

ESSAIS EN ECONOMIE DE LA SANTE ET ASSURANCE

Thèse pour le doctorat en sciences économiques

présentée et soutenue publiquement le 19 décembre 2001 par

Franck BIEN

Directeur : Pierre PICARD, Professeur à l'Université Paris X-Nanterre
Rapporteurs : Christian GOLLIER, Professeur à l'Université de Toulouse I
Lise ROCHAIX, Professeur à l'Université d'Aix Marseille II
Suffragants : Louis EECKHOUDT, Professeur à l'Université de Mons
Pierre-Yves GEOFFARD, Directeur de Recherches CNRS

ESSAIS EN ECONOMIE DE LA SANTE ET DE L'ASSURANCE

Résumé : Cette thèse a pour objet l'étude des relations d'agence bilatérales dont prennent part trois acteurs du système de santé : médecin, patient et assurance-tutelle, uniquement dans le cadre du secteur non conventionné (*secteur 2*) de la médecine ambulatoire et de l'*assurance santé*.

Les deux premiers chapitres proposent une analyse du *secteur 2*. Le service médical est un bien de confiance puisque le patient observe l'action et ne saisit pas le résultat (la qualité de celui-ci). Nous montrons que les médecins opportunistes se construisent une réputation en fournissant une qualité élevée. Puis, celle-ci établie, ils diminuent la qualité du service. Ces comportements opportunistes mis en exergue légitiment une réforme de la fixation des honoraires du secteur 2. Nous établissons que les propriétés des contrats dépendent du nombre de médecins honnêtes dans l'économie et de la valeur du seuil de la qualité appréhendée par la tutelle.

Les quatre derniers chapitres traitent de l'*assurance santé* en introduisant une dimension sanitaire inobservable par l'assureur : l'utilité dépend de deux variables (richesse et santé). Tout d'abord, nous proposons une définition de la prime de risque qui corrobore l'équivalence de Pratt et nous montrons qu'elle est une fonction décroissante de l'état sanitaire. Ensuite, nous caractérisons les contrats d'assurance exclusifs en présence d'aléa moral discret. Nous obtenons, pour une faible désutilité de l'effort, que la pleine assurance ne désincite pas à effectuer un effort. Nous établissons également, pour des contrats séparateurs, que la couverture du dommage de l'individu en bonne santé est supérieure à celui en mauvaise santé. Enfin, dans un cadre d'antisélection, nous établissons que si la classification du risque est fondée sur l'espérance mathématique du dommage, alors les résultats standards ne sont plus vérifiés : présence de contrats de premier rang, inversion des résultats Rosthchild-Stiglitz. Ces résultats théoriques suggèrent l'adoption d'une très grande prudence quant à la vérification empirique de la présence d'antisélection sur les marchés d'assurance santé.

Mots – clés : Assurance santé, Asymétries d'information, Bien de confiance, Economie de la santé, Médecine ambulatoire, Honoraires libres, Risque sanitaire, Théorie des contrats.

ESSAYS ON HEALTH ECONOMICS AND INSURANCE

Abstract : The purpose of this PhD is to study agency relationship in health economics. We consider three actors : patient, doctor and regulatory-insurance in two frameworks : *French ambulatory medicine* and *health insurance*.

The first two chapters aim at analysing the free fees in *French ambulatory medicine*. Medical service is a credence good because the patient does not observe result but only action. We establish doctors are opportunist because they build their reputation on offering best quality and after they supply bad quality. The regulation of free fees of ambulatory medicine is based on doctors' opportunist actions. We can show that the properties of contracts depend on the number of honest doctors and the value of information's regulatory.

The last chapters examine *health insurance*. We use a bidimensional utility function (wealth and health state). First, we define the risk premium which verifies the usual property of the Arrow-Pratt risks premiums. However, we show negative link between risk premium and health status. We also characterise exclusive insurance contracts with discrete hazard moral. For low values of effort, the complete insurance is allowed and with the second-best contracts the patients with good health obtain a better coverage of damage than patients with bad health. Finally, with adverse selection, a new classification of risk, based on expectation of damage, disprove the Rosthchild-Stiglitz results. First-best contracts and opposite Rosthchild-Stiglitz results are possible. This conclusions aims at to be careful for testing adverse selection on the health insurance markets.

Keywords : Ambulatory medicine, Asymetric informations, Credence good, Economics of contracts, Free fees, Health economics, Health insurance, Health risk.

Classification : D81, D82, I11, I18, L14, L15.

Laboratoire d'accueil : THEMA (UMR CNRS 7536).

L'université Paris-X Nanterre n'entend donner aucune approbation ni improbation aux opinions émises dans cette thèse ; ces opinions doivent être considérées comme propres à leurs auteurs.

A Éliane

REMERCIEMENTS

J'exprime ma profonde gratitude à mon directeur de thèse, Monsieur le Professeur Pierre Picard, pour son attention, les conseils et les critiques qu'il a su me prodiguer.

