial V_s}{\partial q} \Big|_{(q_m, W - P_A - W^r)} = -m_s^\#(q_m, W^n - W^r) \cdot \frac{\partial V_s}{\partial W} \Big|_{(q_m, W - P_A - W^r)}$$

De plus, la dérivée de la prime par rapport à q_m est donnée par :

$$\frac{dP_A}{dq_m} = -\bar{m}^\#(q_m, W^n - W^r)^{29}$$

Ces dernières égalités nous permettent d'écrire :

Équation 1.6-xi :

$$\sum_{s \in \mathcal{S}} \pi_s \cdot \left[m_s^\#(q_m, W^n - W^r) - \bar{m}^\#(q_m, W^n - W^r) + \frac{dW^r}{dq_m} \right] \cdot \frac{\partial V_s}{\partial W} \Big|_{(q_m, W - P_A - W^r)} = 0$$

Notons $v_{x,s}^\#(q_m, W)$ l'utilité marginale du revenu dans l'état s :

$$v_{x,s}^\# = \frac{\partial V_s}{\partial W} \Big|_{(q_m, W - P_A)} = \frac{\partial U_s}{\partial x} \Big|_{(x_s^\#, m_s^\#)} = u'_x(x_s^\#)$$

Au moyen des $v_{x,s}^\#(q_m, W^n - W^r)$, l'équation 1.6-xi peut être réécrite sous la forme suivante :

Équation 1.6-xii :

$$\sum_{s \in \mathcal{S}} \pi_s \cdot \left[m_s^\#(q_m, W^n - W^r) - \bar{m}^\#(q_m, W^n - W^r) + \frac{dW^r}{dq_m} \right] v_{x,s}^\#(q_m, W^n - W^r) = 0$$

C'est-à-dire :

Équation 1.6-xiii :

$$\frac{dW^r}{dq_m} = - \frac{1}{\bar{v}_x^\#(q_m, W^n - W^r)} \cdot cov \left(\tilde{v}_x^\#(q_m, W^n - W^r), \tilde{m}^\#(q_m, W^n - W^r) \right)$$

Où :

- $\tilde{v}_x^\#$ est la variable aléatoire dont les réalisations sont $(v_{x,s}^\#)_{s \in \mathcal{S}}$;
- $\bar{v}_x^\#$ est l'espérance de $\tilde{v}_x^\#$, donc la moyenne des $(v_{x,s}^\#)_{s \in \mathcal{S}}$;
- $\tilde{m}^\#$ est la variable aléatoire dont les réalisations sont $(m_s^\#)_{s \in \mathcal{S}}$.

Les suites $(v_{x,s}^\#)_{s \in \mathcal{S}}$ et $(m_s^\#)_{s \in \mathcal{S}}$ sont toutes deux croissantes avec s ³⁰. Ainsi, $\tilde{v}_x^\#$ et $\tilde{m}^\#$ covarient dans le même sens, de sorte que :

$$cov \left(\tilde{v}_x^\#(q_m, W^n - W^r), \tilde{m}^\#(q_m, W^n - W^r) \right) > 0$$

²⁹ Nous faisons implicitement l'hypothèse que les individus n'anticipent pas l'effet de la variation du niveau de couverture sur leurs consommations médicales de sorte qu'ils considèrent le coût unitaire $\bar{m}^\#(q_m, W^n - W^r)$ comme constant.

³⁰ Lorsque s augmente (donc lorsque la sévérité de l'état pathologique s'accroît), la consommation de soins $m_s^\#$ augmente. Par ailleurs, les consommations non médicales $x_s^\# = W^n - q_m \cdot m_s^\# - P_A$ diminuent ; par conséquent $\tilde{m}^\#$ et $\tilde{x}^\#$ sont négativement corrélées.

Par conséquent, la prime de risque W^r augmente lorsque le prix résiduel des soins q_m diminue, donc lorsque le niveau de couverture augmente : **plus le niveau de couverture est élevé et plus l'individu retire de l'utilité liée à la réduction du risque financier.**

Enfin, en supposant W^r négligeable devant le revenu total, W^n , nous obtenons pour W^r la formule simplifiée suivante :

Équation 1.6-xiv :

$$\frac{dW^r}{dq_m} = -\frac{1}{\bar{v}_x^\#} \cdot \text{cov}(\tilde{v}_x^\#, \tilde{m}^\#) = -\frac{1}{\bar{v}_x^\#} \cdot \text{cov}(\tilde{v}_x^\#, \tilde{\psi})$$

Enfin, une dernière approximation permet de retrouver la formule de la prime de risque habituellement utilisée. Cette approximation est basée sur l'hypothèse selon laquelle les dépenses de santé $\tilde{m}^\#$, la prime P_A et la prime de risque W^r sont petites devant le revenu W^n ; nous pouvons effectuer alors le développement limité de $\tilde{v}_x^\#$ suivant :

$$\tilde{v}_x^\# = u'_x(\tilde{x}^\#) = u'_x(W^n - q_m \cdot \tilde{m}^\# - P_A) \cong u'_x(W^n) - (q_m \cdot \tilde{m}^\# + P_A) \cdot u''_x(W^n)$$

De plus, nous pouvons alors écrire :

$$\tilde{m}^\#(q_m, W^n) = \tilde{m}^u(q_m, W^n - P_A) \cong \tilde{m}^u(q_m, W^n)$$

Nous avons donc :

$$\tilde{v}_x^\# \cong u'_x(W^n) - (q_m \cdot \tilde{m}^u + P_A) \cdot u''_x(W)$$

Donc :

$$\bar{v}_x^\# \cong \sum_{s \in \mathcal{S}} \pi_s \cdot [u'_x(W^n) - (q_m \cdot \tilde{m}^u + P_A) u''_x(W^n)] \cong \sum_{s \in \mathcal{S}} \pi_s \cdot [u'_x(W^n)] = u'_x(W^n)$$

$$\text{cov}(\tilde{v}_x^\#, \tilde{m}^\#) \cong \text{cov}(u'_x(W^n) - u''_x(W^n) \cdot (q_m \cdot \tilde{m}^u + P_A), \tilde{m}^u) = -u''_x(W^n) \cdot q_m \cdot \text{var}(\tilde{m}^u)$$

Ainsi :

$$\frac{dW^r}{dq_m}(q_m) \cong -R^A \cdot q_m \cdot \text{var}(\tilde{m}^u(q_m, W^n))$$

Où R^A est l'indice d'aversion absolue pour le risque :

$$R^A = -\frac{u''_x(W^n)}{u'_x(W^n)}$$

En supposant enfin que pour tout $s \in \mathcal{S}$, $|m_s^u(q_m, W^n) - m_s^u(p_m, W^n)|$ est petit devant $var(\tilde{m}^u(p_m, W^n)) = \sigma_m^2$, nous avons :

$$\frac{dW^r}{dq_m}(q_m) = -R^A \cdot q_m \cdot \sigma_m^2$$

Ainsi :

$$W^r(q_m) = \frac{1}{2} \cdot R^A \cdot \sigma_m^2 \cdot (p_m^2 - q_m^2)$$

L'ensemble des résultats concernant la prime de risque sont récapitulés par la propriété suivante :

Proposition 1.6-1 : la prime de risque W^r vérifie les propriétés suivantes

- Elle vérifie la formule suivante :

$$\frac{dW^r}{dq_m} = -\frac{1}{\tilde{v}_x^\#(q_m, W^n - W^r)} \cdot cov(\tilde{v}_x^\#(q_m, W^n - W^r), \tilde{m}^\#(q_m, W^n - W^r))$$

- Elle est d'autant plus élevée que q_m est faible, donc que le niveau de couverture est élevée.
- Lorsque la prime de risque est supposée faible devant le revenu, elle vérifie la formule suivante :

$$\frac{dW^r}{dq_m} = -\frac{1}{\tilde{v}_x^\#} \cdot cov(\tilde{v}_x^\#, \tilde{m}^\#) = -\frac{1}{\tilde{v}_x^\#} \cdot cov(\tilde{v}_x^\#, \tilde{\psi})$$

- Enfin, si la prime de risque, la prime P_A et les consommations médicales sont supposées faibles devant le revenu et si la demande de soins a une élasticité prix négligeable devant sa variance σ_m^2 , nous pouvons utiliser l'approximation :

$$W^r(q_m) = \frac{1}{2} \cdot R^A \cdot \sigma_m^2 \cdot (p_m^2 - q_m^2)$$

Où R^A est le coefficient d'aversion absolue pour le risque.

1.6.2.3. Le bilan en termes de bien-être d'une augmentation du niveau d'assurance

Une augmentation du niveau d'assurance génère comme nous venons de le voir une perte de bien-être liée au risque moral, mais aussi un gain en bien-être lié à une diminution du risque financier. Ces deux éléments doivent être mis en balance pour juger de l'effet du niveau d'assurance sur le bien-être individuel. Nous allons estimer la valeur totale d'une assurance imposant un coût résiduel q_m et montrer que lorsque la demande de soins est élastique, cette valeur est plus faible que lorsqu'elle est

inélastique (ce qui est le cas en information symétrique, sous l'hypothèse d'absence d'auto-assurance et d'autoprotection).

Tenant compte de l'ensemble des gains et perte d'utilité, le bilan global en termes de bien-être s'écrit :

Équation 1.6-xv :

$$WELF(q_m) = W^r(q_m) + \bar{W}^e(q_m) = W^r(q_m) - MHZ(q_m)$$

En tenant compte des équations 2.6 iv et 2.6-xiv, et de l'égalité précédente, nous obtenons l'égalité suivante :

Équation 1.6-xvi :

$$WELF(q_m) = - \int_{q_m}^{p_m} \frac{1}{\bar{v}_x^\#(q, W^n)} \cdot cov(\tilde{v}_x^\#(q, W^n), \tilde{\psi}(q)) dq + \int_{q_m}^{p_m} \bar{\psi}^u(q) dq - (p_m - q_m) \cdot \bar{\psi}(q_m)$$

A partir de l'équation 1.6-xvi nous montrons que le bien-être retiré de l'assurance santé est toujours plus important lorsque nous supposons la demande inélastique au prix des soins que lorsque nous la supposons élastique au prix des soins. Considérons tout d'abord le cas particulier où la demande est inélastique aux prix des soins. Dans ce cas $\bar{\psi}(q) = \bar{\psi}(p_m)$ pour tout q et le risque MHZ_0 s'écrit :

$$MHZ_0(q_m) = (p_m - q_m) \cdot \bar{\psi}(p_m) - \int_{q_m}^{p_m} \bar{\psi}(p_m) dq = 0$$

Le bilan global en termes de bien-être s'écrit :

$$WELF_0(q_m) = W^r(q_m) - MHZ_0(q_m) = W^r(q_m)$$

Considérons maintenant le cas où la demande est élastique aux prix des soins. Dans ce cas $\bar{m}^u(q) < \bar{m}^u(p_m)$ pour tout q :

$$WELF(q_m) = W^r(q_m) + \int_{q_m}^{p_m} \bar{\psi}(q) du - \bar{\psi}(q_m) \cdot (p_m - q_m)$$

Finalement $WELF(q_m) < WELF_0(q_m)$: dès lors que le prix résiduel des soins q_m est strictement inférieur à leur prix de production p_m , le bien-être de l'individu est plus faible lorsque ses consommations médicales sont élastiques aux prix des soins que lorsqu'elles sont inélastiques, en particulier lorsque l'assureur peut observer le comportement des individus (hypothèse d'information symétrique).

1.7. Demande d'assurance en présence de risque moral

1.7.1. Assurance proportionnelle optimale

Nous avons montré qu'en information symétrique, lorsque le contrat est tarifié à son coût actuariel, la couverture préférée est complète. Le phénomène de risque moral modifie la demande de couverture des individus. En effet, lorsque le niveau de couverture augmente, le bien-être généré par la diminution du risque financier peut être contrebalancé par la perte de surplus générée par le risque moral. Comme le soulignent Manning et Marquis (1996) :

“Social choices about health insurance involve a trade-off between the gains from risk reduction and the losses from inappropriate incentives for the purchase of more health care”.

Nous allons voir qu'en information asymétrique, en présence de phénomène de risque moral, la couverture est partielle, même en l'absence de chargements.

Il existe deux approches pour obtenir ce taux optimal de coassurance. La première consiste à analyser l'arbitrage en termes d'augmentation du risque moral et de réduction du risque financier et à rechercher le niveau de couverture qui réalise le meilleur compromis entre ces deux quantités. La seconde consiste à résoudre directement le programme de maximisation de l'utilité. Nous présentons la première méthode, la deuxième est donnée en annexe.

Soit $WELF(q_m)$ l'utilité, sous forme de valeur monétaire, associée au contrat d'assurance laissant un coût résiduel q_m aux individus. Le niveau de couverture optimal est obtenu en recherchant la valeur de q_m qui maximise $WELF(q_m)$.

$$\frac{dWELF(q_m)}{dq_m} = \frac{dW^r(q_m)}{dq_m} - \frac{dMHZ(q_m)}{dq_m} = 0$$

$$\frac{dMHZ(q_m)}{dq_m} = \frac{dW^r(q_m)}{dq_m}$$

Cette égalité peut être interprétée de la manière suivante : partant d'une couverture complète, les individus choisissent leur couverture préférée en arbitrant entre risque moral et réduction du risque financier. Le niveau d'assurance finalement choisi est celui qui égalise les variations marginales de ces deux quantités soit, en reprenant les notations de la partie 1.6 :

Équation 1.7-i :

$$(p_m - q_m) \cdot \frac{d\bar{\psi}}{dq_m}(q_m) = -cov\left(\frac{\tilde{v}_x}{\bar{v}_x}, \tilde{\psi}\right)$$

En multipliant par les parties droites et gauche de l'égalité ci-dessus par $q_m/\bar{\psi}$, nous obtenons :

Équation 1.7-ii :

$$(p_m - q_m) \cdot \eta_{\bar{\psi}, q_m} = -\frac{q_m}{\bar{\psi}^u} \cdot \text{cov}\left(\frac{\tilde{v}_x}{\bar{v}_x}, \tilde{\psi}\right)$$

Soit $q_m^\#$ le prix résiduel solution de l'équation 1.7-ii et $a^\#$ le taux de couverture associé à ce prix résiduel. A partir de l'équation 1.7-ii, nous obtenons la formule suivante du taux de couverture optimal :

Équation 1.7-iii :

$$\frac{a^\#}{1 - a^\#} = \frac{1 - \frac{q_m^\#}{p_m}}{1 - \left(1 - \frac{q_m^\#}{p_m}\right)} = -\frac{1}{\eta_{\bar{\psi}, q_m}} \cdot \text{cov}\left(\frac{\tilde{v}_x}{\bar{v}_x}, \tilde{\psi}\right)$$

Cette équation est similaire à celle de Dixit et Sandmo (1977), liant le taux optimal d'imposition des revenus du travail avec l'élasticité de l'offre agrégée de travail³¹ (voir Auerbach et Hines, 2002). L'offre de travail est ici remplacée par la demande de soins et l'individu paie une somme forfaitaire et reçoit une indemnité proportionnelle à la quantité de soins consommée au lieu de payer une somme proportionnelle à sa quantité de travail et de recevoir un transfert forfaitaire. La problématique est néanmoins analogue : celle d'un arbitrage entre efficacité et équité (la réduction du risque financier aboutissant à gommer les disparités de consommations non-médicales liées à la santé). Dixit et Sandmo notent ainsi à propos de la formule qu'ils ont obtenu :

“Again, we have succeeded in deriving a formula in which the equity element is represented by the numerator and the efficiency element by the denominator”

En se référant à la formule que nous avons obtenue, l'élément d'équité est représenté par le terme de covariance et l'élément d'efficacité par le terme d'élasticité de la demande de soins.

Nous pouvons réécrire l'équation 1.7-iii de manière à obtenir la formule du taux de coassurance optimal :

Équation 1.7-iv

³¹ L'équation de Dixit et Sandmo décrit le lien entre le taux de taxation optimal τ et l'élasticité prix $\eta_{q_m, \bar{L}}$ de l'offre de travail par l'équation suivante :

$$\frac{\tau}{1 - \tau} = \frac{1}{\eta_{q_m, \bar{L}}} \cdot \text{cov}\left(\frac{\tilde{v}_x}{\bar{v}_x}, \frac{\tilde{L}}{\bar{L}}\right)$$

$$c^{\#} = \frac{1}{1 - \frac{1}{\eta_{q_m, \bar{\psi}}} \cdot \text{cov}\left(\frac{\tilde{v}_x}{\bar{v}_x}, \frac{\tilde{\psi}}{\bar{\psi}}\right)}$$

Cette formule nous permet de constater que :

- le taux de coassurance optimal $c^{\#}$ est compris entre 0 et 1 ;
- plus l'élasticité prix de la demande de soins agrégée $\bar{m}^{\#}$ est importante et négative et plus le taux de coassurance est élevé ;
- le taux de coassurance tend vers 1 lorsque $\eta_{q_m, \bar{\psi}^u}$ tend vers $-\infty$ (lorsque les soins sont infiniment élastique) et tend vers 0 lorsque $\eta_{q_m, \bar{\psi}^u}$ tend vers 0. Nous retrouvons ainsi le fait que lorsque l'élasticité prix de la demande de soins est nulle, l'individu se couvre complètement ;
- ces deux cas extrêmes mis à part, $c^{\#} \in]0,1[$.

Intéressons-nous maintenant au terme de covariance. Pour cela, réécrivons l'équation 1.7-iv sous la forme suivante :

Équation 1.7-v :

$$c^{\#} = \frac{1}{1 - \frac{1}{\eta_{q_m, \bar{\psi}}} \cdot \text{cv}(\tilde{v}_x) \cdot \text{cv}(\tilde{\psi}) \cdot \text{corr}(\tilde{v}_x, \tilde{\psi})}$$

Où $\text{cv}(\tilde{v}_x)$ et $\text{cv}(\tilde{\psi})$ sont les coefficients de variation associés respectivement à \tilde{v}_x et $\tilde{\psi}$ et $\text{corr}(\tilde{v}_x, \tilde{\psi})$ est le coefficient de corrélation entre \tilde{v}_x et $\tilde{\psi}$. L'équation 1.7-v permet de constater que le taux de coassurance diminue lorsque :

- les coefficients de variation $\text{cv}(\tilde{v}_x)$ ou $\text{cv}(\tilde{\psi})$ augmentent ; autrement dit lorsque la variabilité de l'utilité marginale du revenu ou celle de la demande de soins s'accroissent ;
- lorsque la corrélation entre \tilde{v}_x et $\tilde{\psi}$ augmente ; autrement dit plus l'utilité marginale du revenu est liée à la demande de soins et plus le niveau de couverture préféré par l'individu est élevé.

Notons enfin que lorsque la demande hicksienne est approximée par la demande marshalienne, lorsque la prime P_A et les dépenses médicales sont petites devant le revenu W^n et lorsque l'élasticité-prix des consommations médicales est petite devant leur variance alors la formule du taux de coassurance optimal devient :

Équation 1.7-vi :

$$c^{\#} = \frac{\frac{d\bar{m}^u}{dq_m}}{\frac{d\bar{m}^u}{dq_m} - R^A \cdot \sigma_m^2}$$

1.7.2. Les mesures empiriques du niveau optimal de couverture

Plusieurs travaux ont cherché à déterminer le niveau de couverture optimal, à l'aide d'hypothèses ou de mesures empiriques concernant les paramètres nécessaires à ce calcul :

- Keeler *et al* (1988) déterminent le taux de coassurance optimal à partir des résultats de leurs estimations concernant les épisodes de soins. Ils considèrent un ensemble de contrats composés d'une franchise (0\$, 100\$, 200\$ ou 300\$), puis d'une assurance proportionnelle (taux de coassurance de 0%, 25%, 50% et 100%) et enfin d'un *stop-loss* qui vaut 1000\$. Les auteurs calculent la perte en bien-être liée à chaque contrat. Ce calcul nécessite d'estimer la dépense totale M sur un an, la valeur V que les individus attribuent à l'ensemble des épisodes de soins sur l'année et la prime de risque ρ du contrat (indiquant la perte liée à l'augmentation du risque financier, par rapport à un contrat de couverture complète). La perte de bien-être s'écrit :

$$WELF = V - M - \rho$$

La valeur que les individus attribuent aux soins consommés à la somme du reste à charge annuel RAC et du surplus attribué aux soins consommés $SURPL$ ³². Ce surplus est calculé en supposant une courbe de demande de soins linéaire. La prime de risque est obtenue en calculant la variance des restes à charges puis en multipliant cette valeur par 0,0005³³.

Les valeurs des différents paramètres dépendent du contrat considéré. Puisque ces contrats ne correspondent pas à ceux imposés par la RAND, les auteurs simulent des épisodes de soins sur une année en s'appuyant sur des lois négatives binomiales. Pour cette simulation, on tient compte du fait que le taux de coassurance appliqué à chaque épisode varie suivant le niveau de reste à charge résultant des épisodes précédents : en dessous de la franchise, le taux de coassurance est de 100% ; entre la franchise et le *stop-loss*, le taux de coassurance est de 0%, 25%, 50% et 100% ; au-delà du *stop-loss*, le taux de coassurance est de 100%. Le nombre d'épisodes est simulé selon une loi négative binomiale dont le paramètre $\lambda(c)$

³² Voir Keeler *et al* (1988), p95 pour le détail de ce calcul.

³³ Rappelons que la prime de risque ρ peut être approximée par :

$$\frac{1}{2} \cdot R^A \cdot Var(RAC)$$

Où R^A est le coefficient d'aversion absolue pour le risque. La présente étude prend pour valeur de ce coefficient 0,001.

dépend du taux de coassurance dans la zone considérée. $\lambda(c)$ est ajusté à partir des données de la RAND sur les épisodes de soins.

La couverture optimale, c'est-à-dire celle conduisant à la plus faible perte de bien-être collectif, comporte une franchises de 275\$ à 410\$, puis un taux de coassurance de 25% jusqu'à 1370\$, puis une couverture complète au-delà de 1370\$³⁴.

- Au moyen des données de l'expérience de la RAND, Manning et Marquis (1996) recherchent le niveau de couverture optimal en présence de risque moral, en se basant sur une spécification paramétrique de l'utilité indirecte apportée par les consommations médicales et non-médicales. Cette fonction détermine les choix de consommations de soins des individus *ex-post* la survenue d'une pathologie et le choix du niveau d'assurance *ex-ante* la survenue de la pathologie. La fonction d'utilité indirecte utilisée est la transformée d'une fonction indirecte « de base » IU par une fonction croissante concave g . La fonction d'utilité indirecte « élémentaire » s'écrit :

$$IU = -p_m^\delta e^{z\beta + \epsilon} + (W^n)^\theta$$

Où ϵ suit une loi normale d'espérance nulle et de variance σ^2

Et la fonction g s'écrit :

$$g(IU) = A \cdot (IU)^\psi$$

La demande de soins est obtenue par l'identité de Roy, à partir de la fonction d'utilité indirecte. Elle ne dépend pas de la transformation effectuée, et peut donc être calculée à partir de la fonction d'utilité indirecte « élémentaire » :

$$m = -\frac{\frac{\partial IU}{\partial p_m}}{\frac{\partial IU}{\partial W}} = \frac{\delta}{\theta} \cdot p_m^{\delta-1} e^{z\beta + \epsilon} \cdot (W^n)^{1-\theta}$$

Soit une fois passé en logarithme :

$$\ln(m) = (\delta - 1) \cdot \ln(p_m) + (1 - \theta) \ln(W^n) + z\beta + \epsilon$$

La demande d'assurance est obtenue comme l'espérance de la fonction d'utilité indirecte :

$$E(g(IU)) = A \cdot E((IU)^\psi)$$

³⁴ Pour assurer au mieux la comparabilité avec les autres résultats, les valeurs ont été converties en dollars 1991.

La demande d'assurance nécessite cette fois de connaître les paramètres de la fonction g , en particulier ψ qui détermine l'aversion au risque des individus. La stratégie d'estimation adoptée est la suivante :

- Tout d'abord, θ , δ , β et σ sont déterminés au moyen d'une régression probit sur la probabilité de dépasser le « *stop loss* », imposé par la RAND. Soit MDE ce *stop-loss*. L'individu choisit de le dépasser si :

$$(W^n - MDE)^\theta \geq -p_m \delta e^{z\beta + \epsilon} + (W^n)^\theta$$

A l'aide d'un développement limité, cette inéquation s'écrit :

$$-(W^n)^\theta \cdot \theta \cdot \frac{MDE}{W} \geq -p_m \delta e^{z\beta + \epsilon}$$

C'est-à-dire :

$$\delta \cdot \ln(p_m) + z\beta - \theta \cdot \ln(W^n) - \ln(\theta) - \ln\left(\frac{MDE}{W^n}\right) + \epsilon \geq 0$$

Le modèle probit associé s'écrit :

$$\alpha \cdot \ln(p_m) + z \cdot \mu + \omega \cdot \ln(W^n) + \lambda \cdot \ln\left(\frac{MDE}{W^n}\right) + u \geq 0$$

Avec $u \sim \mathcal{N}(0,1)$

L'estimation est menée sur les familles qui faisaient tous face à un taux de coassurance non nul dans l'expérience de la RAND. Ces familles étant suivies pendant plusieurs années, on dispose de 2131 familles-années (soient 6066 individus-années) qui constituent l'unité d'observation. Les trois variables clés introduites sont le prix des soins, le revenu et le niveau de *stop loss* de la famille. Le prix des soins est entré sous la forme du niveau de coassurance lorsque le *stop-loss* n'est pas dépassé. Les principales variables de contrôle sont les quatre variables d'état de santé collectées lors de l'expérience de la RAND : la perception générale de l'état de santé, la présence de limitations dans les activités quotidiennes résultant de problèmes de santé, le nombre de maladies chroniques et l'état de santé mental. L'âge, le sexe, l'appartenance ethnique, le nombre de personnes dans le foyer. Ces variables et celles d'état de santé sont moyennées au niveau du foyer. Enfin, les auteurs considèrent la somme des dépenses anticipées par chaque individu du ménage et régressent cette variable sur l'ensemble des variables explicatives précédemment présentées. Ils introduisent les résidus obtenus comme variable explicative du fait de dépasser le MDE. Ces résidus introduits en variable explicative permettent de contrôler des facteurs de consommations de soins inobservables du point de vue des variables collectées mais connues de l'assuré.

Dans une régression probit, les coefficients sont en principe estimés à un facteur d'échelle près et la variance du résidu est fixée à 1. Néanmoins, le fait d'imposer que le

coefficient de $\ln\left(\frac{MDE}{W^n}\right)$ soit égal à 1 permet de lever ce problème d'identification : θ , δ , β sont obtenus en divisant respectivement α , μ et ω par λ . Quant à σ , il vaut l'inverse de λ .

Les auteurs trouvent une élasticité prix égale à -0,18 ce qui est proche de celle estimée par Keeler et al (1988) à partir des épisodes de soins. Par ailleurs, l'élasticité revenu est égale à +0,22.

Une fois estimés θ , δ , β et σ , l'identité de Roy permet de déterminer la demande de soins m correspondante :

$$\ln(m) = (\delta - 1) \cdot \ln(p_m) + (1 - \theta)\ln(W^n) + z\beta + \epsilon$$

- Les auteurs estiment ensuite les paramètres ψ et A en s'appuyant sur les réponses à une question concernant une offre fictive d'assurance. Il est demandé aux individus s'ils seraient prêts à souscrire une assurance supplémentaire réduisant le niveau de leur *stop-loss* de 1/3, de 2/3 ou de 100% (l'assurance est alors complète) en échange d'une prime qui est générée aléatoirement. Soient MDE^{BASE} , $MDE^{1/3}$, $MDE^{2/3}$, MDE^{FULL} les quatre niveaux de *stop-loss* que l'individu peut choisir (MDE^{BASE} est le contrat imposé par la RAND experiment et $MDE^{FULL} = 0$) et h la variable qui indique si l'individu choisit le contrat. Les auteurs examinent si les individus préfèrent ou non souscrire l'offre d'assurance supplémentaire ou conserver le contrat de base seul. L'individu choisit le contrat de couverture j ($j \in \{BASE, \frac{1}{3}, \frac{2}{3}, FULL\}$) plutôt que le contrat de base si :

$$E(g(UI_j)) - E(g(UI_{BASE})) \geq 0$$

L'estimation de ψ et A est effectuée en s'appuyant sur une grille de valeurs possibles de ces deux paramètres. Pour chaque point de cette grille, les utilités indirectes associées à chaque contrat sont déterminées en simulant des valeurs aléatoire de ces utilités : ces valeurs aléatoires sont générées en effectuant plusieurs tirage de ϵ selon une loi $\mathcal{N}(0, \sigma^2)$. Les espérance d'utilité n'étant pas estimées de manière précise (notamment parce-que le nombre de tirage est fini) il existe un aléa sur la mesure de $E(g(UI_j)) - E(g(UI_{BASE}))$. Pour tenir compte de cet aléa, les auteurs se donnent une fonction de lien G , la fonction de répartition de la loi normale) ce qui permet de considérer la souscription de l'assurance supplémentaire en termes de probabilité. Le modèle devient donc :

$$p(h = 1) = G \left[E(g(UI_j)) - E(g(UI_{BASE})) \right]$$

Les valeurs de ψ et A retenues correspondent au point de la grille tel que l'ajustement entre choix prédits et choix réels soit le meilleur. L'analyse est menée auprès de 698 familles, ce qui permet de disposer de 3078 réponses aux offres hypothétiques. L'estimation de ψ vaut 0,425 et celle de A vaut 6,583. Connaissant maintenant l'ensemble des paramètres de la fonction d'utilité indirecte, il est possible de calculer l'élasticité prix de la demande d'assurance et coefficient d'aversion absolue pour le risque. L'élasticité prix vaut -0,54 et l'aversion absolue pour le risque +0,00021.

- La dernière étape consiste à calculer la prime de risque pour différentes formes de contrat indexés k , c'est-à-dire le montant maximum ρ que l'individu est prêt à payer pour souscrire l'assurance supplémentaire qui rend la couverture complète plutôt que d'être couvert que par le contrat k :

$$E(g(UI_{FULL})) = E(g(UI_k))$$

$$(W^n - \rho)^\psi = E((UI_k)^\psi)$$

$$\rho = W^n - [E((UI_k)^\psi)]^{\frac{1}{\psi}}$$

Pour estimer $[E((UI_k)^\psi)]$, les auteurs procèdent de nouveau par simulation : ils effectuent 25 tirages de ϵ selon la loi $\mathcal{N}(0, \sigma^2)$ puis, en faisant la moyenne empirique, calculent $E((UI_k)^\psi)$. En comparant cette prime de risque avec celle d'un contrat de couverture nulle, ils obtiennent le gain offert par chaque alternative en termes de réduction du risque. Par ailleurs, les Manning et Marquis utilisent la demande compensée³⁵ de soins pour estimer la perte en termes de risque moral pour chaque alternative. En faisant la différence entre gain lié à la réduction du risque financier et perte lié au risque moral, ils obtiennent le bilan en termes de bien-être collectif.

Deux formes de contrats sont testées : des contrats d'assurance purement proportionnelle (le seul paramètre est un taux de coassurance que l'on fait varier) et des contrats d'assurance proportionnelle assorties *stop-loss* (on fait varier le niveau de coassurance et la limite sur les dépenses à la charge du patient). Dans le cadre d'un contrat d'assurance purement proportionnel, le taux de coassurance préféré est de 45%. Dans le cadre d'un contrat d'assurance proportionnelle assortie d'un *stop-loss*, la couverture préférée comporte un taux de coassurance de 25% et un *stop-loss* supérieur à 25 000€.

³⁵ Dans l'article, les auteurs n'expliquent pas comment ils obtiennent cette demande compensée.

1.7.3. Effet d'un subventionnement de la couverture sur les niveaux de garanties et les niveaux de consommations de soins

Nous allons maintenant étudier l'effet d'un subventionnement de la couverture sur le niveau de couverture souscrit et sur le bien-être des individus. Les subventionnements prennent plusieurs formes :

- En France comme aux Etats-Unis, une partie des employeurs offrent une couverture à leurs salariés et financent tout ou partie de la prime. En contrepartie cette cotisation est exonérée de cotisations sociales patronales, ce qui en fait un élément attractif de rémunération des salariés. Les individus bénéficient ainsi d'un prix apparent de la couverture plus faible que sur le marché de l'assurance individuelle en échange d'un niveau de salaire plus faible (plusieurs travaux ont tenté de montrer que l'assurance santé se substitue au salaire, avec des résultats contradictoires ; cf Currie et Madrian, 1999, Gruber, 2000). Si l'hypothèse de substitution est bien vérifiée, les individus bénéficient d'une réduction du coût de leur couverture contre un prélèvement sur leur salaire.
- En France, les indépendants peuvent déduire leurs cotisations de couverture santé de leurs bénéfices imposables. Il s'agit bien d'une forme de subventionnement de la couverture santé puisque le coût de la couverture est amoindri par l'exemption de taxes.
- En France toujours, les individus peuvent, sous conditions de ressources, bénéficier de l'Aide complémentaire santé (ACS) qui est une aide forfaitaire ajustée à l'âge. L'ACS est financée par un prélèvement sur les contrats de couverture complémentaire santé. Le coût des contrats est réduit pour les personnes qui bénéficient de ce dispositif et augmenté pour les autres individus.

De la même manière que le subventionnement des soins par la couverture complémentaire santé conduit à une situation sous-optimale, nous pouvons nous attendre à ce que le subventionnement de la couverture complémentaire santé réduise l'efficacité du système de soins. Nous allons nous intéresser essentiellement à un subventionnement proportionnel à la prime. Nous supposons que pour un contrat de coût total $P_A = (p_m - q_m) \cdot \bar{\psi}(q_m)$, une part t de ce coût est à la charge des individus, l'autre part $1 - t$ est financée par un autre agent (employeur, Etat, collectivité locale...). En reprenant les calculs effectués dans la partie 2.6, nous obtenons que le bien-être retiré d'un contrat de niveau de couverture q_m est donné par la formule suivante, analogue à l'équation 1.6-xvi :

Équation 1.7-vii :

$$WELF_{sub}(q_m) = - \int_{q_m}^{p_m} \frac{1}{\bar{v}_x^\#(q, W)} \cdot cov\left(\tilde{v}_x^\#(q, W^n), \tilde{\psi}(q, W^n)\right) dq + \int_{q_m}^{p_m} \bar{\psi}(q) dq - t \cdot (p_m - q_m) \cdot \bar{\psi}(q_m)$$

En recherchant $WELF_{sub}$, nous trouvons que le niveau optimal de q_m est donné par :

$$\frac{dWELF_{sub}}{dq}(q_m) = 0$$

Donc par :

$$-\frac{1}{\bar{v}_x^\#(q_m, W^n)} \cdot cov\left(\tilde{v}_x^\#(q_m, W^n), \tilde{\psi}(q_m)\right) - (1-t) \cdot \bar{\psi}^u(q_m) = t \cdot (p_m - q_m) \cdot \frac{d\bar{\psi}}{dq_m}(q_m)$$

En multipliant les parties gauche et droite de cette égalité par $q_m/\bar{\psi}^u$, nous obtenons :

$$-q_m \cdot cov\left(\frac{\tilde{v}_x^\#}{\bar{v}_x^\#}, \frac{\tilde{\psi}}{\bar{\psi}}\right) - q_m \cdot (1-t) = t \cdot (p_m - q_m) \cdot \eta_{\bar{\psi}, q_m}$$

Par conséquent, le prix résiduel optimal est solution de l'équation 1.7 viii ci-dessous :

Équation 1.7-viii

$$-\frac{q_m}{p_m} \cdot cov\left(\frac{\tilde{v}_x^\#}{\bar{v}_x^\#}, \frac{\tilde{\psi}}{\bar{\psi}}\right) - \frac{q_m}{p_m} \cdot (1-t) = t \cdot \left(1 - \frac{q_m}{p_m}\right) \cdot \eta_{\bar{\psi}, q_m}$$

Soit $q_{m,sub}$ la solution de cette équation. Nous obtenons ainsi le niveau de coassurance c^{sub} en présence de risque moral et de subventionnement de la couverture :

Équation 1.7-ix

$$c^{sub}(t) = \frac{q_{m,sub}(t)}{p_m} = \frac{1}{1 - \frac{1}{\eta_{\bar{\psi}, q_m}} \cdot \left[\left(\frac{1}{t} - 1\right) + \frac{1}{t} \cdot cov\left(\frac{\tilde{v}_x^\#}{\bar{v}_x^\#}, \frac{\tilde{\psi}}{\bar{\psi}}\right)\right]}$$

Nous remarquons que le taux de coassurance préféré décroît lorsque la part de la prime à la charge des individus diminue : plus le niveau de subventionnement est élevé et plus les individus souhaitent être couverts. En particulier, $c^{sub}(t) < c^{sub}(1) = c^\#$ avec $c^\#$ le niveau de coassurance préféré en l'absence de subventionnement : Le niveau de coassurance préféré en présence d'un subventionnement de l'assurance santé est inférieur au taux de coassurance préféré en l'absence d'un tel subventionnement : le subventionnement de la couverture santé accroît la demande de couverture.

Qu'en est-il du bien-être des individus ? Pour répondre à cette question, constatons tout d'abord que la portion de la prime qui n'est pas payée directement par les individus, à savoir :

$$(1 - t) \cdot (p_m - q_{m,sub}(t)) \cdot \bar{\psi}(q_{m,sub}(t))$$

Cette part est payée indirectement au travers de hausse de taxes sur d'autres biens et services que l'assurance santé. Le bilan $WELF(q_{m,sub})$ en termes de bien-être des individus s'écrit donc :

$$\begin{aligned} WELF(q_{m,sub}) &= WELF_{sub}(q_{m,sub}) - (1 - t) \cdot (p_m - q_{m,sub}) \cdot \bar{\psi}(q_{m,sub}) \\ WELF(q_{m,sub}) &= - \int_{q_{m,sub}}^{p_m} \frac{1}{\bar{v}_x^\#(q, W^n)} \cdot cov(\tilde{v}_x^\#(q, W^n), \tilde{\psi}(q, W^n)) dq + \int_{q_{m,sub}}^{p_m} \bar{\psi}(q) dq \\ &\quad - t \cdot (p_m - q_{m,sub}) \cdot \bar{\psi}(q_{m,sub}) - (1 - t) \cdot (p_m - q_{m,sub}) \cdot \bar{\psi}(q_{m,sub}) \\ WELF(q_{m,sub}) &= - \int_{q_{m,sub}}^{p_m} \frac{1}{\bar{v}_x^\#(q, W^n)} \cdot cov(\tilde{v}_x^\#(q, W^n), \tilde{\psi}(q, W^n)) dq + \int_{q_{m,sub}}^{p_m} \bar{\psi}(q) dq \\ &\quad - (p_m - q_{m,sub}) \cdot \bar{\psi}(q_{m,sub}) \end{aligned}$$

Puisque $q_{m,sub} = c^{sub} \cdot p_m$ est inférieur à $q_m^\#$, le prix résiduel qui maximise $WELF(q)$, $WELF(q_{m,sub}) < WELF(q_m^\#)$: il y a une perte de bien-être par rapport à la situation où le subventionnement est nul ; autrement dit, le subventionnement de la couverture santé conduit à une perte d'optimalité. Pauly (1986) décrit le mécanisme qui conduit à cette perte d'optimalité :

- 1) Les subventionnements de l'assurance santé distordent le processus de choix d'une telle assurance, ce qui a pour conséquence l'achat d'assurance dont le niveau est trop élevé.
- 2) L'accroissement du niveau d'assurance a pour effet de distordre la demande de soins.
- 3) Cette distorsion de la demande de soins rentre en interaction avec l'offre de soins, ce qui conduit à un niveau de coût trop élevé.

Il convient d'être prudent quant au caractère purement néfaste d'un subventionnement des contrats. En effet, nous n'avons pas pris en compte dans le bien-être des individus l'effet bénéfique de tels subventionnements en termes de réduction de la sélection adverse (le contrat subventionné est plus attractif pour tous, y compris les individus ayant un faible risque santé) et en termes d'amélioration de l'accès à la couverture santé, donc aux soins. De plus, nous avons supposé dans l'ensemble de ce travail que les assureurs sont en concurrence pure et parfaite. Cette hypothèse est loin d'être évidente ; ainsi Franck et Lamiraud (2009) ont montré qu'en Suisse, malgré un grand nombre d'opérateurs et des conditions favorables à la concurrence, il existe d'importants écarts de

prix entre organismes. Dans ces conditions, nous pouvons faire l'hypothèse qu'en réponse à un subventionnement de quelque forme que ce soit, les organismes augmentent les prix des contrats et/ou la quantité d'assurance offerte tant que leur utilité reste supérieur à celle retirée d'un contrat d'un autre organisme.

Application empirique

Feldstein et Friedman (1977) s'intéressent à l'effet des exonérations de taxes appliquées aux contrats collectifs sur le niveau de couverture choisi par les individus en présence de risque moral. Les auteurs développent un modèle de demande de couverture en présence de risque moral. Dans ce modèle, les auteurs prennent en compte les coûts non monétaires associés à la consommation de soins ; ces coûts non monétaires sont notés p_m^0 . Le prix total des soins en l'absence d'assurance vaut $p_m + p_m^0$; en présence d'assurance, il vaut $q_m + p_m^0$. La demande de soins m des individus est supposée décrite par l'équation suivante :

$$m(q_m) = m_0 \cdot \left(\frac{q_m + p_m^0}{p_m + p_m^0} \right)^\eta$$

Où m_0 est une variable aléatoire suivant une loi de densité f . m_0 représente la demande de soins en l'absence d'assurance ; η est l'élasticité-prix de la demande de soins. La demande de soins étant supposée élastique, l'assurance augmente la quantité de biens et services médicaux consommés. La valeur de ce supplément de soins est mesurée par :

$$V_{m_0} = \frac{\eta \cdot m_0}{(\eta + 1)(p_m + p_m^0)^\eta} \cdot [(p_m + p_m^0)^{\eta+1} - (q_m + p_m^0)^{\eta+1}]$$

Le reste à charge de l'individu correspond à la dépense à la charge de l'individu, c'est-à-dire :

$$E_{m_0}(q_m) = q_m \cdot m_0 \cdot \left(\frac{q_m + p_m^0}{p_m + p_m^0} \right)^\eta$$

La cotisation est donnée par :

$$P_I(q_m) = (p_m - q_m) \cdot (1 + \tau) \cdot (1 - t) \cdot \left(\frac{q_m + p_m^0}{p_m + p_m^0} \right)^\eta \int_0^{+\infty} x \cdot f(x) \cdot dx$$

Où $(p_m - q_m) \cdot p_m \cdot m_0 \cdot \left(\frac{q_m + p_m^0}{1 + p_m^0} \right)^\eta \int_0^{+\infty} x \cdot f(x) \cdot dx$ correspond à l'espérance des prestations versées ($\int_0^{+\infty} x \cdot f(x) \cdot dx$ est l'espérance mathématique de m_0), τ le taux de chargement de la compagnie et t le subventionnement de l'assurance par l'Etat.

La variation d'utilité espérée liée à l'augmentation de la cotisation et à la diminution des restes à charge (donc du risque financier) s'écrit :

$$\int_0^{+\infty} [u(W^n - E_x(q_m) - P_I(q_m)) - u(W^n - E_x(p_m))] \cdot x \cdot f(x) \cdot dx$$

La valeur que l'individu attribue à l'accroissement des consommations médicales peut être traduit en termes d'utilité en multipliant V_{m_0} par la dérivée de u appliquée à la moyenne du revenu disponible avec et sans assurance. Ainsi, l'espérance d'utilité attribuée au supplément de consommation est donnée par :

$$\int_0^{+\infty} V_x \cdot u' \left(W^n - \frac{E_x(q_m) + E_x(1) + P_I(q_m)}{2} \right) \cdot x \cdot f(x) \cdot dx$$

$WELF(q_m)$, le bilan en termes de bien-être s'écrit :

$$\int_0^{+\infty} \left[u(W^n - E_x(q_m) - P_I(q_m)) - u(W^n - E_x(p_m)) \right. \\ \left. + V_x \cdot u' \left(W^n - \frac{E_x(q_m) + E_x(p_m) + P_I(q_m)}{2} \right) \right] \cdot x \cdot f(x) \cdot dx$$

Feldstein et Friedman obtiennent une forme explicite de cette intégrale en prenant comme fonction d'utilité $u(w) = -\exp(-R \cdot w)$ (R est l'aversion absolue pour le risque ; il est constant pour cette fonction d'utilité) et comme fonction de distribution f , une loi gamma de paramètres (α, β) :

$$f(x) = k_0 \cdot x^\alpha \cdot \exp\left(-\frac{x}{\beta}\right)$$

Avec $\alpha = -0,61$ et $\beta = 1573$. Outre le fait que la loi gamma simplifie ici les calculs, elle s'adapte bien à la forme de la distribution des dépenses de santé.

D'autre part, se basant sur les résultats des travaux antérieurs à celui-ci, Feldstein et Friedman considère trois valeurs possibles d'élasticité prix de la demande de soins, $\eta = 0,3$, $\eta = 0,5$ et $\eta = 0,7$ et trois d'aversion absolue pour le risque, $R = 0,0001$, $R = 0,0003$ et $R = 0,0005$. Le taux de chargement τ est égal à 0,09, soit ce qui est observé en moyenne chez les assureurs. Les auteurs calculent le taux de coassurance optimal pour chacune des trois valeurs des paramètres η et R et pour différents niveaux de revenu. A chacun de ces niveaux de couverture correspond un taux de taxation du revenu, ce qui fournit t . La recherche du niveau de couverture optimal est fait avec et

sans les taxes. Le niveau de couverture optimal varie 0,5 à 1 sans exonération des taxes et de 0,37 à 0,79 avec exonérations des taxes. Les exonérations de charges fiscales baissent le coût des contrats et incitent donc les individus à acheter une couverture plus élevée que sans les exonérations. Dans les deux cas de figure, le taux de coassurance augmente avec l'élasticité-prix de la demande de soins et le degré d'aversion absolue pour le risque.

Feldstein et Friedman déterminent ensuite les taux effectifs de coassurance en se basant sur les données de 200 000 salariés de l'administration fédérale. Ces salariés ont le choix de se couvrir ou non et s'ils se couvrent, ils ont le choix entre quatre contrats qui peuvent être hiérarchisés selon leur niveau de franchise et leur niveau de coassurance pour les soins hospitaliers et ambulatoires (les contrats offrant les plus bas niveaux de coassurance sont également ceux proposant les plus faibles franchises). La population est découpée en classes de revenu auxquelles est associé le taux d'imposition correspondant. Pour chaque classe de revenu les taux de coassurance optimaux théoriques sont calculés pour chaque combinaison d'une grille de paramètres d'aversion absolue au risque et d'élasticité-prix. Au final, ne sont sélectionnées que les combinaisons de paramètres telles que taux théoriques et taux réels (calculés par le ratio « reste à charge/dépense totale ») soient proches. Pour chaque classe de revenu et chacune des combinaisons « aversion pour le risque × élasticité-prix » conservées, le modèle établi précédemment permet également de déterminer quelle option serait choisie en absence d'exonérations de taxes. A partir des options choisies avec exonération et en absence d'exonérations de charges il est possible de calculer les taux réel de coassurances, pour l'ambulatoire et l'hôpital, avec et sans exonérations. Sans exonérations, le taux est de 0,6 pour le total des dépenses, 0,44 pour l'hôpital et 0,8 pour l'ambulatoire. Avec exonérations, le taux est de 0,4 pour le total des dépenses, 0,35 environ pour l'hôpital et tombe à 0,45 / 0,5 pour l'ambulatoire soit près de la moitié de ce qui était observé sans exonération.

Enfin, Feldstein et Friedman reprennent le modèle théorique en intégrant l'effet d'un accroissement des prix lié à une meilleure qualité des soins. Un tel accroissement s'accompagne de deux effets contradictoires sur les consommations de soins: un effet à la hausse puisque les soins proposés sont d'une meilleure qualité, et un effet à la baisse puisque les coûts à la charge des patients augmentent.

Une dernière analyse consiste à simuler les consommations médicales et les taux de coassurance optimaux en supposant que le contrat est subventionné à hauteur de 1/3 de son coût total. En effet, comme nous l'avons déjà mentionné, ce type de subventionnement est supposé accroître la perte de bien-être collectif liée au risque moral puisque les individus ne paient pas la totalité de la prime (ce qui accroît leur niveau de couverture et de consommations de soins par rapport au contrat optimal). La simulation fait apparaître une hausse de la dépense moyenne de santé de 8%. La perte de bien-

être par rapport au contrat de couverture complète obtenu en information symétrique est de 63,5\$. Comparé à la perte de 52,6\$ obtenue dans la première simulation, cela représente 10,88\$ en plus. A l'échelle des 67 millions de familles américaines dont le chef de famille est âgé de moins de 65 ans, cela représente une perte de 730 millions d'euros, perte beaucoup plus faible que celle obtenue par Feldman et Dowd (1991) ou Manning *et al* (1987) lorsqu'ils examinaient le passage d'une assurance complète à une assurance avec un taux de coassurance de 95%.

Le mécanisme de risque moral *ex-post* décrit par Pauly a été discuté, voire critiqué, sur deux aspects :

- Ce mécanisme ne tient pas compte de l'effet revenu de l'assurance sur les consommations de soins ; une partie de la hausse des consommations de soins résulte de cet effet revenu qui est neutre d'un point de vue du bien-être collectif et positif en termes d'accès aux soins. Cet effet revenu a été mis en exergue par de Mezza (1983) puis plus tard par Nyman. Nous l'aborderons dans la partie suivante.
- D'autre part, le rôle des producteurs de soins n'est pas pris en compte. Concernant cet aspect, il faut distinguer deux types de discussion : Feldstein (1973) souligne que les producteurs de soins réagissent à la modification de la demande de soins liée à l'assurance santé en accroissant la qualité des soins offerts (par exemple pour les cliniques, en augmentant leur personnel) ; chaque unité de soins apporte donc un plus haut niveau d'utilité, ce qui conduit à revoir à la baisse la perte de bien être liée au risque moral. Rice émet une critique plus radicale en remettant le fait qu'on puisse dériver des consommations de soins des individus leur demande de soins et donc l'utilité qu'ils attachent à ces soins. En effet, pour une pathologie donnée, le patient ne connaît en général ni la nature, ni la quantité de soins à administrer.

C'est l'ensemble de ces phénomènes impliquant le rôle des professionnels de santé que nous allons examiner ici.

1.8. Risque moral et offre de soins

Nous nous sommes placés jusqu'à présent dans un cadre économique où la décision concernant le traitement revient au patient. Le professionnel de santé intervient comme expert : il décrit l'état de santé du patient au moment de la consultation et l'état de santé dans lequel il se trouvera en fonction de la quantité de soins consommés. Il intervient également comme fournisseur de soins: au regard de l'expertise du médecin, le patient choisit la quantité de soins qu'il juge optimal en arbitrant entre l'utilité apporté par chaque unité supplémentaire et le coût résiduel (c'est-à-dire une fois pris en compte les remboursements opérés par l'assurance) de cette unité ; le professionnel de santé

exécute tous les soins demandés par le patient. Ces hypothèses nous ont permis de nous affranchir des phénomènes d'inefficience liés à l'offre.

1.8.1. Contexte économique

Dans le chapitre vingt du « handbook of health economics » consacré à l'organisation industrielle du marché des soins de santé, Dranove et Satterthwaite(2000) soulignent que le modèle qui décrit de manière acceptable le mode de concurrence entre la plupart des producteurs de soins est celui de la concurrence monopolistique. Dans le même ouvrage, McGuire (2000) met également en avant cette organisation de l'offre pour décrire la relation entre patients et médecins :

In virtually all characterizations of physicians in economics journals and textbooks, the physician is portrayed as having some market power. Monopolistic competition includes an element of monopoly (downward-sloping demand) and an element of competition (large number of competitors - each firm ignoring strategic interactions).

La notion de concurrence monopolistique entre producteurs a été introduite par Chamberlin au début des années 1930³⁶. Elle repose sur deux hypothèses : d'une part, les producteurs font face à un coût fixe non nul et, de ce fait, on a une courbe de coût moyen en U. D'autre part, il y a libre entrée des producteurs sur le marché. Enfin, du point de vue des consommateurs, l'offre de chaque producteur présente des spécificités et n'est donc pas parfaitement substituable d'un producteur à l'autre. Les individus valorisent donc de manière différente l'offre de chaque producteur, certains préférant l'offre d'un producteur j_0 , d'autres l'offre d'un producteur j_1 . McGuire (2000) souligne ainsi :

The basic conception of the physician-patient relationship embodied by monopolistic competition is that physicians are imperfect substitutes in the eyes of patients. A patient has a demand for the services of a particular physician, as opposed to demand for "physicians' services" in general.

Les attributs qui rendent l'offre différente d'un producteur sont par exemple l'implantation géographique, la qualité du produit ou du service fourni, la spécialisation du praticien. McGuire met également en avant l'insuffisance d'information que peut avoir le patient concernant les substituts aux praticiens qu'il consulte régulièrement :

Information imperfections could also generate a monopolistically competitive structure. Substitutes for a given physician, but if the patient doesn't know who these are, the patient may be willing to pay more for the services of the familiar doctor, an idea proposed by Pauly and Satterthwaite (1981).

³⁶ Chamberlin E., (1933). Theory of Monopolistic competition

Au regard de ces hypothèses, déterminons les caractéristiques des consommateurs et des producteurs sur ce marché puis les équilibres qui en découlent, en l'absence puis en présence d'assurance santé.

Les caractéristiques des consommateurs

Pour tenir compte du fait que les individus valorisent différemment les soins apportés par chaque professionnel de santé, nous supposons que leur utilité a la forme suivante :

$$U_{s,k}(x_s, m_s) = u_x(x_s) + u_m(s, m_s) + u_j(k)$$

Avec j le producteur de soins préféré par l'individu, $u_j(k)$ le supplément d'utilité apporté par ce producteur j . Nous supposons ainsi que $u_j(k) = u_0 > 0$ si $k = j$ et $u_j(k) = 0$ sinon. Par souci de simplicité, nous supposons de plus que :

- j est une variable aléatoire indépendante et identiquement distribuée parmi les individus et suivant une loi uniforme.
- Les individus ne peuvent être que dans deux états : $s = 0$, en bonne santé, $s = 1$, malade.
- Les individus accordent une utilité nulle aux soins lorsqu'ils sont en bonne santé (autrement dit $u_m(0, m_s) = 0$) et ne consomment donc aucun soin dans cet état.

Les caractéristiques des producteurs de soins

Chaque producteur de biens et services médicaux j transforme une quantité $z_{m,j}$ de l'input z en une quantité m_j de soins selon la fonction de production $g_{m,j}$:

$$g_{m,j}(z_{m,j}) = m_j$$

La fonction de coût $C_{m,j}$ des soins, correspondant au coût minimum lié à la production d'une quantité m de biens et services médicaux, est définie au moyen de la fonction $g_{m,j}$ comme suit :

$$C_{m,j}(m_j) = \min_{g_{m,j}(z_{m,j})=m_j} p_z \cdot z_{m,j}$$

Nous nous intéressons maintenant aux producteurs de biens et services non-médicaux. Ayant supposé qu'ils présentent des caractéristiques identiques et produisent le même panier de biens et services, nous pouvons considérer un seul producteur représentatif (il n'y a donc pas indexation du producteur). Ils transforment une quantité z_x de l'input z en une quantité $x(z_x)$ de soins selon la fonction de production g_x :

$$g_x(z_x) = x(z_x)$$

La fonction de coût C_x , correspondant au coût lié à la production d'une quantité x de biens et services non médicaux, est définie au moyen de la fonction g_x comme suit :

$$C_x(x) = \min_{g_x(z_x)=x} p_z \cdot z_x$$

Les fonctions de production $g_{m,j}$ et g_x sont supposées croissantes, concaves. $C_{m,j_0}(\cdot)$ et $C_x(\cdot)$ sont donc strictement croissantes et convexes : $C'_{m,j_0}(\cdot) > 0$ et $C''_{m,j_0}(\cdot) > 0$ (respectivement $C'_x(\cdot) > 0$ et $C''_x(\cdot) > 0$).

Les caractéristiques des producteurs

Nous notons \hat{p}_m et \hat{p}_x les profits réalisés par les producteurs dans le secteur des biens et services médicaux et dans celui des biens et services non-médicaux, en fonction de la quantité d'input. Nous notons p_m et p_x les mêmes profits, en fonction de la quantité d'output.

- Pour les producteurs de biens et services médicaux :

$$p_m(m_j) = p_m \cdot m_j - C_m(m_j)$$

- Pour les producteurs de biens et services non médicaux :

$$p_x(x) = x - C_x(x)$$

1.8.2. Quantité et prix des soins prescrits par les professionnels de santé dans le cadre d'une concurrence monopolistique, sans assurance.

Cas où il y a liberté tarifaire

Dans le cadre d'un marché des biens et services médicaux imparfaitement concurrentiel, les producteurs peuvent agir à la fois sur le prix et la quantité de soins qu'il vend pour maximiser leur profit. Ils sont néanmoins confrontés à une contrainte de participation, puisque les individus qui accordent un supplément de valeur à leurs soins (ceux pour qui $u_j(j) = u_0$) et qui constituent leur clientèle potentielle ne doivent pas retirer plus d'utilité en consommant les soins d'autres producteurs. Sous les hypothèses précédentes, l'offre de soins de chaque producteur j est donc solution du programme suivant :

$$\max_{m_{1,j}, q_{m,j}} p_{m,j} \cdot m_{1,j} - C_{m,j}(m_{1,j})$$

$$sc: U_{1,j}(x_{1,j}, m_{1,j}) \geq V_{1,j-}$$

$$sc: x_{1,j} + q_{m,j} \cdot m_{1,j} = W^n$$

Où $V_{1,j-}$ est l'utilité maximum retirée lorsque l'individu choisit de consulter un autre professionnel que j . La contrainte de budgets est supprimée en remplaçant $x_{1,j}$ par $W^n - q_{m,j} \cdot m_{1,j}$ dans la contrainte de participation. Le problème devient ainsi :

Programme 1.8-i

$$\max_{m_{1,j}, p_{m,j}} p_{m,j} \cdot m_{1,j} - C_{m,j}(m_{1,j})$$

$$sc: U_{1,j}(W^n - p_{m,j} \cdot m_{1,j}, m_{1,j}) \geq V_{1,j-}$$

Le Lagrangien associé à ce problème d'optimisation s'écrit :

$$\mathcal{L} = p_{m,j} \cdot m_{1,j} - C_{m,j}(m_{1,j}) + \lambda_{1,j} \cdot [U_{1,j}(W^n - p_{m,j} \cdot m_{1,j}, m_{1,j}) - V_{1,j-}]$$

En dérivant le lagrangien par rapport à $m_{1,j}$ et $p_{m,j}$, nous aboutissons respectivement à :

Équations 1.8-i

(a)

$$p_{m,j} - C'_{m,j}(m_{1,j}) + \lambda_{1,j} \cdot \left[\frac{\partial U_{1,j}}{\partial m} \Big|_{(W^n - p_{m,j} \cdot m_{1,j}, m_{1,j})} - p_{m,j} \cdot \frac{\partial U_{1,j}}{\partial x} \Big|_{(W^n - p_{m,j} \cdot m_{1,j}, m_{1,j})} \right] = 0$$

(b)

$$m_{1,j} - \lambda_{1,j} m_{1,j} \cdot \frac{\partial U_{1,j}}{\partial x} \Big|_{(W^n - p_{m,j} \cdot m_{1,j}, m_{1,j})} = 0$$

La deuxième équation indique que $\lambda_{1,j} > 0$, donc que la contrainte de participation est saturée. Ainsi, les valeurs de $m_{1,j}$ et $p_{m,j}$ parmi lesquels le producteur effectue ses choix sont liées par les équations :

Équations 1.8-ii

(a)

$$U_{1,j}(W^n - p_{m,j} \cdot m_{1,j}, m_{1,j}) = V_{1,j-}$$

(b)

$$C'_{m,j}(m_{1,j}) = \frac{\frac{\partial U_{1,j}}{\partial m} \Big|_{(W^n - p_{m,j} \cdot m_{1,j}, m_{1,j})}}{\frac{\partial U_{1,j}}{\partial x} \Big|_{(W^n - p_{m,j} \cdot m_{1,j}, m_{1,j})}}$$

Par le théorème des fonctions implicites, pour tout $m_{1,j}$, il existe une fonction $p_{m,j}(m_{1,j})$ telle que $(m_{1,j}, p_{m,j}(m_{1,j}))$ respecte l'équation 1.8-ii (a). La courbe $(m_{1,j}, p_{m,j}(m_{1,j}))$ est telle que :

$$p_{m,j}(m_{1,j}) = \frac{\frac{\partial U_{1,j}}{\partial m} \Big|_{(W^n - p_{m,j} \cdot m_{1,j}, m_{1,j})}}{\frac{\partial U_{1,j}}{\partial x} \Big|_{(W^n - p_{m,j} \cdot m_{1,j}, m_{1,j})}} - m_{1,j} \cdot \frac{\partial p_{m,j}}{\partial m_{1,j}} > \frac{\frac{\partial U_{1,j}}{\partial m} \Big|_{(W^n - p_{m,j} \cdot m_{1,j}, m_{1,j})}}{\frac{\partial U_{1,j}}{\partial x} \Big|_{(W^n - p_{m,j} \cdot m_{1,j}, m_{1,j})}}$$

Cette courbe, notée « U=Vj- » sur le graphique 1.8-i ci-dessous est donc située au-dessus de la courbe de demande des individus (notée « DEM » dans le graphique 1.8 i) qui vérifie quant à elle l'équation :

$$p_{m,j} = \frac{\frac{\partial U_{1,j}}{\partial m} \Big|_{(W^n - p_{m,j} \cdot m_{1,j}, m_{1,j})}}{\frac{\partial U_{1,j}}{\partial x} \Big|_{(W^n - p_{m,j} \cdot m_{1,j}, m_{1,j})}}$$

Considérons maintenant la courbe de revenu marginal (notée « RM » dans le graphique 1.8-i). Notons que cette courbe est située au-dessous de celle de demande puisque, par décroissance de la fonction de demande $p_{m,j}$:

$$\frac{\partial(p_{m,j} \cdot m_{1,j})}{\partial m_{1,j}} = p_{m,j} + m_{1,j} \cdot \frac{\partial p_{m,j}}{\partial m_{1,j}} < p_{m,j}$$

De plus, sous l'hypothèse que l'élasticité prix de la demande de soins $\eta_{m_{1,j}, p_{m,j}}$ est comprise entre -1 et 0³⁷, la courbe de revenu marginal est décroissante puisque :

$$\frac{\partial(p_{m,j} \cdot m_{1,j})}{\partial m_{1,j}} = p_{m,j} \cdot \left(1 + \frac{m_{1,j}}{p_{m,j}} \cdot \frac{\partial p_{m,j}}{\partial m_{1,j}} \right) = p_{m,j} \cdot \left(1 + \frac{1}{\eta_{m_{1,j}, p_{m,j}}} \right) < 0$$

A partir des équations 1.8-ii (a) et (b), nous constatons que :

³⁷ Cette hypothèse d'une élasticité prix supérieure à -1 est confirmée par différents résultats empiriques américains : Ringel et al (2002) et Liu et Chollet (2006) notent tous deux que les élasticités prix trouvées par ces travaux sont négatives mais d'assez faible intensité, quasiment toujours comprise entre -1 et 0 : concernant l'élasticité prix totale, Ringel fait état de valeurs tournant autour de 0,17 et Liu et Chollet de valeurs autour de -0,2. Les élasticités varient fortement d'un type de soins à l'autre mais restent comprises entre -1 et 0.

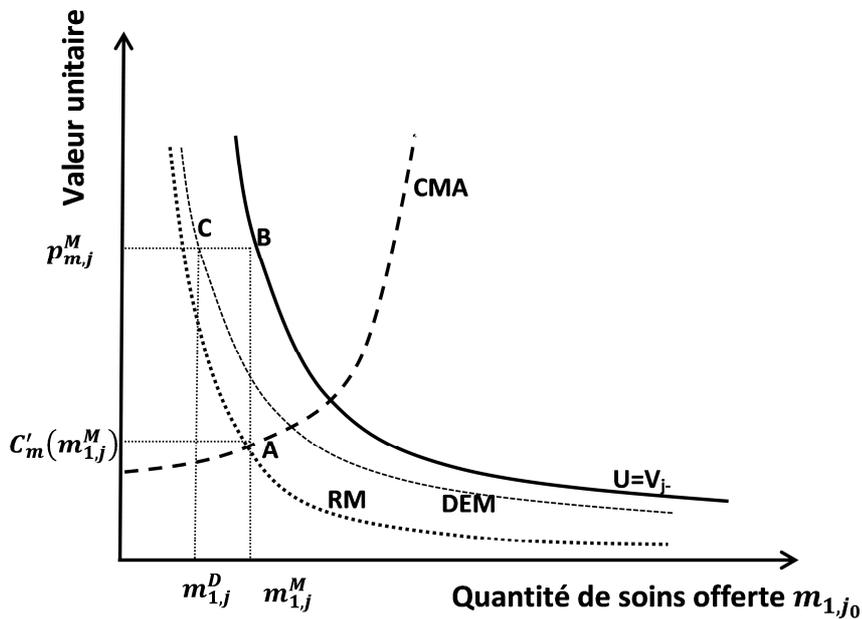
Équation 1.8-iii

$$\frac{\partial(p_{m,j} \cdot m_{1,j})}{\partial m_{1,j}} = \frac{\frac{\partial U_{1,j}}{\partial m} \Big|_{(W^n - p_{m,j} \cdot m_{1,j}, m_{1,j})}}{\frac{\partial U_{1,j}}{\partial x} \Big|_{(W^n - p_{m,j} \cdot m_{1,j}, m_{1,j})}} = C'_{m,j}(m_{1,j})$$

Autrement dit, le producteur fixe la quantité de soins $m_{1,j}^M$ qu'il va distribuer aux individus dans l'état 1 (c'est-à-dire les individus malades) en égalisant son revenu marginal et son coût marginal (situation repérée par le point A dans le graphique 1.8-i ci-dessous). Le prix des soins appliqué par le professionnel de santé est alors $p_{m,j}(m_{1,j}^M)$ (situation repérée par le point B dans le graphique 1.8-i ci-dessous).

Nous considérons enfin la courbe de demande de soins. Nous constatons donc qu'au prix $p_{m,j}(m_{1,j}^M)$ (situation repérée par le point C dans le graphique 1.8 i ci-dessous), les individus auraient consommé une quantité de soins $m_{1,j}^D$ inférieure à celle prescrite par le professionnel de santé.

Graphique 1.8-i : quantité et prix des soins consommés lorsque les offreurs sont en concurrence monopolistique : pas de régulation des prix



Grille de lecture du graphique :

- CMA : coût marginal de production des soins ;
- RM : revenu marginal du producteur ;
- DEM : courbe de demande du consommateur ;
- $U=V_j$: courbe des $(m_{1,j}, p_{m,j})$ tels que :

$$U_{1,j}(W^n - p_{m,j} \cdot m_{1,j}, m_{1,j}) = V_{1,j}$$

- $q_{m,j}^D$: coût des soins imposé par le professionnel de santé
- $m_{1,j}^M$: quantité de soins offerts par le professionnel de santé ;
- $m_{1,j}^D$: quantité de soins qui serait consommé par l'individu au prix $p_{m,j}^M$;

Cas où le prix de vente est fixé par une autorité régulatrice

Le cas qui vient d'être décrit est celui où le médecin a une liberté tarifaire. Supposons que cette liberté soit limitée par une autorité (Etat, ...) qui impose un tarif p_m^R suffisant pour faire un profit non nul, mais inférieur au prix $p_{m,j}^M$ que le producteur de soins aurait fixé s'il avait la liberté de tarification. Dans ce cas, le problème se réduit à un programme d'optimisation selon la seule variable de quantité de soins :

Programme 1.8-ii

$$\max_{m_{1,j}} p_m^R \cdot m_{1,j} - C_{m,j}(m_{1,j})$$

$$sc: U_{1,j}(W^n - p_m^R \cdot m_{1,j}, m_{1,j}) \geq V_{1,j}$$

Soient :

- $m_{1,j}^{Rd}$ la solution de l'équation :

$$p_m^R = C'_{m,j}(m_{1,j})$$

- $m_{1,j}^{Rv}$ la solution de l'équation :

$$U_{1,j}(W^n - p_m^R \cdot m_{1,j}, m_{1,j}) = V_{1,j}$$

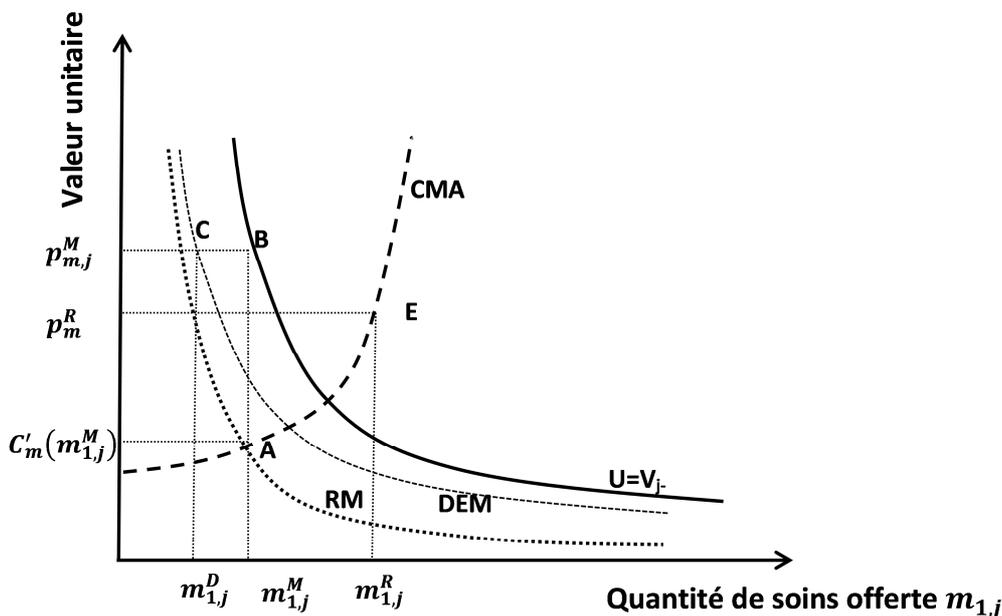
La solution $m_{1,j}^R$ du programme 1.8-ii est telle que :

Situation 1 :

$$m_{1,j}^R = m_{1,j}^{Rd} \text{ si } U_{1,j}(W^n - p_m^R \cdot m_{1,j}^{Rd}, m_{1,j}^{Rd}) \geq V_{1,j}$$

Cette situation est représentée par le point E dans le graphique 1.8-ii ci-dessous.

Graphique 1.8-ii : quantité et prix des soins consommés lorsque les offreurs sont en concurrence monopolistique : régulation des prix (situation 1)

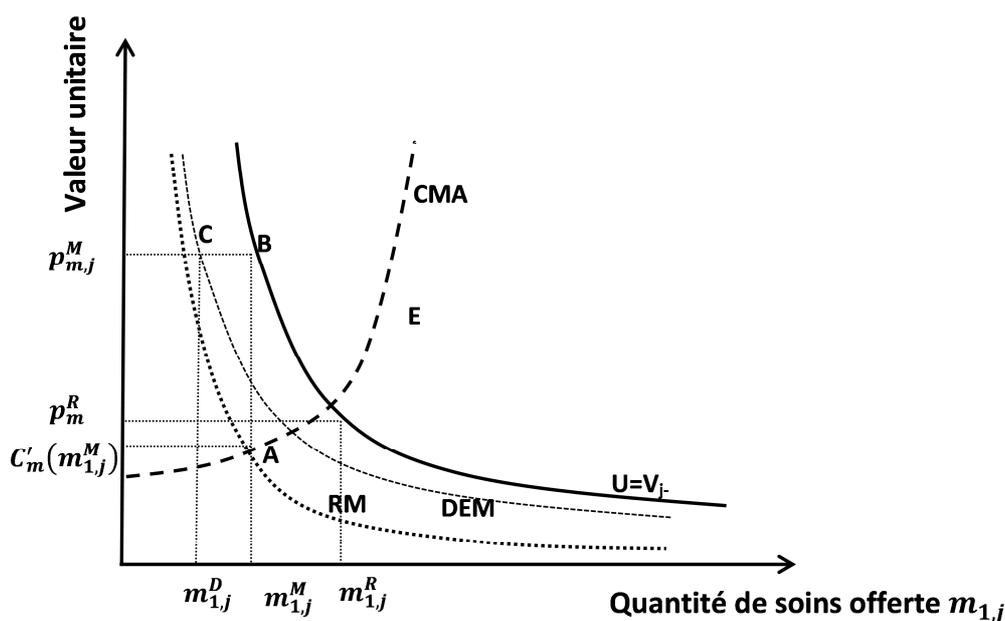


Situation 2 :

$$m^R_{1,j} = m^{Rv}_{1,j} \text{ si } U_{1,j}(W^n - p^R_m \cdot m^R_{1,j}, m^R_{1,j}) < V_{1,j}$$

Cette situation est représentée par le point E dans le graphique 1.8-iii ci-dessous.

Graphique 1.8-iii : quantité et prix des soins consommés lorsque les offreurs sont en concurrence monopolistique : régulation des prix (situation 2)



Dans les deux situations, la quantité de soins offerte est supérieure à ce que le professionnel de santé offrirait s'il était libre d'appliquer le tarif de son choix, et supérieure à la quantité de soins demandée par l'individu au prix p_m^R : le producteur augmente son revenu non plus par accroissement des prix, mais par augmentation de la quantité de soins administrés.

1.8.3. Effet de l'assurance sur la quantité et le prix des soins prescrits par les professionnels de santé dans le cadre d'une concurrence monopolistique

Examinons maintenant l'effet de l'assurance sur la demande de biens et services médicaux dans ce contexte de concurrence monopolistique. Supposons plus exactement que les individus bénéficient d'un contrat laissant à la charge des individus la somme c pour un euro dépensé.

Lorsque le producteur de soins dispose d'une liberté tarifaire, le programme de maximisation est analogue à celui étudié dans la partie 1.8.2 s'écrit :

Programme 1.8-iii

$$\max_{m_{1,j}, p_{m,j}} p_{m,j} \cdot m_{1,j} - C_{m,j}(m_{1,j})$$

$$sc: U_{1,j}(x_{1,j}, m_{1,j}) \geq V_{1,j-}(c)$$

$$sc: x_{j,1} + c \cdot p_{m,j} \cdot m_{j,1} = W^n - P_A$$

Où $V_{1,j-}(c)$ est l'utilité retirée auprès des autres professionnels de santé. $V_{1,j-}(c)$ est croissante lorsque c décroît.

Autrement dit :

$$\max_{m_{1,j}, p_{m,j}} p_{m,j} \cdot m_{1,j} - C_{m,j}(m_{1,j})$$

$$sc: U_{1,j}(W^n - P_A - c \cdot p_{m,j} \cdot m_{1,j}, m_{1,j}) \geq V_{1,j}(c)$$

Dans le cas du modèle à deux risque utilisés ici, $P_A = (1 - c) \cdot \pi_1 \cdot p_{m,j} \cdot m_{1,j}$. Le programme s'écrit finalement :

Programme 1.8-iv

$$\max_{m_{1,j}, p_{m,j}} p_{m,j} \cdot m_{1,j} - C_{m,j}(m_{1,j})$$

$$sc: U_{1,j}(W^n - P_A - c \cdot p_{m,j} \cdot m_{1,j}) \geq V_{1,j}(c)$$

$$sc: P_A = (1 - c) \cdot \pi_1 \cdot p_{m,j} \cdot m_{1,j}$$

La résolution de ce problème est effectuée comme dans la partie 1.8.2.

- La courbe de coût marginal est inchangée puisque l'assurance ne modifie pas le coût marginal auquel le producteur fait face.
- La courbe de revenu marginal du producteur, nommée maintenant $RM(c)$ est donnée par l'équation :

Équation 1.8-iv

$$\frac{\partial(p_{m,j} \cdot m_{1,j})}{\partial m_{1,j}} = \frac{1}{c} \cdot \frac{\frac{\partial U_{1,j}}{\partial m} \Big|_{(W^n - c \cdot p_{m,j} \cdot m_{1,j} - P_A, m_{1,j})}}{\frac{\partial U_{1,j}}{\partial x} \Big|_{(W^n - c \cdot p_{m,j} \cdot m_{1,j} - P_A, m_{1,j})}}$$

Cette courbe est décalée vers le haut par rapport au cas où l'individu ne possède pas d'assurance.

- La courbe de demande DEM(c) de soins est donnée par l'équation :

$$p_{m,j} = \frac{1}{c} \cdot \frac{\frac{\partial U_{1,j}}{\partial m} \Big|_{(W^n - c \cdot p_{m,j} \cdot m_{1,j} - P_A, m_{1,j})}}{\frac{\partial U_{1,j}}{\partial x} \Big|_{(W^n - c \cdot p_{m,j} \cdot m_{1,j} - P_A, m_{1,j})}}$$

Cette courbe est située au-dessus de celle de demande de soins en l'absence d'assurance.

- Soit « $U(c)=V_j(c)$ » la courbe qui donne l'ensemble des $(m_{1,j}, p_{m,j}(m_{1,j}))$ tels que :

$$U_{1,j}(W^n - c \cdot p_{m,j} \cdot m_{1,j} - P_A, m_{1,j}) = V_{1,j}(c)$$

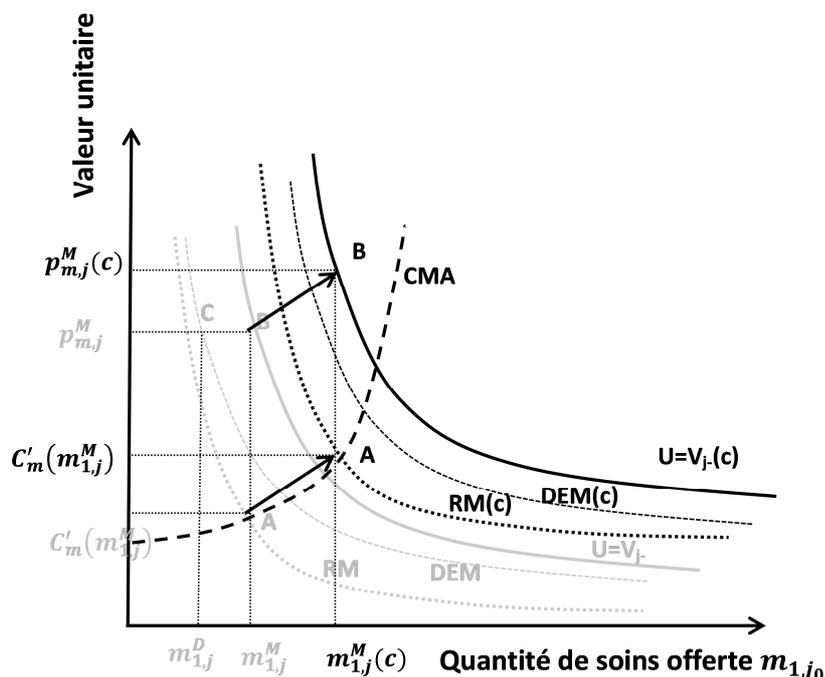
Cette courbe respecte l'équation suivante :

$$p_{m,j}(m_{1,j}) = \frac{1}{c} \cdot \frac{\frac{\partial U_{1,j}}{\partial m} \Big|_{(W^n - c \cdot p_{m,j} \cdot m_{1,j} - P_A, m_{1,j})}}{\frac{\partial U_{1,j}}{\partial x} \Big|_{(W^n - c \cdot p_{m,j} \cdot m_{1,j} - P_A, m_{1,j})}} - m_{1,j} \cdot \frac{\partial p_{m,j}}{\partial m_{1,j}}$$

Cette courbe est située au-dessus de celle de demande de soins et au-dessus de la courbe « $U=V_j$ ».

La représentation graphique des différentes courbes et leur comparaison avec celles obtenues en absence d'assurance (cf graphique 1.8 vi, ci-dessous) permet de constater qu'en présence d'assurance, les médecins offrent une quantité $m_{1,j}^M(c)$ plus importante de soins, à un prix $p_{m,j}^M(c)$ plus élevé.

Graphique 1.8-iv : quantité et prix des soins consommés lorsque les offreurs sont en concurrence monopolistique : équilibre de court terme



Grille de lecture du graphique :

- CMA : coût marginal de production des soins ;
- RM(c) : revenu marginal du producteur lorsque le niveau de coassurance est c ;
- DEM(c) : courbe de demande du consommateur lorsque le niveau de coassurance est c ;
- $U=V_j(c)$: courbe des $(m_{1,j}, p_{m,j})$ tels que :

$$U_{1,j}(W^n - c \cdot p_{m,j} \cdot m_{1,j}, m_{1,j}) = V_{1,j-}(c)$$

- $p_{m,j}^D(c)$: coût des soins imposé par le professionnel de santé lorsque le niveau de coassurance est c ;
- $m_{1,j}^M(c)$: quantité de soins offerts par le professionnel de santé lorsque le niveau de coassurance est c ;

Comment évolue la demande de soins lorsque le médecin ne dispose pas de liberté tarifaire, parce que l'état, l'assureur public ou privé impose un tarif opposable p_m^R . Dans ce cas, l'assureur augmente la quantité de soins au-delà de ce qu'il aurait prescrit s'il avait la possibilité de choisir son tarif et au-delà également de la quantité de soins demandée par l'individu.

Pour synthétiser l'ensemble de ces résultats, lorsque l'offre des professionnels de santé n'est pas parfaitement substituable d'un professionnel à l'autre il y a concurrence monopolistique. Dans le contexte de cette concurrence imparfaite, l'assurance conduit les professionnels de santé à accroître

leur prix et/ou la quantité de soins offerte aux individus. Lorsque le producteur de soins ne dispose pas de liberté tarifaire et que le prix des soins est fixé par l'Etat ou le tiers-payeur, le professionnel de santé augmente la quantité de soins offerte au-delà de la quantité demandée par l'individu : dans ce cas, la concurrence imparfaite amplifie donc le phénomène de risque moral.

Ainsi, dans un contexte où c'est le médecin qui exerce un pouvoir sur la quantité de soins prescrits et/ou le prix des soins, l'assurance entraîne une augmentation de la dépense en biens et services médicaux via une augmentation du prix et de la quantité des soins prescrits. Notons que notre analyse est très voisine de celle menée par McGuire (2000), la différence étant que dans notre cas, les courbes d'offre sont directement construites à partir l'utilité des consommateurs et non à partir de leur surplus. Cette approche permet de disposer d'un peu plus de détails sur les mécanismes conduisant les médecins à augmenter leur prix ou leur offre de soins. Notons enfin que nous n'avons pas considéré un paramètre essentiel sur lequel le professionnel de santé peut jouer qui est la qualité des soins. Nous nous intéresserons à l'effet de l'assurance sur la qualité des soins administrés dans la partie 1.8.4.

1.8.3.1. Effet de l'assurance sur la concurrence par les prix entre professionnels de santé

L'assurance joue un rôle ambigu sur le comportement des individus vis-à-vis des tarifs pratiqués par les professionnels de santé. Dranove et Satterthwaite (2000) signalent que les individus couverts peuvent être moins incités à changer de médecin, la diminution du coût des soins réduisant la propension des individus à rechercher un médecin au meilleur coût. Mais à contrario, le fait que les professionnels de santé soient imparfaitement substituables peut réduire l'effet que l'assurance sur le fait de substituer le professionnel de santé préféré par l'individu par un autre plus cher. En effet, le professionnel de santé dont les prix sont plus élevés peut certes présenter des avantages sur certains attributs (par exemple la qualité) mais qui ne compense pas d'autres avantages (tels que la proximité géographique, la connaissance du patient qu'a le professionnel de santé) que l'individu valorisait chez le médecin consulté habituellement.

1.8.4. Assurance santé et qualité des soins administrés

Feldstein (1973) s'intéresse à la manière dont les producteurs de soins adaptent la qualité et le prix des biens et services médicaux qu'ils offrent en fonction du niveau de couverture de la population et à la manière dont les individus réagissent, en termes de demande d'assurance à la modification du prix des soins. Son analyse porte plus exactement sur les soins hospitaliers offerts par les hôpitaux à but non lucratif. Feldstein suppose que les gestionnaires de ces hôpitaux arbitrent entre quantité de soins administrés et qualité de ces soins ; cela se traduit par une fonction d'utilité de la forme $U(BDD, QH)$ où BDD est le nombre de jours passés à l'hôpital et QH est la qualité des soins.

D'autre part, ces organismes à but non lucratif tarifient les journées d'hospitalisation au coût moyen par patient ; ce coût va être d'autant plus élevé que la qualité des soins offerts aux patient est importante.

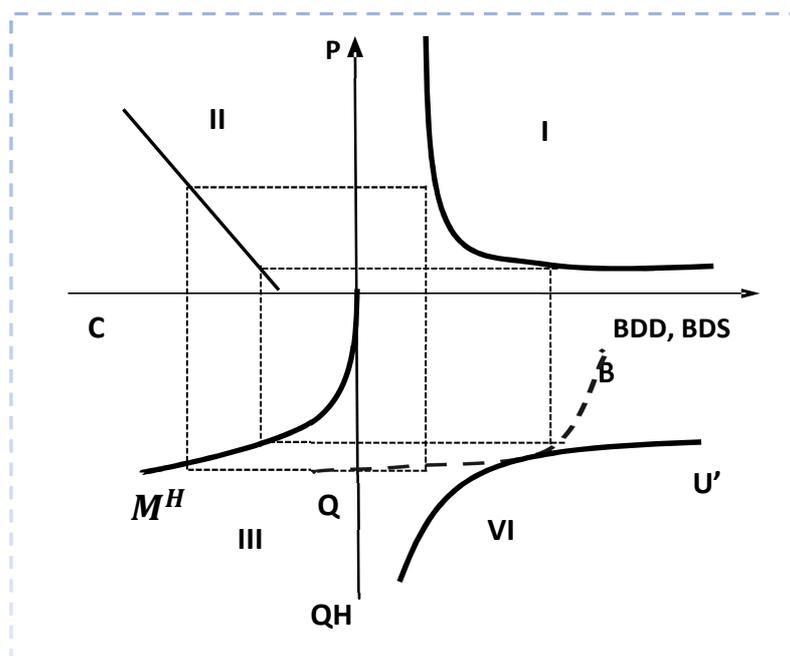
Les individus quant à eux peuvent être ou non assurés. Dans tous les cas, leur demande de journées d'hospitalisation décroît avec le coût des nominal des soins (c'est-à-dire avant remboursement des soins par l'assurance). Plus le prix des soins est important et plus la qualité des soins qu'ils reçoivent est élevée, mais moins ils demandent de journées d'hospitalisation.