Mes remerciements vont également à Philippe Batifoulier, Fabrice Bien, Olivier Biencourt, Emmanuel Boutron, Jeanne Follet, Nathalie Fombaron, Frédéric Gannon, Céline Mongin et Fabrice Tricou. Qu'ils trouvent ici l'expression de ma reconnaissance à la mesure de ce qu'ils ont apporté à ce travail.

Je souhaite également remercier Fayçal Methenni, Irène Schaeffer en raison du professionnalisme avec lequel ils remplissent leur fonction et le laboratoire THEMA pour m'avoir offert un digne cadre de travail.

J'associe également à ces remerciements le Ministère de l'Education Nationale de la Recherche et de la Technologie ainsi que l'Université Paris X-Nanterre pour le soutien financier qu'ils m'ont accordé pendant cette recherche.

Finalement, je remercie l'institution scolaire (dans son acception générale) pour m'avoir éduqué (dans son acception philosophique) et fourni cette chance (dans son acception sociologique) de soutenir publiquement cette thèse.

*A cette heure encore
Je veux croire
Qu'il existerait
Un droit de la voûte,
Cette vérité arquée
De l'espace.
Fléchie par l'œil,
Infinie,
Céleste,
Elle fait fléchir le fer,
La volonté,
Mortel,
D'être un dieu.*

Ernst Meister.

Table des matières

1	Introduction Générale	5
1.1	Economie de la santé ¹	5
1.2	Incertitude et asymétrie d'information	8
1.3	Une application de la théorie des contrats	11
1.3.1	La médecine ambulatoire	13
1.3.2	L'assurance santé	14
1.3.3	Organisation de la thèse	15
2	Une analyse des honoraires libres	19
2.1	Introduction	19
2.2	L'absence d'un signal de qualité	21
2.2.1	Une revue de la littérature	22
2.2.2	Les faits stylisés	24
2.2.3	Le prix, un signal de qualité	26
2.2.4	La durée de consultation, un signal de qualité?	27
2.3	La prestation médicale : un bien de confiance	28
2.3.1	Un bien spécifique	29
2.3.2	Les différentes modélisations	31
2.3.3	L'opportunisme des médecins	34
2.3.4	Réputation individuelle et bien de confiance	36
2.4	Conclusion	46
3	La réforme du secteur 2	49
3.1	Introduction	49
3.2	La rémunération des médecins	50
3.2.1	Comportement de demande	51
3.2.2	Comportement d'offre	51
3.3	Homo economicus versus Homo hippocraticus	53
3.3.1	Le serment d'Hippocrate	53

¹Une définition possible et contestée est : "l'économie de la santé est l'étude la répartition de façon rationnelle des ressources permettant de lutter contre les formes de la morbidité ou de les prévenir, en essayant de trouver les procédés les meilleurs au moindre coût" (Labourdet, 1988). La spécificité de cette discipline tient au fait que les conséquences d'une maladie engendrent une dépréciation de la richesse financière (coût du traitement) et de l'altérité physique de l'individu.

3.3.2	Les implications	54
3.3.3	Comportements des médecins	55
3.4	Honnêteté et bien de confiance	56
3.4.1	Le modèle	57
3.4.2	Antisélection ou bien de confiance	58
3.4.3	Antisélection et bien de confiance	62
3.4.4	Introduction de l'honnêteté des médecins	65
3.5	Conclusion	71
4	Une analyse du risque sanitaire	73
4.1	Introduction	73
4.2	Primes de risque	75
4.2.1	Définition de la prime de risque sanitaire	75
4.2.2	Primes de risques et soins sanitaires	76
4.3	Prudence	79
4.3.1	Définition	80
4.3.2	Risque sanitaire assurable	82
4.3.3	Risque sanitaire non assurable	85
4.4	Conclusion	90
5	Santé, prévention et assurance	91
5.1	Introduction	91
5.2	Le modèle en information parfaite	93
5.3	Le modèle avec une seule asymétrie d'information	96
5.3.1	L'état de santé de l'individu est inobservable	96
5.3.2	L'effort d'autoprotection est inobservable	97
5.4	Inobservabilité de l'effort et de l'état de santé	102
5.4.1	Définition de l'équilibre	102
5.4.2	Le programme	103
5.4.3	Etude des contraintes	104
5.4.4	Les contrats d'équilibre	107
5.5	Conclusion	113
6	Antisélection et assurance : un état des lieux	115
6.1	Introduction	115
6.2	Le modèle de RS (1976)	117
6.2.1	Le cadre	117
6.2.2	L'équilibre	118
6.3	Extensions	123
6.3.1	Tailles différentes de dommage	123
6.3.2	Aversions au risque différentes	124
6.3.3	Une généralisation	132
6.4	Validation empirique de l'antisélection	133
6.5	Conclusion	136