Effet du niveau de couverture sur la qualité et la quantité de journées d'hospitalisation offertes

Le mécanisme décrivant la quantité et la qualité des soins offerts par l'hôpital est décrit dans le graphique 1.8-v ci-dessous. Le cadran I représente la demande de soins hospitaliers par les individus en fonction du prix nominal des soins, le cadran II la relation entre le prix nominal et le coût effectif des soins qui peuvent être alloués à chaque patient. Par rapport au prix nominal, le coût effectif tient également compte des subventionnements publics et privés. Ce coût effectif détermine la quantité d'input (équipement et personnel), donc la qualité des soins qui peut être dédiée à chaque patient : le cadran III lie le coût par patient à la qualité des soins obtenus. Le cadran IV lie qualité et quantité des soins au regard des relations décrites dans les autres cadrans : une hausse du prix des soins hospitaliers engendre une diminution de la demande de soins (cadran I) mais aussi un accroissement du coût pouvant être appliqué par patient (cadran II). Il s'en suit une augmentation de la qualité des soins. Dans le cadran IV, la courbe BQ en traits discontinus représente les différentes combinaisons possibles qualité – quantités de soins, compte tenu de la position de la courbe de demande de soins et du mécanisme qui vient d'être décrit. L'hôpital choisit parmi ces combinaisons possibles celle maximisant l'utilité de ses gestionnaires. Il s'agit du point de tangence entre la courbe BQ et U' , la représentation graphique de l'utilité $U(BDD, QH)$.

Supposons maintenant que le niveau de couverture de l'individu augmente : il s'assure ou accroît ses niveaux de garanties. Pour un prix nominal P donné, le prix à la charge de l'individu diminue, la quantité de soins qu'il consomme augmente donc. Cela se traduit par un déplacement vers le haut de la courbe de demande de soins (celle dans le cadran I). Il en découle que la courbe BQ des combinaisons qualité-quantité se déplace vers le bas du cadran IV ; chacune de ces deux quantités va augmenter. Sous l'hypothèse que qualité et quantité sont des biens normaux, le prix et la quantité de soins vont augmenter.

Graphique 1.8-v : dépendance entre le prix des soins et la qualité des soins



L'effet de l'assurance sur le prix des journées d'hospitalisation a déjà été modélisé et estimé dans l'étude de Feldstein (1971). Feldstein étudie l'effet de l'assurance sur le prix des soins. Il réalise une application empirique en travaillant sur des données agrégées au niveau des Etats américains. Les données vont de 1959 à 1965 et l'unité de temps considéré est l'année. Feldstein construit une mesure du niveau de couverture en calculant l'inverse du taux de coassurance moyen par Etat. Le taux de coassurance moyen est calculé sous la forme d'une moyenne pondérée du taux de coassurance chez les non-couverts et chez les couverts :

$$COINS_{it} = (1 - ENR_{it}) \times COINS_{it}^{NI} + ENR_{it} \times COINS_{it}^I$$

ENR_{it} est la proportion de personnes couvertes par un contrat d'assurance santé dans l'état i durant l'année t . $COINS_{it}^{NI} = 1$ puisqu'il s'agit du taux de coassurance chez les personnes sans assurance santé. Feldstein ne dispose pas des niveaux moyens de coassurance des personnes couvertes $COINS_{it}^I$ par Etat et par année. Il dispose par contre de ces niveaux de coassurance $COINS_t^I$ pour l'ensemble des USA, par année. Il suppose donc les niveaux de coassurance des personnes couvertes identiques d'un Etat à l'autre, égal à $COINS_t^I$. Le taux de coassurance moyen par état est donc :

$$COINS_{it} = (1 - ENR_{it}) + ENR_{it} \times COINS_t^I$$

Et le niveau de couverture est donné par :

$$QINS_{it} = [COINS_{it}]^{-1} = [(1 - ENR_{it}) + ENR_{it} \times COINS_t^I]^{-1}$$

L'auteur suppose ensuite que le prix des soins suit l'équation dynamique suivante :

$$pcare_{it} = \mu \cdot pcare_{it}^* + (1 - \mu) \cdot pcare_{it-1}$$

Où $pcare_{it}$ est le logarithme du prix des soins à la date t et $pcare_{it}^*$ décrit les valeurs d'équilibre de long terme prises par ce prix. L'équation présentée permet de décrire la dynamique autour des valeurs d'équilibre. En particulier, lorsque μ est compris entre 0 et 1, $pcare_{it}$ est la moyenne de $pcare_{it}^*$ et $pcare_{it-1}$ pondérés par $1 - \mu$ et μ . Nous constatons que cette dynamique conduit à ramener les variables de demande de soins en t vers leur valeur d'équilibre, lorsqu'elles s'en sont écartées en $t - 1$.

Feldstein suppose que $pcare_{it}^*$ est une fonction du niveau d'assurance

$$pcare_{it}^* = \varphi \cdot z_{it}$$

Effet du prix des soins sur la demande d'assurance

Feldstein étudie simultanément l'effet du prix des soins sur la demande d'assurance. Il suppose une fonction de demande de soins présente la forme suivante :

$$PDD_{it} = k \cdot [PCARE_{it} \cdot COINS_{it}]^{\eta_1} \cdot [INC_{it}]^{\eta_2} \cdot \prod_{j>2} [X_{it}^j]^{\eta_j} = k \cdot \left[\frac{PCARE_{it}}{QINS_{it}} \right]^{\eta_1} \cdot [INC_{it}]^{\eta_2} \cdot \prod_{j>2} [X_{it}^j]^{\eta_j}$$

$\{X_{it}^j | j > 2\}$ correspond à un ensemble de facteurs qui déplacent la courbe de demande de journées d'hospitalisation, η_j à l'élasticité de la demande de soins par rapport à chaque facteur η_j et $USEX_{it} = k \cdot \prod_{j>2} [X_{it}^j]^{\eta_j}$ l'effet de l'ensemble de ces facteurs sur cette courbe. Cette variable s'interprète comme la localisation de la courbe de demande de soins.

L'ensemble des effets des facteurs $USEX_{it}$ sont susceptibles d'influencer la demande d'assurance en faisant varier la demande de soins PDD_{it} . Puisque les effets de ces facteurs sur la demande de soins passent par $USEX_{it}$, Feldstein ne cherche uniquement les valeurs de $USEX_{it}$. Il n'estime donc pas les η_j , il ne définit même pas de manière précise l'ensemble des variables $\{X_{it}^j | j > 2\}$. Il extrait les valeurs de $USEX_{it}$ à partir de la formule :

$$USEX_{it} = PDD_{it} \cdot \left[\frac{PCARE_{it}}{QINS_{it}} \right]^{-\eta_1} \cdot [INC_{it}]^{-\eta_2}$$

en prenant pour élasticité prix $\eta_1 = -0,67$ et pour élasticité revenu, $\eta_2 = 0,29$, des valeurs établies dans la précédente étude réalisée par Feldstein (1971). Le calcul de $USEX_{it}$ ne requière ainsi que la connaissance des variables PDD_{it} , $PCARE_{it}$ et $QINS_{it}$.

Par ailleurs, Feldstein suppose l'équation dynamique suivante pour la demande d'assurance :

$$\frac{ENR_{it}}{ENR_{it-1}} = \left(\frac{ENR_{it}^*}{ENR_{it-1}} \right)^\lambda$$

$$\frac{QINS_{it}}{QINS_{it-1}} = \left(\frac{QINS_{it}^*}{QINS_{it-1}} \right)^\mu$$

ENR_{it}^* et $QINS_{it}^*$ décrivent les valeurs d'équilibre de long terme prises par la proportion de personnes couvertes et la quantité d'assurance. Passées en logarithme, les équations présentées ci-dessus s'écrivent :

$$enr_{it} = \delta \cdot enr_{it}^* + (1 - \delta) \cdot enr_{it-1}$$

$$qins_{it} = \lambda \cdot qins_{it}^* + (1 - \lambda) \cdot qins_{it-1}$$

Comme pour le prix des soins, les équations présentées permettent de décrire la dynamique autour des valeurs d'équilibre : enr_{it} et $qins_{it}$ sont les moyennes de enr_{it}^* pondéré par $1 - \delta$ et δ et enr_{it-1} et $qins_{it}^*$ et $qins_{it-1}$ pondérés par $1 - \lambda$ et λ .

Les variables enr_{it}^* et $qins_{it}^*$ sont fonctions du logarithme des variables influençant la demande d'assurance, décrites ci-dessus : $pins_{it}$ le prix de l'assurance, $usex_{it}$ la variation de la demande de soins résultant des facteurs autres que le prix, inc_{it} le revenu, $pcare_{it}$ le prix des soins et $group_{it}$, la proportion de personnes travaillant dans l'industrie et les services gouvernementaux. Cette dernière variable est un proxy de la proportion d'individus ayant accès à une assurance de groupe. Cette catégorie d'assurance tend à être moins chère qu'une assurance individuelle du fait d'économies d'échelle et du pouvoir de négociation des entreprises. Elle bénéficie de plus d'un régime fiscal favorable. L'ensemble de ces éléments sont susceptibles d'accroître la demande d'assurance des individus.

Le modèle dynamique s'écrit donc :

$$enr_{it} = \delta \cdot \beta_0 + \delta \cdot \beta_1 \cdot pins_{it} + \delta \cdot \beta_2 \cdot pcare_{it} + \delta \cdot \beta_3 \cdot inc_{it} + \delta \cdot \beta_4 \cdot usex_{it} + \delta \cdot \beta_5 \cdot group_{it} + (1 - \delta) \cdot enr_{it-1}$$

$$qins_{it} = \lambda \cdot \gamma_0 + \lambda \cdot \gamma_1 \cdot pins_{it} + \lambda \cdot \gamma_2 \cdot pcare_{it} + \lambda \cdot \gamma_3 \cdot inc_{it} + \lambda \cdot \gamma_4 \cdot usex_{it} + \lambda \cdot \gamma_5 \cdot group_{it} + (1 - \lambda) \cdot qins_{it-1}$$

Le prix de l'assurance $p_{ins_{it}}$, des soins $p_{care_{it}}$ et la localisation de la courbe de demande de soins $u_{sex_{it}}$ étant endogènes, Feldstein utilise une procédure d'estimations par variables instrumentales, les instruments choisis étant les variables exogènes ($group_{it}$ et inc_{it}) aux dates t et $t - 1$.

Les estimations font apparaître un effet très significatif du prix des soins $p_{care_{it}}$ sur la proportion de personnes couvertes enr_{it} et sur la quantité d'assurance $q_{ins_{it}}$. La variable $u_{sex_{it}}$ exprimant la localisation de la demande de soins a un effet également significatif, très proche de celui de $p_{care_{it}}$: ainsi, la demande d'assurance semble être sensible à la dépense de soins. L'effet du revenu n'est par contre pas significatif ce qui peut refléter des influences contraires de ce facteur (un plus haut revenu diminue l'aversion au risque mais accroît la demande de soins donc le risque financier ; il permet de plus de bénéficier de d'avantages d'exonération de charges fiscales sur les contrats d'entreprise).

Demande d'assurance et demande de soins à l'équilibre

Par le biais de l'équation de demande d'assurance en fonction du prix des soins et de l'équation de la demande de soins en fonction du niveau d'assurance, Feldstein d'un système dynamique à deux équations qui peuvent se résumer ainsi :

$$q_{ins_{it}} = \lambda \cdot \epsilon_1 \cdot x_{it} + \lambda \cdot \epsilon_2 \cdot p_{care_{it}} + (1 - \lambda) \cdot q_{ins_{it-1}}$$

$$p_{care_{it}} = \mu \cdot \epsilon_3 \cdot z_{it} + \mu \cdot \epsilon_4 \cdot q_{ins_{it}} + (1 - \mu) \cdot p_{care_{it-1}}$$

Feldstein étudie la stabilité de ce système ; celle-ci est assurée si et seulement si :

$$(1 - \mu \cdot \lambda) \cdot (1 - \mu \cdot \lambda \cdot \epsilon_2 \cdot \epsilon_4) < 1$$

L'auteur montre que compte tenu des valeurs issues de l'estimation, le système est stable. La valeur d'équilibre du prix s'écrit :

$$p_{care} = \frac{\epsilon_3}{1 - \epsilon_2 \cdot \epsilon_4} z + \frac{\epsilon_1 \cdot \epsilon_3}{1 - \epsilon_2 \cdot \epsilon_4} x$$

La valeur $1 - \epsilon_2 \cdot \epsilon_4$ est comprise entre 0 et 1 ce qui signifie que l'effet de tout facteur z ayant une influence sur le prix des soins voit son effet amplifié du fait de l'assurance : la hausse du prix des soins accroît le niveau d'assurance, effet qui lui-même augmente le prix des soins. Tout facteur x accroissant le niveau d'assurance entraîne également une hausse du prix des soins.

Calcul de la perte de surplus

L'individu est initialement non couvert. Soit P_0 , le prix qu'il paie par journée d'hospitalisation et Q_0 le nombre de journées d'hospitalisation qu'il demande. Supposons maintenant que l'individu reçoive une assurance qui laisse à la charge de l'individu une fraction c du coût de la journée d'hospitalisation. Comme le montre la sous-partie précédente, cette hausse du niveau d'assurance entraîne un accroissement du coût par journée d'hospitalisation, de P_0 à P_1 . Cette hausse est associée à une meilleure qualité des soins, ce qui se traduit graphiquement par un déplacement vers la droite de la courbe DD de demande de soins (elle devient D'D').

Le graphique 1.8-vi ci-dessous permet de constater que par rapport au calcul traditionnel de la perte de surplus, deux effets contraires surviennent :

- D'une part, la hausse du prix des soins, de P_0 à P_1 , accroît plus fortement le niveau de prime.
- D'autre part, le déplacement de la courbe de demande de soins, liée à une amélioration de la qualité des soins, engendre une valorisation plus forte des soins consommés.

Le prix de la journée d'hospitalisation à la charge de l'individu passe de P_0 à $c \cdot P_1$. Le niveau de consommations hospitalières augmente, passant de Q_0 à Q_1 . Le surcoût total à payer, sous forme de prime et de reste à charge, est égal à :

$$Q_0 \cdot (P_1 - P_0) + P_1 \cdot (Q_1 - Q_0)$$

La valeur du surcroît de consommation est la somme des surfaces des trois zones grisées dans le graphique 1.8-vi ci-dessous. Elle s'écrit donc³⁸ :

$$(Q_1 - Q_0) \cdot c \cdot P_1 + \frac{1}{2} \cdot (P_2 - c \cdot P_1) \cdot (Q_1 - Q_0) + (P_2 - P_0) \cdot Q_0$$

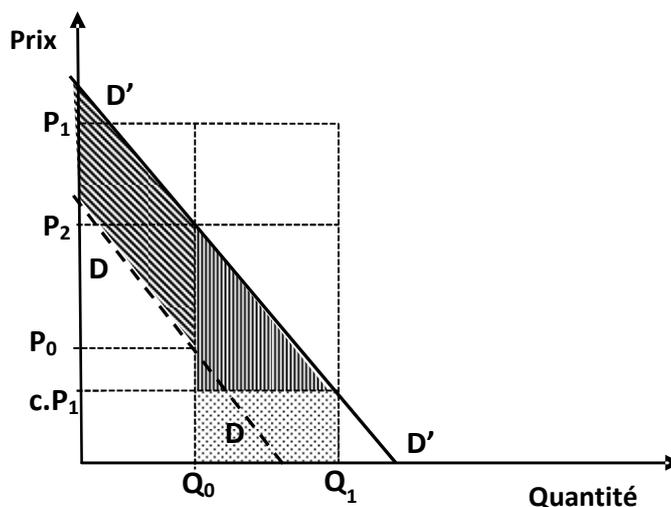
Où P_2 serait le coût des soins avec la courbe de demande D'D' si l'individu avait conservé le niveau de consommation Q_0 .

La perte de surplus est calculée dans un premier temps sans tenir compte du gain en bien être liée à la réduction du risque financier ; elle vaut donc :

$$\Delta WELF = (P_1 - P_2) \cdot Q_0 + (Q_1 - Q_0) \cdot \left[(1 - c) \cdot P_1 - \frac{1}{2} \cdot (P_2 - c \cdot P_1) \right]$$

³⁸ La surface du parallélogramme hachuré obliquement s'écrit comme la base (c'est-à-dire $P_2 - P_0$) multipliée par la hauteur (c'est-à-dire Q_0).

Graphique 1.8-vi : Effet de l'assurance sur la demande de journées d'hospitalisation (avec modification de la qualité des soins)



La perte de surplus peut se réécrire de manière plus opérationnelle :

$$\Delta WELF = \left\{ (1 - \alpha) \cdot (\Pi - 1) + \frac{|\eta|}{2} \cdot [\Pi \cdot (1 - c) + (1 - \alpha) \cdot (\Pi - 1)] + \left[\frac{1 + \alpha \cdot (\Pi - 1) - c \cdot \Pi}{1 + \alpha \cdot (\Pi - 1)} \right] \right\} P_0 Q_0$$

Où $\Pi = P_1/P_0$, l'accroissement du pris des journées d'hospitalisation liée à l'assurance, $\alpha = (P_2 - P_0)/(P_1 - P_0)$ mesure le changement de qualité des soins perçue par le patient ($P_2 - P_0$ représente le changement de disposition du patient, du fait de l'accroissement de la qualité des soins et $P_1 - P_0$ le coût total associé à cet accroissement de la qualité des soins. Feldstein précise que Π est lié à c par une relation du type $\Pi = c^{-\omega}$.

Ne disposant pas par ses données de la dépense $P_0 Q_0$ en l'absence d'assurance, Feldstein l'établit en fonction de la dépense $P_a Q_a$ engendrée par le niveau de couverture effectif, celui dont disposent les individus à la fin des années 60 (donc au moment de l'étude) :

$$P_a Q_a = \Pi_a \cdot \left\{ (1 - \eta) + \frac{c_a \cdot \Pi_a \cdot \eta}{1 + \alpha \cdot (\Pi_a - 1)} \right\} \cdot P_0 Q_0$$

Pour estimer comment évolue la perte de surplus lorsque le niveau d'assurance augmente, Feldstein effectue les calculs pour trois taux de coassurance : 0,33 (le taux de coassurance moyen au moment de l'étude), 0,5 et 0,67. En ce qui concerne les autres paramètres, Feldstein considère plusieurs valeurs possibles :

- Pour $|\eta|$, 0,4 et 0,8, ce qui correspond à des élasticités-prix autour -0,37, la valeur estimée par Feldstein (1971).
- Pour α , 0, 0,33 et 0,67, ce qui correspond à un large éventail de valeurs, en l'absence d'informations précises sur ce paramètre.
- Pour ω , les valeurs 1 et 0,5.
- Enfin, $P_a Q_a$, la dépense hospitalière en 1969 vaut 12,6 milliards de dollars.

L'application de chacune de ces combinaisons de paramètres permet de disposer de différents niveaux de pertes de surplus possibles pour chaque niveau de coassurance. Feldstein ainsi estime qu'augmenter le taux de coassurance de 0,33 à 0,5 permet un gain de 1,2 à 3,4 milliards de dollars et augmenter le taux de coassurance de 0,33 à 0,5 permet un gain de 1,9 à 4,8 milliards de dollars.

Feldstein tiens ensuite compte du gain en bien-être lié à la réduction du risque financier. Pour cela, il établit pour tout contrat i de prime P_i^j et de niveau de coassurance c_i la surprime maximum Λ_i que l'individu serait prêt à payer pour obtenir une sur-assurance le débarrassant totalement de son risque :

$$u(W - \Lambda_i - P_i^j) = E[u(W - c_i \cdot \tilde{M}_i - P_i^j)]$$

Où \tilde{M}_i est la dépense liée aux journées d'hospitalisation. L'auteur suppose une fonction d'utilité de la forme $u(Y) = A \cdot (1 - \exp(-R \cdot Y))$, ce qui permet d'exprimer la surprime Λ_i de manière simple :

$$\exp(-R \cdot \Lambda_i) = E(\exp(R \cdot c_i \cdot \tilde{M}_i))$$

La valeur du passage d'un taux de coassurance c_i à un taux de coassurance nul est mesurée par la prime de risque $\rho_i = \Lambda_i - c_i \cdot E(\tilde{M}_i)$.

V_{ij} , la perte de valeur liée à une augmentation du taux de coassurance de c_i à c_j s'exprime comme la différence entre la prime de risque du contrat i et celle du contrat j .

$$\exp(-RV_{ij}) = \exp(-R \cdot (\rho_j - \rho_i)) = \frac{E(\exp(R \cdot c_j \cdot \tilde{M}_j)) \cdot \exp(R \cdot c_i \cdot E(\tilde{M}_i))}{E(\exp(R \cdot c_i \cdot \tilde{M}_i)) \cdot \exp(R \cdot c_j \cdot E(\tilde{M}_j))}$$

$\tilde{M}_i = P_i \cdot \tilde{N}_i$ où \tilde{N}_i est le nombre totale de journées d'hospitalisation par famille. Feldstein suppose que \tilde{N}_i est le mélange de deux lois : une loi de Poisson de paramètre κ_i décrivant le nombre de séjours et une loi gamma de paramètre (ν, ξ_i) décrivant la durée des séjours³⁹. Ainsi :

$$E\left(\exp(R \cdot c_i \cdot \tilde{M}_i)\right) = E\left(\exp(R \cdot c_i \cdot P_i \cdot \tilde{N}_i)\right) = \exp(\kappa_i \cdot [(1 - R \cdot P_i \cdot c_i \cdot \xi_i)^{-\nu} - 1])$$

$$E(\tilde{M}_i) = P_i \cdot \nu \cdot \xi_i \cdot \kappa_i$$

On en déduit que :

$$V_{ij} = \left\{ \kappa_j \cdot [(1 - R \cdot P_j \cdot c_j \cdot \xi_j)^{-\nu} - 1] - \kappa_i \cdot [(1 - R \cdot P_i \cdot c_i \cdot \xi_i)^{-\nu} - 1] \right\} - P_j \cdot \nu \cdot \xi_j \cdot \kappa_j + P_i \cdot \nu \cdot \xi_i \cdot \kappa_i$$

Par ailleurs, la relation entre \tilde{N}_j et \tilde{N}_i (le nombre de séjours moyen en l'absence d'assurance) est donné par la formule :

$$E(\tilde{N}_j) = E(\tilde{N}_i) \cdot \frac{\left\{ 1 - \frac{\eta}{c_j} [\alpha + (1 - \alpha) \cdot c_j^\omega - c_j] \right\}}{\left\{ 1 - \frac{\eta}{c_i} [\alpha + (1 - \alpha) \cdot c_i^\omega - c_i] \right\}}$$

Cette formule permet d'établir le nombre moyen de jours d'hospitalisation (donc $\nu \cdot \xi_j \cdot \kappa_j$) pour le contrat j .

La référence i considérée est la situation en termes de couverture des individus et de consommations de soins à la fin des années 60. Pour cette situation de référence, Feldstein dispose des paramètres de la distribution des nombres moyens de séjour et de la durée des séjours, ce qui lui permet de déterminer ν , ξ_i , κ_i puis ξ_j , κ_j . Enfin, plusieurs valeurs « vraisemblables » de l'aversion au risque sont testées : $R = 0,0001$, $R = 0,0003$ et $R = 0,0007$.

Les calculs de pertes de surplus tenant compte du gain en bien être liée à la réduction du risque financier dépassent, pour la plupart des combinaisons possibles de paramètres, les deux milliards de dollars et vont jusqu'à 9 milliards de dollars. Feldstein estime que pour les valeurs de paramètres qui lui semblent le plus vraisemblable, la diminution du niveau d'assurance occasionnerait des gains importants.

Feldman et Dowd (1991) reprennent l'idée de Feldstein selon laquelle l'assurance modifie le coût des soins achetés : une réduction du niveau d'assurance à grande échelle modifierait la demande de soins et donc le coût des soins ; d'autre part comme l'a décrit Feldstein, une augmentation du niveau

³⁹ Ainsi l'espérance du nombre de séjours est κ_i , l'espérance de la durée des séjours est $\nu \cdot \xi_i$ et la variance de la durée de séjour est $\nu \cdot \xi_i^2$.

d'assurance incite les producteurs de soins à proposer des biens et services médicaux de meilleure qualité. Feldman et Dowd critiquent ainsi directement l'approche de Manning *et al* (1987) qui, dans leur calcul de perte de surplus, considèrent les prix comme constants. Néanmoins, considérant les estimations menées depuis 1973 sur l'élasticité prix de la demande de soins et sur l'aversion absolue pour le risque, ils sont amenés à revoir les hypothèses de Feldstein concernant ces paramètres. Feldman et Dowd soulignent d'une part que les élasticités prix retenues par Feldstein (de -0,4 à -0,8) sont beaucoup trop importantes au regard de celles fournies par la RAND (de -0,1 à -0,2), d'autre part que les indices d'aversion pour le risque qu'il utilise sont beaucoup trop faibles. Le but de leur article est donc de fournir une nouvelle estimation de la perte de bien être lié au risque moral, en tenant compte d'une part de l'effet de l'assurance sur le prix des soins et d'autre part des nouveaux paramètres d'élasticité prix et d'aversion au risque : ils calculent la perte de surplus liée au fait d'être complètement couvert plutôt que d'être couvert par le contrat imposant un taux de coassurance de 95% dans l'expérience de la RAND. Rappelons que ce contrat comprend un *stop-loss* au-delà de 1000€ de dépense à la charge du patient, de sorte qu'en moyenne, les bénéficiaires de ce contrat font face à un taux de coassurance de 31%. Les auteurs effectuent leur analyse en considérant les deux cas de figure : un passage d'un taux de coassurance nul à un taux de coassurance de 31% et d'un taux de coassurance nul à un taux de coassurance de 95%.

Leur modèle économique utilise une fonction d'utilité séparable constituée de la somme de u_x , l'utilité accordée aux consommations non médicales et de u_m , l'utilité retirée des soins. Ils supposent de surcroît que les individus bénéficient d'une couverture assurantielle avec taux de coassurance c constant. Afin de déterminer la valeur retirée de la réduction du risque financier, ils déterminent la prime de risque, ici la somme maximum que l'individu est prêt à payer en plus de l'espérance de ses remboursements pour bénéficier d'une couverture complète plutôt que du contrat de taux de coassurance c^{40} ($c = 0,31$ ou $0,95$). Ainsi, cette prime ρ_{10} est telle que :

$$\begin{aligned} E[u_x(W - c \cdot p_m^0 \cdot m^0 - (1 - c) \cdot E(p_m^0 \cdot m^0))] + E[u_m(m^0)] \\ = E[u_x(W - \rho_{10} - E(p_m^1 \cdot m^1))] + E[u_m(m^1)] \end{aligned}$$

Où m^0 , p_m^0 et m^1 , p_m^1 sont les quantités et les prix des soins consommés, lorsque le taux de coassurance est respectivement égal à 0 et à c .

⁴⁰ La présentation de la modélisation économique diffère un peu de celle de Feldman et Dowd puisque eux considéreraient la prime de risque liée au fait d'être couvert par un contrat de niveau de coassurance c plutôt que de ne pas être couvert. Néanmoins, comme leur étude porte sur une comparaison entre le fait d'être complètement couvert plutôt que de faire face à un taux de coassurance c , il semble plus claire de calculer la prime de risque liée à cette situation.

Une approximation de la prime de risque est obtenue en effectuant un développement limité de l'égalité précédente :

$$\rho_{10} = [E(p_m^0 \cdot m^0) - E(p_m^1 \cdot m^1)] + \frac{E[u_m(m^1) - u_m(m^0)]}{u'_x} + \frac{R^A}{2} \cdot \sigma^2$$

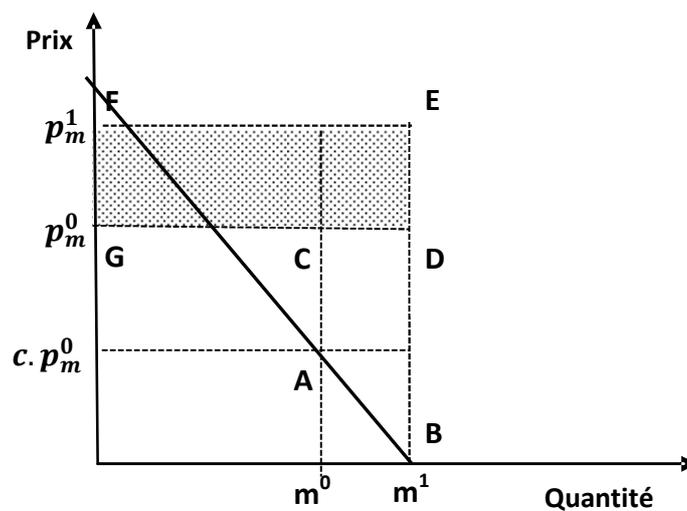
Où σ^2 est la variance des restes à charge $c \cdot p_m^0 \cdot m^0$. R^A est l'indice d'aversion absolue pour le risque :

$$R^A = -\frac{u''_x}{u'_x}$$

Le premier terme de la partie droite de l'égalité précédente représente le coût (sous forme de hausse de la prime et du reste à charge) lié à l'augmentation des consommations de soins et du prix des soins. Le deuxième terme représente la valeur que l'individu attribue au supplément de soins consommés. La somme de ces deux termes est donc la perte de surplus liée au risque moral. Elle peut se réécrire :

$$\underbrace{[E(p_m^0 \cdot m^1) - E(p_m^1 \cdot m^1)]}_{\text{Perte de surplus liée à l'augmentation du prix des soins (DEFG)}} + \underbrace{[E(p_m^0 \cdot m^0) - E(p_m^0 \cdot m^1)] + \frac{E[u_m(m^1) - u_m(m^0)]}{u'_x}}_{\text{Perte de surplus liée à l'accroissement des consommations de soins, à prix constant (CABD)}}$$

Graphique 1.8-vii : Perte de surplus dans le cas où l'assurance modifie le prix des soins



La perte de surplus liée à l'accroissement des consommations médicales, à prix constant, est représentée par la surface ABCD dans le graphique 1.8-vii. Feldman et Dowd supposent que la courbe de demande de soins est approximativement linéaire, la perte de surplus à prix constant, s'écrit donc :

$$\frac{1}{2}(p_m^0 + p_m^0 - c \cdot p_m^0) \cdot (m^1 - m^0) = p_m^0 \cdot \left(1 - \frac{c}{2}\right) \cdot (m^1 - m^0)$$

La perte de surplus liée à l'augmentation du prix des soins est représentée par la surface DEFG. Elle vaut donc :

$$(p_m^1 - p_m^0) \cdot m^1$$

Ne disposant pas d'élasticité du prix des soins au niveau de coassurance, ils effectuent leurs calculs en supposant plusieurs taux de variation : une augmentation de 0%, de 10%, de 20%. Par ailleurs, ils s'appuient sur des résultats issus de l'expérience de la RAND pour tracer la courbe de demande de soins.

Le dernier terme à estimer est :

$$\frac{R^A}{2} \cdot \sigma^2$$

Il s'agit du gain en bien être lié à la réduction du risque financier. Feldman et Dowd prennent comme indice d'aversion absolue pour le risque, la valeur de 0,0036, issue des travaux de Marquis et Holmer (1986). Quant à la variance, elle est estimée en prenant la variance des dépenses parmi les individus sur la couverture laissant un niveau de coassurance de 95%, soit 25 828.

Feldman et Dowd aboutissent aux pertes suivantes (voir tableau 1.8-i ci-dessous). Celles-ci s'évaluent de 33,4 à 109,3 milliards de dollars, suivant les hypothèses en termes de niveau de coassurance et de variation du prix des soins. Les variations de prix sont fictives puisque Feldman et Dowd ne disposent pas d'une estimation du niveau d'assurance sur du ces prix. Les résultats ont donc plutôt un but illustratif. Quoiqu'il en soit, l'étude conduit à conclure que la perte de bien-être liée au risque moral est importante.

Tableau 1.8-i : Perte de bien-être lié à une couverture complète (en milliards de dollars 1984)

Hypothèses sur variation du prix des soins	Perte de bien être lorsque l'individu est bénéficiaire d'une couverture complète plutôt que par un :	
	Contrat avec taux de coassurance de 31%	Contrat avec taux de coassurance de 95%
0%	61	33,4
10%	85	57,4
20%	109,3	81,7

Par rapport aux estimations de Manning (1987), l'étude de Feldman et Dowd a le mérite de tenir compte de l'aversion pour le risque, dans le calcul de la perte de bien-être collectif. Par rapport aux travaux de Feldstein (1973) elle utilise pour l'élasticité-prix et l'aversion au risque, des paramètres estimés et non plus des valeurs fictives « plausibles ». Par contre, leur méthode de calcul de perte de surplus peut-être critiquée dans la mesure où ils ne tiennent pas compte du supplément de bien-être que peut apporter aux individus l'achat de soins plus chers. Or, Feldstein a souligné et pris en compte dans les producteurs de soins réagissent à des niveaux d'assurance plus élevés en améliorant la qualité des soins. Nous avons également suggéré qu'un individu pouvait avoir recours à un professionnel de santé ou à un établissement hospitalier plus cher pour éviter certains coûts d'opportunité (coûts directs et indirects liés aux transports, coût d'attente. Ces éléments sont valorisés par les individus et doivent donc être prise en compte dans le calcul de la perte de bien être des individus. D'autre part, les travaux de Keeler *et al* (1988), à partir des données de la RAND ont montré qu'à un niveau individuel, l'assurance n'avait qu'un faible effet sur le coût des épisodes de soins. Il semble donc que la variation de bien-être liée à l'effet de l'assurance sur le coût des soins ne représente qu'un enjeu mineur, ce qui remet en cause l'approche de Feldstein et celle de Feldman et Dowd. Néanmoins, les données de la RAND sont à un niveau micro : de par son niveau d'assurance individuel, un patient à lui seul est peu susceptible d'influencer le coût d'une grande partie de ses soins (notamment les prescriptions). Par contre, une variation du niveau d'assurance au niveau d'une zone géographique assez large, une variation du niveau de couverture peut induire une adaptation de la qualité et des tarifs de la part des producteurs de soins, à l'instar du mécanisme décrit par Feldstein.

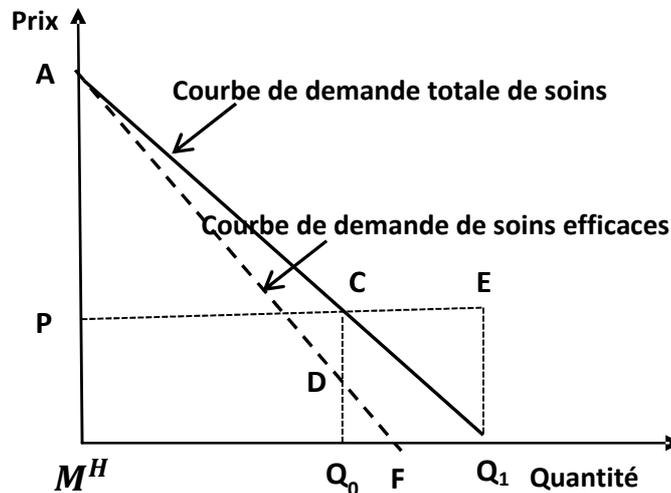
1.8.5. Assurance et demande induite

La critique de Rice (1992) est nettement plus radicale car elle remet en cause la courbe liant prix et demande de soins comme instrument de mesure de la valeur de soins consommés. Cette critique repose sur l'hypothèse que les individus ont une mauvaise connaissance de l'efficacité des soins. Pour cette raison, la décision sur la nature et la quantité de soins à consommer des soins incombe en partie au médecin qui détient cette connaissance. En principe, s'ils sont bien informés, les patients devraient valoriser les soins en fonction de leur efficacité médicale : ainsi les patients dont le niveau de couverture est faible devraient, par rapport à ceux dont le niveau de couverture est élevé, renoncer en priorité aux soins ayant la plus faible efficacité médicale. Or, les travaux de Lohr *et al* (1986) ont montré aux moyens des données de la RAND que les patients les moins couverts renoncent à peu près autant à des soins très efficace qu'à des soins peu efficaces. Selon Rice, ces résultats démontrent que les individus ont une mauvaise connaissance de l'efficacité des biens et services médicaux et que leur courbe de demande ne peut servir à mesurer la « vraie » valeur du surplus de soins consommés. La vraie perte correspond aux consommations de soins médicalement inutiles à dire d'expert. A partir des travaux de Brook et Vaiana (1989), Rice considère qu'environ un quart de la dépense médicale totale (soit 150 milliards de dollars en 1990) est consacrée à des soins inutiles. La responsabilité de cette perte incombe en grande partie aux praticiens, du fait de prescriptions erronées ou de phénomène de demande induite. Les patients quant à eux ne sont pas en mesure de sélectionner parmi les soins qui leur sont prescrits ceux qui sont efficaces et ceux qui ne le sont pas. Au regard de ce constat, Rice juge inadaptés les dispositifs de coassurance visant à responsabiliser le patient ; il préconise plutôt des mesures en direction des producteurs de soins : édition de guides de bonnes pratique, paiement prospectif des hôpitaux, développement des HMO (Health Maintenance Organization) qui, en salariant les praticiens, rapprochent leurs intérêts de ceux de la compagnie, à avoir une offre optimale de soins.

Cette approche a été discutée par Feldman et Dowd (1993). Ils articulent leurs critiques autour de deux points : d'une part, l'ampleur de la perte de surplus à lorsque sont pris en compte les soins inutiles ; d'autre part, la validité des résultats de Lohr *et al* (1986). Concernant le premier point, Feldman et Dowd font un compromis entre la théorie traditionnelle du consommateur et les réserves établies par Rice : ils considèrent la courbe de demande totale de soins et la courbe de demande de soins efficaces. La seconde est bien évidemment située en dessous de la première. Suivant l'approche traditionnelle, Feldman et Dowd considèrent que le triangle au-dessus de la courbe de demande totale représente intégralement une perte de surplus (perte CEQ_1 dans le graphique 1.8-vii) : il reste ainsi le triangle CQ_0Q_1 ; suivant l'approche de Rice, ils retranchent de CQ_0Q_1 , l'ensemble

des soins non efficaces, c'est-à-dire le quadrilatère DCQ₁F. Au final, il ne reste que le triangle DQ₀F, soit moins que le triangle CQ₀Q₁.

Graphique 1.8-viii : la demande de soins (basée sur les travaux de Lohr *et al* (1986))



Ainsi, la perte de surplus à laquelle on aboutit serait bien plus considérable que celle obtenue par l'approche traditionnelle par la théorie du consommateur. Concernant les résultats de Lohr *et al* (1986), Feldman et Dowd mettent en doute leur validité dans la mesure où l'efficacité des soins n'est jugée que lorsqu'il y a eu recours au système médical. Or, bien souvent, ce n'est qu'une fois que le médecin a émis un diagnostic que l'opportunité du recours aux soins peut être révélée. Ainsi, la méthode de Lohr *et al* conduit à surévaluer la proportion de soins inefficaces en comptabilisant comme tels les recours liés à des symptômes qui ne permettent pas au patient de juger de son état *a priori*. Au final, Feldman et Dowd reconnaissent que le patient peut n'avoir qu'une connaissance très imprécise de l'utilité des soins qui lui sont administrés et considèrent donc que les patients doivent être mieux informés. Néanmoins, cette limite doit être vue comme un problème classique d'information imparfaite et ne peut pas remettre en cause l'approche classique par la théorie du consommateur. Bien qu'utile pour éclairer le consommateur, l'expertise de l'efficacité des soins ne peut pas fournir une mesure alternative de l'utilité des soins. En effet, d'une part il n'existe pas dans ce domaine de consensus entre experts. D'autre part, il existe plusieurs facteurs autres que l'efficacité des soins qui influencent l'utilité que les individus retirent de la consommation de soins ; on peut citer par exemple l'aversion au risque. Or, l'approche normative de Rice ignore de tels facteurs. Les auteurs concluent en reconnaissant que l'approche par la théorie du consommateur est relative au degré d'information dont disposent les individus et est susceptible d'évoluer au fur-et-à-mesure que cette information s'affine. Néanmoins, Feldman et Dowd considèrent que le bien-être

des consommateurs doit être examiné à la lumière de la valeur que les individus attribuent aux diagnostics et prescriptions du médecin et non au regard de la valeur à dire d'experts de ces diagnostics et prescriptions.

En réponse à ces critiques, Rice (1993) déplore le fait que Feldman et Dowd s'appuient sur la théorie du consommateur pour invalider son approche, supposant ainsi que cette théorie est vérifiée. Pour preuve, Rice met en exergue le fait que la perte de surplus calculée par Feldman et Dowd et censée tenir compte des soins inefficaces s'appuie en grande partie sur la théorie classique du consommateur : ainsi, dans le graphique ci-dessus le triangle CEQ_1 au-dessus de la courbe de demande est considérée intégralement comme une perte de surplus collectif. Le supplément de soins consommés par les personnes couvertes serait donc intégralement constitué de biens et services inefficaces au regard de ceux consommés sans assurance. Or, les résultats de Lohr *et al* montrent que le surplus de soins consommé est à la fois constitué de soins inefficaces et efficaces. Concernant la validité des résultats de Lohr *et al*, Rice reconnaît certes qu'ils peuvent être entachés d'erreurs méthodologiques mais considère qu'ils sont importants car il n'existe quasiment pas d'autres travaux sur le lien entre assurance et efficacité des soins.

La réponse de Rice sur la méthode de calcul de surplus employée par Feldman et Dowd apparaît justifiée. En effet, si on considère que la valeur que la courbe de demande de soins ne représente pas la valeur réelle des soins (celle correspondant à leur efficacité), il n'y a pas de raison de considérer que le triangle CEQ_1 au-dessus de la courbe de demande totale de soins représente une perte de bien-être. Par contre, la critique de la méthodologie adoptée par Lohr *et al* pose une vraie question de fond à laquelle Rice ne répond pas : les médecins n'ont pas pour unique fonction de prescrire un traitement, ils fournissent également aux individus un service d'expertise sur leur état de santé et sur la nature des soins nécessaires. Devant un ensemble de symptômes, la consultation d'un praticien peut être *a priori* justifiée même si la pathologie s'avère *a posteriori* totalement bénigne : le patient en retire des informations sur l'origine de ces symptômes, la nécessité ou non de traitements et la nature de ces traitements et valorise plus ou moins ces informations en fonction de son goût pour la santé, de son aversion pour le risque... Ces considérations et celles de Rice renvoient à la remarque de Grossman sur le fait que les individus ne demandent pas des soins mais de la santé, plus exactement de l'information sur leur santé et, au besoin, une amélioration de leur état de santé. Le médecin répond à cette demande en fournissant un diagnostic, des conseils et des soins. Ainsi, le choix de la nature et de la quantité de bon nombre de soins relèvent du médecin mais la décision d'initier un épisode de soins revient aux patients. Contrairement à ce que suggère Rice (*"the fact that demand curve exists only indicates that consumers are forced to make choices about*

amount of and types of services to purchase”) il existe donc une part de choix individuels dans la quantité de soins achetés. Dans la mesure où les individus effectuent ces choix, ils lui attribuent une valeur en fonction de différents critères. L’enjeu n’est donc pas seulement de s’interroger sur ce que reflètent les courbes de demande de soins (pour reprendre le titre de la réponse de Feldman et Dowd), mais également de déterminer quelle courbe reflète les choix individuels. Au regard des précédentes remarques, la courbe basée sur la fréquence des épisodes de soins semble fournir une réponse adéquate à cette question. Ainsi, *a priori*, il n’y a pas de raisons de remettre en cause les pertes de surplus issus des travaux de Keeler *et al* sur l’élasticité prix de la fréquence des épisodes de soins.

L’approche de Rice apparaît également trop normative lorsque le patient a une certaine latitude dans le choix des soins. Un exemple caricatural est le choix d’une paire de lunettes prescrite par un ophtalmologiste. Le patient peut choisir entre une paire de lunettes « basique » ou « sophistiquée ». Le surcoût lié à la seconde apparaît médicalement superflus mais l’individu peut pourtant accorder une forte utilité à l’aspect esthétique, par exemple parce que sa profession l’oblige à être en contact avec le public. De même, la nécessité de prendre certains médicaments antidouleur (par exemple de l’Advil suite à des soins dentaires) peut être très relative à la façon dont l’individu ressent la douleur : un patient peut les demander à juste titre aux praticiens, d’autres non. Au contraire, un patient peut renoncer à un soin, même efficace, parce qu’il va engendrer de la douleur, de la fatigue, des contraintes en termes de temps ou un risque pour la santé. Ces facteurs influençant l’utilité que l’individu attribue aux soins sont soulignés par Pamela Peel (1993) dans une discussion de l’article de Rice. Ce dernier répond que les attributs autres que la qualité des soins ne peuvent expliquer à eux seuls les résultats observés par Lohr *et al*. Néanmoins la remarque de Peel démontre en partie qu’on ne peut pas se baser sur le seul jugement des experts pour décider si un traitement est justifié ou non.

L’approche de Rice garde tout de même une forte pertinence. En soulignant qu’en l’absence de connaissances médicales suffisantes, le patient délègue au médecin le choix du traitement, il interroge sur l’opportunité d’appliquer sur un certain nombre de prescriptions des taux de coassurance visant à responsabiliser le patient. De telles mesures supposent que le patient connaît la qualité des soins, qu’il peut effectuer un arbitrage entre cette qualité et le coût restant à sa charge et qu’il peut influencer cette prescription. Si de telles mesures sont pertinentes lorsque le patient effectue un choix (initiation d’un épisode de soins, décision de consommer certains soins pour lesquels le patient peut exprimer des préférences), elles le sont nettement moins lorsque le choix revient au médecin (prescription de médicaments, d’analyses médicales, d’actes de kinésithérapie).

Ainsi, en France, l'application de franchises non-remboursables sur les boîtes de médicaments depuis 2008 afin de responsabiliser les patients est discutable et semble avoir une efficacité limitée (Kambia Chopin, Perronnin, 2010). Dans ce cas, les préconisations de Rice concernant l'encadrement des pratiques des médecins sont justifiées.

1.9. Couverture assurantielle et accès aux soins

L'interprétation de l'effet de la couverture sur la quantité de soins consommés comme étant purement néfaste a été contestée par certains économistes tels que De Mezza (1983) et Nyman (1999a). Pour ces économistes, le mécanisme de substitution des consommations non-médicales par les consommations médicales décrit par Pauly n'explique qu'une partie du surplus de consommation généré par l'assurance. L'autre partie résulte d'un effet revenu : les remboursements des soins accroissent le pouvoir d'achat des individus malades, leur permettant d'acheter plus de biens et services médicaux et non-médicaux que s'ils n'avaient pas été assurés (Bardey et al, 2002).

L'argument de Nyman s'appuie sur le fait que la variation de consommation générée par une hausse du niveau d'assurance résulte de l'addition de deux effets :

- Un effet substitution : la baisse du coût apparent des soins augmente leur attractivité par rapport aux consommations non médicales.
- Un effet revenu : la baisse du coût apparent des soins permet de dégager des ressources pour des consommations médicales et non médicales.

L'augmentation résultant de l'effet revenu est neutre du point de vue du bien-être collectif : le pouvoir d'achat des individus en bonne santé diminue, mais le pouvoir d'achat des individus en mauvaise santé augmente. Seule l'augmentation de consommation résultant de l'effet « substitution » entraîne une perte de surplus.

1.9.1. Mesurer l'effet revenu induit par une variation du niveau d'assurance

Un enjeu important est de mesurer ce qui dans la hausse de la quantité de soins consommés, relève de l'effet substitution de l'effet revenu, afin de déterminer lequel des deux effets l'emporte. Pour distinguer ces deux effets, Nyman considère les consommations médicales et non-médicales dans le cas d'une couverture nulle, puis dans le cas d'un contrat d'assurance classique remboursant *ex-post* l'achat des soins une fraction $1 - c$ des dépenses médicales. Il considère enfin le cas d'un contrat « idéal » (nommé contrat « contingent claims ») qui verse à l'individu *ex-ante* l'achat des soins une fraction $1 - c$ du montant des dépenses médicales nécessaires à son traitement ; l'individu peut utiliser tout ou partie de cette somme pour acheter ses soins ou les allouer à d'autres types de

consommations. Ce contrat est idéal dans la mesure où l'individu a la possibilité de réaliser l'allocation optimale entre consommations médicales et non-médicales qui va lui permettre de maximiser son utilité. Dans ce cas, il n'y a pas de perte de surplus collectif. Nyman compare alors le surplus de dépenses médicales engendré par un contrat « classique » au surplus de dépense lié à un contrat « contingent claim ». Le second est forcément inférieur au premier car dans le cas d'un contrat « contingent claim », l'individu a la possibilité d'allouer tout ou partie de la somme que lui verse l'assureur à des dépenses non médicales. La part de dépenses « inefficaces » est la différence entre le surplus de dépenses lié à au contrat « classique » et le surplus de dépenses lié à un contrat « contingent claim ».

L'approche que nous avons utilisée dans la partie 1.5.3 nous a permis d'aboutir directement à une décomposition suivante de l'élasticité prix de la demande marshallienne :

$$\eta_{m_s^{\#}, q_m} = \eta_{\psi_s^u, q_m} + \eta_{W, m_s^u} \left[\theta_M \cdot \frac{c - (1 - c) \cdot \eta_{\bar{m}^u, q_m}}{1 - \theta_M \cdot (1 - c) \cdot (1 - \theta_M \cdot \eta_{\bar{m}^u, W})} - \frac{c \cdot \theta_{M_s}}{1 - \theta_M \cdot (1 - c)} \right]$$

Cette formule qui fait apparaître les effets substitution et revenu :

- L'effet substitution est égal à $\eta_{\psi_s^u, q_m}$
- L'effet revenu vaut :

$$\eta_{W, m_s^u} \left[\theta_M \cdot \frac{c - (1 - c) \cdot \eta_{\bar{m}^u, q_m}}{1 - \theta_M \cdot (1 - c) \cdot (1 - \theta_M \cdot \eta_{\bar{m}^u, W})} - \frac{c \cdot \theta_{M_s}}{1 - \theta_M \cdot (1 - c)} \right]$$

Ainsi :

- lorsque θ_M , la part de la dépense de santé moyenne dans le revenu des ménages, sont faibles devant θ_{M_s} , la part de la dépense de santé des individus dans l'état s , l'élasticité revenu est positive.
- lorsque θ_{M_s} , c et η_{W, m_s^u} sont élevés, proches de 1, l'effet revenu l'emporte sur l'effet substitution η_{ψ, m_s^u} .

En guise d'illustration, Nyman cite plusieurs types de soins (pontages coronariens, transplantations) dont le coût se chiffre à plusieurs dizaines, voire plusieurs centaines de milliers d'euros. Ces soins sont donc inabordables pour la plupart des individus, dans la mesure où leur prix excède le revenu annuel de la majeure partie de la population.

Le supplément de consommations de soins lié à cet effet revenu correspond à des soins que les individus jugent « utiles »⁴¹, mais qu'ils n'auraient pas eu les moyens d'acheter s'ils pas été assurés. Ce supplément ne doit donc pas être comptabilisé dans la perte de surplus. Utilisant les résultats de l'expérience de la RAND, Nyman (1999b) estime ainsi qu'après correction par l'effet revenu, la mesure de la perte surplus est réduite de 2/3, ce qui conduit à revoir à la baisse les niveaux de reste à charge optimaux à imposer aux assurés.

1.9.2. L'accès aux soins comme motif de l'assurance

Pour Nyman, la prise en compte de l'effet revenu ne remet pas seulement en cause l'ampleur de la perte de bien-être collectif. Il conduit également à revoir le motif de la souscription d'une couverture. Rappelons que selon les modèles classiques de demande d'assurance, les individus souscrivent une couverture parce qu'ils sont averse au risque et cherchent donc à diminuer l'aléa sur leurs dépenses de santé. Nyman voit deux problèmes à cette approche :

- Tout d'abord, le fait que dans le modèle traditionnel, l'individu compare le risque qu'il subit lorsqu'il n'est pas assuré à celui qu'il subit lorsqu'il est assuré. Or, dans le cas de procédures très onéreuses, dont le coût excède les capacités financières de l'individu, ce dernier a une dépense nulle lorsqu'il n'est pas assuré. Au contraire, lorsqu'il est assuré, sa dépense va fortement augmenter puisque l'assurance lui permet d'avoir accès à la procédure médicale : son « risque » est donc plus élevé lorsqu'il est assuré que lorsqu'il n'est pas assuré.
- Ensuite, des expériences menées sur les choix en incertain ont remis en cause l'utilité espérée comme mode de décision des individus. Ces expériences appuient la théorie des perspectives (*prospect theory*) qui postule que les individus basent leur choix sur des gains et des pertes par rapport à une référence et non sur les résultats finaux de la loterie. Un individu peut ainsi préférer une loterie offrant un gain aléatoire élevé plutôt qu'un gain sûr relativement faible même si l'espérance de gain est plus importante dans la seconde loterie que dans la première. Nyman s'appuie ainsi les résultats d'une expérience dans laquelle des individus sont confrontés à deux cas de figure : dans le premier, il est demandé aux individus `faisant face à une perte possible de 200\$ survenant avec une probabilité de 0,25 s'ils préfèrent payer 50\$ pour couvrir totalement cette perte ou s'ils préfèrent ne pas s'assurer; les individus préfèrent s'assurer. Par contre si on demande aux mêmes individus s'ils préfèrent perdre 50\$ de manière certaine ou 200\$ de manière aléatoire, ils vont préférer le second choix. Pourtant, du point de vue de l'utilité espérée, les deux choix sont équivalents.

⁴¹ C'est-à-dire que les individus leur accordent une valeur monétaire supérieure à leur coût.

Pour Nyman, les individus ne choisissent pas une assurance pour réduire leur risque financier, mais pour bénéficier d'un transfert de revenu en cas de maladie, de la part des individus en bonne santé. Ce transfert de revenu va permettre de solvabiliser la demande de soins de l'individu lorsque la pathologie nécessite un traitement très onéreux. Considérons par exemple une pathologie ayant une chance sur 100 000 de survenir, mais qui nécessite des soins coûtant 100 000€. Un individu peut choisir un contrat couvrant cette dépense. Si le contrat est tarifé à son coût actuariel, la prime sera de $100\,000 \times 1/100\,000 = 1\text{€}$. En cas de maladie, l'individu bénéficiera ainsi d'un transfert de revenu net de $100\,000 - 1 = 99\,999\text{€}$, ce qui lui permet de consommer un soin qu'il n'aurait pas pu obtenir s'il n'avait pas été assuré. Dans la pratique, les individus, même aisés, souscrivent des contrats d'assurance qui couvrent également des dépenses de santé modestes. Ainsi, la fonction de réduction du risque financier ne peut pas être écartée. Néanmoins, l'approche de Nyman illustre le fait que l'assurance peut-être également souscrite pour obtenir un meilleur accès aux soins.

1.9.3. Le débat sur l'effet revenu

Les travaux de Nyman ont à leur tour fait l'objet de critiques sur le plan technique et théorique. Une d'entre-elles porte sur le mode de calcul de l'effet revenu. Selon Blomqvist (2001), les modalités de calcul de Nyman conduisent à surévaluer fortement l'effet revenu (donc à sous-évaluer la perte de surplus) du fait d'une surestimation de l'élasticité revenu et de la part des dépenses à la charge des individus les moins bien couverts⁴². Une critique plus fondamentale adressée par Blomqvist (2001) porte sur le modèle théorique adopté par De Mezza puis Nyman. Selon Blomqvist, ce modèle repose largement sur une comparaison entre couverts et non couverts, mettant en particulier en avant le fait que ces derniers sont dans l'incapacité de payer des dépenses de santé catastrophiques. Si le modèle est pertinent pour justifier la nécessité de couvrir les non-assurés, il est inadapté pour critiquer les restes à charge prévus par les contrats d'assurance. Blomqvist souligne en particulier que la plupart des contrats, y compris ceux de l'expérience de la RAND, prévoient des *stop-loss*, c'est-à-dire une couverture complète des dépenses catastrophiques. Ainsi, sans nier la nécessité de considérer l'effet revenu, Blomqvist insiste sur le fait que l'analyse de l'influence de la couverture doit se faire à la marge, c'est-à-dire qu'il faut étudier l'effet du fait d'être plus ou moins bien couvert chez des personnes assurées.

⁴² En réponse, Nyman (2008) a jeté le doute sur les résultats issus de l'expérience de la RAND, faisant valoir le potentiel biais d'attrition auquel était soumise cette enquête. Cette remise en cause a suscité une vive réaction de Newhouse et al, (2009) ceux-ci démontrant que la critique de Nyman était infondée.

1.9.4. Effet à la marge du niveau de couverture : analyses empiriques exploitant les couvertures complémentaires

Ettner (1997) analyse le lien entre le fait de bénéficier d'une couverture Medigap, complétant les remboursements du programme Medicare, et les consommations de soins. Le programme Medicare est une couverture publique destinée aux personnes âgées aux Etats-Unis. En 1991, date à laquelle ont été collectées les données utilisées par Ettner, Medicare offrait une couverture des soins hospitaliers (Medicare part A) et des soins de médecins (Medicare part B). Pour ces soins, ce programme laisse à la charge des patients des dépenses substantiels (sous la forme de franchises et de taux de coassurance). De plus, en 1991, Medicare ne remboursait pas les médicaments et offrait une couverture limitée des soins de long terme. La couverture Medigap permet de couvrir tout ou partie des coûts laissés à la charge des patients sur les soins dans le panier de biens et services médicaux de Medicare. Elle peut également offrir une couverture des médicaments et des séjours en institution. Les contrats Medigap diffèrent donc selon le niveau de couverture et le panier de soins couverts. Les contrats Medigap peuvent être obtenus par l'intermédiaire de l'ancien employeur, par Medicaid ou souscrits par une démarche individuelle.

Bien que le but principal de l'étude d'Ettner soit l'analyse de l'antisélection dans le choix de souscrire ou non Medigap et dans le niveau de couverture choisi, elle fournit également des résultats sur l'aléa-moral chez les souscripteurs de ce type de couverture. Les données utilisées sont issues du Medicare current beneficiary survey (MCBS), réalisée en 1991. Cette enquête recueille auprès des bénéficiaires de Medicare, des données sur leurs caractéristiques sociodémographiques, leur état de santé, leur couverture santé et des données administratives sur les consommations de soins présentées au remboursement de Medicare. Ces données ne portent donc que sur des types de soins couverts par le dispositif de base : sont ainsi disponibles les remboursements totaux par Medicare, les remboursements par Medicare Part A, Part B, les dépenses et remboursements pour les hospitalisations, les soins de médecins et les soins ambulatoires autres que ceux de médecin ; pour les soins de médecin, on dispose également du nombre de visites et pour les soins hospitaliers du nombre de jours d'hospitalisation.

Ettner régresse les différents indicateurs Y de consommations médicales sur le niveau de couverture Medigap et les variables de contrôle (état de santé, caractéristiques sociodémographiques et socio-économique) au moyen d'une procédure en deux étapes : tout d'abord la probabilité $p(Y > 0|X)$ d'avoir au moins un soin dans l'année est estimée au moyen d'un modèle probit puis $E(Y|Y > 0, X)$ est estimée chez les consommant par une régression linéaire. Ceux-ci sont appréhendés par le fait de bénéficier de ne pas bénéficier de contrat Medigap, (la référence) de bénéficier d'un contrat Medigap basique (variable « BASIC ») et le fait de bénéficier d'un contrat Medigap « généreux »

(variable « BASIQUE+ »). Un contrat Medigap est considéré comme généreux s'il rembourse les médicaments et ou les séjours en institution. On peut en effet faire l'hypothèse que le fait de couvrir ces soins non-assurés par Medicare est corrélé au niveau de remboursement des restes à charges sur les soins couverts par Medicare : cette variable est donc un proxy du niveau de remboursement des soins analysés. Le phénomène d'antisélection est testé et contrôlé en s'appuyant sur les individus couverts par l'intermédiaire de leur employeur, a priori moins sujets à l'antisélection que les individus couverts par une démarche individuelle. En effet, ils sont moins sensibles au prix du contrat puisque l'employeur paie une partie de la prime, ils ont un choix plus restreint de niveaux de couverture et bénéficient d'une meilleure péréquation des risques au sein de leur entreprise (Couffinhal, 1999)⁴³. Concrètement, la sélection adverse est prise en compte en introduisant dans la régression la variable « ADVERSE » et « ADVERSE+ » qui indiquent tous deux si le contrat a été souscrit par une démarche individuelle et respectivement si le contrat est basique ou généreux. La mesure du risque moral est ainsi donnée par la valeur des coefficients « BASIQUE » et « BASIQUE+ ».

Les résultats font apparaître que le fait de bénéficier d'une couverture Medigap basique plutôt que de ne pas bénéficier de couverture Medigap n'a un effet significatif que sur le nombre de visites chez le médecin. En revanche, le fait de bénéficier d'une couverture Medigap généreuse a un effet significatif sur les remboursements totaux, de Medicare Part B, les remboursements de soins de spécialistes, les remboursements des autres soins ambulatoires, ainsi que sur le nombre de visites chez le médecin.

En France, plusieurs travaux sur l'effet des couvertures complémentaires santé sur les consommations médicales ont été menés. Etant donné que les couvertures complémentaires réduisent fortement les restes à charges laissées par l'Assurance maladie obligatoire, elles sont susceptibles d'avoir une forte influence sur les consommations de soins. Plusieurs études se sont donc attachées à mesurer l'ampleur de cet effet, entre autres Caussat et Glaude (1993), Chiappori et al (1998), Genier (1998) et Albouy et Crépon (2007). Les analyses ont porté sur différents types de soins observés sur différentes périodes. Ces travaux ne mettent pas en évidence d'effet significatif de la complémentaire santé sur les consommations de soins hospitaliers, ce qui peut s'expliquer d'une part par le fait que ces soins sont déjà fortement remboursés par l'Assurance maladie obligatoire et d'autre part présentent un caractère essentiel voire vital (les individus, mêmes non couverts sont donc peu enclins à y renoncer). En ce qui concerne les soins ambulatoires, la complémentaire santé a un effet significatif sur la probabilité de recourir au moins une fois aux soins sur une période donnée ;

⁴³ Par contre, les contrats d'entreprise aux Etats-Unis peuvent en général être souscrits de manière facultative, à la différence de la France où la majeure partie des couvertures complémentaires santé collectives sont à caractère obligatoire.

il est par contre moins net sur le volume des soins ou la dépense engagée lorsqu'il y a recours. Il faut noter que ces travaux ne cherchent pas à évaluer si le surplus de consommation de soins ambulatoire occasionné par la complémentaire santé reflète, au moins partiellement, un meilleur accès aux soins, où seulement de la surconsommation. Si on considère qu'il est nécessaire d'avoir régulièrement un contact avec le système de santé à titre préventif, alors l'effet de la complémentaire santé sur la probabilité de recours peut être interprété comme positif. Néanmoins, ce raisonnement est plus normatif et qu'économique. D'autre part, à l'exception de Chiappori et al (1998) qui analysaient l'effet d'avoir un coassurance plus ou moins élevé chez des personnes couvertes, les autres travaux utilisent comme variable d'assurance le fait d'être couvert plutôt que non-couvert par une complémentaire santé. Or, la non-couverture par une complémentaire santé ne concerne que 7% environ de la population, le reste étant couvert à des degrés très variables (cf Couffinal, Perronnin, 2004). Au sein des personnes assurées, l'effet de la couverture peut être très différent suivant le niveau de couverture choisi. Enfin, sur un plan purement technique, notons qu'à l'exception de l'étude de Chiappori et al qui utilise une expérience naturelle et des données longitudinales, l'ensemble des travaux cités utilisent des données en coupe transversale. Elles contrôlent donc le biais d'autosélection (lié au choix des individus d'être plus ou moins assurés selon leur niveau de consommation de soins) en utilisant des hypothèses instrumentales qui peuvent être contestées.

Au-delà de ces limites, l'intérêt qu'il peut y avoir à traiter les problématiques de risque moral et d'accès aux soins dans le cadre français n'est a priori pas évident. En effet, les problèmes d'accès aux soins chez les plus pauvres semblent avoir été traités grâce aux dispositifs à destination des ménages pauvres (AMG puis CMU-C et Aide complémentaire santé). Mais il faut souligner que malgré ces dispositifs, il reste selon l'enquête ESPS une part non négligeable d'individus déclarant renoncer à des soins pour raison financière. En ce qui concerne la question de l'efficience, elle ne semble à première vue pas se poser. En effet, les individus ayant en principe le libre choix de leur couverture complémentaire, ils peuvent sélectionner le contrat qui est optimal pour eux en termes d'arbitrage entre couverture du risque et réduction du risque moral. Néanmoins, ce raisonnement suppose que chaque individu paie la totalité du surcoût lié au niveau de couverture qu'il a choisi. Or, ce n'est que partiellement vrai puisqu'une partie des dépenses est socialisée par le biais de l'Assurance maladie obligatoire : l'éventuelle surconsommation de soins liée aux complémentaires santé peut se répercuter sur le Régime obligatoire et participer ainsi à son déficit. De plus, le coût de nombreux contrats de couverture complémentaire (les contrats de groupe obligatoire, les contrats ouvrant le droit à l'ACS) est réduit du fait d'un subventionnement public, ce qui peut conduire à un arbitrage inapproprié entre risque moral et réduction du risque financier.

Chapitre 2.

**Access to physician services:
does supplemental insurance matter?
Evidence from France**

Tomas C. BUCHMUELLER (University of California, Irvine and NBER, USA, IRDES), Agnès COUFFINHAL (IRDES), Michel GRIGNON (IRDES), Marc PERRONNIN (IRDES)

2.1. Abstract

In France, public health insurance is universal but incomplete, with private payments accounting for roughly 25% of all spending. As a result, most people have supplemental private health insurance. We investigate the effects of such insurance on the utilization of physician services using data from the 1998 Enquête sur la santé et la protection sociale, a nationally representative survey of the non-institutionalized French population. Our results indicate that insurance has a strong and significant effect on the utilization of physician services. Individuals with supplemental coverage have substantially more physician visits than those without. While French patients have greater freedom than patients in other countries to choose to see a specialist rather than a general practitioner, we find no evidence that supplemental insurance affects this decision.

2.2. Introduction

In France, as in many other European countries, a publicly financed system provides universal health insurance coverage, but that coverage is not complete. France's Social Security system covers roughly 75% of overall individual health expenditures. The remaining 25% consists of copayments for basic services, spending on services that are either not covered or poorly covered by the public system and provider charges in excess of official reimbursement amounts. In response to these gaps in the public system, 85% of the French population has supplemental health insurance.

The effects of supplemental insurance coverage on medical care utilization are of considerable interest to health economists and policy makers, as they have significant implications for both economic efficiency and equity. The issue of efficiency relates to how supplemental insurance interacts with the public system. Like public financing schemes in other countries, the French system imposes cost-sharing on patients as a strategy for controlling utilization (Couffinhall, Paris, 2001). Because most insurance policies reimburse patients for their out-of-pocket payments, supplemental insurance is likely to render ineffective the desired incentive effect, raising both total and public sector health spending, and reducing economic welfare.

The utilization effects of supplemental insurance also have implications for equity. While the incidence of the financing of the public system is roughly proportional, private health expenditures in France are more regressive than in other countries, with the exception of the US and Switzerland (Van Doorslaer & Wagstaff, 1993). While France rates very highly in terms of life expectancy and other health indicators, the aggregate figures obscure large differences related to income and social class (Haut Comité de la santé publique, 2002). The fact that both supplemental insurance coverage and access to medical care are positively related to income and social class raises the concern that

the limitations in the coverage provided by the public system may contribute to the inequity in health outcomes.

Seeking to reduce these inequities in access, in 2000 the French government introduced a new means-tested program, the Couverture maladie universelle (CMU), to supplement the health insurance coverage provided by the Social Security system and reduce the out-of-pocket costs faced by low-income patients. The CMU provides free supplemental insurance to households whose income is below a certain threshold (533€ per month for single individuals; the threshold amount increases with family size). The coverage is designed to cover all out-of-pocket costs for a standard market basket of medical services.

As a starting point for understanding the impact of this policy on health spending and access to care, it is useful to examine the relationship between supplemental insurance coverage and utilization in the period just prior to the implementation of the CMU. Using a nationally representative survey of the non-institutionalized French population conducted in 1998, we investigate the two most important ways that health insurance is likely to affect health spending. First, by lowering the cost of care, insurance may cause a general increase in the quantity of medical care demanded, thus increasing the number of visits. Second, we consider whether supplemental coverage influences the choice of provider, conditional on seeking care. Specifically, we examine the effect of insurance on the choice between seeing a specialist rather than a general practitioner (GP).

The greatest challenge with this type of analysis is distinguishing a causal effect of health insurance i.e., moral hazard – from the effect of unobserved heterogeneity – i.e., adverse selection. The most obvious concern is that individuals with greater expected medical care use – say, for reasons of poor health – may be more likely to purchase supplemental insurance. If this is the case, the empirical relationship between insurance and utilization will overstate the moral hazard effect. We address this issue in several ways. First, the fact that our data includes detailed information on individual health status allows us to investigate the relationship between health risk and insurance coverage. While it does not rule out the possibility of selection based on unobservable characteristics, the richness of the health variables in our data provides a strong test for health-related adverse selection. The inclusion of these variables in our utilization regressions also mitigates concerns that our estimated insurance effects are driven by differences in health status between people with and without private insurance. Second, using a strategy previously employed by Ettner (1997) and Genier (1998), we contrast the effects of insurance that is obtained through the workplace with individually purchased coverage. To the extent that adverse selection is an issue in the market for supplemental insurance, the problem should be limited to policies purchased directly by individuals. We also

estimate several bivariate probit models that treat insurance coverage and the probability of having a physician visit as jointly determined outcomes.

Our results point to a strong effect of supplemental health insurance on the use of physician services. Probit models of insurance coverage provide no support for the hypothesis that people who are 'sicker' are more likely to purchase supplemental insurance coverage. Models of the effect of insurance on physician visits indicate a large and statistically significant effect and the magnitude of the effect is comparable for individually purchased and employer-provided insurance. Similarly, the point estimates from bivariate probit models are qualitatively similar to those from simpler models that treat supplemental insurance as exogenous. Taken together, these results support a moral hazard interpretation of the effect of the relationship between supplemental health insurance and the use of physician services.

2.3. Health insurance in France

All legal residents of France are covered by public health insurance, which is one of the Social Security system's entitlement programs. Between 1980 and 1995, the public sector's share of spending on health care services declined from 79.4% to almost 75%; it has stayed roughly at that level since then. This trend has coincided with an increase in the proportion of the population with private supplemental health insurance. Whereas in 1970 roughly one-half of the population had such coverage, the figure increased to 85% in the 1990s, with private insurance accounting for half of all private health expenditures.

Because roughly half of all supplemental insurance contracts are provided through the workplace, industrial/occupational affiliation is an important factor determining health insurance coverage. Coverage tends to be higher for managers and highly educated professionals than for semi-skilled and unskilled workers, and is lowest for the unemployed (Bocognano et al, 2000). The vast majority of individually purchased policies is provided by nonprofit organizations known as *mutuelles*, which are organized largely on occupational, professional or regional lines and which neither exclude potential subscribers on the basis of health nor use risk-rated premiums. Some institutions increase the premium when individuals retire but the adjustment is highly regulated and far from actuarial. About one-quarter of individually purchased coverage is through for-profit insurers who can legally use health status in the underwriting process.

Whereas in other countries supplemental coverage provides access to providers operating outside the public system or a means of 'jumping the queue' to obtain hospital or specialty care, this is not the case in France. Patients with and without private coverage see the same physicians and neither

face any explicit constraints on access. Patients can see as many physicians as they like, as often as they like, and do not need a referral to see a specialist. Thus, in the French system, the main benefit provided by private insurance is to reduce out-of-pocket payments.

In the case of physician visits, there are two sources of out-of-pocket payments: the *ticket modérateur*, an official co-payment set by the government, and the difference between a provider's fee and the 'standard tariff', which is the maximum amount reimbursed by the public system. Out-of-pocket costs are higher for specialists in part because the *ticket modérateur* is defined as a fixed percentage of the standard tariff, which is higher for specialists, and because specialists are more likely to balance bill. Thirty-eight percent of specialists are allowed to charge more than the standard tariff, compared to 11% of GPs.

In principle, supplemental insurance may induce substitution away from GPs toward specialists by reducing the price difference between the two types of providers. This may be viewed as an effect of insurance on patient demand for quality, particularly in France where the hierarchical nature of medical education is likely to contribute to a perception that specialists provide a higher quality of care. However, the magnitude of this effect in the current French context may be small as the price difference between specialists and generalists is not generally large. At the time of our data, the standard tariffs were 22.9€ for specialists and 17.5€ for GPs, which translates to less than a 2h difference in the "*ticket modérateur*). The amount by which specialist fees exceeds the amount reimbursed by the public system varies. In our data, average total charges for specialist and generalist visits would translate to out-of-pocket payments of roughly 15.5€ and 6.6€; respectively.

2.4. Previous research on supplemental insurance and utilization

Although there is a large literature on the effects of health insurance on medical care utilization, most studies are from the US and, for various reasons, are of limited applicability to France and other European countries. The best-known research comes from the Rand Health Insurance Experiment (Manning WG et al, 1987). Individuals randomly assigned to a plan with first dollar coverage had 37% more physician visits than those facing co-insurance rates of 25% and 50% more than individual with 50% coinsurance. These are the results that are most relevant to the French experience since French patients with comprehensive supplemental insurance face co-insurance rates for ambulatory that are close to zero, compared to rates of between 25% and 50% for individuals without supplemental coverage. More recent US studies use large, cross-sectional data sets that, similar to the French data we use, include considerable information on health status as well as economic and socio-demographic variables (Spillman, 1992, Hahn, 1994, Marquis & Long, 1994-1995). Controlling for observable individual characteristics, these studies find that uninsured adults have roughly half as

many physician visits per year as the insured. Perhaps more relevant to the French system are studies on the effects of private “Medigap” insurance, which supplements Medicare coverage for elderly Americans. A number of studies find that Medigap coverage is positively associated with utilization (Cartwright *et al.*, 1992) though it is unclear how much of this effect represents moral hazard and how much is due to self-selection.

Research from several European countries indicates a significant positive effect of supplemental private health insurance on outpatient utilization (for instance Harmon and Nolan (2001) for Ireland, Vera-Hernandez (1999) for Spain, and Schellhorn (2001) for Switzerland). Most relevant to our work are several prior studies focusing on France. Chiappori *et al.* (1998) analyze a natural experiment resulting from the responses of a private supplemental insurer to a change in cost-sharing required by the public system. They compare changes in utilization for a reference group for whom coverage was nearly complete and constant over time and a ‘treatment group’ that experienced a slight increase in required cost sharing. While this approach mitigates the problem of selection bias, it raises other problems relating to the generalizability of the results. All the individuals analyzed had supplemental insurance and even after the regime change the cost sharing required of the treatment group was quite modest. These facts combined with the likelihood that individuals in both groups had higher than average incomes, may explain their finding that coinsurance had little effect on use. Differences in utilization between individuals with and without any supplemental coverage may be greater, especially in a broader population.

Three other French studies use nationally representative data. Mormiche (1993) presents descriptive evidence on the relationships among social class, medical care utilization, and insurance coverage. Individuals from higher income occupations have more physician visits overall and, conditional on having any visits, are more likely to receive care from specialists. Access to supplemental insurance coverage also varies with income and social class. However, because they do not control for other factors, these correlations are difficult to interpret. Two other studies (Caussat & Glaude, 1993, Genier, 1998) use data from 1980 and 1991–1992 to examine the relationship between supplemental insurance and utilization with regressions that control for income and a limited set of health and demographic variables. Both studies find significant effects of insurance on the probability of having any physician visits, but no significant effect on the number of visits conditional on having any.

As noted in the introduction, insurance coverage may be endogenous with respect to utilization. Many studies treat insurance as exogenous and rely on extensive health-related control variables to mitigate the bias due to unobserved heterogeneity. Ettner (1997) takes an innovative approach to

disentangling moral hazard and adverse selection effects of Medigap insurance. First, she estimates models of the determinants of Medigap coverage to test for adverse selection based on observable health status. Then, in her utilization regressions she distinguishes between Medigap coverage provided by a former employer and individually purchased policies. The idea is that adverse selection is unlikely to be a major issue for employer-provided coverage, but may be for individually purchased plans. A third set of studies uses a variety of multiple equation strategies (e.g. bivariate probit, instrumental variables). In our work, we use all of these methods.

There has been much less research on the effect of insurance on the choice between specialists and generalists. Marquis (1994-1995) addresses this question using data from the Rand experiment and generally finds no effect. Using survey data from the US, Cooper et al (1996) find that among the nonelderly, uninsured patients are no less likely to see specialists than those with private insurance. For patients over age 65, they find a significant and positive relationship between supplemental health insurance and the decision to see a specialist. Apart from the descriptive results presented by Mormiche (1993) this question has not been examined for France or other European countries.

2.5. Data and descriptive evidence

Our data are from the Enquête sur la santé et la protection sociale (ESPS), a national household survey conducted by CREDES every other year. The ESPS sample is drawn from a permanent sample of persons insured by the three main sickness funds which cover roughly 95% of the French population. Using a combination of telephone interviews and self-administered questionnaires, the ESPS collects individual and household level data, including information on demographic and socioeconomic characteristics, health status, supplementary insurance coverage and medical care utilization. We use data from the 1998 survey, for which 7996 households were interviewed. In this study, we focus on the behavior of adults ages 25 and above. A small fraction of the French population is exempted from the payment of *ticket modérateur* either because of high cost chronic conditions or extremely low incomes. It is difficult to classify their insurance status because this exemption provides less financial protection than typical supplemental policies, and yet these individuals do not face the same prices as the uninsured. For this reason, and the fact that eligibility for these exemptions is linked to variables that could influence both utilization and the demand for insurance (most notably poor health), we exclude these individuals from our analysis. It should be noted that because people who receive an exemption are less likely than average to hold private supplemental insurance, the percentage of people in our sample without coverage is less than in the population. Since health status is a crucial control variable for analyzing medical care utilization, we also exclude respondents who did not complete the part of the survey pertaining to health. Having

excluded the individuals who did not complete the consumption record, our final sample consists of 8161 individuals.

ESPS respondents are asked to keep a record of all medical consumption during a month, including information on the nature of the good or service consumed, the medical condition that prompted the utilization, the type of professional providing the service, the content of the visit, and the amount charged for the care. The main advantage of this survey methodology is that the utilization data are highly accurate and complete. The cost of this accuracy, however, is that in a one-month period the vast majority of respondents have no utilization, and among users the modal number of visits is one.

Table 2.5-i presents the distributions for the utilization outcomes that we analyze: visits to any physician, visits to GPs and visits to specialists. The data are presented for the full sample and separately by insurance status. For all three measures, the mean is significantly higher for subsample with supplemental coverage. Considering all types of physicians, those with coverage have nearly twice as many visits as those without (0.46 vs 0.23). A closer investigation reveals that this difference is driven by a difference at the extensive margin: insurance is associated with a doubling of the probability of positive utilization (31.8% vs 16.5%). For GP and specialists visits, the insurance-related difference in having any visits is 10 and 7 percentage points, respectively, which are large effects given sample means for these outcomes. Conditional on having at least one visit, there is not a statistically significant difference in the mean number of visits for the two groups. Given this pattern, our econometric analysis will focus mainly on the effect of insurance on the probability of having any visits.

Summary statistics for our independent variables are reported in table 2.5-ii.

The ESPS provides information on the source of each person's insurance coverage. Slightly more than half of the sample receives supplemental insurance through the workplace. Because employer-provided health insurance represents a relatively small share of total compensation in France and comparable insurance is readily available outside the workplace, there is no reason to expect workers to choose jobs based on their demand for insurance. As a result, to the extent that adverse selection is an issue in the French system it should be limited to individually purchased contracts. Therefore, following Ettner (1997) and Genier (1998), we can test for adverse selection by comparing the effect of these two types of policies on utilization. The validity of this test rests on the comparability of the two types of coverage. While limited, the information on the quality of coverage that is available in the ESPS does not point to large differences in quality. When asked about the quality of reimbursement for specialized care, 45% of respondents with group insurance and 43% of persons with individual contracts said their coverage was complete.

Table 2.5-i : Distribution and Mean Number of Physician Visits Over a Month Period by Insurance Status, Adults over 25

	Full Sample	No Supplemental Insurance	With Supplemental Insurance
All Visits			
0	69.4%	83.5%	68.2%
1	21.8%	13.2%	22.6%
2	6.0%	1.8%	6.4%
3 or More	2.8%	1.5%	2.8%
Mean	0.44	0.23	0.46
(std. dev.)	(0.83)	(0.65)	(0.85)
GP Visits			
0	77.7%	87.1%	76.9%
1	19.0%	12.0%	19.6%
2	2.8%	0.6%	3.0%
3 or More	0.5%	0.3%	0.5%
Mean	0.26	0.14	0.27
(st. dev.)	(0.54)	(0.40)	(0.55)
Specialists Visits			
0	87.4%	94.2%	86.8%
1	9.5%	4.4%	9.9%
2	2.2%	0.9%	2.3%
3 or More	0.9%	0.5%	1.0%
Mean	0.18	0.09	0.18
(std. dev.)	(0.58)	(0.43)	(0.59)
Number of Observations	8,161	667	7,494
(% of full sample)	(100%)	(8%)	(92%)

Table 2.5-ii : Sample Characteristics

	Full Sample	No Supplemental insurance	With supplemental insurance
Socio-Demographic Characteristics			
Age (Years)	46.3	44.4	46.5
Male	48.4%	52.9%	48.0%
Married	82.5%	64.6%	84.1%
Primary Education	19.3%	21.2%	19.1%
Secondary Education	37.1%	37.8%	37.1%
High School	15.8%	13.6%	15.9%
Higher Education	24.8%	22.8%	25.0%
Unknown	3.0%	4.6%	2.9%

	Full Sample	No Supplemental insurance	With supplemental insurance
Household Income per Consumption Unit			
Less than € 305	3.1%	8.4%	2.6%
€ 305 to 457	8.3%	19.3%	7.3%
€ 457 to 610	10.9%	13.8%	10.6%
€ 610 to 762	12.8%	10.0%	13.0%
€ 762 to 915	12.6%	12.1%	12.6%
€ 915 to 1,220	16.4%	8.7%	17.1%
More than € 1,220	23.4%	14.2%	24.3%
Unknown	12.5%	13.5%	12.5%
Health Status			
Self-Assessed Health Status			
Fair and Poor	12.3%	13.9%	12.2%
Good	61.5%	55.4%	62.0%
Excellent	21.4%	24.7%	21.1%
Unknown	4.8%	6.0%	4.7%
Moderate to High Mortality Risk	18.7%	17.1%	18.8%
Disabled	18.7%	16.3%	18.9%
Experiencing Frequent Pain	15.7%	16.2%	15.6%
Pregnant Women	1.1%	0.4%	1.1%
Tobacco Consumption			
Former Smoker	21.7%	14.7%	22.4%
Light Smoker	11.3%	14.2%	11.0%
Heavy Smoker	14.4%	22.3%	13.7%
Non-smoker	52.6%	48.8%	52.9%
Labor Market Status			
Active	66.7%	57.2%	67.5%
Inactive	8.6%	12.4%	8.3%
Unemployed	5.6%	16.0%	4.6%
Retired	18.3%	12.0%	18.9%
Students	0.8%	2.4%	0.7%
Occupational Profession			
Farmers	4.4%	5.2%	4.3%
Artisans Retailers	5.3%	9.7%	4.9%
Executives and Intellectual Professions	12.7%	9.4%	13.0%
Technical Professions	20.7%	10.0%	21.6%
Other Employees	29.0%	28.0%	29.2%
Skilled Workers	16.5%	15.7%	16.6%
Unskilled Workers	8.4%	15.4%	7.7%
Unknown	3.0%	6.6%	2.7%
Public employee	24.8%	13.9%	25.8%
Location of Residence			
High Density of Physicians	22.3%	38.2%	20.9%
Low Density of Physicians	7.3%	7.5%	7.2%
Number of Observations			
	8,161	667	7,494

The potential for bias from unobserved heterogeneity is also reduced by the fact that the ESPS includes a rich set of health status measures and other controls. Self-assessed global health status is measured using a 10-point scale. We use this information to construct indicator variables for people in excellent health and those whose health is fair or poor. Respondents are also asked to list their medical conditions, to report information about risk factors (smoking) and to say whether or not they frequently experience pain or difficulties in performing daily activities. Based on responses to these questions, physician researchers have constructed categorical variables reflecting mortality risk and levels of disability. In our analysis, we use this information to construct indicators for individuals with a high mortality risk and for those reporting some type of disability. Our regressions also include three dummy variables relating to smoking, a dummy for individuals who report experiencing frequent pain, and one for women who are pregnant at the time of the survey.

Demographic and economic control variables include age, gender, family income (reported in intervals and normalized by the number of consumption units in the family), education, employment status and occupation/labor force activity. Two additional control variables describe whether the individual lives in a region where the density of physician is high or low.

2.6. Econometric analysis

2.6.1. The determinants of supplemental insurance coverage

An important first step in testing for adverse selection is to examine the effect of observed health status on insurance coverage. We do this in the context of a probit model where coverage by private supplemental health insurance (C) is a function of economic and demographic variables (X), health status (H) and labor market status (L) :

Equation 2.6-i

$$C^* = X'\beta_c + H'\gamma_c + L'\delta_c + u_c$$

$$C = 1 \text{ if } C^* > 0$$

Results for this model are reported in the first column of table 2.6-i. In columns 2 and 3 we report coefficients from a multinomial logit model with three outcomes: no supplemental insurance (the base category), employer-provided insurance and individually purchased coverage. First and foremost, we find no evidence that individuals in poor health are more likely to obtain supplemental insurance. Some of the probit coefficients indicate a positive relationship between poor health and coverage, though others have the opposite sign. Most importantly none of the health coefficients are statistically significant individually and they are insignificant as a group as well. This result is consistent with the findings of Hurd and McGarry (1997) for Medigap insurance in the US, Shmueli (2001) for supplemental insurance in Israel and Genier (1998) who analyzes a different French

survey. Our results are stronger than those studies because we have substantially more information on health status. The multinomial logit results also indicate no significant relationship between health and insurance coverage.

Table 2.6-i : Determinants of supplemental insurance coverage

Independent Variables	Model :	Multinomial Logit	
	Outcome :	Employer-Provided Coverage	Individually-Purchased Coverage
Demographic Characteristics			
Age	Probit Any Supplemental Insurance	0.122** (0.022)	0.022 (0.020)
Age ²		-0.0013** (0.0002)	-0.0003 (0.0002)
Male		-0.42** (0.11)	-0.39** (0.11)
Married		1.45** (0.11)	0.81** (0.11)
Secondary Education		0.08 (0.14)	0.11 (0.14)
High School Education		0.10 (0.18)	0.10 (0.18)
Higher Education		-0.11 (0.19)	-0.10 (0.19)
Education Unknown		-0.59* (0.25)	-0.29 (0.24)
Household Income			
€ 305 to 457		0.35 (0.24)	-0.01 (0.21)
€ 457 to 610		1.13** (0.24)	0.46* (0.21)
€ 610 to 762		1.62** (0.25)	0.87** (0.22)
€ 762 to 915		1.42** (0.24)	0.55* (0.22)
€ 915 to 1,220		2.16** (0.25)	1.12** (0.23)
More than € 1,220		2.10** (0.24)	1.00** (0.22)
Income Unknown		1.40** (0.24)	0.65** (0.21)
Health Variables			
Self-Assessed Health = Fair/Poor		-0.12 (0.15)	-0.11 (0.15)
Self-Assessed Health = Excellent		-0.10 (0.11)	-0.13 (0.11)
Self-Assessed Health = Unknown		-0.21 (0.21)	0.00 (0.20)
Moderate to High Mortality Risk		0.14 (0.14)	0.16 (0.14)
Disabled		0.25 (0.15)	0.24 (0.15)
Experiencing Frequent Pain		-0.06	-0.04

Independent Variables	Model :	Multinomial Logit		
	Outcome :	Probit Any Supplemental Insurance	Employer-Provided Coverage	Individually- Purchased Coverage
Pregnant Woman		(0.07) 0.42 (0.30)	(0.14) 0.59 (0.63)	(0.14) 0.98 (0.62)
Tobacco Consumption				
Former Smoker		0.03 (0.06)	0.11 (0.13)	0.08 (0.13)
Light Smoker		-0.16* (0.07)	-0.33* (0.14)	-0.28* (0.14)
Heavy Smoker		-0.27** (0.07)	-0.45** (0.13)	-0.48** (0.13)
Labor Market Status				
Inactive		-0.31** (0.09)	-0.65** (0.18)	-0.52** (0.18)
Unemployed		-0.58** (0.08)	-1.40** (0.16)	-0.77** (0.15)
Retired		0.34** (0.11)	0.61** (0.23)	0.85** (0.22)
Students		-0.05 (0.23)	-0.32 (0.50)	0.14 (0.43)
Professional Occupation				
Farmers		-0.14 (0.11)	-1.81** (0.28)	0.25 (0.22)
Artisans Retailers		-0.35** (0.09)	-2.07** (0.23)	0.003 (0.18)
Executives and Intellectual Professions		0.09 (0.09)	0.37 (0.20)	-0.12 (0.20)
Technical Professions		0.29** (0.08)	0.70** (0.17)	0.45** (0.17)
Skilled Workers		0.14 (0.08)	0.42** (0.16)	0.07 (0.16)
Unskilled Workers		-0.22** (0.08)	-0.39* (0.16)	-0.38* (0.16)
Occupation Unknown		-0.10 (0.14)	-0.30 (0.29)	-0.12 (0.28)
Employed in the public sector		0.17** (0.06)	0.36** (0.13)	0.29* (0.13)
Other Variables				
High Density of Physicians		-0.51** (0.05)	-0.90** (0.10)	-1.18** (0.10)
Low Density of Physicians		-0.12 (0.09)	-0.22 (0.18)	-0.37* (0.17)
Constant		0.26 (0.27)	-2.77** (0.59)	0.61 (0.55)
Number of Observations		8,161	7,902	
Log-Likelihood		-1,964	-6,314	
Pseudo R ²		0.150	0.128	

** : significant at the 1% level, * : 5% level. Standard errors in parentheses.

For the probit model, the dependent variable is equal to 1 for individuals with private supplemental insurance and zero for those without. No insurance is also the reference category for the multinomial logit.

In contrast, we do find a significant relationship between smoking behavior and coverage. Evaluating the probit function at the sample means implies that heavy (light) smokers are 3.5 (2.0) percentage points less likely than non-smokers to have private insurance. One possible interpretation is that smokers have a higher tolerance for risk and/or place a lower value on their health than non-smokers. However, answers to subjective questions that were addressed to only a subset of ESPS respondents seem to argue against the idea that self-selection on the basis of ‘tastes’ is an important factor more generally. These questions ask respondents to state what influence they believe medical care and other factors (work conditions, family tensions and the environment) have on health. Individuals with and without supplemental coverage answer these questions quite similarly. In particular, there is no statistically significant difference in the percentage saying that medical care is either important or very important.

The other coefficients confirm that socioeconomic status has a significant impact on the probability of having supplemental insurance. Coverage increases with income and, relative to working adults, is lower for the unemployed and higher for retirees. The marginal effect of age on any supplemental insurance coverage is quadratic, with the probability of coverage increasing until age 43 and declining thereafter. However, this result is somewhat deceiving given the large positive effect of being retired. The multinomial logit results reveal that age is not significantly related to individual purchases of insurance relative to having no supplemental coverage. Since medical care utilization increases with age, this also runs counter to the notion that this market is subject to adverse selection.

2.6.2. The effect of supplemental insurance on physician visits

The probability of having any physician visits over a one month period is modeled as a function of whether or not the individual has supplemental insurance coverage plus the same demographic and health status variables that entered the coverage equation:

Equation 2.6-ii

$$V^* = X'\beta_v + H'\gamma_v + \theta.C + u_v$$

$$V = 1 \text{ if } V^* > 0$$

$$V = 0 \text{ Otherwise}$$

The parameter of interest is γ , the coefficient on the supplemental insurance variable. As noted, we take two approaches to account for potential bias related to adverse selection. The first is to allow for separate effects of supplemental insurance that is received through an employer and coverage that is individually purchased. In this case, the latent propensity of having at least one physician visit is given by:

Equation 2.6-iii

$$V^* = X'\beta_v + H'\gamma_v + \theta_E \cdot C_E + \theta_I \cdot C_I + u_v$$

where the subscripts E and I refer to employer provided and individually purchased coverage, respectively. If adverse selection is an issue, we should see $\theta_E < \theta_I$.

The other approach is to estimate equation 2.6-i and equation 2.6-ii jointly as a bivariate probit model. Convincing identification of this model requires variables that can reasonably be assumed to influence coverage, but to have no direct impact on utilization. Variables related to occupation and labor force status are the most viable candidates in the ESPS, with the most plausible instrument being an indicator of employment in the public sector. All public employees are able to participate in mutual insurance contracts designed for civil servants, and most of them take up these contracts; but there is no reason to expect employment in the public sector to have an impact on health status or the consumption of care. In fact, in our data, there is no significant difference in the observed health of public employees and other adults and when this indicator variable is included directly in the utilization regression (controlling for insurance coverage) its coefficient is not statistically different from zero. Variables describing employment status (five categories) and professional occupation (eight categories) are also candidate instruments which have been used previously (Vera-Hernandez, 1999). table 2.6 ii presents the results for two single-equation probit models (Equation 3.6 ii and equation 3.6 iii). Turning first to the controls, all of the health status variables have a statistically significant effect on utilization, which suggests that these measures capture different aspects of health. All else equal, individuals with the poorest self-assessed health are 9 percentage points more likely to have a visit than those in the baseline health category, who are themselves 10 percentage points more likely to have a visit than individuals with the best self-assessed health. Heavy smokers have slightly lower utilization than non-smokers (a 3 percentage points difference), while former smokers have more (by 5 percentage points). Utilization is significantly higher for women than for men and pregnancy substantially increases utilization by women. Utilization increases slightly with income and education.

Table 2.6-ii: The Probability of Any Physician Visit, Full Regression Results

Coverage		
Covered supplemental insurance	0.423** (0.065) [0.129]	-----
Employment Based Coverage	-----	0.451** (0.068) [0.153]
Individually Purchased Insurance	-----	0.397** (0.067) [0.138]
Demographic Characteristics		
Age	-0.02** (0.01)	-0.02** (0.01)
Age ²	0.00027** (0.0001)	0.00029** (0.0001)
Male	-0.36** (0.04)	-0.35** (0.04)
Married	0.03 (0.04)	0.02 (0.04)
Secondary Education	0.07 (0.05)	0.08 (0.05)
High School Education	0.10 (0.06)	0.10 (0.06)
Higher Education	0.18** (0.06)	0.20** (0.07)
Education Unknown	-0.11 (0.10)	-0.07 (0.10)
Household Income		
€ 305 to 457	0.12 (0.11)	0.12 (0.11)
€ 457 to 610	0.15 (0.10)	0.15 (0.11)
€ 610 to 762	0.14 (0.10)	0.14 (0.11)
€ 762 to 915	0.22* (0.10)	0.21* (0.10)
€ 915 to 1,220	0.17 (0.10)	0.17 (0.10)
More than € 1,220	0.17 (0.10)	0.16 (0.10)
Income Unknown	0.02 (0.10)	-0.00 (0.11)
Health Variables		
Self-Assessed Health = Fair/Poor	0.25** (0.05)	0.27** (0.05)
Self-Assessed Health = Excellent	-0.33** (0.04)	-0.32** (0.04)
Self-Assessed Health = Unknown	-0.14 (0.07)	-0.16* (0.08)
Moderate to High Mortality Risk	0.28** (0.04)	0.27** (0.04)
Disabled	0.26** (0.05)	0.27** (0.05)

Experiencing Frequent Pain	0.12** (0.04)	0.11* (0.05)
Pregnant Woman	1.27** (0.16)	1.27** (0.16)
Tobacco Consumption		
Former Smoker	0.16** (0.04)	0.15** (0.04)
Light Smoker	-0.04 (0.05)	-0.04 (0.05)
Heavy Smoker	-0.10 (0.05)	-0.09 (0.05)
Labor Market Status		
Inactive	-0.08 (0.06)	-0.07 (0.07)
Unemployed	-0.04 (0.07)	-0.06 (0.07)
Retired	0.09 (0.06)	0.07 (0.06)
Students	-0.09 (0.22)	-0.09 (0.22)
Professional Occupation		
Farmers	-0.13 (0.08)	-0.11 (0.09)
Artisans Retailers	-0.04 (0.08)	-0.02 (0.08)
Executives and Intellectual Professions	0.08 (0.06)	0.07 (0.06)
Technical Professions	0.01 (0.05)	0.00 (0.05)
Skilled Workers	-0.01 (0.05)	-0.02 (0.05)
Unskilled Workers	-0.01 (0.06)	-0.01 (0.06)
Occupation Unknown	-0.04 (0.12)	-0.03 (0.12)
Employed in the public sector	0.03 (0.04)	0.03 (0.04)
Other Variables		
High Density of Physicians	-0.04 (0.04)	-0.04 (0.04)
Low Density of Physicians	-0.05 (0.06)	-0.05 (0.06)
Constant	-0.72** (0.22)	-0.69** (0.22)
Number of Observations	8,161	8,161
Log-Likelihood	-4,565	-4,562
Pseudo R ²	0.09	0.09

Notes: The dependent variable is equal to 1 for individuals who have any visits over a month and zero otherwise (probit regression). Standard errors in parentheses. Marginal effect on the probability of utilization in brackets (insurance coefficient only).

** : significant at the 1% level, * : 5% level.

The effect of insurance is large and statistically significant. The results in column 1 show that controlling for individual characteristics, supplemental insurance increases the probability of having at least one physician visit by nearly 13 percentage points, which is roughly 80% of the raw difference reported in table 2.5-i. To put this in some perspective, this effect is approximately twice the magnitude of the regression-adjusted difference in utilization between the highest and lowest income categories, but smaller than the difference in utilization between the least healthy and the healthiest individuals in the sample.

In column 2, the coefficient on employer-provided insurance is slightly larger than that for individually purchased coverage. This is the opposite of what we should find if the relationship between insurance and utilization were being driven by adverse selection. However, like in Genier (1998), the difference between the two coefficients is not different from zero at standard levels of significance. The coefficients on the control variables are essentially identical across the two specifications. The sample size is slightly smaller for the column 2 model because we exclude people for whom there is no information on the source of coverage. The results are qualitatively the same if we include those individuals in the sample and add a dummy variable for source of coverage unknown. Table 2.6-iii presents the main results for bivariate probit models that exclude different combinations of variables from the utilization equation, while maintaining them in the equation for insurance coverage (Note that we include all these variables in the utilization models reported in and none of their coefficients are significant). Since the results for the control variables are essentially the same as in table 2.6-i and Table 2.6-ii, we report only the supplemental insurance coefficient, the estimate of the correlation between the error terms of the coverage and utilization equations and the likelihood ratio for testing the bivariate probit model to the simpler (restricted) model in which coverage and utilization are estimated separately. We report these results for several versions of the model to show the sensitivity of the results to the assumed exclusion restrictions.

Because we believe the theoretical basis of the public employment dummy is stronger than similar across the different specifications. For all versions of the bivariate probit model, the insurance coefficient is larger than the one from the single equation probit model, though less precisely estimated. The correlation coefficient is negative for all specifications, though it is never statistically significant. Taken at face value, these results would suggest that there is favorable selection into supplemental insurance coverage, which would imply that treating insurance as exogenous leads to an understatement of the true moral hazard effect. However, given the imprecision of the point estimates, a more prudent interpretation is that there is no significant difference between the model that treats insurance as exogenous and the bivariate probit specification.

This interpretation is supported by likelihood-ratio tests as well as by an alternative test proposed by Lollivier (2001), which is constructed by estimating separate probit models of insurance coverage and utilization and then computing the score statistic under the null assumption that ρ is equal to zero. The results of this test reject the hypothesis that the error terms from the two probit equations are correlated.

The pattern of results from the bivariate probit models as well as from the models that use the source of coverage to test for adverse selection are consistent with the idea that in France supplemental insurance increases utilization by lowering the out-of-pocket cost of care. This view is also supported by responses to a question administered to a subsample of ESPS respondents about whether in the year prior to the survey the person had foregone seeking care for financial reasons. Individuals without supplemental insurance are substantially more likely to report having foregone care than those with insurance (32% vs 12%) and the difference remains significant after controlling for other individual characteristics. We also estimated several other regressions which are not reported, but which we will discuss briefly here. First, we estimated separate probit models of the probability of having any GP or specialist visits. As with the results for any visits, for both of these outcomes, the regression adjusted effect of insurance is nearly as large as the raw difference reported in table 2.5-i. The results for the control variables are also qualitatively similar to those in table 2.6-ii. In particular, all the coefficients on the health status variables are statistically significant with the expected sign. We find no statistically significant difference between the effects of employer-provided and individually purchased insurance. Likewise, bivariate probit results for the probability of having at least one GP or specialist visit contradict the adverse selection hypothesis and give the same qualitative results as the ones described above.

Table 2.6-iii: Partial Results for Alternative Probit Bivariate Models

Variables Excluded from Utilization Equation	Insurance Coeff. (std. error)	ρ (std. error)	Likelihood Ratio
1. None	0.423 (0.065)	-----	-----
2. Public employee	0.690 (0.354)	- 0.142 (0.192)	0.554
3. Public employee + Labor market status	0.847 (0.275)	- 0.229 (0.158)	2.089
4. Public employee + Professional occupation	0.768 (0.304)	-0.184 (0.170)	1.178
5. Public employee + Labor market status + Professional occupation	0.920 (0.237)	- 0.271 (0.142)	3.636

Notes: The sample size for all models is 8,161

The labor market status variables are indicators for the following categories: not in the labor force, unemployed, retired, student (the omitted category is actively employed). The professional occupation variables are indicators for the following categories: executive, technical professional, skilled worker, artisan retailer, agricultural worker, unskilled worker, profession unknown (the omitted category is other occupations). The model reported in row 1 includes all the potential instruments in the utilization equation. The likelihood ratios are calculated using restricted models that correspond to the bivariate models

We also estimated several other regressions which are not reported, but which we will discuss briefly here. First, we estimated separate probit models of the probability of having any GP or specialist visits. As with the results for any visits, for both of these outcomes, the regression adjusted effect of insurance is nearly as large as the raw difference reported in Table 2.5-i. The results for the control variables are also qualitatively similar to those in Table 2.6-ii. In particular, all the coefficients on the health status variables are statistically significant with the expected sign. We find no statistically significant difference between the effects of employer-provided and individually purchased insurance. Likewise, bivariate probit results for the probability of having at least one GP or specialist visit contradict the adverse selection hypothesis and give the same qualitative results as the ones described above.

2.6.3. Insurance and the choice of physician

As noted above, since cost sharing is generally greater for specialty as compared to primary care, supplemental insurance might also influence the type of physician a patient chooses to see. We test for this type of effect by taking the physician visit as the unit of observation and estimating a probit model of the decision to see a specialist rather than a GP.

Equation 2.6-iv

$$S^* = X' \beta_s + H' \gamma_s + \delta_s \cdot R + \theta_s \cdot C + u_s$$

$$S = 1 \text{ If } S > 0$$

$$S = 0 \text{ otherwise}$$

In (4) the observed dependent variable, S , equals one if a visit was to a specialist and zero if it was to a GP. The independent variables include the demographic and health variables that appear in the utilization regression plus a dummy variable, R , for whether the visit was referred by a physician. The logic behind this variable is that many visits to specialists will occur because of a referral from a primary care physician, but GPs do not tend to refer to other GPs.

A key issue here is that the choice between the two types of providers is not relevant for all medical conditions that prompt a visit. In many situations visiting a GP is the only reasonable option, whereas other conditions lie clearly within the domain of a specialist. In these extreme cases, while financial considerations may affect the decision of whether or not to seek care, conditional on seeking care there is little reason to expect an effect of insurance on the type of provider chosen. Thus, if (4) is

estimated on a sample of all visits the estimate of θ_s represents an average of a zero effect for some conditions and a possibly non-zero effect for others.

A more informative test, therefore, is to examine the effect of supplemental insurance on the choice of physician type for conditions where either a GP or specialist could provide the care. We identified such cases according to observed patient behavior in earlier ESPS surveys (1992–1995). In those data, we group all visits by the ICD-9 code for the condition prompting the visit. For each ICD subchapter we calculated the proportion of visits that took place in a specialist’s practice and sorted the data according to this measure (excluding subchapters for which the number of visits was too small to get a reliable estimate). Then, we distinguished three groups of medical conditions: those which are treated by a specialist in less than 30% of cases, those treated by a specialist in more than 70% of cases, and those in between, for which we considered the possibility of substitution to be most likely. (Results do not vary when these arbitrary thresholds are changed.) Thus, in addition to analyzing the full sample of visits (3591 over a month), we estimated a model on the sub-sample that falls in this middle category (1402 visits).

The full results are reported in table 2.6-iv. As expected, referred visits are much more likely to take place at a specialist’s. For the most part, other control variables have limited impact on the type of provider chosen. The effect of insurance is stronger when we limit the analysis to visits where substitution between specialists and GPs seems more plausible. However, even there, the impact of insurance is not significantly different from zero.

Table 2.6-iv: insurance coverage and the choice of provider, full regressions results

	Probability that the Visit Takes Place at a Specialist’s	
	All Visits	Visits with possible substitution
Covered by supplemental insurance	0.04 (0.12)	0.19 (0.21)
Visit referred by a physician	1.22 ** (0.06)	1.18 ** (0.10)
Demographic Characteristics		
Age	0.028** (0.011)	0.01 (0.018)
Age ²	-0.0004** (0.0001)	-0.0003 (0.0002)
Male	-0.20** (0.06)	-0.29 ** (0.10)
Married	-0.14* (0.06)	-0.23* (0.10)
Secondary Education	-0.01 (0.07)	0.11 (0.12)
High School Education	-0.00	0.18

Probability that the Visit Takes Place at a Specialist's		
	All Visits	Visits with possible substitution
	(0.09)	(0.14)
Higher Education	0.05	0.11
	(0.10)	(0.16)
Education Unknown	-0.05	-0.35
	(0.16)	(0.30)
Household Income		
€ 305 to 457	-0.11	-0.06
	(0.17)	(0.29)
€ 457 to 610	0.01	-0.10
	(0.17)	(0.28)
€ 610 to 762	0.12	-0.06
	(0.16)	(0.27)
€ 762 to 915	0.05	-0.12
	(0.16)	(0.27)
€ 915 to 1,220	0.20	0.07
	(0.16)	(0.27)
More than € 1,220	0.30	0.24
	(0.16)	(0.27)
Income Unknown	-0.04	-0.11
	(0.17)	0.27
Health Variables		
Self-Assessed Health = Fair/Poor	0.03	-0.08
	(0.06)	(0.10)
Self-Assessed Health = Excellent	-0.01	-0.08
	(0.07)	(0.13)
Self-Assessed Health = Unknown	-0.03	0.13
	(0.11)	(0.19)
Moderate to High Mortality Risk	-0.12	-0.30**
	(0.06)	(0.10)
Disabled	0.09	-0.11
	(0.06)	(0.11)
Experiencing Frequent Pain	0.02	0.11
	(0.06)	(0.10)
Pregnant Woman	0.77**	0.09
	(0.14)	(0.58)
Tobacco Consumption		
Former Smoker	0.13*	0.25*
	(0.06)	(0.10)
Light Smoker	0.08	-0.18
	(0.08)	(0.13)
Heavy Smoker	-0.08	-0.18
	(0.08)	(0.14)
Labor Market Status		
Inactive	0.03	0.06
	(0.09)	(0.14)
Unemployed	0.04	-0.14
	(0.11)	(0.18)
Retired	0.09	0.12
	(0.09)	(0.15)
Students	-0.47	-1.09
	(0.32)	(0.63)

Probability that the Visit Takes Place at a Specialist's		
	All Visits	Visits with possible substitution
Professional Occupation		
Farmers	0.04 (0.13)	0.12 (0.20)
Artisans Retailers	0.00 (0.12)	-0.14 (0.21)
Executives and Intellectual Professions	0.11 (0.09)	0.04 (0.16)
Technical Professions	0.05 (0.07)	-0.01 (0.12)
Skilled Workers	-0.08 (0.09)	-0.04 (0.15)
Unskilled Workers	-0.23* (0.10)	-0.25 (0.16)
Occupation Unknown	0.19 (0.17)	0.32 (0.25)
Other Variables		
High Density of Physicians	0.26** (0.06)	0.37** (0.10)
Low Density of Physicians	-0.10 (0.10)	-0.09 (0.16)
Constant	-0.97** (0.33)	0.01 (0.56)
Number of Observations	3591	1402
Log-Likelihood	-2042	-755
Pseudo R ²	0.16	0.21

The lack of a significant effect of insurance on choice of provider may not be surprising given that, as noted previously, even for people without supplemental insurance the difference in out-of-pocket cost for a GP and a specialist is often very small. However, this 'non-result' is subject to several important caveats. First, because it is difficult to identify exactly those cases where a patient could see either a GP or a specialist, our estimates are necessarily imprecise. When we limit the sample to visits where substitution seems most plausible, the point estimate is actually quite large, but is insignificant because of a large standard error. Thus, while the 95% confidence interval includes zero, it also includes very large effects.

Second, since we only observe the choice of provider for people who seek care, these results need to be interpreted in light of the strong effect of insurance on that preceding decision. In particular, since out-of-pocket costs are higher for specialists (if only slightly so) the tendency of uninsured patients to forego care for financial reasons may be greatest in cases where the services in question are those of a specialist. Thus, our results for this conditional decision may be biased toward zero due to self-selection. In principle, one could account for this selection with a Heckman type model, though in our

data there are no variables that can plausibly be argued to affect then decision to seek care but not the choice of provider.

2.7. Discussion

Our main results can be summarized as follows. We find a large and statistically significant difference between French adults with and without supplementary health insurance in the probability of seeing a physician during a one-month period. In contrast, conditional on having at least one visit, we find no statistically significant effect of insurance on the number of visits or the type of physician chosen. However, both of these latter results are subject to important caveats. Given the short time period over which utilization is measured in our data, very few people have multiple visits. With respect to the choice of physician, insurance may have a stronger effect in settings where the difference in out-of-pocket costs for specialists and GPs is more pronounced. Even in the French context, our estimates may be biased downward by the fact that we cannot model the decision to seek care and the choice of provider as joint decisions.

While differences in data preclude precise comparisons, the impact of insurance on the probability of visiting a physician is large relative to the findings of previous studies. In percentage terms, the difference we find between individuals with and without supplemental insurance is of a comparable magnitude of the gap in utilization between Americans with and without any insurance and is larger than the percent effects found in the Rand experiment. This might seem to suggest that our estimates of the causal effect of insurance on utilization are biased by adverse selection.

However, considerable evidence argues against this interpretation. If adverse selection were an important factor, we would expect individuals whose observed health status was poor to be more likely to purchase insurance. Despite having several different measures of health status, we find no evidence of this kind. Similarly, while these health status measures and other demographic variables are strongly related to utilization, only a small part of the insurance-related difference in utilization can be explained by differences in observed characteristics of individuals with and without coverage. Two distinct tests for the effect of unmeasured characteristics also provide no support for the adverse selection hypothesis, while auxiliary regressions point to out-of-pocket costs as a factor reducing the utilization of individuals without private insurance.

Because the public system covers the majority of the cost of outpatient care, the moral hazard effect of private insurance leads to higher public expenditures for medical care. Since private insurance typically covers the cost of the *“ticket modérateur”*, it largely negates the impact of that policy, which was introduced to control spending by reducing utilization. Because in France private health

insurance coverage increases with income those patients who must pay the “*ticket modérateur*” tend to have lower than average incomes. Thus, the moral hazard effect we find has implications not only for efficiency, but also contributes to income-related inequality in access to care.

The magnitude of this effect is illustrated in table 2-7-i, which reports actual utilization for four income groups and predicted utilization for each group under the assumption that everyone had private supplemental insurance. The latter figures are estimated based on the regression results reported in table 2.6-ii. The figures show that in 1998, individuals in the highest income category were 27% more likely to have at least one physician visit during a one-month period than individuals in the lowest income group (33.6% vs 26.5%). Extending supplemental insurance to the entire population would increase mean utilization for all groups, but the change would be larger for the lower income groups because of their lower baseline level of insurance. Thus, achieving universal coverage in this way would reduce the difference in utilization between the highest and lowest income groups from 7 percentage points to 5.6 percentage points. The results of this simulation suggest that the CMU, which was implemented in 2000, should have increased utilization for the target population, though the effect on overall inequalities in access will be relatively modest. The actual impact of the CMU is an important subject for future research.

Table 2.7-i: The Effect of Extending Supplemental Insurance Coverage on Income-Related Differences in the Utilization of Physician Services

	% of Sample	Actual Utilization: [% with ≥ 1 visit]	Predicted Utilization: Universal Coverage	Change (Predicted-Actual)
A. Income/CU ≤ 457€	13.0%	26.5%	28.6%	2.1%
B. 457€ < Income/CU ≤ 762€	27.0%	29.1%	30.1%	1.0%
C. 762€ < Income/CU ≤ 1220€	33.2%	32.3%	33.0%	0.7%
D. Income > 1220€	26.8%	33.6%	34.2%	0.6%
E. All Adults		30.6%	31.5%	0.9%
D – A		7.1%	5.6%	
D/A		1.268	1.196	

Notes: Actual utilization represents the percent of individuals in each income category with at least one physician visit. Predicted utilization is calculated by predicting for each individual the probability of having at least one visit if covered by supplemental insurance and then calculating the mean of these predictions for each category. Predictions are based on the regression model reported in Column 1 of Table 4.

2.8. Acknowledgements

The authors would like to thank for their helpful comments on earlier versions of this paper: Franck Durand, Randall Ellis, Pierre-Yves Geoffard, Ulf Gerdtham, Christine Meyer, Carol Propper, Lise Rochaix, Martin Schellhorn, as well as the participants to the CREDES research seminar, to the ECuity II Project Workshop in Rome, to the International Health Economics Conference organized at the Hong Kong University of Science and Technology, and SEMINARIS in Paris. An earlier study on a close topic was realized with Karine Szwarzensztein and her input should be acknowledged.

Chapitre 3.

Does free complementary health insurance help the poor to access health care? Evidence from France

Michel GRIGNON (McMaster University: Department of Economics, Department of Health, Aging and Society, Centre for Health Economics and Policy Analysis), Marc PERRONNIN (IRDES), John LAVIS (McMaster University: , Centre for Health Economics and Policy Analysis, Department of Clinical Epidemiology and Biostatistics, Department of Political Sciences).

3.1. Abstract

The French government introduced a free complementary health insurance plan in 2000, which covers most of the out-of-pocket payments faced by the poorest 10% of French residents. This plan was designed to help the nonelderly poor to access health care. To assess the impact of the introduction of the plan on its beneficiaries, we use a longitudinal data set to compare, for the same individual, the evolution of his/her expenditures before-and-after enrolment in the plan. This before-and-after analysis allows us to remove most of the spuriousness due to individual heterogeneity. We also use information on past coverage in a difference-in-difference analysis to evaluate the impact of specific benefits associated with the plan. We attempt at controlling for changes other than enrolment through a difference-in-difference analysis within the eligible (rather than enrolled) population. Our main result is the plan's lack of an overall effect on utilization. This result is likely attributable to the fact that those who were enrolled automatically in the free plan (the majority of enrollees), already benefited from a relatively generous plan. The significant effect among those who enrolled voluntarily in the free plan was likely driven by those with no previous complementary coverage.

Keywords: health-care financing; equity and inequalities; health-care systems; health insurance; natural experiment

JEL codes: I38

3.2. Introduction

The consumption of health care is subject to price and income effects as well as to health-care providers' influence. Awareness of demand-side price effects has led some decision makers in countries with universal health insurance to increase user charges as a way to curb expenditure growth: increasing user charges not only redistributes the financial burden of payment for health care away from the insurer and to the insured, but also reduces the total demand for health-care. A reduction of the average per capita health-care expenditure has often been described as 'beneficial' (Pauly, 2004), though the impact on social welfare ultimately depends on the price and income elasticities in effect (Nyman, 2003) and on the impact of reduced utilization on health outcomes. Increasing user charges among the poorest individuals is particularly likely to lead to sub-optimal levels of medical consumption (Mossialos and Thomson, 2003; van Doorslaer *et al.*, 2004), especially when they cannot afford complementary insurance that covers user charges and is often more available to higher-income individuals (Thomson and Mossialos, 2004). High out-of-pocket expenditures, such as those faced by many Canadians for prescription drugs (Romanow, 2002) and by many working-aged Americans for all types of health care in the US (CBO, 2002), have recently prompted recommendations or legislation to reduce these financial barriers to health care.

Designing a scheme to remove financial barriers to accessing health care raises two kinds of issues: The first is the positive challenge of predicting the impact that a change in insurance coverage has on health-care consumption among the poor (known as the 'access gap' in the US literature (Long and Marquis, 1994). The second is the normative issue of whether the scheme raises the consumption of the poor to the desired level of health care or to levels detrimental to overall welfare.

3.3. Previous literature

This study addresses the first of these two issues: how does coverage influence utilization among the poor? It is standard to simulate the access gap based on a cross-section comparison of health-care utilization among the uninsured and insured populations, with dummies describing individuals' health insurance status (e.g. none, public, employer-sponsored, or individual market), controlling for relevant observable characteristics (health status, race, ethnicity, education level, residence) (Long and Marquis, 1994; Short *et al.*, 1997; Hadley and Holahan, 2003). The set of characteristics defines a propensity score for utilization.

This study contributes to the literature on the access gap by using a quasi-natural experiment that took place in France. In 2000, the French government introduced a free income-tested health insurance plan that differentially affected the coverage of different population sub-groups. We employ a before-and-after design with control to assess the impact of the health insurance plan on

utilization. The use of before-and-after comparisons in longitudinal, quasi-experimental setting better controls for selection bias. Controlling only for observable characteristics as in cross-sectional studies cannot control for any endogeneity of insurance status (or selection effect). If unobservable (to the econometrician) factors drive utilization of the uninsured toward lower levels, the propensity score over-estimates the access gap and the likely impact of removing financial barriers. Controlling for such endogeneity in cross-section data requires an instrumental variables or a full-information maximum likelihood strategy where insurance status is predicted in a first step (Newhouse et al., 1980). However, such a strategy is often not feasible given sample sizes and the complexity of the health-care utilization model (Short et al., 1997; Newhouse et al., 1980). Long and Marquis (1994) recommend experiment designs where previously uninsured individuals are enrolled in a plan and utilization of health-care services is measured before-and-after enrolment. Ideally subjects would be randomly assigned to various plans, such as in the Rand Health Insurance Experiment (Newhouse et al., 1993), but natural quasi-experiments in which a free plan is implemented and its beneficiaries observed before-and-after implementation can also be used to estimate the impact of insurance among the poor, controlling for unobserved heterogeneity (Currie and Thomas, 1995; Evans et al., 2002; Tilford et al., 1999).

The importance of selection-induced endogeneity in health-care utilization is disputed. Shellhorn (2001) and Gardiol et al. (2005) demonstrated that the choice of plan in Switzerland is subject to selection biases (healthier individuals choose higher deductibles and lower premiums), whereas Buchmueller et al. (2004) do not find any strong selection effect in the choice of complementary health insurance in France. To the best of our knowledge, little is known about the insurance choices of the poor in rich countries. Kilbreth et al. (1998) studied the working poor in two US states (Maine and Washington) before and after their inclusion in a highly subsidized state-sponsored plan and could not find evidence of selection bias: enrollees of the state-sponsored plan were not in poorer health before enrolment than a control group comprising individuals covered under their employer-sponsored health insurance plan.

While much existing work compares utilization of the uninsured to that of a population with a mixture of (often unknown) coverage levels and qualities (Long and Marquis, 1994; Short et al., 1997; Hadley and Hohalan, 2003), we study the change in utilization for individuals benefiting from a standard package in the after period. As described below, we also know what type of coverage these individuals had in the before period, albeit less accurately, and, even less accurately, who was eligible to the plan but did not take it up. As a result, we can assess the impact of various levels of

incremental coverage, as well as compare the enrolled population with a control group of similar individuals who did not enroll.

In summary, while the quasi-experiment we exploit is not perfect, since it was subject to some selection bias, we believe it improves importantly on cross-section studies.

3.4. The design of a quasi-experiment: removing financial barriers to access in France

France's universal health insurance system relies heavily on user charges (Thomson and Mossialos, 2004). Patients face user charges when they visit a general practitioner or specialist, stay in hospital, or buy prescription drugs, eyewear glasses, and dental prostheses (Couffinhal and Paris, 2003). Copayment rates vary according to the type of service and overall the social insurance scheme covers only 75% of total health-care expenditures. Individuals can purchase private health insurance on a voluntary basis to cover some of these user charges. Indeed, the primary role for private insurance in France is to cover such user charges (i.e. it is complementary to the social insurance system rather than supplementary or substitutive (Thomson and Mossialos, 2004). Such private complementary insurance is usually funded through flat-rate premiums, which are sometimes subsidized by employers. As a consequence, the poor and the unemployed are less often covered by private insurance and face out-of pocket expenditures when using health-care services.

In 2000, the French Government introduced a free complementary health insurance plan (i.e. Couverture Maladie Universelle Complémentaire, hereafter the "free plan", (van Ingen, 2003; Polton, 2004) to help the non-elderly poor to access health care. The free plan requires no out-of-pocket payments by beneficiaries at the point of use and all physicians, dentists, and opticians must accept patients covered by the plan (though there are concerns that some providers discourage patients covered by the plan to use their services). Individuals with a household income of less than 535€ per adult equivalent per month (587€ in 2005), slightly below France's 'unofficial' poverty line of half the median income, were eligible for the plan in 2000. The equivalence scale weights the first individual in a household as 1, the second 0.5, the third and fourth 0.3, and the fifth and all other individuals 0.4.

The plan does not cover the elderly (i.e. those aged 65 or more) or the disabled because the minimum income they receive from the government is higher than the cut-off.

Almost 4 million people were automatically enrolled when the plan began operation on 1 January 2000. The number of enrollees increased to 5.3 million in June 2001, due to voluntary enrolment. It decreased to 4.7 million in September 2001, when some individuals automatically enrolled in January

2000 failed to re-apply. Enrolment stabilized thereafter and the free plan now covers roughly 7.8% of the population. This is less than the 10% of the population estimated to live under the poverty line.

Three types of transitions to free complementary health insurance were possible for the poor. The first transition, which applies to the very poor, is from an existing means-tested assistance plan to the free plan. The very poor enrolled in the existing means-tested assistance plan (i.e. Aide Médicale Générale, hereafter AMG) moved from facing only a few types of user charges (e.g. the difference between market prices and the agreed-upon fees of the social sickness funds for some specialists, as well as for dental prostheses and prescription eyewear) to facing no user charges for any type of care. The very poor are defined as those who received France's minimum-income benefit and accounted for about 3% of the population at the time of the introduction of the free plan. This group, unlike other potentially eligible citizens, was enrolled automatically in the free plan. After 6 months in the free plan, their eligibility was re-assessed.

The second transition is for those who previously purchased a commercially available complementary health insurance plan and who enrolled in the free plan. Such individuals moved from facing a variety of user charges and then being partially or fully reimbursed by private plan in which they were enrolled to facing no user charges (and paying no insurance premium). Another transition was from having to pay up-front and being reimbursed to no payment at all at the point of use. A survey conducted in 1998 before the introduction of the free plan found that only 52% of individuals in the lowest income bracket (i.e. the poorest 5% of the population) were covered through a commercial plan compared to 85% of the entire population (Bocognano et al., 1999). Private plans often include some forms of cost-sharing (Paris et al., 2003), and it is likely that the coverage purchased by the poor was not one of the best (lowest cost-sharing) plan. Consequently, the free plan improved coverage even for those already covered by a private plan.

The third transition was the most dramatic: those who were not enrolled in either a commercial plan or the means-tested plan and who enrolled in the free plan. Such individuals moved from facing all forms of user charges and being only partially reimbursed later from the social sickness fund to facing no use charges and no up-front payment. As noted above, the user charges included co-insurance on physicians' fees (30%), laboratory tests (40%), prescription drugs (up to 65% depending on the drug), and medical devices (e.g. 35% for dental prostheses and prescription eyewear); co-payments on nights spent in hospital (e.g. h10.70 for one night in a general medicine ward); and the difference between market prices and the agreed-upon fees of the universal social insurance plan for some specialists as well as for dental prostheses and prescription eyewear. On average, user charges

amount to 23% of total expenditures (all types), 8% on hospital care, 28% on physicians' services, and 35% on prescription drugs (Couffinhal and Paris, 2003).

In summary, we can observe the impact of automatic enrolment (without any selection bias) on the very poor, for whom the change in coverage was small, as well as the impact of voluntary enrolment (with potential selection bias) for two groups for which free plan expanded coverage to differing degrees and who had exhibited differing attitudes toward insurance. The analysis examines the behavioural and financial consequences of the free plan on the use of health care (ambulatory, drugs, and hospital) with specific analyses of expenditures on general practitioners, specialists, and prescription drugs.

3.5. Methods and econometric strategy

The study design is built around a series of before-and-after and difference-in-difference comparisons.

For each individual we first calculate average quarterly expenditures for each of three periods:

- period a comprising the first four pre-enrolment quarters;
- period b comprising the last four pre-enrolment quarters;
- period c comprising the post-enrolment quarter(s), which could range from one to three quarters depending on the date of enrolment in the free plan.

Before–after on whole sample: First, we do a simple before–after comparison on the full sample. For this comparison, the change between periods a and b serves as our baseline insofar as it constitutes the 'natural' trend for this population. Following standard practice [e.g. Kilbreth et al. (1998)], the net effect of the free plan is measured as the difference between the change between periods b and c and the change between periods a and b (i.e. $[(c - b)/b] - [(b - a)/a]$).

Comparisons across sub-groups: We conduct separate before–after comparisons for each sub-group defined above by their transition type: automatically enrolled population (smallest expected impact), voluntary enrollees with prior complementary insurance coverage (intermediary expected impact), and voluntary enrollees who had no prior complementary insurance coverage (largest expected impact).

We compare changes in utilization among voluntary enrollees who had no prior coverage and for a comparator group of eligible individuals with no prior coverage but who chose not to enroll. If the difference between these two groups is significant, it means that enrolment had an impact on consumption, independent of any period effect affecting the whole population (Winkelman, 2004).

Finally, we compare changes across regions of residence. Van de Voorde et al. (2001) showed that supply plays a significant role in the consumption of ambulatory care by patients faced with varying degrees of cost-sharing. This region-based comparison is designed to take potential supply effects into account.

Chart 1 (a) and Chart 1 (b) summarize our design and the associated analyses.

Chart 1 (a): Voluntary enrolment starting March 2nd, 2000 - Number of quarters the individual was observed: 11

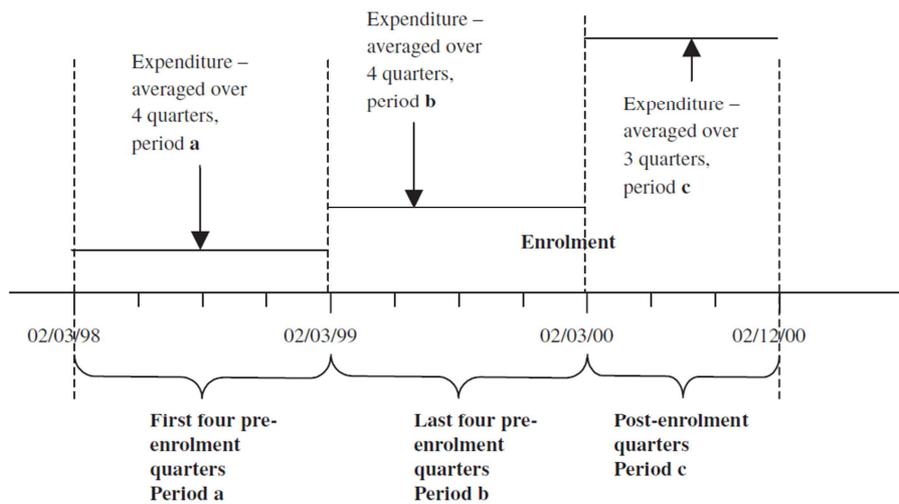
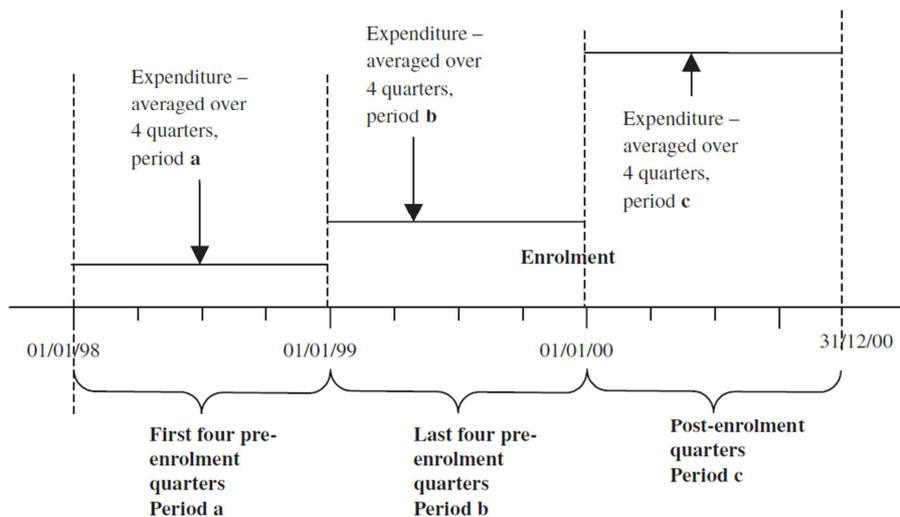


Chart 1 (b): Automatic enrolment starting January 1st, 2000 - Number of quarters the individual was observed: 12



3.5.1. Econometric strategy

As is standard, we employ a two-part model for our utilization equations: part 1 examines the probability of using any health care; and part 2 examines health-care expenditures conditional on having used health care. Although the two-part model offers many advantages, in our context it creates considerable complexity in computing the effect of the free plan on expected total costs and the standard deviations of the estimated differences, a subject to which we return below.

Since our goal is simply to assess the impact of enrolment, we adopt a fixed effect strategy: all determinants of utilization (participation and conditional use) are captured in the fixed individual effect, and we compare utilization across periods a, b, and c. We use the fixed effect model to predict the average quarterly consumptions over all observed individuals, taking the fixed effect into account: $E(Y|t = i)$ for $i = a, b$ or c . In a last step we calculate two rates of increase in these expected values: post-enrolment (c) versus pre-enrolment (b), and pre-enrolment (b) versus pre-enrolment (a). The only control we add to the fixed effect model is the season, which can vary with periods and affect health-care utilization. Although it would be advantageous to include a time-varying measure of health status, information on health is observed only once (in the 1998 survey). We therefore must assume that health status is constant over the observation period, an assumption made more plausible by the fact that the sample includes only young adults and children (individuals over the age of 65 are not eligible for the free plan) and we follow people over only three-year span.

Part 1 regarding participation entails a comparison of probabilities across groups, standardized by season. We employ a linear probability model (LPM) rather than a probit or logit estimation for two reasons. The fixed effects from the LPM are easier to interpret than from a logit. Furthermore, a logit model with fixed effects based on the Chamberlain method would not allow estimating predicted values, which precludes calculating partial effects for utilization or total expenditure. The LPM certainly suffers from two weaknesses. First, it predicts some probabilities outside the unit interval. However, since our purpose here is to estimate the partial effect of the discrete variable 'period' (see above), this is not too important an issue (Wooldridge, 2001). Following Wooldridge's recommendation (Wooldridge, 2001) we checked the validity of our LPM by comparing the partial effects with those estimated from a logit; we also run some validity checks described in the results section. Second, the variances of the estimated coefficients are affected by heteroskedasticity, but since we bootstrap the confidence intervals (see below) this is not a matter of concern.

Analyses based on those who voluntarily enrolled are potentially subject to selection bias. We reduce such bias by including a fixed effect in the regressions to capture the effect of all fixed variables, both those that are observed (e.g. gender) and those that are not observed (e.g. propensity to use health

care). This allows us to measure the impact of unobserved characteristics on changes in utilization. While this is not totally satisfying because the enrolment bias entails an endogeneity issue (the fact that someone enrolls is endogenous to the explained variable) and not simply an issue of heterogeneity, it does offer some control. A better way of controlling it would have been to use an instrument predicting voluntary enrolment that is uncorrelated with planned health-care utilization. This is not feasible, however, because the free plan was implemented simultaneously all over France. Following Buchmueller et al. (2004), we compare the baseline (self-reported) health status of those eligible (according to their self-reported income in 1998) who chose to enroll with that of those eligible who chose not to.

We present the results as the partial effects of the change between two time periods on the probability of using any health care, conditional expenditures, and expected total expenditures. Expected total expenditure is the conditional expenditure times the probability of using any health care. The calculation of the expected total expenditure relies on an estimation of what would have been the conditional expenditure for those individuals who did not use any health care during the period covered. This estimation is not straightforward because the fixed effect of the second-part regression is not known for non-users and is possibly correlated with explanatory variables for those who used health care. Ignoring the fixed effect might bias the estimation of the coefficients in the regressions but integrating it into the smearing estimate may produce an inconsistent estimate of this quantity due to heteroskedasticity. To cope with this problem, we used an auxiliary regression to project the fixed effect on the explanatory variables (season and a dummy variable for two of the three periods) for those who used health care. By applying the linear expression of the projection to the whole population, we were able to calculate an estimate of the fixed effect for all individuals, including those who did not use any health care. Further details of the calculation are available upon request.

Finally, because the estimator is complex (a two-part model with fixed effect), we used a bootstrap method to estimate standard errors and confidence intervals.

3.5.2. Data sources

We use two main sources: the before-and-after on the whole population as well as the comparison between automatic and voluntary enrolment are run on Echantillon permanent des assurés sociaux (EPAS), an administrative database of claims records. When the analysis is conducted on sub-populations defined by prior coverage (AMG, private, none), we rely on a restricted sample where EPAS is merged to the Santé protection sociale (ESPS) survey: we use information in the survey to define prior coverage. The difference in difference analysis comparing enrolled with eligible

individuals who did not enroll is run on a subsample of the merged ESPS–EPAS sample (individuals with no complementary coverage in 1998).

Claims database

The data were extracted from a claims database for a panel of 50 000 enrollees in France's main social sickness funds EPAS for the years 1998, 1999, and 2000. Panel members are drawn randomly from the total pool of enrollees and remain in the sample until their death. New panel members are added every year so that the panel continues to reflect the current population. For each individual, we extracted the following data from the database: health-care utilization, including hospitalizations and expenditures on each of general practitioner services, specialist services, prescription drugs, and prescription eyewear; birth date, area of residence, whether enrolled in the free plan, and (if applicable) date of enrolment in the free plan. The database includes overall expenditure, including out-of-pocket payments (including those reimbursed by a complementary health insurance plan).

Individuals who enrolled in the free plan after 1 October were excluded so that we have at least one post-enrolment quarter for each individual. We also exclude outliers (3% of sample) who had an average quarterly expenditure over 2500€ for at least several consecutive quarters). The rationale for excluding extreme expenditure levels is that we did not want our measure of impact of enrolment to be driven by an extraordinary change on a small number of individuals. The cut-off has been chosen as it is the equivalent of a yearly expenditure of 10 000€ or five times the average expenditure per capita in France in 1999 and it lies two standard deviations above the mean expenditure in our sample (see Table I(b)). Because we use the logarithm of expenditures in the analysis, the exclusion does not affect the findings of our study (which we confirmed by comparing the findings to analyses including the outliers).

Using expenditure data offers several advantages over self-reported utilization. Expenditures reported in claims data provide a much more fine grained and reliable indicator of utilization than self-reported numbers of visits or drugs (physicians are paid fee-for-service in ambulatory care, according to a national fee schedule and prices for prescribed drugs are set nationally (Sandier et al., 2004). For hospital care, private expenditures in claims data derive from bills sent by private hospitals (which constitute approximately 20% of total hospital expenditures). Public hospitals receive a global budget; however, hospitals are required to report on the amount of care they provide to individual patients and claims data contain estimated costs per stay, based on the type of care received (specialty department who treated the patient) and the length of stay using a national cost schedule. This is a standard method for cost calculation under global budget. Importantly for our study, the

costing method remained constant over the period of study and is applied similarly across all hospitals in the country, independent of region or patients' insurance status; hence, it is not a source of bias for our study.

It was not possible to adjust explicitly for inflation because the relevant observations for each individual begin the quarter in which they enroll, which can occur at any point in the year. Because we use a before-and-after design, inflation is only a potential problem for the full-sample before–after design if the rate of inflation differs substantially across the comparison periods (in the difference-indifference analyses, even differential inflation does not pose a problem). Medical prices in France are not subject to considerable central control and track closely with the consumer price index. According to Eco-Sante France (2007), the overall rate of medical inflation remained essentially constant during the period 1998–2000 (+0.8% in 1999 and +1.1% in 2000). Inflation for prescription drugs was 0% in both 1999 and 2000; inflation for physician services was 0.6% in 1999 and 0.7% in 2000; inflation for hospital services was 2.2% in 1999 and 1.3% in 2000. Hospital expenditures is the only cause for possible concern, but even in this case the absolute difference in rates is less than 1%.

Descriptive statistics on utilization of health care by types of services are given in Table 3.5-i.

Table 3.5-i: descriptive statistics on utilization by members of the EPAS panel

(a) Probability of any utilization (quarterly rates) – standard error of the probability are in brackets

	1998	1999	2000
All types of care	0.56 (0.50)	0.62 (0.49)	0.71 (0.46)
Generalist care	0.41 (0.49)	0.46 (0.50)	0.52 (0.50)
Specialist care	0.17 (0.38)	0.19 (0.39)	0.24 (0.43)
Pharmacist	0.48 (0.50)	0.53 (0.50)	0.62 (0.49)
Hospital	0.06 (0.23)	0.06 (0.24)	0.06 (0.24)

(b) Utilization (expenditures in €, quarterly averages) - standard error of the probability are in brackets

	1998	1999	2000
All types of care	219 (845)	262 (982)	292 (998)
Generalist care	21 (42)	23 (40)	26 (41)
Specialist care	14 (53)	15 (61)	18 (65)
Pharmacist	47 (136)	56 (138)	75 (185)
Hospital	102 (760)	125 (900)	126 (903)

Survey data.

The data used in the comparisons across those with differing prior coverage and difference in difference analyses come from a general population survey on health and health insurance conducted in 1998 (ESPS): the sample for the survey is drawn at random from the claims database and we use the data set merging claims with survey information. Three features of the survey are worth noting. First, the survey covers continental France only; it excludes residents of the overseas territories and Corsica, which account for 12–13% of total enrolment into the free plan. In a separate analysis not reported herein, we found that compared to continental residents, individuals living in overseas territories used less health care in 1998 and 1999 and their use of health care was more strongly influenced by the free plan. In the absence of overseas residents from the analyses comparing the effect of the free plan by the nature of prior supplemental insurance coverage will therefore make the net effect of the free plan appear smaller than it actually was.

Second, because only about 1 in 5 panel members in the claims database were surveyed the sample for analyses involving complementary coverage is small and therefore the power to detect differences is lower than we would have preferred.

Third, as with any survey, ESPS potentially suffers from a non-random non-response. Overall, 40% of the random sub-sample completed the survey and can be used in the present analysis. Previous studies have found that those under 30, the poor and those who do not use health care are among those who are less likely to participate in the survey (Auvray et al., 2003), but non-response to the survey would bias these comparisons in the unlikely situation where the non-response rate varied systematically with the type of transition to the plan.

To determine the eligible non-enrolled population (for the difference-in-difference analysis with a control group), we use self-reported household yearly income and the equivalence scale used to determine eligibility (1 for the first individual in the household, 0.5 for the second one, 0.3 for the two following ones, and 0.4 for each individual beyond the fourth one). Using self-reported income yields some imprecision in our sub-population (some individuals we deem eligible are not and vice versa). Moreover, since we use complementary coverage in 1998 to determine our eligible sub-population, we must also rely on their income as reported in the 1998 survey (using the 2000 survey, we would know their income for the year 1999, but not their previous coverage).

Descriptive statistics of sampled populations are given in Table 3.5 ii.

Table 3.5-ii: Descriptive statistics, sampled populations

	Sample : EPAS – enrolled in the free plan	Sample: EPAS – enrolled in the free plan and live in metropolitan France (overseas excluded)	Sample: EPAS and ESPS (general population survey), enrolled in the free plan
Age			
< 20	36.6%	36.4%	36.1%
20-29	14.5%	14.6%	16.4%
30-39	17.0%	17.7%	19.2%
40-49	15.0%	15.9%	16.0%
50 +	17.0%	14.5%	11.5%
Mean Age (in years)	30.3	29.2	28.1
% female	52.7%	52.6%	56.2%
N (size)	2,824	2,283	489
Complementary coverage 1998			
AMG	N/A	N/A	47.7%
Private plan	N/A	N/A	26.0%
No coverage	N/A	N/A	29.9%

Physician supply data. In the analysis of potential supply side effects, we use data on physician supply in each French administrative region drawn from a database on France's health-care system (Eco-Santé France, 2007).

3.6. Results

3.6.1. Baseline: before-and-after analysis (Table 3.6 i)

Before–after analysis: full sample and by enrolment status. The before–after analysis of the full sample reveals that enrolling in the free plan had a positive (+3.2% points) and almost significant effect on the probability of using any health care. The impact on total health-care expenditures is also positive, but the confidence interval of the net effect ($[c - b] - [b - a]$) includes 0, so the effect is not statistically significant. Enrolment leads to a significant and large (+15% points) increase in the probability of using specialist care, but a negative and not significant effect on expenditures for specialist care conditional on seeing a specialist. Enrolment has a small negative and non-significant impact on utilization of GP services. Finally enrolment had a large positive, but again not significant impact on prescription drugs use.

Table 3.6-i: Changes in the probability of using any health care and in health-care expenditure, by type of care received

	Change between periods a and b	Change between periods b and c	Net effect of enrollment
All types of care			
Probability of any use	11.8% [8.0%, 15.2%]	14.9% [11.3%, 18.5%]	3.1%
Conditional expenditure	2.7% [4.4%, 10.2%]	4.8% [-2.7%, 12.8%]	2.1%
Unconditional expenditure	14.6% [5.4%, 23.7%]	20.1% [9.6%, 31.9%]	5.5%
Generalist care			
Probability of any use	17.3% [12.7%, 21.9%]	14.9% [10.3%, 19.5%]	-2.4%
Conditional expenditure	-0.2% [-4.3%, 4.0%]	-0.5% [-4.8%, 3.9%]	-0.3%
Unconditional expenditure	17.0% [10.2%, 23.3%]	14.3% [7.4%, 20.9%]	-2.7%
Specialist care			
Probability of any use	4.4% [-3.9%, 13.1%]	18.8% [9.4%, 28.9%]	14.4%
Conditional expenditure	2.5% [-6.9%, 12.9%]	-7.3% [-16.1%, 2.5%]	-4.8%
Unconditional expenditure	6.8% [-5.0%, 20.0%]	9.3% [-2.9%, 23.5%]	2.5%
Prescription drugs			
Probability of any use	12.1% [8.3%, 16.2%]	17.2% [13.1%, 21.3%]	5.1%
Conditional expenditure	0.7% [-5.1%, 6.8%]	6.9% [0.5%, 13.7%]	6.2%
Unconditional expenditure	12.7% [4.9%, 20.9%]	24.7% [15.8%, 34.6%]	12.0%

Sources: EPAS 1998-2000

Notes: N = 2,686 individuals or 31,593 quarters for all analyses. Period a corresponds to the first four pre-enrollment quarters, period b to the last four pre-enrollment quarters, and period c to the post-enrollment quarters. The probability of any use is the probability of having any type of health-care expenditure during the period covered. The conditional expenditure is the average quarterly expenditure of those individuals who used any health care during the period covered. The unconditional expenditure is the product of the probability of any use and the conditional expenditure. For an elaboration of the issues involved in calculating the conditional expenditure, see Note 23. A change of x% between two periods indicates that the difference between the values in the two periods is divided by the value in the first period. The 95% confidence interval is provided in brackets. None of the net effects of enrollment are statistically significant.

The fixed individual effect, which captures all individual characteristics except season and enrolment in the free plan, is significant in all four analyses and explains 20–50% (depending on the type of services) of individual variance in expenditures. Therefore, unobserved individual characteristics influence the use of health care more than enrolment in the free plan.

3.6.2. Influence of automatic versus voluntary enrolment (Table 3.6-ii)

Table 3.6-ii: Changes in the probability of using any health care and in health-care expenditure, by type of care received and by whether enrollment was automatic or voluntary

	Change between periods a and b	Change between periods b and c	Net effect of enrollment
All types of care			
Automatic enrollment			
Probability of any use	14.3% [10.1%, 18.7%]	14.4% [10.7%, 18.4%]	0.1%
Conditional expenditure	3.8% [-4.6%, 13.1%]	0.4% [-8.2%, 9.9%]	-3.4%
Unconditional expenditure	18.5% [8.3%, 33.1%]	14.7% [3.3%, 28.2%]	-3.8%
All types of care			
Voluntary enrolment			
Probability of any use	0.1% [-6.6%, 7.0%]	29.2% [19.0%, 41.4%]	29.1%
Conditional expenditure	-9.7% [-23.1%, 6.1%]	46.1% [23.5%, 72.8%]	55.8%
Unconditional expenditure	-9.6% [-25.9%, 8.6%]	86.2% [53.9%, 123.6%]	95.8%
Generalist care			
Automatic enrollment			
Probability of any use	20.5% [14.0%, 27.4%]	14.4% [8.8%, 21.8%]	-6.1%
Conditional expenditure	-0.0% [-4.9%, 4.9%]	-3.1% [-8.0%, 20.6%]	-3.1%
Unconditional expenditure	20.4% [12.5%, 29.2%]	10.8% [2.6%, 20.5%]	-9.6%
Generalist care			
Voluntary enrolment			
Probability of any use	2.2% [-7.8%, 14.7%]	33.9% [20.1%, 52.7%]	31.7%
Conditional expenditure	-8.3% [-17.2%, 1.6%]	18.2% [6.2%, 31.6%]	26.5%
Unconditional expenditure	-6.3% [-19.9%, 9.4%]	57.5% [34.5%, 84.7%]	63.8%
Specialist care			
Automatic enrollment			
Probability of any use	6.8% [-3.6%, 17.4%]	13.7% [3.4%, 25.6%]	6.9%
Conditional expenditure	-0.5% [-11.2%, 11.6%]	-10.5% [-3.6%, 17.4%]	-10.0%
Unconditional expenditure	6.0% [-6.5%, 20.8%]	1.2% [-20.5%, 0.8%]	-4.8%
Specialist care			
Voluntary enrolment			
Probability of any use	-9.5% [-27.0%, 14.9%]	87.0% [44.8%, 151.5%]	96.5%
Conditional expenditure	9.5% [-11.1%, 34.9%]	0.8% [-19.6%, 26.4%]	-8.7%
Unconditional expenditure	-0.3% [-25.2%, 36.3%]	81.4% [28.4%, 160.9%]	81.1%
Prescription drugs			
Automatic enrollment			
Probability of any use	14.4% [9.3%, 19.4%]	16.1% [10.7%, 21.6%]	1.7%
Conditional expenditure	0.9% [-6.0%, 8.2%]	4.8% [-2.7%, 12.9%]	3.9%
Unconditional expenditure	15.3% [5.5%, 27.8%]	21.2% [9.7%, 34.9%]	5.9%
Prescription drugs			
Voluntary enrolment			
Probability of any use	-0.0% [-8.2%, 8.6%]	33.7% [22.5%, 50.0%]	33.7%
Conditional expenditure	-4.0% [-15.8%, 9.3%]	26.0% [9.9%, 44.4%]	30.0%
Unconditional expenditure	-4.0% [-15.9%, 9.5%]	64.0% [40.5%, 92.2%]	68.0%

Sources: EPAS 1998-2000

Notes: N = 2,329 individuals or 27,948 quarters for analyses involving individuals who were automatically enrolled in the free plan. N = 357 individuals or 3,645 quarters for those who enrolled voluntarily in the free plan. Period a corresponds to

the first four pre-enrollment quarters, period b to the last four pre-enrollment quarters, and period c to the post-enrollment quarters. The probability of any use is the probability of having any type of health-care expenditure during the period covered. The conditional expenditure is the average quarterly expenditure of those individuals who used any health care during the period covered. The unconditional expenditure is the product of the probability of any use and the conditional expenditure. For an elaboration of the issues involved in calculating the conditional expenditure, see Note 23. A change of x% between two periods indicates that the difference between the values in the two periods is divided by the value in the first period. The 95% confidence interval is provided in brackets. The net effects of enrollment that are statistically significant are shown in bold.

Enrolment has a significant impact when it is voluntary only: it increases the probability of any use, conditional expenditure (except for specialist care) and expected total expenditure on all types of care and for prescription drugs as well as for physicians' services.

This somewhat small overall impact of the plan stems from the fact that those previously covered by AMG who experienced only a minor change in coverage (who constitute 87% of the sample) were not strongly affected: differences in rates of change across enrolment are never significantly different from 0; the only marginally significant impact is a decrease in the probability of using GPs' services. Since we also observe a suggestive (but non-significant) increase in the probability of using specialists' services, we can infer that the only impact of the free plan on the sub-population of those already covered by AMG was a small substitution of specialist care for GP care. This is consistent with the fact that a primary benefit of the free plan for this sub-population was the removal of any extra billing from specialists.

By contrast, among those who enrolled voluntarily, the free plan had a very strong and significant effect on the probability of using any health care and on total health-care expenditure: post-enrolment utilization is twice that of pre-enrolment utilization. The increase is larger for specialists' services and prescription drugs than for GP services. The impact on specialists' services is channeled through the probability of any contact rather than by average expenditure conditional on any contact. Since the voluntarily enrolled population includes individuals with no previous complementary coverage at all, these findings suggest that the co-payments on which the French health-care system relies dramatically reduced health-care utilization by the poor.

3.6.3. Difference-in-difference analyses

Influence of previous supplemental insurance coverage: difference-in-difference analysis within the enrolled population (Table 3.6-iii)

Table 3.6-iii: Differences across nature of complementary insurance coverage before enrolment in the free plan (1998) in changes in health care utilization due to enrolment (changes between periods b and c)

	Change between periods b and c – previously enrolled in a commercial plan	Change between periods b and c – previously without any coverage
	compared to change between periods b and c – previously enrolled in AMG	compared to change between periods b and c – previously enrolled in AMG
All types of care		
Probability of any use	+18.1 (S)	+20.7 (S)
Conditional expenditure	+7.3 (NS)	+41.5 (S)
Expected total expenditure	+25.3 (NS)	+68.7 (S)
Generalist care		
Probability of any use	+18.9 (S)	+30.6 (S)
Conditional expenditure	+15.1 (NS)	+6.3 (NS)
Expected total expenditure	+35.3 (S)	+36.4 (NS)
Specialist care		
Probability of any use	+17.5 (NS)	+56.7 (S)
Conditional expenditure	-0.6 (NS)	+20.6 (NS)
Expected total expenditure	+14.3 (NS)	+79.8 (S)
Prescription drugs		
Probability of any use	+13.4 (NS)	+20.9 (S)
Conditional expenditure	+13.3 (NS)	+23.6 (MS)
Expected total expenditure	+28.6 (NS)	+49.5 (S)

Reference = AMG (means-tested plan) in 1998 – all values are expressed in percentage points: a difference of X means that if the relative change for those enrolled in AMG in 1998 was A%, then the relative change for those in the column was (A+X)%; NS stands for non-significant, S for significant and MS for marginally significant (confidence intervals at the 95% level overlap but the overlap is smaller than 1 percentage point).

Sources: EPAS 1998-2000, ESPS 1998

Notes: N = 101 individuals or 1,166 quarters for analyses involving individuals with a commercial plan in 1998. N = 237 individuals or 2,835 quarters for analyses involving individuals with a means-tested plan (AMG) in 1998. N = 137 individuals or 1,607 quarters for analyses involving individuals with neither a commercial plan nor a means-tested plan in 1998.

According to their self-reported status in 1998, the proportion of voluntary enrolled (versus automatically enrolled beneficiaries of AMG) should have been higher in 1998 than what we observed in 2000 $((101+137)/357 = 0.67)$. We offer four explanations for this difference: 1) survey respondents may not have been aware that they were covered by the means-tested plan (AMG) in 1998; 2) survey respondents were drawn from continental France only and these individuals had lower rates of eligibility for the means-tested plan than individuals living in overseas territories; 3) survey respondents were less likely to include the poor, who are eligible for the means-tested plan; 4) a higher proportion of the population were eligible for the free plan than were eligible for the means-tested plan; and 5) the annual rate of entry for individuals who were enrolled automatically in the means-tested plan (i.e., those who became eligible for France's minimum-income benefit) is 25%. Caisse Nationale des Allocations Familiales, Direction des Etudes et de la Recherche, "Légère hausse des bénéficiaires du RMI au 30 juin 2002," *L'Essentiel* (October 2002), <http://www.caf.fr/coupdœil.htm> (accessed 15 December 2003). Period a corresponds to the first four pre-enrollment quarters, period b to the last four pre-enrollment quarters, and period c to the post-enrollment quarters. The probability of any use is the probability of having any type of health-care expenditure during the period covered. The conditional expenditure is the average quarterly expenditure of those individuals who used any health care during the period covered. The expected total expenditure is the product of the probability of any use and the conditional expenditure. A change of x% between two periods indicates that the difference between the values in the two periods is divided by the value in the first period. The 95% confidence interval is provided in brackets.

The comparison of changes among those automatically enrolled against the voluntary enrollees reveals substantial differences in responses. This difference-in-difference analysis clearly demonstrates a large and significant impact on the utilization of specialists' care (the impact of enrolment is 80% points greater for those without coverage than for those with AMG) and prescription drugs (50% points). Enrolment in the plan has approximately the same impact on expenditures on GPs' services for those without coverage and those with AMG, but increases the probability of any visit to a GP larger among the former group.

We checked that such a difference-in-difference did not exist before (between periods a and b, results not reported here): there was a positive, small, but significant difference between these two sub-populations in the probability of using GPs' services and of having any prescription drug. However, these differences are very small compared to those observed between the same sub-populations for period b and c.

Coverage from commercial plans for the poor appears to be lower than what AMG provided: almost all differences are positive, suggesting that enrolling in the free plan from a commercial one had a stronger impact than enrolling from AMG. However, the only significant differences are for GPs' services. Our results suggest that some commercial plans do not guarantee access to GPs' services.

Difference-in-difference analysis within the eligible population

We observe a strong and significant impact of enrolment on the population with no complementary coverage before being enrolled in the free plan. However, such a change could be coincidental and stemming from changes in the period 1999-2000 affecting this population with no complementary coverage and low income (e.g. specialists offering lower fees to these individuals).

To control for such a possibility, we run a difference-in-difference analysis within the eligible population: we compare the rates of change in health care consumption of individuals with no

previous coverage who enrolled in the plan to that of individuals with no previous coverage, a comparable income level (eligible to the free plan according to their income level in 1998) but who did not enroll in 2000. For these individuals, period a is calendar year 1998, period b is calendar year 1999 and period c is calendar year 2000.

Table 3.6-iv: Differences between eligible (in 1998) not enrolled individuals and eligible (in 1998) enrolled individuals, not covered by any complementary insurance in 1998 in changes in health care utilization due to enrolment (changes between periods b and c)

All types of care	
Probability of any use	+26.0 (S)
Conditional expenditure	+36.8 (MS)
Expected total expenditure	+69.3 (S)
Generalist care	
Probability of any use	+47.2 (S)
Conditional expenditure	+ 2.9 (NS)
Expected total expenditure	+49.0 (S)
Specialist care	
Probability of any use	+55.0 (NS)
Conditional expenditure	- 0.5 (NS)
Expected total expenditure	+58.1 (NS)
Prescription drugs	
Probability of any use	+32.1 (S)
Conditional expenditure	+30.3 (MS)
Expected total expenditure	+66.6 (S)

Notes: N = 137 individuals or 1,607 quarters for analyses involving enrolled individuals with neither a commercial plan nor a means-tested plan in 1998. N = 133 individuals or 1,596 quarters for analyses involving non enrolled but eligible individuals with no previous coverage.

As table 3.6-iv shows, the differences in the rates of change between these two groups are always large for probability of any use and total expenditures. It is large also for conditional expenditure on all types of services, specialists and prescription drugs. The difference is small on conditional expenditures on GPs' services (for which we did not observe any impact of enrolment) and zero on conditional expenditures on specialists' services (for which the impact of enrolment was strong but imprecise).

The differences for specialists' services are never significant, even though they are strong and positive for probability of any use and total expenditure.

Selection bias: influence of baseline health status on enrolment

We compare health status (self-assessed health) in 1998 of those who chose to enroll (for all previous coverage status) with that of those eligible in 1998 who chose not to.

Table 3.6-v: self-assessed health according to enrolment status in 2000, among individuals in the survey deemed eligible in 1998

Self-assessed Health (from 0 to 10)*	enrolled in 2000	not enrolled in 2000	T-VALUE	TTEST
Unknown	20.6%	24.9%	1.4	NS
note<=5	10.1%	7.0%	-1.5	NS
note=6,7	9.7%	9.3%	-0.2	NS
note=8,9	33.5%	24.4%	-2.8	1%
note=10	26.2%	34.5%	2.4	5%

* 0 stands for the lowest level of self-assessed health, 10 for the highest level of self-assessed health.

Table 3.6-v clearly shows that the distribution of self-assessed health status is the same in both subpopulations, indicating that selection bias is not likely to be linked to need. The only significant differences are across very good and excellent health, but there is no reason to think that this would affect dramatically the decision to enroll in the plan.

Supply-side effects

Last, we checked that our findings were not the results of opportunistic inducement of utilization on the supply side. To explore how the net effect of enrollment in the free plan varied by the regional supply of physicians, we examined the net effect of enrollment separately for those who lived in regions with a high density of physicians and for those who lived in regions with a low or medium density of physicians. Our findings (not presented here but available upon request) show that this should not be the case since enrollment significantly increases the probability of any use of specialist care as well as expenditures (any use and conditional expenditures) on prescription drugs in regions with low or medium density of physicians, but not in regions with high density of physicians as would be expected under the supplier-induced demand were present.

Validity of the participation approximation

We verified the validity of the LPM in our setting by checking the proportion of predictions outside the 0–1 range: according to the sub-population, the model predicts between 11 and 15% probabilities out of the unit interval, the only exception being specialist care (22% out of the unit interval). However, the proportion of predicted probabilities out of the [0.05; 1.05] range is never above 10%. We also re-ran the whole estimation setting the predicted participation probabilities at 0 when the fitted value was negative and at 1 when it was greater than 1. This does not change our conclusions (tests results are available on request).

3.7. Discussion

The lack of a strong net effect from a free complementary health insurance plan for the average non-elderly poor person in France can likely be attributed largely to the fact that the majority of enrollees (87% of our sample) already benefited from a relatively generous means-tested plan, therefore enrollment meant for them an incremental change only. For those individuals already covered by the means-tested plan, the marginal benefit of being automatically enrolled in the free plan -- the removal of a few forms of out-of-pocket payments at the point of use -- appears not to have been large enough to influence their care-seeking behaviour. The only exception might be the case of specialists' care for which an increase in utilization can be seen on the overall enrolled population.

Individuals who enrolled voluntarily in the free plan had a significantly higher probability of using all types of care, generalist care, specialist care and prescription drugs after enrollment, as well as increased health-care expenditure for all types of care, generalist and specialist care, and prescription drugs.

The net effect of enrollment among those who enrolled voluntarily in the free plan was likely driven by those with no previous supplemental coverage prior to enrollment rather than by those previously covered by a commercial plan. Individuals with no supplemental coverage constituted 58% of our sample of voluntary enrollees and they moved from facing all forms of out-of-pocket payments at the point of use and being only partially reimbursed later from the social sickness fund to facing no out-of-pocket payments at the point of use. Because coverage under the free plan is near the median for coverage under commercial plans, the marginal benefit of voluntary enrolment for individuals with a commercial plan was much smaller: no longer having to pay the premium for the commercial plan and no longer being temporarily out of pocket for health-care payments until they were reimbursed by the social sickness fund and their commercial plan.

This impact of the free plan on health care utilization of those previously not covered has three causes: (1) A true elasticity of demand for health care to its price among the poor: faced with a lower (indeed zero) price, individuals use much more care, mostly specialist visits and drugs than when faced with a variety of co-payments averaging 23%. (2) Pent-up demand: the change in utilization among those previously not covered reflects the slope of their demand as well as the stock of past unmet needs and can therefore overestimate the true elasticity of demand. Of course, pent-up demand is really another (long term) component of the effect of co-payment on health care utilization among the poor. Its effect is not expected to last in the long run. (3) Enrollment bias: those who voluntarily enroll may be those who expect to use health care more. Because we use repeated observations of the same individual (from 10 to 12 quarters) and a fixed individual effect, we can control for any idiosyncratic propensity to use health services differently, as long as this propensity is constant across enrolment. One strong reason why the decision to enroll could correlate with a change in consumption across enrolment (therefore not accounted for by our fixed effect) could be initial health: those in poor initial health could simultaneously decide to enroll more often (because they benefit more from the plan) and to increase their consumption; if such were the case we would overestimate the causal impact of the plan on health care utilization and expenditure. We are able to show that enrolment is not linked to health status in 1998. However, we cannot rule out that another individual characteristic (other than health and not observed in our study) could explain both the decision to enroll and to change health care utilization.

There are three main limitations to our study: having no data on health outcomes, having only one year of data on both ambulatory and hospital care after the free plan was introduced, and having to rely on merged data from a general population survey to examine how the net effect of enrollment in the free plan varied by the nature of previous supplemental insurance coverage. With no data on health outcomes, we could not examine the health consequences of introducing the free plan. With only one year of data after the free plan was introduced, we could not examine whether the enrollment bias has a transient impact only, as would be the case if individuals remain enrolled in the free plan as long as they remain eligible for it, or a permanent impact, as would be the case if individuals did not maintain their enrollment when they no longer expected to use health care. As well, we could not examine the full effect of the free plan on dental prostheses and prescription eyewear because (for technical as well as political reasons) many dentists and opticians were not ready to welcome those enrolled in the free plan until the end of the plan's first year. Relying on merged data from a general population survey for one type of analysis created problems because the resulting sample was small and excluded individuals living overseas, whose use of health care was much more influenced by the free plan than individuals living in continental France.

Chapitre 4.
**Supplemental health insurance
and health care consumption:
A dynamic approach to
the moral hazard phenomenon**

Carine FRANC (Cermes3, Cnrs UMR 8211, Inserm U988, EHESS), Marc PERRONNIN (IRDES), Aurélie PIERRE (IRDES).

4.1. Abstract

We analyze the existence and persistence of moral hazard over time to test the assumption of pent-up demand. We consider the effects of supplemental health insurance provided by a private insurer when added to compulsory public insurance already supplemented by private insurance. Using panel data from a French mutuelle, we compute error component models with the Chamberlain specification to control for adverse selection. By separating outpatient care consumption into (1) the probability of healthcare use, (2) the number of uses conditional on use and (3) the per-unit cost of care, we provide evidence that supplemental insurance is significantly and positively associated with (1), (2) and (3). However, these effects decrease significantly over time. This pattern supports the existence of pent-up demand, the magnitude of which varies greatly and depends on the dimensions (1), (2) and (3) and the type of care (physician care, prescription drugs, dental care or optical care).

Keywords: Supplemental health insurance, moral hazard, health care expenditures, longitudinal analysis

JEL codes: D82, I13

4.2. Introduction

Economic theory predicts that, by reducing the price of healthcare at the point of use, increased health insurance coverage encourages individuals to increase their consumption of care (assuming that healthcare is an ordinary good). This prediction results from changes in the individual tradeoffs that can be illustrated utilizing two classical, conceptual effects highlighted by Slutsky (1915). Both of these effects lead to an increase in consumption even if they have different welfare effects (Pauly, 1968). First, the substitution effect results from a change in relative prices: when a good is less expensive than other goods, individuals face incentives to purchase this good more frequently (given that their income remains unchanged). Second, the income effect, which results from an increase in purchasing power, allows consumers to acquire larger quantities of all types of goods, including the good for which the price has decreased. These two effects are often associated with moral hazard. However, in the healthcare context, increased purchasing-power (for instance, due to a copayment reduction) enables sick individuals to purchase valuable healthcare that they could not afford without health insurance (De Meza, 1983; Nyman, 1999a). Thus, the income effect of improving access to healthcare can positively affect equity.

Numerous empirical studies have estimated the effect of health insurance on inpatient and/or outpatient healthcare consumption. However, endogeneity bias may occur as a result of adverse selection; indeed, individuals who decide to purchase health insurance may also have had higher

healthcare expenditures prior enrollment. Estimates obtained from the Rand health insurance experiment (HIE) are widely acknowledged. Newhouse and the Insurance Experiment Group (1993) obtained price elasticity values of approximately -0.2 for inpatient and outpatient care. In addition to this experimental research, different modeling strategies have been used to control for adverse selection. Several studies suggest using instrumental variables (for example, Ettner, 1997; Albouy and Crepon, 2007), whereas others suggest observing the exogenous change in copayments in healthcare plans (Cherkin *et al.*, 1989; Chiappori *et al.*, 1998). Regardless of the strategy used to control for adverse selection, nearly all of the US studies have found negative price elasticities for inpatient and/or outpatient care (Cutler and Zeckhauser, 2000; Zweifel and Manning, 2000). This consensus is not observed in European studies. For example, Schokkaert *et al.*, (2010) and Bolhaar *et al.*, (2008) analyze the influence of complementary health insurance on hospitalization and outpatient care in Belgium and Ireland, respectively, and do not observe significant effects (except in dental care). In France, Caussat and Glaude (1993) and Albouy and Crepon (2007) demonstrate that coverage by complementary health insurance has no influence on hospital care consumption but has a significant, but weak, influence on the probability of using outpatient care. This result appears to be consistent with the results of previous studies using French data (Genier, 1998; Chiappori *et al.*, 1998; Buchmueller *et al.*, 2004; Raynaud, 2005). However, moral hazard and its expected magnitude depend on both the operationalization of the concept itself and the institutional context of health insurance provision.

Different approaches have been adopted to analyze the relevance and magnitude of moral hazard. Some studies have analyzed the level of expenditure conditional on utilization (for instance in France, Caussat and Glaude, 1993); others consider the probability of initiating an episode and the frequency of episodes (in France, Genier, 1998). Some studies have analyzed the persistence of a potential increase in healthcare consumption over time (Long *et al.*, 1998; Keeler *et al.*, 1982; Newhouse *et al.*, 1982). However, considering the effects of health insurance from a dynamic perspective is crucial because even if health insurance improves access to care in a sustainable manner, its effects on consumption patterns can change over time. This is the assumption incorporated into the theory of pent-up demand: a strong, immediate and temporary increase in healthcare expenditures occurs just after enrollment, assuming that the effect of health insurance decreases over time. This theory also assumes that newly insured individuals may have either postponed or avoided healthcare consumption prior to obtaining the new plan, which explains the difference in consumption immediately before and after subscription. Empirical studies have failed to give strong evidence to this theory. For instance, Long *et al.*, (1998) find little support for this behavior among the newly insured. As early as 1982, Keeler *et al.* find relatively little evidence of pent-up demand after a year,

whereas Newhouse *et al.*, (1982) find evidence of this pent-up demand for medical care at both the beginning and the end of the three-year Health Insurance Experiment.

In France, a universal public health insurance (PHI) program, which is compulsory and uniform for all individuals living legally in the territory, covers a broad range of medical products. Since the institution of Social Security (1945), re-insurable copayments for any service have been implemented and are computed as percentages of regulated prices. Moreover, several providers are allowed to charge extra fees in addition to the regulated fees (Buchmueller and Couffinhal, 2004). The sum of copayments and extra fees may be high, particularly for dental and optical care. Private complementary health insurance (CHI), offered by *mutuelles*, provident institutions and private insurance companies, is primarily purchased to cover these costs. Even though CHI is not compulsory, more than 90% of individuals eligible for PHI benefit from CHI. In the French context, most studies have analyzed moral hazard by focusing on the CHI coverage effects while CHI contracts are very heterogeneous in terms of both prices and warranties.

In this paper, we study the existence and persistence of moral hazard over time. We analyze the effects of an extra complementary health insurance (ECHI) contract provided by a French private health insurer (*mutuelle*) in addition to PHI (uniform and compulsory) and to basic complementary health insurance (BCHI) optional and uniform for all insured of this *mutuelle*. To understand the manner in which healthcare expenditures are eventually modified by ECHI, we determine the influence of insurance on all of the determinants of healthcare expenditures: (1) do the ECHI insured have an increased probability of utilizing healthcare; (2) do they increase the number of uses conditional on use; and (3) do they purchase more expensive care? To conduct this study, we use error component models by introducing the Chamberlain specification that enables us to control for adverse selection to the greatest extent possible. This paper is organized as follows: in the second section, we present the data and the context. In an extensive third section, we present the economic model and the different residual assumptions that must be defined to ensure that adverse selection is controlled for as much as possible. The fourth section presents the results. In the final section, we discuss the results and conclude.

4.3. Data

4.3.1. Context of dataset

The dataset originates from the administrative records of a non-profit mutuelle, which primarily insures civil servants and their relatives. This mutuelle manages both PHI and CHI. Until July 2003, this mutuelle only offered one uniform and basic complementary health insurance contract (BCHI). Since July 2003, a voluntary extra complementary health insurance (ECHI) contract has been offered in addition to BCHI such that any BCHI enrollee has the opportunity to individually (within a household) purchase ECHI without a specific underwriting process. During the six months preceding the ECHI offer, an informational campaign was launched to spread awareness among the BCHI insured. Table 4.3-i presents the reimbursements respectively provided by PHI, BCHI and ECHI. The additional coverage provided by ECHI mostly concerns physicians' extra fees and copayments for dental and optical care. The ECHI plan was community rated, and the premium was set at 11€ per individual per month (with no charge for additional children beyond two) and remained unchanged during our observation period. Subscribing to ECHI had no effect on the BCHI contract, including the premium.

Table 4.3-i : PHI, Pooling BCHI and ECHI: coverage for different types of healthcare expenses

	PHI	BCHI in addition to PCHI	ECHI in addition to BCHI
Physician care			
Medical practitioners fees (GPs and specialists)	70% RP	+30% RP	+30% RP
Dental care			
Preventive dentistry	70% RP	+30% RP	/
Dental prostheses / orthodontia	70% RP	+140% RP	+105% RP
Other fees (surgical acts, radiology procedures))	70% RP	+30% RP	+30% RP
Drugs			
White vignette drugs	65% RP	+35% RP	/
Blue vignette drugs	35% RP	+60% RP	/
Optical			
Eyeglasses	65% RP	+900% RP	+92€ or +31€ (1)
Frame	65% RP	+55	+61€
Contact lenses	0€ or 65€ RP	+115	+115€
Other prescription			
Auxiliary care	60% RP	+40% RP	/
Orthopedia	65%	+130% RP	/
Acoustic prosthetics	65%	+260% RP	/
Hospitalization			
Medical practitioners fees	80% or 100% RP	+20% or +0% RP	+30% RP
Stay costs	80% or 100% RP	+20% or +0% RP	/
Per diem copayment	/	14€*	/
Other types of care			
Medical transportation	65% RP	+35% RP	/
Spa care (2)	65% or 70% RP	between +20% and +35% RP	/

*10€ for a psychiatric hospitalization and 14€ otherwise

Note: RP = Regulated price

White / blue vignette: this indicator permits the segmentation of drugs. White vignette represents medical utility drugs, blue one represents moderate medical utility

PHI reimburse 100% of RP for medical practitioners fees and stay costs when they are linked to surgery acts whose cost exceed 91€

(1) ECHI reimburse 31€ for unifocal eyeglasses and 92€ for multifocal eyeglasses

(2) For Spa care, PHI reimburse 65% or 70% of RP depending of the type of care

4.3.2. Sample and dataset

We analyze a representative sample of BCHI customers that includes 18,126 individuals. We observe each insured for 10 semesters, from January 2001 to December 2005. By the end of our observation period (5 semesters after ECHI implementation), 20% of all BCHI enrollees subscribed the ECHI plan: 42% enrolled immediately (during the first semester), 23% during the second (first semester of 2004), 13% during the second half of 2004, 11% during the first semester of 2005 and 10% during the last semester of 2005. No subscribers abandoned the ECHI plan during the observation period.

The dataset includes many individual socio-demographic characteristics known to influence health insurance demand, such as age, gender, labor market status, residence, insured status and a proxy for income⁴⁴. These characteristics are known for December 2005, and we assume that they did not change from January 2001. The dataset also provides detailed information for individual outpatient healthcare expenses from 2001 to 2005 for specific expenses: the date of utilization; the type of care (physician care, dental care, prescription drugs and optical care)⁴⁵; and the entire cost of care that was partly reimbursed by PHI, BCHI and eventually ECHI. Tables 4.2 ii (a) and (b) present the main characteristics of our sample. We analyze the effect of ECHI for each type of care by successively considering (1) the increase in the probability of using healthcare, (2) the number of uses conditional on use and (3) the cost per use.

⁴⁴ Policyholder Income is approximated using the national grid of wage indices for civil servants. Multiplied by the value at that point (53.9€ in December 2005), we computed the values to obtain a monthly gross wage.

⁴⁵ We focus on outpatient care and exclude hospital care data that may be incomplete (several services provided during hospitalization are not reported in the *mutuelle* patient file).

Table 4.3-ii (a): Sample demographic characteristics in 2005

	Full sample	covered by the ECHI	not covered by the ECHI
Age	49.5	56.3	47.7
gender			
Man	62.6%	59.4%	63.4%
Woman	37.4%	40.6%	36.6%
Wage index brackets (in points)			
<= 1350€	6.1%	7.5%	5.7%
From 1350€ to 1790€	57.8%	51.9%	59.3%
From 1790€ to 2240€	20.5%	22.6%	19.9%
From 2240€ to 3110€	10.2%	11.6%	9.8%
> 3110€	5.5%	6.5%	5.2%
Administrative situation			
Active	40.1%	43.7%	39.1%
Retired	31.2%	38.5%	29.3%
Student	1.5%	0.9%	1.6%
No professional activity	27.3%	16.9%	29.9%
Family situation			
CHI policyholder	65.6%	76.0%	63.0%
Spouse	14.8%	16.7%	14.3%
Child	19.6%	7.3%	22.8%
Residential location (grouping of regions)			
Ile de France (IdF = Paris region)	7.8%	10.4%	7.1%
Parisian basin regions (excluded IdF)	16.8%	15.1%	17.3%
Northern regions	4.4%	4.5%	4.4%
Eastern regions: Alsace Lorraine	6.0%	9.2%	5.2%
Eastern regions: Franche Comté	2.6%	2.6%	2.6%
Western regions	16.9%	12.7%	18.0%
South-western regions	15.4%	15.6%	15.3%
Center-Eastern regions	11.7%	12.8%	11.5%
Mediterranean regions	14.3%	13.1%	14.6%
Number of observations	18126	8668	9458
Number of pool CHI policies	12173	6780	5393
Average number of beneficiaries per pool CHI policy	1.52	1.42	1.56

(b) Healthcare utilization in 2005.

	Full sample		covered by the ECHI		not covered by the ECHI	
	before July 2003	/ after	before July 2003	/ after	before July 2003	/ after
Average expenditure per semester (in euros)						
outpatient care	554.6	/ 659.5	673.2	/ 860.5	524.1	/ 607.7
Physician care	81.5	/ 87.8	100.1	/ 109.6	76.7	/ 82.1
Dental care	63.8	/ 72.4	76.1	/ 123.2	60.6	/ 59.3
drugs	201.2	/ 242.4	255.3	/ 308.8	187.2	/ 225.2
Optical care	31.7	/ 39.9	40.5	/ 63.1	29.4	/ 33.9
Probability of using conditional on use per semester						
outpatient care	0.881	/ 0.900	0.887	/ 0.930	0.880	/ 0.892
Physician care	0.813	/ 0.819	0.834	/ 0.866	0.808	/ 0.807
Dental care	0.244	/ 0.248	0.270	/ 0.290	0.237	/ 0.235
drugs	0.800	/ 0.811	0.827	/ 0.861	0.793	/ 0.798
Optical care	0.109	/ 0.117	0.125	/ 0.158	0.105	/ 0.106
Number of uses conditional on use per semester						
outpatient care	19.6	/ 26.1	23.0	// 29.0	18.8	/ 25.3
Physician care	4.8	/ 5.0	5.4	// 5.7	4.6	/ 4.9
Dental care	2.0	/ 2.5	2.1	// 2.8	2.0	/ 2.3
drugs	8.7	/ 9.3	10.4	// 11.2	8.2	/ 8.8
Optical care	2.2	/ 2.9	2.2	// 2.9	2.2	/ 2.8
Average cost use conditional on use per semester (in euros)						
outpatient care	34.5	/ 33.8	35.1	/ 37.2	34.3	/ 32.8
Physician care	21.2	/ 21.6	22.0	/ 22.4	21.0	/ 21.4
Dental care	127.6	/ 121.4	122.2	/ 142.3	129.2	/ 114.5
drugs	24.1	/ 26.3	26.2	/ 27.7	23.6	/ 25.9
Optical care	142.5	/ 119.5	159.2	/ 136.9	137.4	/ 112.9
Number of observations	18126		8668		9458	

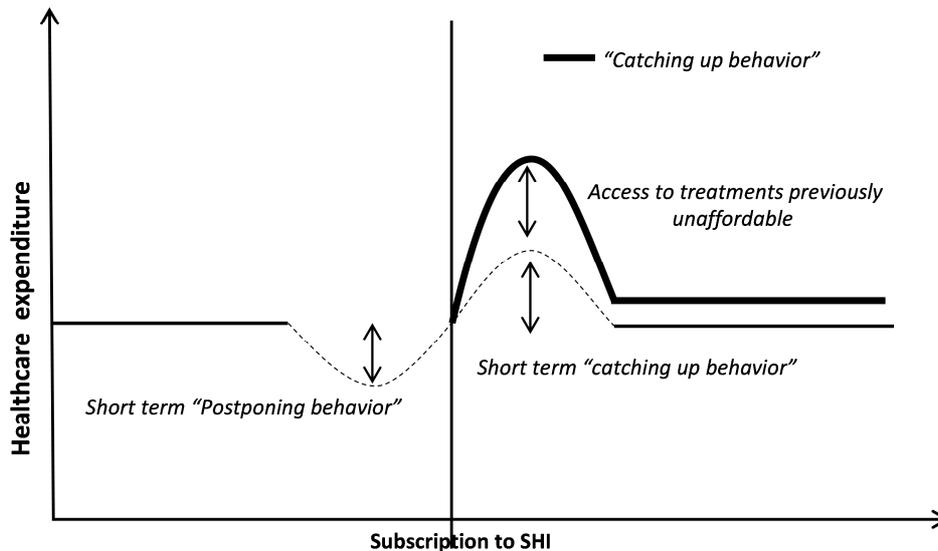
4.4. Economic model and methods

4.4.1. The assumptions of pent-up demand

The pent-up demand theory assumes that the effect of health insurance on health expenditures does not remain constant over time. Long et al. (1998) argued that an increase in insurance coverage may have two types of consequences: a “postponing behavior” prior to enrollment and a “catching up behavior” immediately after. “Catching up behavior” results in a sharp increase in health expenditure just after the rise of the insurance coverage that decreases over time (graph 4.4 i). “Postponing behavior” before subscribing to health insurance results from an economic tradeoff: individuals transitory postpone non-urgent treatments to benefit from higher reimbursements (Chen et al., 2004) (dotted thin line). This short term “postponing behavior” reflects an inter-temporal substitution of healthcare consumption in order to take profit of lower prices and leads to a short term catching up effect (dotted thin line). This effect has to be added to the expenses associated to access to treatments previously unaffordable to constitute the overall “catching up behavior” (solid thick line). Indeed individuals may have foregone to unaffordable treatments before the increase of insurance coverage.

This concept can be brought closer to moral hazard as one does not completely exclude it in the overall “catching up effect” (solid thick line). However, validating the pent-up demand assumption means that even if there is moral hazard, it does not last over time and does not reflect a permanent change in the consumption pattern.

Graph 4.4-i: Variation of health care consumption according to the pent-up demand theory



In this paper, we test the pent-up demand assumption by considering the existence of the “catching-up” effect and its evolution over time.

4.4.2. Health care expenditures

An increase in healthcare consumption caused by an increase of insurance coverage has to be divided into three dimensions: the probability of using care, the number of uses of care conditional on use and the cost of care per use.

Équation 4.4-i

$$E(\text{exp}) = E(\text{cost}/\text{nbuse}) \cdot E(\text{nbuse} | \text{use} = 1) \cdot p(\text{use} = 1)$$

- The probability of using care reflects access to healthcare or to a type of care.
- The number of uses conditional on use reflects the frequency of healthcare consumption when individuals have access to healthcare or to a type of healthcare.
- The cost per use reflects the average expense per use. This price per use is the ratio of the total cost and the number of uses that both vary over time. An increase in the cost per use may reflect several factors (for example, an increase in the cost and/or a decrease in the number of uses). If the number of uses increases, the cost increases more sharply illustrating

a change in the consumption pattern (expecting a higher level of quality; Feldstein, 1971) or a change in the supply side (providers charge higher fees in response to the higher level of coverage; Delattre and Dormont, 2000).

4.4.3. The pent-up demand assumption

According to the pent-up demand assumption, the effect of health insurance on health expenditures does not remain constant over time. Indeed, because individuals postpone health care consumption before they obtain health insurance, the effect of insurance can be greater immediately following subscription. Postponing health spending before obtaining health insurance (or foregoing care) may reflect various economic tradeoffs: individuals may have discontinued treatment for financial reasons, or they may have conveniently decided to postpone non-urgent treatments to benefit from higher guarantees (Chen et al., 2004). In the long term, health insurance may also have a negative effect on health care consumption as a result of improvements in the health statuses of patients.

As a result of a change in cost sharing (following a health insurance subscription), health care consumption can appear to increase significantly immediately after the change occurs because newly insured individuals may have postponed or discontinued their health care treatment before obtaining insurance. The effect of health insurance may appear to be strong immediately after subscription and thus may decrease over time a fortiori when considering the improvements in health status in the long-term.

In this paper, we test the pent-up demand assumption on health care consumption (with the breakdown as in Eq. 1) using panel data that enable us to perform this test using appropriate methods to control for the endogeneity that may result from adverse selection.

4.4.4. Empirical models

Notations

According to the equation of healthcare consumption (Eq. 1), we test pent-up demand for outpatient care and each type of care by studying the effects of ECHI through three models:

- (1) the probability of using care: $p(use_{is} = 1)$,
- (2) the conditional number of uses of care: $E(nbuse_{is} | use_{is} = 1)$,
- (3) the cost per use: $E(ucost_{is})$,

where i ($i=1, \dots, N$) is the individual index and s ($s=1, \dots, 10$) corresponds to a semester from January 2001 to December 2005.

For each ECHI subscriber, we define five dichotomous variables ($LI_{is}^1 \dots LI_{is}^5$) corresponding to the length of enrollment (into ECHI) at any semester s ($s > 6$, ECHI was implemented in July 2003 $s=6$); for instance, an individual who subscribed at semester seven ($s=7$), the length of enrollment at the end of the period will be four semesters such that $LI_{is}^4 = 1$.

In each model, we introduce covariates linked to both health insurance demand and healthcare consumption: X_{is} represents a vector of age and age squared (for individual i at semester s) and Z_i represents a vector of individual fixed variables (gender, number of individuals covered by BCHI within the household, wages of the subscribers of BCHI, labor market status, family situation and location).

Random effect models with Chamberlain specification

For (1), (2) and (3), we use random effects models that consider two levels of unobservable heterogeneity; the first level refers only to individuals and the second level allows heterogeneity between individuals and semesters.

$$y^{1,2,3} = f(X_{is}, Z_i, LI_{is}, sem_s, v_{is}^{1,2,3}, \varepsilon_{is}^{1,2,3})$$

$v_{is}^{1,2,3}$ represents the unobserved heterogeneity parameters that are fixed in time and $\varepsilon_{is}^{1,2,3}$ is the unobserved heterogeneity parameters that vary over time.

Without the Chamberlain specification, random effect models are based on the assumption that the residuals, whether fixed or variable over time, are not correlated with the explanatory variables. This strict exogeneity assumption is rather strong because individual health status and risk aversion are both unobservable. As a result, even if the panel structure of the data enables us to account for a significant part of adverse selection, unobservable factors can result in biased estimators. For instance, a risk-adverse (or ill) individual may decide to subscribe to a supplemental health insurance contract and consume more because of risk aversion (or health status).

The Chamberlain specification (Chamberlain, 1984) takes into account for the correlation between unobserved constant individual characteristics (such as risk aversion and/or the constant part of an individual's health status) and the decision to subscribe to an ECHI plan and the length of enrollment in the ECHI. This specification decomposes the fixed residuals (v_i^1, v_i^2, v_i^3) into two components: the first one corresponds to the correlation with the length of enrollment in ECHI (LI) and the second component corresponds to exogenous factors.

$$v_i^1 = \sum_{\tau=1}^{10} \Psi_{i\tau}^1 \cdot LI_{i\tau} + w_i^1, \quad v_i^2 = \sum_{\tau=1}^{10} \Psi_{i\tau}^2 \cdot LI_{i\tau} + w_i^2, \quad v_i^3 = \sum_{\tau=1}^{10} \Psi_{i\tau}^3 \cdot LI_{i\tau} + w_i^3.$$

Under these specifications, the three economic models are as follows:

Model (1), the probability of using care

$$use_{is}^* = \alpha_1 + \gamma_1 \cdot X_{is} + \beta_1 \cdot Z_i + \delta_1 \cdot LI_{is} + \sum_{\tau=1}^{10} \Psi_{i\tau}^1 \cdot LI_{i\tau} + \eta_1 \cdot sem_s + w_i^1 + \varepsilon_{is}^1$$

$$\left| \begin{array}{l} use_{is} = 1 \text{ if } use_{is}^* \geq 0 \\ use_{is} = 0 \text{ if } use_{is}^* < 0. \end{array} \right.$$

Model (2), the number of uses conditional on use

$$nbuse_{is} = f(\alpha_2 + \gamma_2 \cdot X_{is} + \beta_2 \cdot Z_i + \delta_2 \cdot LI_{is} + \sum_{\tau=1}^{10} \Psi_{i\tau}^2 \cdot LI_{i\tau} + \eta_2 \cdot sem_s + w_i^2 + \varepsilon_{is}^2).$$

Model (3), the cost per use

$$ucost_{is} = \exp(\alpha_3 + \gamma_3 \cdot X_{is} + \beta_3 \cdot Z_i + \delta_3 \cdot LI_{is} + \sum_{\tau=1}^{10} \Psi_{i\tau}^3 \cdot LI_{i\tau} + \eta_3 \cdot sem_s + w_i^3 + \varepsilon_{is}^3).$$

Following the Chamberlain specification, we assume that the vectors of residuals $(v_i^1, \varepsilon_{is}^1)$, $(v_i^2, \varepsilon_{is}^2)$ and $(v_i^3, \varepsilon_{is}^3)$ are independent in pairs. For instance, the decision to use care and the expenditure for a specific use are assumed to be independent. Thus, the equations can be estimated separately in a three-part model⁴⁶.

The method of generalized estimating equations (GEE)

Two different methods are available to estimate three-part models, depending on the assumptions made about the residuals. First, the GEE method is a semi-parametric method and requires minimal assumptions regarding residuals. This method is an extension of generalized linear models (Liang and Zeger, 1986) that consists of minimizing the square of the generalized residuals⁴⁷ that are normalized by the matrix of their variance. The variance of the generalized residuals is written as a function of an unknown working matrix that describes the time correlation and must also be estimated. All estimations are obtained in a step-by-step approach by estimating the coefficients and determining the generalized residuals to provide the required matrix. This process continues until all of the parameters converge. For (1) (the probability of using care), we assume that $v_i^1 + \varepsilon_{is}^1$ is normally distributed; for (2) and (3), we assume that f is exponential without making assumptions about v_i^2 and ε_{is}^2 , v_i^3 and ε_{is}^3 . The GEE method provides more robust estimations, but the estimators are less efficient than those of parametric methods.

⁴⁶ Without this assumption, we would have used sample selection models that were adapted to the panel data (Wooldridge, 1995) and identified independent variables to explain each equation separately.

⁴⁷ These residuals correspond to the difference between the explained variables and its expectancy conditional on the explanatory variables.

Second, a maximum-likelihood random effect model (RE-MLE) may be applied to determine the estimators. For (1), we assume that v_i^1 and ε_{is}^1 are normally distributed. For (2), v_i^2 is normally distributed, and ε_{is}^2 has a truncated negative binomial law. Finally, for (3), v_i^3 and ε_{is}^3 are normally distributed.

The GEE and RE-MLE methods provide similar estimators and standard errors. However, the RE-MLE method does not converge for all types of care in model (2). In the following section, we present and discuss the results obtained using the GEE method⁴⁸.

4.5. Results

4.5.1. Effects of ECHI

The effects of ECHI on the dimensions of healthcare consumption (1), (2) and (3) according to the length of enrollment are presented in table 4.5-i.

Table 4.5-i : The effect of supplemental health insurance on healthcare consumption according to the length of enrollment (GEE estimates)

Length of enrollment	1) Probability of use			2) Number of uses conditional on use			3) Cost per use conditionnal on use		
	ME (points)	Pr. (ME=0)	Pr. (ΔME=0)	ME (%)	Pr. (ME=0)	Pr. (ΔME=0)	ME (%)	Pr. (ME=0)	Pr. (ΔME=0)
Outpatient care									
1 st semester	+4.3	***		+7.6	***		+12	***	
2 nd semester	+3.9	***	ns	+3.4	**	**	+11.5	***	ns
3 rd semester	+3.4	***	*	+4.3	**	*	+7	***	***
4 th semester	+2.6	***	***	+1.8	ns	***	+6.7	***	***
5 th semester	+1.6	*	***	+2.7	ns	**	+3	ns	***
<i>Ref. Without ECHI (predicted value)</i>	<i>90.4 pts</i>			<i>22.9 uses</i>			<i>34.1 €</i>		
Physician care									
1 st semester	+4.3	***		+3.5	***		+0.7	*	
2 nd semester	+4.1	***	ns	+1.3	ns	ns	+0.1	ns	ns
3 rd semester	+3.4	***	ns	+0.9	ns	ns	0	ns	ns
4 th semester	+2.6	***	**	-0.1	ns	*	-0.4	ns	*
5 th semester	+1.8	*	***	+0.8	ns	ns	-1.0	ns	**
<i>Ref. Without ECHI (predicted value)</i>	<i>81.6 pts</i>			<i>4.9 uses</i>			<i>21.4 €</i>		
Dental care									
1 st semester	+7.4	***		+20.0	***		+21.3	***	
2 nd semester	+4.7	***	**	+8.8	***	***	+34.6	***	**
3 rd semester	+3.7	***	***	+7.3	***	***	+12.1	***	*
4 th semester	+2.8	***	***	+6.2	***	***	+21.5	***	ns
5 th semester	+1.0	ns	***	+7.7	***	***	+13.7	**	ns
<i>Ref. Without ECHI (predicted value)</i>	<i>24.6 pts</i>			<i>2.2 uses</i>			<i>124.5 €</i>		

⁴⁸ The results obtained using the maximum likelihood method are available upon request.

Length of enrollment	1) Probability of use			2) Number of uses conditional on use			3) Cost per use conditionnal on use		
	ME (points)	Pr. (ME=0)	Pr. (ΔME=0)	ME (%)	Pr. (ME=0)	Pr. (ΔME=0)	ME (%)	Pr. (ME=0)	Pr. (ΔME=0)
Prescription drugs									
1 st semester	+3.4	***		+1.6	***		-1.5	ns	
2 nd semester	+3.7	***	ns	+2.2	ns	ns	-0.7	ns	ns
3 rd semester	+3.2	***	ns	+4.1	ns	*	0	ns	ns
4 th semester	+2.3	***	ns	+1.2	ns	ns	-1.4	ns	ns
5 th semester	+1.2	Ns	**	+4.8	**	*	-1.5	ns	ns
<i>Ref. Without ECHI (predicted value)</i>	80.5 pts			9.0 uses			25.2 €		
Optical care									
1 st semester	+5.9	***		+6.3	***		+10.1	***	
2 nd semester	+3.8	***	**	+6.3	***	ns	+12.3	***	ns
3 rd semester	+3.5	***	**	+3.4	ns	ns	+7.4	**	ns
4 th semester	+2.5	***	***	+1.4	ns	*	+6.3	*	ns
5 th semester	+2.3	***	***	+3.7	ns	Ns	+7.4	*	ns
<i>Ref. Without ECHI (predicted value)</i>	11.3 pts			2.5 uses			130.6 €		

ME: marginal effect (in points or as a percentage)

Pr.: probability that the marginal effect is equal to 0; p<0.01: ***; 0.01<p<0.1: **; 0.1<p<=0.5: *; ns: p>0.5

Pr. (ΔME = 0): Probability that the ME of the first semester of enrollment is equal to the ME of semesters

4.5.1.1. The probability of using health care

ECHI has significant effects on the probability of using at least one type of outpatient care. This effect appears to decrease significantly with the length of enrollment but remains highly significant throughout the five semesters: from +4.3 points during the first semester of enrollment to +2.6 points during the fourth semester and +1.6 points during the fifth semester.

A similar trend is observed for nearly all types of care (except for prescription drugs). For instance, ECHI offers important additional reimbursements for dental and optical care, and the effect is significantly high during the first semester of enrollment: +7.4 points (dental care) and +5.9 points (optical care) (cf. Table 4.5-i). The probability of purchasing prescription drugs is also significantly higher for the ECHI insured (and does not significantly decrease over time, except during the fifth semester) despite the fact that ECHI does not directly increase the insurance coverage for drugs.

4.5.1.2. The number of uses conditional on use

ECHI appears to have a significant influence on the number of times a subscriber uses outpatient care. This effect is strong and significant only during the first three semesters of enrollment; the number of uses increases by nearly 8% during the first semester after enrollment. This effect decreases over time and is non-significant after the fourth semester of enrollment.

Interestingly, the decreasing marginal effect is sharp for physician and optical care as ECHI effects are only significant during the first semester of enrollment (respectively +3.5% and +6.3%). For prescription drugs, no trend is clearly observed (the effects range from +1.2% to +4.8%). For dental

care, ECHI effect is significantly large (+20.0%) during the first semester of enrollment and remains significant during the fifth semester (+7.7%) although this effect decreases over time.

4.5.1.3. *The cost per use conditional on use*

ECHI significantly affects the cost of outpatient care particularly during the first year (+12%) and remains significant during the first two years of enrollment.

The effect of ECHI on the cost of care is primarily driven by its effects on dental and optical care. For dental care, ECHI significantly increases the cost per unit (+21.3% for the first semester of enrollment and +34.6% for the second semester). The increase between the two first semesters of enrollment is significant and may result from constraints such as delays in obtaining dental appointments and treatments duration. During the following semesters, the effect of ECHI becomes smaller but remains significant (between 12% and 21%). For optical care, the effect is significant and rather high during the two first semesters of enrollment (+10% and +12%, respectively), then becomes smaller but remains significant during subsequent periods.

4.5.2. Effect of main other covariates

The effects of the main covariates that were introduced in addition to SHI are presented in Table 4.5-ii.

Table 4.5-ii: The effect of the main explanatory variables

	Probability of utilization		Number of uses conditional to use		Cost per use conditional on use	
	ME (points)	Pr. (ME=0)	ME (%)	Pr. (ME=0)	ME (%)	Pr. (ME=0)
Outpatient care						
Age	-0.0843	ns	-2.4620	***	+3.8589	***
Age ²	+0.0037	***	+0.0420	***	-0.0293	***
Women (ref: men)	+5.22	***	+34.05	***	-2.63	***
income (ref: <1350€ per month)						
From 1350€ to 1790€	+3.73	***	+8.75	**	-0.63	ns
From 1790€ to 2240€	+3.61	***	+5.53	ns	+1.34	ns
From 2240€ to 3110€	+1.78	*	+1,26	ns	+5.61	**
> 3110€	+0.86	ns	-8.40	*	+8.84	***
Ref. Without ECHI (predicted value)	90.4 pts		22.9 uses		34.1 €	
Physician care						
Age	-0.4167	***	-2.2972	***	+0.3798	***
Age ²	+0.0073	***	+0.0326	***	-0.0045	***
Women (ref: men)	+7.85	***	+26.63	***	+3.84	***
income (ref: <1350€ per month)						
From 1350€ to 1790€	+6.66	***	+0.33	ns	+3.12	***
From 1790€ to 2240€	+6.12	***	-3.78	ns	+5.72	***
From 2240€ to 3110€	+3.70	***	-6.33	*	+7.85	***
> 3110€	+2.47	*	-13.71	***	+15.52	***
Ref. Without ECHI (predicted value)	81.6 pts		4.9 uses		21.4 €	
Dental care						
Age	+1.2689	***	-0.2172	ns	+6.1230	***
Age ²	-0.0115	***	+0.0010	ns	-0.0447	***
Women (ref: men)	+4.67	***	-2.45	*	+0.61	Ns
income (ref: <1350€ per month)						
From 1350€ to 1790€	+2.78	***	-0.87	ns	-4.02	Ns
From 1790€ to 2240€	+5.61	***	-1.46	ns	-3.29	Ns
From 2240€ to 3110€	+5.90	***	-4.32	ns	-6.91	Ns
> 3110€	+11.05	***	-3.03	ns	-3.87	Ns
Ref. Without ECHI (predicted value)	24.6 pts		2.2 uses		124.5 €	
Drugs						
Age	-0.6324	***	-1.7494	***	+1.3459	***
Age ²	+0.0106	***	+0.0346	***	+0.0020	Ns
Women (ref: men)	+8.37	***	+29.89	***	-7.84	***
income (ref: <1350€ per month)						
From 1350€ to 1790€	+5.40	***	+7.80	**	-1.40	Ns
From 1790€ to 2240€	+4.40	***	+4.85	ns	-6.60	**
From 2240€ to 3110€	+1.38	ns	+0.35	ns	-6.18	*
> 3110€	-1.01	ns	-10.74	**	-14.80	***
Ref. Without ECHI (predicted value)	80.5 pts		9.0 uses		25.2 €	
Optical care						
Age	+0.6120	***	-0.1497	ns	+4.3711	***
Age ²	-0.0051	***	-0.0015	ns	-0.0297	***
Women (ref: men)	+4.00	***	-0.62	ns	+12.05	***
income (ref: <1350€ per month)						
From 1350€ to 1790€	+1.29	**	-0.46	ns	+0.24	Ns
From 1790€ to 2240€	+2.83	***	+1.25	ns	+4.02	Ns
From 2240€ to 3110€	+4.09	***	+0.63	ns	+14.68	***
> 3110€	+3.33	***	-0.99	ns	+22.76	***
Ref. Without ECHI (predicted value)	11.3 pts		2.5 uses		130.6 €	

The effects of ECHI are expressed in percentage points for the probability of use and as a percentage for the number of uses conditional on use and the cost per use. For the probability of use, the number of uses conditional on use and the cost per use, the first column presents the effect of ECHI, and the second column indicates the significance of the effect as follows: *** significant at the 0.1% level; ** significant at the 1% level; and *: significant at the 5% level

4.5.2.1. Effect of age and gender

For outpatient care, squared age has a positive effect on the probability of using at least one type of outpatient care: the probability of use increases sharply with age. However, the effect of age differs by type of care.

For physician care, the probability of use and the number of uses follow a U-shaped trend (with the minimum at 29 years and 35 years, respectively), whereas the cost per use conditional on use varies according to an inverse U shape (maximum at 42 years). For prescription drugs, we observe a similar pattern in the probability of use and the number of uses of these drugs, whereas age has a linear positive effect on the cost per use. For optical and dental care, both the probability of use and the cost per use vary following an inverse U pattern, whereas age has no effect on the number of uses conditional on use.

Being a woman significantly increases the probability of using care, regardless of the types of care. However, the gender effect on the number of uses and the cost per use varies according to the type of care; women use physician care and prescription drugs more frequently but use dental care significantly less, and there is no significant difference for optical care. Women have significantly higher cost per use for physician and optical care and lower costs for drugs.

4.5.2.2. Effect of income

The probability of using at least one type of outpatient care varies with income following an inverse U shape: compared to the poorest individuals in our population (<1350€ per month), the probability of using care is significantly higher of +3.7 points for those in the second income bracket (between 1350€ and 1790€ per month) and of +1.8 points for those in the fourth income bracket (between 2240€ and 3110€ per month). For the highest income bracket, the income effect is no longer significant.

A similar pattern is observed for the probability of using physician care and drugs. Indeed, the largest effect is observed for the second income bracket (+6.7 points for physician care and +5.4 points for prescription drugs) and differences observed between the wealthiest and the poorest individuals are weak or insignificant. For optical and dental care, the probability of using care always increases with income.

An inverse U shape is also observed for the number of uses of outpatient care, including prescription drugs; individuals from the second income bracket have, *ceteris paribus*, the highest number of uses. In contrast, individuals who earn more than 3110€ per month have the lowest number of uses, and the difference is significantly lower than that for the poorest individuals. For physician care, the

number of uses decreases with income, whereas income has no significant effect on dental and optical care number of uses conditionally to use.

Finally, the cost per use of physician and optical care increases with income, and for prescription drugs, the cost per use decreases.

4.6. Discussion

In this paper, we estimate the marginal effect of additional health insurance coverage on healthcare consumption, and more interestingly, we examine the persistence of this effect over time separating healthcare consumption into three dimensions: the probability of using care, the number of uses conditional on use and the cost per use. This study, the first utilizing French data, is also interesting due to both the length of our observation period (more than two years for the first subscribers) and the design of our study; we consider the marginal effects of supplementary health insurance coverage (Blomqvist, 2001), while previous research such as Nyman (1999b) compared the insured and uninsured. This difference is particularly relevant in France, where more than 90% of individuals have purchased CHI in a market where contracts are highly heterogeneous.

Our paper presents several findings: first, we highlight that ECHI has an immediate and positive effect on the three dimensions of healthcare consumption and that the relevance and the magnitude of this effect vary depending on the type of care. Second, our analyses indicate that this positive effect on outpatient consumption significantly decreases over time, supporting the pent-up demand assumption.

4.6.1. The effect of ECHI according to the type of care

For physician care, little to no effect is observed on the number of uses and the cost per use conditional on use, whereas the probability of using physician care increases significantly. As the average cost per use for physician care (maximum +0.7%) increased less rapidly than the number of uses (maximum +3.5%), we conclude that the increase in insurance coverage has not led subscribers to choose more expensive doctors but has increased access to physicians.

While the contract does not offer additional coverage for prescription drugs, it is interesting to note the significance of ECHI its effect on drugs expenses. This effect likely results indirectly following the increase in physician care: in France, the ratio of visits resulting in the prescription of drugs is higher than in any other European country: 90% of consultations produce a prescription, compared to 83.1% in Spain, 72.3% in Germany and 43.2% in the Netherlands (HCAAM, 2006).

For dental and optical care, ECHI strongly and positively influences the consumption regardless of the dimension (1), (2), and (3) considered. These effects reflect high copayments remaining after the reimbursements of both PHI and BCHI for dental and optical care. ECHI reduces these copayments and facilitates access to care in different ways. First, one can consider the short term “postponing behavior” and/or the access to treatments previously unaffordable. Second, one can also consider individuals already using optical and/or dental care prior to the ECHI subscription and who decide to buy more expensive care either to improve quality (assuming that the price per unit is a proxy of the quality for dental prostheses, for instance) or to opt for more expensive goods (optical frames).

4.6.2. The pent-up demand assumption

Another important result of the paper is the study of moral hazard over time. Our results emphasize that after an initial increase in healthcare consumption, the effect of health insurance significantly decreases over time, illustrating the catching up phenomenon (Long et al., 1998). The pent-up demand effect is particularly relevant for the probability of using care (1), whatever the type of care but especially for dental and optical care. The pent-up demand effect is also relevant to the number of uses conditional on use (2), and is arguable for the cost per use conditional on use (3).

The overall effect of ECHI on the probability to use outpatient care corresponds to an increase in access to care mainly driven by an improvement in access to physician care but also by a high increase in access for dental and optical care. For these types of care, high extra fees are the most common, and individuals probably exhaust their healthcare needs that they could not afford prior to enrollment. It is also possible that the announcement of the ECHI may have resulted in a short term “postponing effect” even if we did not actually distinguish these effects.

Another important result is that there is still a significant and positive effect of ECHI at the end of our observation period. Indeed, the increase in insurance coverage has persistent effects on the probability to use care (1), the number of dental care uses conditional on use (2) and the dental and optical costs per use conditional on use (3). These results can reflect three types of behaviors: first, the increase in insurance coverage may permit sustainable access to previously unaffordable healthcare; this could be especially true for the access of optical and dental care. For instance, vision problems, such as some chronic diseases, lead to regular predictable (to some extent) and rather high expenses which could discourage the initiation of care prior to subscribing to ECHI and this may explain why no ECHI subscribers gave up the contract (during the observation period). In other words, the increase in the coverage did not lead to an opportunistic behavior which would correspond to wait until becoming acutely ill to purchase insurance, use it for the duration of the illness and then drop it. Second, the remaining significant effect of ECHI over time may also reflect

moral hazard due to an effective reduction in the price of care. Concerning the optical care per unit cost, the permanent effect may reflect a preference for expensive goods like optical frames revealed by the price effect. Third, but less likely, the length of our observation period may be too short to observe convergence of ECHI healthcare consumption to the level of those who have not subscribed.

Despite these interesting results, our research remains limited in some respects. Even if the robust econometric model with the Chamberlain specification used on panel data allowed us to control for a large part of adverse selection bias, it did not allow us to control for the additional potential bias resulting from the insurance decision and unobserved heterogeneity that varies over time. Consequently, the effect of ECHI on healthcare consumption may be overestimated. This could be the case if a health shock occurred during our observation period leading to both the enrollment in ECHI and an increase of healthcare consumption. Another limitation of our study concerns the observed population (i.e., civil servants and their relatives), which is not perfectly representative of the general population. However, this specificity is not likely to significantly affect the trends in the results if the observed population does not have a specific change in healthcare consumption from ECHI enrollment.

Finally, we observe the effect over time of a shift from one specific health insurance plan to another and show that an increase in health insurance coverage has transitory positive and significant effect for the access to some type of care like physician care and drugs while the effect appeared to be more persistent in terms of access and per unit cost for dental or optical care. To go further, future research could analyze the effect of a shift from any level of coverage to any other to estimate the price elasticities per type of care.

Conclusion

L'assurance complémentaire santé en France permet de compléter les remboursements de l'Assurance maladie obligatoire ; elle est offerte par les entreprises, pour les salariés ou leurs proches ou est souscrite par une démarche personnelle. Elle couvre une grande majorité de la population, plus de neuf français sur dix. Plusieurs travaux menés dès le milieu des années 1990 ont montré que l'assurance complémentaire santé accroît les dépenses de santé en médecine de ville. Néanmoins, cet effet fait l'objet d'interprétations contradictoires. L'indicateur de renoncement aux soins pour raisons financières indique que bénéficiaire d'un contrat d'assurance complémentaire santé améliore l'accès aux soins. Des travaux tels que ceux de Desprès et al (2011) ont montré que le renoncement aux soins est significativement plus faible chez les personnes couvertes. La théorie économique indique qu'en augmentant le niveau couverture assurantielle, les contrats d'assurance santé peuvent aussi induire des phénomènes de risque moral « néfastes ». La France, comme de nombreux pays de l'OCDE, fait face à un poids croissant des dépenses de santé dans le revenu national et le Régime public d'assurance maladie subit un déficit chronique depuis le début des années 80. Dans ce contexte, il est tentant d'imputer la croissance des dépenses au fait que les individus seraient trop couverts, donc déresponsabilisés. Des politiques publiques telles que la mise en place des participations forfaitaires et des franchises médicales non-remboursables par les contrats dits « responsables » s'appuient sur cette hypothèse pour la mise en place de restes à charges visant à modérer les consommations médicales du patient.

Analyser l'effet de l'assurance complémentaire santé représente donc un enjeu crucial tant en termes d'équité que d'efficacité du système de santé. A l'heure où les pouvoirs publics s'interrogent sur le contenu que doivent avoir les contrats responsables pour concilier accès aux soins et responsabilisation du patient, ce travail vise à faire un point aux travers de rappels théoriques et d'une analyse empirique sur l'effet global de l'assurance santé sur les consommations médicales d'une part et sur la nature de cet effet d'autre part. Concernant ce dernier aspect, il faut tout de suite signaler que les résultats obtenus sont issus de preuves indirectes et appellent une recherche plus avancée.

5.1. Les enseignements de la théorie économique

Des rappels théoriques que nous avons développés dans le premier chapitre, nous pouvons tirer les enseignements suivants :

- Tout d'abord l'assurance modifie la demande de soins par un effet substitution et un effet revenu. S'il y avait symétrie de l'information sur l'assuré entre assureur et assurés, seul l'effet revenu existerait. L'effet substitution apparaît du fait de l'impossibilité pour l'assureur d'observer précisément l'état de santé des agents ; l'assureur ne peut alors qu'offrir des

contrats dont les remboursements sont contingents à la dépense, c'est-à-dire qui effectuent les remboursements au moment où l'individu décide de ses consommations médicales.

- L'effet substitution entraîne un phénomène d'inefficience en termes d'allocation des ressources : le supplément de soins achetés est valorisé par les individus en deçà de son coût total, lequel se répercute sur l'ensemble des assurés via une augmentation de la prime du contrat lui-même et sur celles du régime d'assurance santé de base (s'il existe), si le volume des consommations médicales s'accroît. Ainsi, il y a accroissement des cotisations pour les individus qui bénéficient du supplément de couverture mais aussi pour ceux qui n'en bénéficient pas.
- L'effet revenu accroît la capacité des individus à accéder aux soins en solvabilisant la demande de biens et services médicaux. Il permet également de consommer plus d'autres biens et services non-médicaux. Cet effet n'entraîne pas d'inefficience ; il est de plus susceptible d'augmenter l'utilité collective de l'ensemble des individus bénéficiant d'une même couverture, en diminuant les disparités de consommation entre ces individus.
- L'effet revenu lié à un accroissement du niveau de couverture est d'autant plus élevé que le niveau de couverture de base est faible, et que l'effet du revenu sur la consommation médicale agrégée est élevé et que le poids des dépenses de santé est élevé. De ce résultat, nous déduisons qu'un accroissement du niveau de couverture n'a certainement pas les mêmes conséquences en termes d'effet revenu, suivant le niveau de ressources des individus et suivant le niveau de couverture dont ils bénéficiaient avant l'accroissement.
- Au regard du rôle positif joué par la couverture santé sur l'accès aux soins, le subventionnement des contrats par les pouvoirs publics via des exonérations de charges sociales et fiscales permet d'accroître ce niveau de couverture. Mais en amplifiant le phénomène de risque moral, cette forme de subventionnement conduit à une situation sous-optimale, du moins lorsqu'elle est proportionnelle au niveau de couverture souscrit. Néanmoins, il a été signalé qu'outre une amélioration l'accès à l'assurance santé, ce type de couverture peut corriger partiellement une autre inefficience, la sélection adverse. Inciter les individus à se couvrir peut également être bénéfique si nous supposons qu'ils sous évaluent le risque financier lié à la santé.
- L'effet de la couverture assurantielle sur les consommations de soins peut se faire via l'offre de soins du fait d'une induction de la demande par l'offre ou du fait d'une concurrence imparfaite des professionnels de santé. Dans le premier cas, nous supposons que l'individu

n'a qu'une connaissance très imparfaite de sa santé et de l'effet des soins. Il construit sa courbe de demande en partie au moyen de l'expertise du médecin, lequel peut manipuler l'information dont il dispose pour accroître la demande de soins de ses patients. Dans le second cas, il n'y a pas asymétrie d'information entre patient et professionnel de santé. Néanmoins, parce qu'un professionnel de santé possède des attributs qui sont valorisés par certains patients et qui lui sont propre, son offre n'est pas parfaitement substituable par celle d'un autre praticien. Le professionnel peut profiter de cet avantage pour accroître le prix des soins ou, lorsque les prix sont régulés, le volume de soins. Etant donné que le professionnel de santé accroît les dépenses de santé en fonction de la disposition à payer du patient, la couverture assurantielle incite le praticien à accroître le volume ou le prix des dépenses médicales. La demande induite et la concurrence monopolistique indiquent qu'une partie de l'effet de la couverture assurantielle sur les consommations n'est pas de la responsabilité du patient mais du médecin, les mesures de régulation doivent alors aussi concerner les professionnels de santé.

5.2. Les principaux résultats des trois articles présentés

Les trois articles présentés dans cette thèse permettent d'une part de mesurer l'effet d'une hausse du niveau de couverture santé sur la demande de soins et d'autre part à tenter de déterminer la nature de cet effet, en cherchant notamment si il reflète un pur effet substitution néfaste ou un effet revenu bénéfique en termes d'amélioration de l'accès aux soins. Pour mémoire, les travaux présentés ont les caractéristiques suivantes :

- Dans le premier article, nous travaillons en population générale et examinons l'effet du fait d'être couvert plutôt que non-couvert sur la consommation de soins de médecins.
- Dans le deuxième article, nous nous focalisons sur la population des individus ayant adhéré à la CMU-C, une population pauvre donc, et examinons l'effet de l'entrée dans le dispositif sur plusieurs types de consommations de soins de ville. Nous considérons d'une part les populations qui étaient couverts auparavant par l'AMG ou une couverture complémentaire « privée », la CMU-C représente alors en général une amélioration du niveau de couverture ; d'autre part, les individus qui n'étaient pas couverts, la CMU-C leur permet de passer à la situation de personne couverte.
- Le dernier article analyse l'effet d'une hausse du niveau de couverture chez des personnes déjà couvertes par un contrat de couverture complémentaire santé de base. La population considérée est constituée de fonctionnaires, ayant différents niveaux de revenus. L'analyse

est menée sur la globalité des soins de ville puis quatre types de soins : les soins de médecins, de dentiste, d'optique et les médicaments. En analysant l'évolution de l'effet de la couverture dans le temps, il est possible de déterminer 1) s'il existe un effet de rattrapage, traduisant un éventuel effet revenu. 2) si la couverture complémentaire a un effet qui persiste dans le temps, traduisant une modification durable des préférences des individus.

Effet de la couverture complémentaire sur la consommation de soins de médecins

Les trois articles mettent en évidence un effet très sensible d'une élévation du niveau de couverture par une complémentaire santé sur la probabilité de recourir à un médecin, qu'il soit généraliste ou spécialiste.

Le deuxième article montre que l'effet de la CMU-C est significatif sur la probabilité de recours et la dépense conditionnelle au recours, uniquement chez les individus qui n'étaient pas couverts auparavant par l'AMG, notamment ceux qui n'avaient pas de couverture complémentaire. Ce que nous pouvons interpréter comme un effet de rattrapage doit cependant être considéré avec précaution. Nous observons en effet une baisse de la dépense conditionnelle en soins de généraliste l'année précédant l'adhésion à la CMU-C chez ceux qui ont adhéré volontairement au dispositif, ce qui peut indiquer que ces individus ont repoussé dans le temps une partie de leurs consommations pour pouvoir bénéficier de la gratuité des soins. Le fait que cet effet de report ne soit pas observé sur les soins de spécialistes, pourtant plus chers, vient tempérer l'hypothèse selon laquelle l'effet de rattrapage serait dû uniquement à un effet de report. Il semble que grâce à la CMU-C, les individus précédemment non-couverts aient pu recourir à des soins de médecins, en particulier des soins de spécialistes, auxquels ils devaient renoncer auparavant. Cette interprétation est confortée par les analyses économétriques de Desprès et al (2011) qui montrent que le renoncement aux soins de généralistes et de spécialistes est significativement plus faible chez les individus couverts par la CMU-C que chez les non couverts.

Le troisième article montre que le fait d'être couvert pour une (petite) partie des dépassements augmente significativement la probabilité de recours et le nombre de recours conditionnellement au fait d'en avoir. Cet effet est cependant transitoire, ce qui paraît indiquer que d'une part il y a eu effet de rattrapage et d'autre part qu'il n'y a pas de risque moral à long terme. Ce supplément de couverture semble donc avoir permis à ses bénéficiaires de consommer des soins qu'ils ne pouvaient pas financer auparavant.

Enfin, notons que le premier et le troisième article suggèrent que la couverture assurantielle n'a pas d'effet sur le prix des soins de médecins consommés, ce qui peut s'expliquer par le fait qu'il n'y a

somme toute que de faibles différence de prix entre les deux types de praticiens. Une autre interprétation est que les professionnels de santé ne sont que très imparfaitement substituables les uns aux autres. Les bénéfices que peut tirer un patient à consulter un praticien dont les prix sont plus élevés peuvent être contrebalancés par la perte d'avantages (la proximité géographique, la connaissance du patient, la sollicitude) offerts par le professionnel de santé consulté habituellement et que l'individu valorise.

Effet de la couverture complémentaire sur la consommation de médicaments

Le deuxième article permet d'analyser l'effet de la CMU-C sur les consommations de médicaments chez les personnes qui ne disposaient pas de couverture auparavant. Les résultats font apparaître une hausse très nette de la probabilité de recourir au moins une fois aux médicaments et de la dépense de médicament chez ceux qui ont recouru. Outre le fait que les individus ont une meilleure capacité financière à acheter les médicaments qui leur sont prescrits, nous pouvons supposer que cet effet résulte également d'un meilleur accès aux médecins, entraînant une plus forte probabilité de se voir prescrire des médicaments. Cette interprétation est renforcée par les résultats du troisième article. La complémentaire santé de base complétait déjà les remboursements de l'Assurance maladie pour tous les médicaments dans le panier de soins de la Sécurité sociale. La surcomplémentaire n'apporte aucun supplément de garantie. Pourtant la souscription de la surcomplémentaire s'accompagne d'une hausse du recours aux médicaments, effet qui peut là encore s'expliquer par le fait que la surcomplémentaire donne un meilleur accès aux prescripteurs.

Effet de la couverture complémentaire sur la consommation d'optique et de soins dentaires

L'effet d'un supplément de couverture sur la consommation de soins d'optique et dentaires est étudié dans le troisième article de la thèse (chapitre 5). Nous avons constaté que ce supplément de garanties a un impact très significatif sur la probabilité de recourir à ces types de soins au moment de la souscription. L'effet est également important sur le nombre de recours conditionnellement au fait d'en avoir eu un et sur la dépense par recours. Sur ce type de soins dont le prix est élevé, il semble bien y avoir eu un effet de rattrapage, les individus utilisant le supplément de garanties offert par la surcomplémentaire pour acheter des soins plus chers, qu'ils ne pouvaient pas se payer auparavant. L'effet de la hausse de garanties persiste dans le temps en ce qui concerne le coût moyen des soins achetés : les individus paraissent donc avoir substitués une partie des soins qu'ils consommaient auparavant par des soins plus chers. Ainsi, la présence d'aléa moral néfaste ne peut être écartée pour ces types de soins. Pour étudier de manière plus approfondie l'existence d'une amélioration de l'accès aux soins et de risque moral, il serait intéressant de croiser l'effet du surplus de garanties avec

le niveau de revenu des individus. Un pur effet revenu devrait se traduire en principe par un effet élevé de la couverture chez les individus ayant un faible niveau de ressources, effet qui décroît ensuite lorsque le niveau de revenu augmente.

5.3. Mise en perspective des résultats avec l'organisation du système d'assurance maladie

Les résultats de ces travaux s'inscrivent dans la problématique générale sur le rôle de l'assurance dans l'équité d'accès aux soins et sur l'efficacité du système assurantiel. Ils entrent ainsi dans la réflexion sur les modifications du système assurantiel visant à la fois à garantir l'équité d'accès aux soins et l'efficacité du système de santé. Ces réflexions portent sur la redéfinition de la notion de contrat responsable d'une part et sur la mise en place de plafonnements des tickets modérateurs laissés par l'Assurance maladie obligatoire d'autre part (cf. Geoffard et Lagasnerie, 2012).

Pour l'instant, le cahier des charges des contrats responsables fixe des seuils minimum de remboursements pour certains soins, à savoir le ticket modérateur pour les consultations de médecin et les analyses médicales dans le cadre du parcours de soins et une fraction du ticket modérateurs pour les médicaments à vignette blanche. Il fixe également des limites de remboursements (sous la forme de pénalités) lorsque le parcours de soins n'est pas respecté et stipule que les participations forfaitaires et franchises médicales ne doivent pas être couvertes. Il ne pose par contre aucune limite sur la couverture des dépassements d'honoraire chez le médecin, et sur les remboursements des frais d'optique et de dentaire. Ce cahier des charges apparaît ainsi insuffisant pour assurer une bonne accessibilité à l'ensemble des soins du panier de biens et services médicaux de l'assurance maladie mais aussi pour limiter les comportements d'aléa moral. La fixation de seuils planchers et plafonds plus contraignants semble donc nécessaire. Ainsi, lors du 40^{ème} congrès de la mutualité à Nice François Hollande a abordé la question de la révision du cahier des charges des contrats responsables, soulignant qu'elle devait aboutir à un meilleur accès à l'optique et aux soins dentaires et favoriser le parcours de soins. Par ailleurs, les députés ont voté le 16 octobre 2013 un amendement plafonnant la prise en charge des dépassements d'honoraire à 150% du tarif conventionnel. Ainsi, se dessine un système d'encadrement des remboursements.

Ce système de remboursements encadrés appelle plusieurs remarques : tout d'abord, se pose la question du mode de fixation des seuils. La réponse suggérée par la partie de contexte théorique est qu'ils doivent être établis de manière à opérer un compromis entre efficacité et équité. Cet objectif semble difficile à tenir dans le cadre d'un plafond unique par type de soin. Le troisième article montre ainsi qu'en ce qui concerne l'optique et le dentaire que l'hypothèse d'aléa moral ne pouvait pas être écartée mais également que des différences d'accessibilité pouvaient subsister du fait des restes à charges laissés par la surcomplémentaire. Un même niveau de remboursement peut donc

constituer un plancher pour certains individus et un plafond pour d'autres. Il apparaît donc nécessaire de moduler les remboursements en fonction des niveaux de ressources des individus.

Un autre point important concerne le statut des dépassements d'honoraires par les remboursements planchers. D'un côté, ces dépassements freinent l'accès aux soins, notamment dans certaines zones géographiques où ils sont fréquemment pratiqués. Mais imposer la couverture d'une certaine dose de dépassements pose problème : un tel système de remboursements crée *de facto* un ensemble de dépenses reconnues par les pouvoirs publics, en plus de celui existant pour l'assurance maladie obligatoire, dans la mesure où les planchers seraient supérieurs aux tarifs conventionnels. Se pose alors la question du sens qu'il y a à laisser ces deux régimes coexister, le système ainsi créé étant complexe et schizophrène. D'autre part, face à ces deux régimes, les professionnels de santé peuvent être incités à élever leurs tarifs en s'appuyant sur le système le plus avantageux pour eux. La fixation d'un tel seuil rend également caduques les tarifs opposables imposés par les pouvoirs publics concernant les bénéficiaires de la CMU-C.

Enfin, la définition de niveaux de remboursements minima que doivent assurer les complémentaires santé pour assurer un accès correct aux soins nous conduit à nous demander si ce socle de base ne devrait pas être plutôt intégré dans les remboursements de l'Assurance maladie obligatoire, dans la mesure où ces garanties minimales correspondent aux besoins de tous. En effet, la raison d'être de l'Assurance maladie complémentaire semble plutôt liée à l'hétérogénéité des préférences en termes de couverture. A partir du moment où la couverture d'un reste à charge est demandé par tous pour assurer un accès aux soins, il semble plus logique d'intégrer cette couverture dans le système d'Assurance public de base.

La mise en place d'un système de franchise sur les remboursements de l'Assurance maladie obligatoire introduit une alternative intéressante. Rappelons que ce système consisterait à prendre en charge l'intégralité des tickets modérateurs de médecine de ville au-delà d'un certain niveau de franchise dépendant du niveau de revenu. Cette nouvelle architecture de l'Assurance maladie obligatoire est présentée par Geoffard et Lagasnerie (2012). Une telle restructuration permettrait de mieux prendre en charge les restes à charge catastrophiques ; elle permettrait également aux individus de mieux appréhender leur risque santé. Enfin, elle aboutirait à une assurance de base plus égalitaire. Le rôle des organismes complémentaires serait concentré sur la couverture de la franchise, des dépassements d'honoraires et des prestations hors panier de soins de l'Assurance maladie. Comme le soulignent Geoffard et Lagasnerie, les individus pourraient concentrer leurs ressources sur la couverture de ces prestations supplémentaires. Nous pouvons imaginer la subsistance de systèmes d'aides (mais allégés) permettant aux individus d'obtenir un contrat couvrant suffisamment les dépassements d'honoraire, l'optique et le dentaire pour accéder à ces soins. En même temps, un

caractère responsable peut être maintenu pour ces contrats, avec un plafonnement des remboursements. Cette solution apparaît séduisante mais serait probablement difficile à mettre en place et pas forcément évidente à faire accepter pour les personnes les plus aisées, celles-ci payant alors des cotisations plus élevées pour un service plus faible. D'autre part, se pose la question de l'incitation que pourraient avoir les organismes à participer à la régulation de l'offre.

Quelle que soit l'approche retenue, il semble important d'un point de vue statistique d'analyser l'effet de l'assurance selon le niveau de revenu afin d'analyser les différences d'accessibilité selon le niveau de ressources.

5.4. Quelques remarques additionnelles

Un obstacle potentiel à l'accès aux soins et qui n'a pas été abordé est l'avance de frais par les patients. Le deuxième travail permet de constater que le passage à une exonération totale d'avance de frais pour les bénéficiaires de la CMU-C ne s'est pas traduit par une inflation des dépenses. Ce résultat apparaît intéressant au moment où des débats portent sur l'application du tiers payant intégral pour les consultations de médecins. Une telle exonération lèverait un obstacle supplémentaire à l'accès aux soins.

Concernant les prescriptions, il faut signaler que le choix de leur nature et de leur volume incombe en principe en grande partie au professionnel de santé. L'application de copaiements visant à responsabiliser le patient pose alors question. Il serait intéressant d'analyser plus en profondeur le lien entre la couverture complémentaire et le volume et la nature ainsi que la quantité d'actes ou de médicaments prescrits. Cette remarque conduit à souligner qu'une régulation de la demande soit également 'accompagner comme nous l'avons vu d'une régulation de l'offre, l'équilibre entre les deux dépendant de la nature des soins.

Annexe technique

6.1. Niveau d'assurance préféré en présence de risque moral (résolution par maximisation de l'utilité espérée)

Le programme de maximisation reflétant les choix des individus s'écrit :

$$\begin{aligned} & \max_{\substack{(x_s) \in \mathbb{R}_+^{S+1} \\ (m_s) \in \mathbb{R}_+^{S+1} \\ (q_m) \in \mathbb{R}_+^{S+1}}} \sum_{s \in \mathcal{S}} \pi_s \cdot U_s(x_s, m_s) \\ \text{sc: } & x_s + p_m \cdot m_s = W + (p_m - q_m) \cdot m_s - P_A \quad \forall s \in \mathcal{S} \\ & \text{sc: } P_A = (p_m - q_m) \cdot \bar{m}^\#(q_m, W) \\ & \text{sc: } q_m \leq p_m \end{aligned}$$

En remplaçant P_A dans la première condition, par sa valeur donnée dans la deuxième condition, le programme se réécrit :

$$\begin{aligned} & \max_{\substack{(x_s) \in \mathbb{R}_+^{S+1} \\ (m_s) \in \mathbb{R}_+^{S+1} \\ (q_m) \in \mathbb{R}_+^{S+1}}} \sum_{s \in \mathcal{S}} \pi_s \cdot U_s(x_s, m_s) \\ \text{sc: } & x_s + q_m \cdot m_s = W - (p_m - q_m) \cdot \bar{m}^\#(q_m, W) \quad \forall s \in \mathcal{S} \\ & \text{sc: } 0 \leq q_m \leq p_m \end{aligned}$$

$$\mathcal{L} = \sum_{s \in \mathcal{S}} \pi_s \cdot U_s(x_s, m_s) + \sum_{s \in \mathcal{S}} \lambda_s \cdot [W - (p_m - q_m) \cdot \bar{m}^\# - x_s - q_m \cdot m_s] + \mu \cdot [p_m - q_m] + \zeta \cdot q_m$$

Les multiplicateurs de Lagrange respectant les contraintes de positivités suivantes : $\lambda_s > 0 \forall s \in \mathcal{S}$ (la contrainte associée étant saturée), $\mu \geq 0$ et $\zeta \geq 0$. En dérivant \mathcal{L} par rapport à (x_s, m_s) pour tout $s \in \mathcal{S}$ et par rapport à q_m , nous obtenons les conditions de premier ordre de ce problème :

Équations 6.1-i

(a)

$$\pi_s \cdot \left. \frac{\partial U_s}{\partial x} \right|_{(x_s, m_s)} - \lambda_s = 0 \quad \forall s \in \mathcal{S}$$

(b)

$$\pi_s \cdot \frac{\partial U_s}{\partial m_s} \Big|_{(x_s, m_s)} - \lambda_s \cdot q_m = 0 \quad \forall s \in \mathcal{S}$$

(c)

$$\sum_{s \in \mathcal{S}} \lambda_s \cdot \left[\bar{m}^\#(q_m, W) - (p_m - q_m) \cdot \frac{\partial \bar{m}^\#}{\partial q_m} - m_s \right] = \mu - \zeta$$

Ces conditions sont équivalentes à :

Équations 6.1-ii

(a)

$$\pi_{s_1} \cdot \frac{\partial U_{s_1}}{\partial x} \Big|_{(x_{s_1}, m_{s_1})} = \frac{\lambda_{s_1}}{\lambda_{s_0}} \cdot \pi_{s_0} \cdot \frac{\partial U_{s_0}}{\partial x} \Big|_{(x_{s_0}, m_{s_0})} \quad \forall s_0, s_1 \in \mathcal{S}$$

(b)

$$\frac{\partial U_s}{\partial m_s} \Big|_{(x_s, m_s)} = q_m \cdot \frac{\partial U_s}{\partial x} \Big|_{(x_s, m_s)} \quad \forall s \in \mathcal{S}$$

(c)

$$\sum_{s \in \mathcal{S}} \lambda_s \cdot \left[m_s + (p_m - q_m) \cdot \frac{\partial \bar{m}^\#}{\partial q_m} - \bar{m}^\#(q_m, W) \right] = \zeta - \mu$$

La condition de premier ordre CPO2 et les contraintes budgétaires impliquent que les solutions en termes de consommations de soins sont telles que $(x_s, m_s) = (x_s^\#, m_s^\#)$. Autrement dit, les consommations de soins anticipées par les individus au moment de la souscription du contrat sont égales aux consommations de soins exécutées après la souscription. Après avoir remplacé m_s par m_s^{CA} et λ_s par sa valeur issue de l'équation 7.1 (a) :

$$q_m \cdot \sum_{s \in \mathcal{S}} \pi_s \cdot \frac{\partial U_s}{\partial x} \Big|_{(x_s^\#, m_s^\#)} \cdot \left[(p_m - q_m) \cdot \frac{\partial \bar{m}^\#}{\partial q_m} + m_s^\# - \bar{m}^\#(q_m, W) \right] = \zeta - \mu$$

Nous notons :

- $v_{x,s}$ l'utilité marginal du revenu dans l'état s , c'est-à-dire la dérivée de U_s par rapport à x ou, de manière équivalente, la dérivée de l'utilité indirecte V_s par rapport au revenu W :

$$v_{x,s} = \frac{\partial U_s}{\partial x} \Big|_{(x_s^{\#}, m_s^{\#})} = \frac{\partial V_s}{\partial W} \Big|_{(q_m, W)}$$

- $v_{m,x}$ la dérivée de U_s par rapport à m :

$$v_{m,s} = \frac{\partial U_s}{\partial x} \Big|_{(x_s, m_s)} = \frac{\partial V_s}{\partial W} \Big|_{(q_m, W)}$$

L'équation 6.1-ii (c) devient :

Équation 6.1-iii

$$\sum_{s \in \mathcal{S}} \pi_s \cdot v_{x,s} \cdot \left[(p_m - q_m) \cdot \frac{\partial \bar{m}^{\#}}{\partial q_m} + m_s^{\#}(q_m, W) - \bar{m}^{\#}(q_m, W) \right] = \zeta - \mu \quad \forall s \in \mathcal{S}$$

Nous allons nous appuyer sur cette dernière équation pour analyser le niveau de couverture préféré par les individus. Nous menons cette analyse en distinguant selon que l'élasticité-prix de la demande agrégée de soins est nulle ou strictement négative.

Cas 1 : l'élasticité de la demande agrégée de soins est nulle ($\partial m^{\#} / \partial q_m = 0$).

En incorporant cette hypothèse dans l'Équation 6.1-iii, cette dernière devient :

Équation 6.1-iv

$$\sum_{s \in \mathcal{S}} \pi_s \cdot v_{x,s} \cdot [m_s^{\#}(q_m, W) - \bar{m}^{\#}(q_m, W)] = \zeta - \mu$$

Nous remarquons que la somme à gauche de l'équation 6.1-iv correspond à la covariance $cov(\tilde{m}^{CA}, \tilde{v}_{x,s})$ avec \tilde{m}^{CA} et $\tilde{v}_{x,s}$ les variables aléatoires de demande de biens et services non médicaux et d'utilité marginale du revenu. De plus, $\tilde{m}^{\#}$ et $\tilde{v}_{x,s}$ covarient donc dans le même sens⁴⁹ :

$$\zeta - \mu = cov(\tilde{m}^{\#}(q_m, W), \tilde{v}_{x,s}) \geq 0$$

Nous avons donc $0 \leq \mu \leq \zeta$. Le cas où $0 < \mu \leq \zeta$ ne peut pas se produire car sinon, les contraintes $0 \leq q_m$ et $0 \leq q_m - p_m$ seraient toutes les deux saturées ; cela impliquerait $p_m = q_m = 0$, ce qui est impossible puisque $p_m > 0$. Donc $0 = \mu \leq \zeta$ de sorte que la contrainte $q_m = 0$ est saturée. Le coût résiduel q_m des soins est nul, autrement dit la couverture est complète.

⁴⁹ $m_{s_0}^{\#} \leq m_{s_1}^{\#}$ implique $x_{s_0}^{\#} \geq x_{s_1}^{\#}$ d'après l'équation d'équilibre budgétaire, ce qui entraîne, par décroissance de la fonction u'_x :

$$v_{x,s_0} = u'_x(x_{s_0}^{\#}) \leq u'_x(x_{s_1}^{\#}) = v_{x,s_1}$$

Cas 2 : l'élasticité de la demande agrégée de soins est strictement négative ($\partial m^\# / \partial q_m < 0$).

Supposons tout d'abord que $q_m = 0$. Dans ce cas, l'utilité marginale du revenu est constante selon les états de la nature, égale à v_x .

Équation 6.1-v

$$v_x \cdot \sum_{s \in \mathcal{S}} \pi_s \cdot \left[(p_m - q_m) \cdot \frac{\partial \bar{m}^\#}{\partial q_m} + m_s^\#(q_m, W) - \bar{m}^\#(q_m, W) \right] = \zeta - \mu \quad \forall s \in \mathcal{S}$$

Or :

$$\sum_{s \in \mathcal{S}} \pi_s \cdot [m_s^\#(q_m, W) - \bar{m}^\#(q_m, W)] = 0$$

L'équation 6.1-v devient donc :

$$\zeta - \mu = v_x \cdot (p_m - q_m) \cdot \frac{\partial \bar{m}^\#}{\partial q_m} < 0$$

Nous avons donc $0 \leq \zeta < \mu$. Si les deux inégalités étaient strictes, cela signifierait que $p_m = q_m = 0$, ce qui est impossible puisque $p_m > 0$. Le seul cas de figure possible est donc $\mu > 0$ et $\zeta = 0$, ce qui signifie que $0 < q_m$ et $q_m = p_m$: le coût résiduel des soins est donc égal à p_m autrement dit la couverture est nulle ce qui est en contradiction avec l'hypothèse d'une couverture complète. La couverture optimale est donc incomplète, ce qui est contradictoire avec l'hypothèse $q_m = 0$.

Supposons maintenant que $p_m = q_m$. La couverture étant incomplète, le multiplicateur ζ est nul. En remplaçant q_m par p_m dans l', nous obtenons :

$$\sum_{s \in \mathcal{S}} \pi_s \cdot v_{x,s} \cdot [m_s^\#(p_m, W) - \bar{m}^\#(p_m, W)] = -\mu \leq 0$$

En écrivant la partie de gauche de cette égalité sous la forme de la covariance $cov(\tilde{m}^\#(p_m, W), \tilde{v}_{x,s})$ et en se rappelant que $\tilde{m}^\#$ et $\tilde{v}_{x,s}$ covarient dans le même sens, nous obtenons :

$$0 \leq cov(\tilde{m}^\#(p_m, W), \tilde{v}_{x,s}) = -\mu \leq 0$$

Ainsi, $\mu = cov(\tilde{m}^{CA}(p_m, W), \tilde{v}_{x,s}) = 0$, la contrainte $p_m - q_m \geq 0$ n'est donc pas saturée. La couverture optimale n'est donc pas nulle, ce qui est contradictoire avec l'hypothèse $p_m = q_m$. Finalement nous avons :

$$0 < q_m < p_m$$

$$\mu = \zeta = 0$$

De la nullité des multiplicateurs μ et ζ , nous tirons :

$$\sum_{s \in \mathcal{S}} \pi_s \cdot v_{x,s} \cdot \left[(p_m - q_m) \cdot \frac{\partial \bar{m}^\#}{\partial q_m} + m_s^\#(q_m, W) - \bar{m}^\#(q_m, W) \right] = 0$$

Dans le cas où l'élasticité-prix de $\bar{m}^\#$ est strictement négative, le niveau optimal de couverture est donc déterminé par l'égalité :

Équation 6.1-vi

$$\sum_{s \in \mathcal{S}} \pi_s \cdot v_{x,s} \cdot [\bar{m}^\#(q_m, W) - m_s^\#] = (p_m - q_m) \cdot \frac{\partial \bar{m}^\#}{\partial q_m} \cdot \sum_{s \in \mathcal{S}} \pi_s \cdot v_{x,s} = (p_m - q_m) \cdot \frac{\partial \bar{m}^\#}{\partial q_m} \cdot \bar{v}_x$$

De cette égalité, nous tirons la formule suivante du taux de couverture optimal⁵⁰ :

$$\frac{a^\#}{1 - a^\#} = \frac{1 - \frac{q_m}{p_m}}{1 - \left(1 - \frac{q_m}{p_m}\right)} = -\frac{1}{\eta_{\bar{\psi}^\#, q_m}} \cdot \text{cov} \left(\frac{\tilde{v}_x}{\bar{v}_x}, \frac{\tilde{m}^\#}{\bar{m}^\#} \right)$$

⁵⁰ Nous écrivons l'équation 6.1-vi sous la forme suivante :

$$-\text{cov}(\tilde{v}_x, \tilde{m}^\#) = (p_m - q_m) \cdot E(\tilde{v}_x) \cdot E \left(\frac{\partial \tilde{m}^\#}{\partial q_m} \right) = \frac{p_m}{q_m} \cdot \left(1 - \frac{q_m}{p_m}\right) \cdot \bar{v}_x \cdot \frac{q_m}{\bar{m}^\#} \cdot \frac{\partial \bar{m}^\#}{\partial q_m}$$

Bibliographie

- Albouy V., Crepon B. (2007). Aléa moral en santé : une évaluation dans le cadre du modèle causal de Rubin. *Institut National de la Statistique et des Etudes Economiques. (I.N.S.E.E.). Direction des Etudes et Synthèses Economiques*. Document de travail 2007/12.
- Auvray L., Doussin A., Le Fur P. (2003). Santé, Soins et Protection Sociale en 2002. *Questions d'Economie de la Santé, N°78*.
- Arrow K.-J. (1963). Uncertainty and the welfare of medical care. *American Economic Review*, 53 (5), pp. 941-973.
- Auerbach, A.-J., Hines Jr, J.-R. (2002). Taxation and economic efficiency. *Handbook of public economics*, Volume 3, pp. 1347-1421.
- Bardey., D., Couffinhal A, Grignon G. (2002). Trop d'assurance peut-il être néfaste ? Théorie du risque moral ex post en santé. *Questions d'économie de la Santé, N°53*.
- Blaxter, M. (1990). Health and lifestyles. London and New York: *Tavistock/Routledge Publication*.
- Blomqvist, A. (2001). Does the economics of moral hazard need to be revisited? A comment on the paper by John Nyman. *Journal of Health Economics*, 20 (2), pp. 283-288.
- Bocognano A, Couffinhal A, Dumesnil S, Grignon M (2000). La complémentaire maladie en France: qui bénéficie de quels remboursements? *CREDES Report 1317*.
- Bocognano, A., Dumesnil, S., Frérot, L., Grandfils, N., Le Fur, P., Sermet, C. 1999. Santé , Soins et Protection Sociale en 1998. *IRDES Report 1282*.
- Bolhaar J, Lindeboom M, Van der Klaauw B. (2008). A dynamic analysis of the demand for health insurance and healthcare. *IZA Discussion Papers 3698. Institute for the Study of Labor (IZA)*.
- Brook, R. H., Vaiana, M. E. (1989). Appropriateness of Care. *National Health Policy Forum*.
- Buchmueller T.-C., Couffinhal A. (2004). Private Health Insurance in France. *OECD health working paper*, 12. Available at <http://www.oecd.org/dataoecd/35/11/30455292.pdf>.
- Buchmueller T.-C., Couffinhal A., Grignon M., Perronnin M. (2004). Access to physician services: does supplemental insurance matter? Evidence from France. *Health Economics*, 13 (7), pp. 669-687.
- Cartwright W, Hu T, Huang L. (1992). Impact of varying Medigap insurance coverage on the use of medical services of the elderly. *Applied Economics*; 24: 529–539.
- Caussat L., Glaude M. (1993). Dépenses médicales et couverture sociale. *Economie et Statistique*, 265, pp. 31-43.
- Chamberlain G. (1984). Panel data. *In: Handbook of Econometrics*, vol. 2. *North-Holland: Amsterdam*.
- Chamberlin E. (1933). The theory of monopolistic competition. *Cambridge: Harvard University Press*.

- Cherkin, Daniel C., Louis Grothaus and Edward H. Wagner (1989). The effect of office visit copayments on utilization in a health maintenance organization. *Medical Care*, 27(7), pp. 669-79.
- Chiappori P.-A., Durand F., Geoffard P.-Y. (1998). Moral Hazard and the Demand for Physician Services: First Lessons from a French Natural Experiment. *European Economic Review*, 42, pp. 499–511.
- Colle, A.-D., Grossman, M. (1978). Determinants of pediatric care utilization. *Journal of Human Resources*, Vol. 13 (Supp. 1978), pp. 115-53.
- Cooper P.-F., Nichols L.-M., Taylor A.-K. (1996). Patient choice of physician: do health insurance and physician characteristics matter? *Inquiry*, 33(3), pp. 237–246.
- Congressional Budget Office. (2002). Issues in Designing Prescription Drugs Benefit for Medicare. Congressional Budget Office, Washington.
- Couffinhal A. (1999). Concurrence en assurance santé : entre efficacité et sélection. *Thèse de doctorat. Sous la direction de Rochaix L., Université Paris-IX Dauphine.*
- Couffinhal A., Paris V. (2001) Utilization fees imposed to public health systems users in France. Proceedings of a workshop organized for the Commission on the Future of Health Care in Canada.
- Couffinhal A., Perronnin M., avec l'appui de Chevalier J. et Lengagne P. (2004). Accès à la couverture complémentaire maladie en France : une comparaison des niveaux de remboursement. Enquête ESPS 2000-2002. *Rapport IRDE,S N° 1521.*
- Currie, J., Madrian, B. C. (1999). Health, health insurance and the labor market. *Handbook of labor economics*, Volume 3, pp. 3309-3416.
- Currie, J., Thomas, D. (1995). Medical care for children: Public insurance, private insurance, and racial differences in utilization. *Journal of Human Resources*, pp. 135-162.
- Cutler D.-M., Zeckhauser R.-J. (2000). The anatomy of Health insurance. *Handbook of Health Economics*, Volume 1, Chapter 11, pp. 563-643
- Delattre, E., Dormont, B. (2000). Induction de la demande de soins par les médecins libéraux français. *Etude micro-économétrique sur données de panel. Économie & prévision*, 142(1), pp. 137-161.
- Després C., Dourgnon P., Fantin R., Jusot F. (2011), « Le renoncement aux soins pour raisons financières : une approche économétrique », *Irdes, Questions d'économie de la santé*, N°170.
- De Meza, D. (1983). Health insurance and the demand for medical care. *Journal of Health Economics*, 2, pp. 47–54
- Dionne, G. (1981). Moral hazard and search activity. *Journal of Risk and Insurance*, pp. 422-434.

- Dixit A., Sandmo A. (1977). Some simplified formulae for optimal income taxation. *The Scandinavian Journal of Economics*, 417-423.
- Dowd, B. E. (1982). The Logic of Moral Hazard: A game theoretic illustration. *The Journal of Risk and Insurance*, 49(3), 443-447.
- Dranove, D., & Satterthwaite, M. A. (2000). The industrial organization of health care markets. *Handbook of health economics*, 1, 1093-1139.
- Ehrlich I., Becker G.-S (1972). Market Insurance, Self-Insurance and Self-Protection. *The Journal of Political Economy*, 80 (4), pp. 623-648.
- Ettner S. (1997). Adverse selection, the purchase of medigap insurance by the elderly. *Journal of Health Economics*, 16(5), pp. 543–562.
- Evans W.-N., Schoenman J.-A., Schur L.-C. (2002). Primary care case management for medicaid recipients: evaluation of the Maryland access to care program. *Inquiry* 34, pp. 155–170.
- Feldman R., Dowd B. (1991). A New Estimate of the Welfare Loss of Excess of Health Insurance? *The American Economic review*, 81 (1), pp. 297-301.
- Feldman R., Dowd B. (1993). What does the demand curve for medical care measure? *Journal of Health Economics*, 12 (2), pp. 193-200.
- Feldstein, P., Severson, R. (1964). Report of the Commission on the Cost of Medical Care. *The American Medical Association, Chicago, IL*.
- Feldstein M.-S. (1971), Hospital Cost Inflation: A Study in Nonprofit Price Dynamics. *A.E.R.* 61 (5), pp. 853-872.
- Feldstein M.-S. (1973), The Welfare Loss of Excess Health Insurance. *The Journal of Political Economy*, Vol. 81, No. 2, Part 1., pp. 251-280.
- Feldstein, M., Friedman, B. (1977). Tax subsidies, the rational demand for insurance and the health care crisis. *Journal of Public Economics*, 7(2), 155-178.
- Frank, R. G., Lamiraud, K. (2009). Choice, price competition and complexity in markets for health insurance. *Journal of Economic Behavior and Organization*, 71(2), 550-562.
- Geoffard, P. Y., de Lagasnerie, G. (2012). Réformer le système de remboursement pour les soins de ville: Une analyse par micro simulation. *Economie et Statistiques*, N° 455-456.
- Genier P. (1998). Assurance et recours aux soins. Une analyse microéconomique à partir de l'enquête Santé 1991-1992 de l'INSEE. *Revue économique*, 49(3), pp 809-819.
- Grignon M., Perronnin M., Lavis J. (2007). Does free complementary health insurance help the poor to access health care? Evidence from France. *Health Economics*, 17(2), pp 203-219.

- Grossman, M. (1972). On the concept of health capital and the demand for health. *The Journal of Political Economy*, 80(2), 223-255.
- Gruber, J. (2000). Health insurance and the labor market. *Handbook of health economics*, 1, pp. 645-706.
- Gardiol, L., Geoffard, P. Y., Grandchamp, C. (2005). Separating selection and incentive effects in health insurance. *Working Paper 2005-38, PES – Ecole Normale Supérieure de Paris, Paris*
- Hadley J, Holahan J. (2003). Covering the uninsured: how much would it cost? *Health Affairs, Web Exclusive W3*, pp. 260–265.
- Hahn B. (1994) Health care utilization: the effects of extending insurance to adults on medicaid or uninsured. *Medical Care*, 32(3), pp. 227–239.
- Harmon C, Nolan B. (2001). Health insurance and health services utilization in Ireland. *Health Economics*, 10(2), pp. 135–45.
- Haut Conseil pour l’Avenir de l’Assurance Maladie (HCAAM). 2006. Avis sur le médicament, adopté par le Haut Conseil pour l’Avenir de l’Assurance Maladie.
- Haut Comité de la santé publique (2002). La santé en France 2002. *La documentation française* (ed).
- Henriet D., Rochet J.-C. (1991). Microéconomie de l’assurance. *Economica, Collection « Economie et statistiques avancées »*.
- Hume, D. (1740). A treatise of human nature.
- Hurd MD, McGarry K. (1997). Medical insurance and the use of health care services by the elderly. *Journal of Health Economics*, 16(2), pp. 129–154.
- Kambia-Chopin B., Perronnin M. (2013) Deductibles and the Demand for Prescription Drugs: Evidence from French Data. *IRDES working paper n° 54*.
- Kaplan, S., Garrick, B.- J. (1981). On the quantitative definition of risk. *Risk analysis*, 1(1), pp. 11-27.
- Keeler E.-B, Rolph J., Duan N., Hanley J., Manning W., Jr. (1982). The demand for episodes of medical treatment: interim results from the health insurance experiment. *Monograph R-2829-HHS. Rand Corporation: Santa Monica, CA*.
- Keeler, E. B., & Rolph, J. E. (1988). The demand for episodes of treatment in the health insurance experiment. *Journal of Health Economics*, 7(4), pp. 337-367.
- Kilbreth E.-H., Coburn A.-F., McGuire C., Martin D.-P. *et al.* (1998). State-sponsored programs for the uninsured: is there adverse selection? *Inquiry*, 35, pp. 250–265.

- Laffont J.-J. (1985). Cours de Théorie économique. Volume II: Economie de l'Incertain et de l'Information. *Economica, Collection « Economie et statistiques avancées »*.
- Laffont, J. J., Martimort, D. (2009). The theory of incentives: the principal-agent model. *Princeton University Press*.
- Liang K.-Y., Zeger S.-L. (1986). Longitudinal data analysis for discrete and continuous outcomes. *Biometrics*, 16(5), pp. 121-130.
- Long, S.-H., Marquis, M.-S. (1994). The uninsured 'access gap' and the cost of universal coverage. *Health Affairs*, 2, pp. 211–220.
- Lohr, K.-N., Brook, R.-H., Kamberg, C.-J., *et al.* (1986). Use of medical care in the RAND Health Insurance Experiment: diagnosis-and service-specific analyses in a randomized controlled trial. *Medical care*, pp. 1-87.
- Lollivier S. (2001). Endogénéité d'une variable explicative dichotomique dans le cadre d'un modèle probit bivarié. *Annales d'Economie et de Statistique*, 62, pp. 251–269.
- Long S.-H., Marquis M.-S., Rodgers J. (1998). Do people shift their use of health services over time to take advantage of insurance? *Journal of Health Economics*, 17(1), pp. 105–115.
- Liu S., Chollet D. (2006). Price and Income Elasticity of the Demand for Health Insurance and Health Care Services: A Critical Review of the Literature. *Mathematica Policy Research*.
- Mc Guire, T. G. (2000). Physician agency. *Handbook of health economics*, 1, 461-536.
- Manning W. G., Newhouse J. P., Duan N., Keeler E. B., Leibowitz A., et Marquis S. M. (1987), « Health Insurance and the Demand for Medical Care: Evidence from a Randomized Experiment », *American Economic Review*, 77 (3), pp. 251-277.
- Manning, W.-G, Marquis, S. (1996). Health insurance: the tradeoff between risk pooling and moral hazard. *Journal of Health Economics*, 15, pp. 609–640.
- Marquis M.-S., Long S.-H. (1994–1995). The uninsured access gap: narrowing the estimates. *Inquiry*, 31(4), pp. 405–414.
- Marquis M.-S., Long S.-H. (1994–1995). The uninsured access gap: narrowing the estimates. *Inquiry*; 31(4), pp. 405–414.
- Mormiche P. (1993). Les disparités de recours aux soins en 1991. *Economie et Statistiques*, 265, pp. 45–52.
- Mossin, J. (1968). Aspects of rational insurance purchasing. *The Journal of Political Economy*, 76(4), 553-568.

- Newhouse, J.-P. (1978). Insurance benefits, out-of-pocket payments, and the demand for medical care: a review of the literature. *Rand Corporation*.
- Newhouse, J.-P., Phelps, C.-E., & Marquis, M.-S. (1980). On having your cake and eating it too: Econometric problems in estimating the demand for health services. *Journal of Econometrics*, 13(3), pp. 365-390.
- Newhouse, J.-P., Manning, W.-G., *et al* (1982). Some interim results from a controlled trial of cost sharing in health insurance. Monograph R-2847-HHS. Rand Corporation: Santa Monica, CA.
- Newhouse J.-P. (1993), *Free for All? Lessons from the RAND Health Insurance Experiment*, *Harvard University Press*.
- Nyman, J. (1999a). The value of health insurance: the access motive. *Journal of Health Economics*, 18, 141–152.
- Nyman, J. (1999b). The economics of moral hazard revisited. *Journal of Health Economics*, 18, 811–824.
- Nyman, J.-A. (2003). The theory of demand for health insurance. *Stanford University Press*.
- Nyman, J. (2008). Health Plan Switching and Attrition Bias in the RAND Health Insurance Experiment. *Journal of Health Politics, Policy and Law*, 33(2), pp 309-317.
- Paris V, Polton D, Sandier S. 2003. Recent developments in France. *Euro Observer: Newsletter of the European Observatory on Health Care Systems*, 5, pp. 5–6.
- Pauly M.-V., (1968). The Economics of Moral Hazard: Comment. *American Economic Review*, 58, pp. 531-537
- Pauly, M., Satterthwaite, M. (1981), The pricing of primary care physicians' services: A test of the role of consumer information. *Bell Journal of Economics*, (12) pp. 488-506.
- Peele P.-B. (1993). Evaluating welfare losses in the health care market. *Journal of Health Economics*, 12(2), pp. 205-208.
- Phelps C.-E., Newhouse J.-P. (1972). The effects of coinsurance on the demand for physician services. *Rand report prepared for Office of Economic Opportunity*, RE-976-OEO.
- Raynaud D. (2005). Les déterminants individuels des dépenses de santé: l'influence de la catégorie sociale et de l'assurance maladie complémentaire. *Etudes et Résultats* N° 378, DREES. French Ministry of Health.
- Rice T. (1992). An alternative framework for evaluating welfare losses in the health care market. *Journal of Health Economics*, 11 (1), pp.85-92

- Rice T. (1993). Demand curves, economists, and desert islands: a response to Feldman and Dowd. *Journal of Health Economics*, 12 (2), pp. 201-204.
- Rice T. (1993). A model is only as good as its assumptions: a reply to Peele. *Journal of Health Economics*, 12 (2), pp. 209-211.
- Ringel, J.-S., Hosek, S.-D., Vollaard B.-A., Mahnovsky, S. (2002). The Elasticity of Demand for Health Care A Review of the Literature and Its Application to the Military Health System. *Rand Monograph Reports*, N° 1355.
- Romanow, R. (2002). Building on Values. The Future of Health Care in Canada, Chapter 9. Health Canada: Ottawa.
- Rosett, R.-N., Huang, L- F. (1973). The effect of health insurance on the demand for medical care. *The Journal of Political Economy*, pp. 281-305.
- Scitovsky, A.-A., Snyder N.-M. (1972). Effect of coinsurance on the use of physician services. *Social Security Bulletin*, 35, pp. 3-19.
- Schellhorn, M. (2001). The effect of variable health insurance deductibles on the demand for physician visits. *Health Economics*, 10(5), pp. 441–456.
- Schokkaert E., Van Ourti T., De Graeve D., Lecluyse, A., Van de Voorde, C. (2010). Supplemental health insurance and equality of access in Belgium. *Health Economics* 19(4), pp. 377–395.
- Shmueli A. (2001). The effect of health on acute care supplemental insurance ownership: an empirical analysis. *Health Economics*, 10(4), pp. 341–350.
- Short P.-F., Hahn, B.-A, Beauregard K, Harvey, P.-H, Wilets, M.-L. (1997). The effect of universal coverage on health expenditures for the uninsured. *Medical Care*, 35(2), pp. 95–113.
- Smith, A., (1776). *The Wealth of Nations*. The Modern Library, New York.
- Spillman B.-C. (1992). The impact of being uninsured on utilization of basic health care services. *Inquiry*; 29(4), pp. 457–466.
- Tilford J.-M., Robbins J.-M., Shema S.-J., Farmer F.-L.. (1999). Response to health insurance by previously uninsured rural children. *Health Services Research*, 34(3), pp. 761–775.
- Van de Voorde, C., Van Doorslaer, E, Schokkaert, E. (2001). Effects of cost sharing on physician utilization under favourable conditions for supplier-induced demand. *Health Economics*, 10(5), pp. 457–471.
- Van Doorslaer E, Wagstaff A. (1993). Equity in the finance of health care: methods and findings. *In Equity in the Finance and Delivery of Health Care. An International Perspective*. Van Doorslaer E, Wagstaff A and Rutten F (eds). Oxford University Press: Oxford.

- Van Doorslaer, E., Koolman, X., Jones, A.-M.. (2004). Explaining income-related inequalities in doctor utilization in Europe. *Health Economics*, 13(7), pp. 629–647.
- Van Ingen F. (2003). Couverture Maladie Universelle: Reaching the Poorest in France. <http://www.euro.who.int/document/sed/07case2.pdf>.
- Vera-Hernandez A.-M (1999). Duplicate coverage and demand for health care. The Case of Catalonia. *Health Economics*, 8(7), pp. 579–598.
- Wicksell, K. (1896). New principle in just taxation.
- Winkelmann, R. (2004). Co-payments for prescription drugs and the demand for doctor visits – evidence from a natural experiment. *Health Economics*, 13(11), pp. 1081–1089.
- Wooldridge J.-M. (1995). Selection corrections for panel data models under conditional mean independence assumptions. *Journal of Econometrics*, 68, pp. 115–132.
- Wooldridge J.-M. (2001). *Econometric Analyses of Cross-section and Panel Data*, Chapter 15. MIT Press: Cambridge, MA.
- Zweifel P, Manning WG. (2000). Moral hazard and the consumer incentives in healthcare. *Handbook of Health Economics*, vol. 1, pp. 409–459.

Effet de l'assurance complémentaire santé sur les consommations médicales : entre risque moral et amélioration de l'accès aux soins

Résumé : La théorie économique conduit à supposer que l'assurance santé accroît les consommations médicales en générant du risque moral ex-post, c'est-à-dire l'achat par les individus de soins dont la valeur est faible au regard de leur coût total. Certains économistes soulignent que tout ou partie de ce surplus de consommation peut résulter également d'une solvabilisation de la demande de soins du fait des remboursements, traduisant ainsi un meilleur accès aux soins. Comprendre lequel de ces effets prédomine représente un enjeu important en termes d'équité et d'efficacité du système de santé. Cette thèse vise à mesurer l'ampleur de l'effet de l'assurance complémentaire santé sur les consommations médicales en France et à en comprendre la nature en s'appuyant sur trois articles. Le premier article analyse l'effet du fait d'être couvert par une complémentaire santé sur le recours aux médecins, le second évalue l'impact de la CMU-C sur les consommations médicales de ses bénéficiaires, enfin le troisième étudie l'effet d'une surcomplémentaire santé permettant de compléter les remboursements du contrat de base d'une mutuelle de fonctionnaire. En se focalisant sur l'assurance complémentaire et surcomplémentaire, ces articles permettent d'étudier l'effet d'une variation à la marge du niveau de couverture sur différentes catégories de population.

Mots clés : assurance santé (JEL I13), allocation efficiente, analyse coût-bénéfice (JEL D61), information asymétrique (JEL D82).

Effect of complementary health insurance on medical care consumptions: risk moral and better access to health care components

Abstract: According to economic theory, health insurance raises medical care consumptions by inducing ex-post moral hazard behavior, it is to say the purchase of health care that individual value below their production cost. Nevertheless, among the economists community, some suggest that these additional consumptions may be the consequence of an increase of financial resources in case of illness due to reimbursements, reflecting a better access to health care. Understanding which of these effects dominates is a crucial issue to assess equity and efficiency of health care system. Based on three articles, this thesis aims to estimate how great the effect of complementary health insurance on health care consumptions is in France, and to understand the nature of this effect. The first article analyzes the effect of being covered by a complementary health insurance on the use of physician care, the second article assess how the CMU-C scheme affected health care consumptions of its beneficiaries, and the third article examines the effect of a supplementary health insurance contract that provides reimbursements in addition of the benefits of a basic complementary health insurance contract that covers civil servants. By focusing on complementary and supplementary health insurance, these articles offer the opportunity to study the effect of health insurance at the margin, on different populations.

Key words: health insurance (JEL I13), allocative efficiency, cost-benefit analysis (JEL D61), asymmetric information (JEL D82).