7	Antisélection et santé	139
7.1	Introduction	139
7.2	Généralisation du modèle RS à un cadre sanitaire	140
7.3	Hétérogénéité de l'utilité	143
7.4	Dommages hétérogènes	146
7.4.1	Le modèle	147
7.4.2	Etudes des contraintes d'incitation	149
7.4.3	Equilibre du marché d'assurance	154
7.5	Conclusion	167
8	Conclusion Générale	169
8.1	La médecine ambulatoire	169
8.2	L'assurance santé	170

Table des figures

6.1	Contrats d'assurance en information parfaite	119
6.2	Inexistence d'un contrat mélangeant	120
6.3	Contrats d'assurance en information imparfaite	121
6.4	Existence de l'équilibre RS	122
6.5	Contrats de premier rang	124
6.6	Existence de profits positifs	128
6.7	Configuration RS	130
6.8	Equilibre mélangeant	131
7.1	Equilibre séparableur en information parfaite	142
7.2	Equilibre semi-séparableur à subventions croisées	145
7.3	Contraintes d'incitation : $p_H > p_L$ et $D_H > D_L$	151
7.4	Contraintes d'incitation : $p_H > p_L$ et $D_H < D_L$	152
7.5	Contraintes d'incitation : $p_L > p_H$ et $D_H > D_L$	153
7.6	Equilibre de premier rang	155
7.7	Equilibre de type RS	157
7.8	Equilibre séparableur	159
7.9	$S_H = \underline{S}_H$	159
7.10	Contrat L à profits positifs et couverture partielle	160
7.11	Contrat L à profits positifs et couverture complète	161
7.12	Existence de l'équilibre	162
7.13	Equilibre de premier rang	163
7.14	Equilibre RS inversé	164
7.15	Contrat H à profits positifs et couverture partielle	165
7.16	Contrat H à profits positifs et couverture totale	166
7.17	Existence de l'équilibre	167

Chapitre 1

Introduction Générale

“Il faut douter plus radicalement encore”, Nietzsche.

1.1 Economie de la santé¹

En 1997, 728.3 milliards de francs furent consacrées en France aux dépenses sanitaires, soit 13 350 francs par habitant. Ces données confirment une certaine stabilisation des dépenses de santé depuis 1993, qui s'établit à 9,6% du PIB. La France se classe quatrième pays de l'OCDE et deuxième dans l'Union Européenne. Un écart de 1.5 point est constaté entre la moyenne de ces zones économiques et la part française. La France dépense plus que les autres pays européens sans que cet écart ne s'explique par des différences sanitaires significatives. Par ailleurs, l'accroissement des dépenses ne s'accompagne pas d'une réduction des inégalités.

La nécessité d'une maîtrise des dépenses de santé, à laquelle les principaux pays industriels semblent se soumettre, n'a pas toujours été un objectif prioritaire des pouvoirs publics. Pendant les Trente Glorieuses, la croissance des dépenses de santé apparaissait au contraire comme légitime afin d'accroître le progrès médical. L'âge d'or de la médecine, qui débuta à l'après-guerre (avec la découverte des antibiotiques et autres vaccins), s'est prolongé jusqu'au milieu des années 70.

Cette croissance des dépenses sanitaires était d'autant plus acceptable que l'activité économique des pays était soutenue et que l'indicateur de bonne santé (espérance de vie) augmentait spectaculairement. Durant cette période, renoncer à l'augmentation des dépenses sanitaires ou les maîtriser était considéré comme nuisible pour le bien-être de la population (Commissariat Général du

¹Une définition possible et contestée est : "l'économie de la santé est l'étude la répartition de façon rationnelle des ressources permettant de lutter contre les formes de la morbidité ou de les prévenir, en essayant de trouver les procédés les meilleurs au moindre coût" (Labourdet, 1988). La spécificité de cette discipline tient au fait que les conséquences d'une maladie engendrent une dépréciation de la richesse financière (coût du traitement) et de l'altérité physique de l'individu.

Plan, 1994). En effet, une augmentation des dépenses de santé s'accompagnait d'une hausse de la qualité des soins. Les progrès médicaux et la croissance économique des Trente Glorieuses ne pouvaient conduire à la prise en considération de contraintes financières.

Toutefois, le déclin de la croissance économique marqua la résurgence des contraintes financières. La France comme les autres pays industriels à cherché à rationaliser l'offre de soins. La crise économique (liée aux crises pétrolières) fit prendre conscience aux Etats de la nécessité d'une stabilisation des prélèvements obligatoires. Ils devaient arbitrer entre les différentes fonctions collectives : la santé, le chômage, l'éducation, etc...

Après le second choc pétrolier, la réaction française se caractérisa par une juxtaposition de différents plans qui ne s'incrivirent pas dans un cadre de politique générale de la santé (Commissariat Général du Plan, 1994). Il est à noter qu'une des ruptures les plus emblématiques fut l'instauration du *numerus clausus* pour les médecins. Le choix explicite était de conserver les principes sur lesquels le système de santé est fondé et de l'adapter à un environnement économique moins favorable.

Les plans d'assainissement ont été abondants (en fréquence un tous les 18 mois en moyenne entre 1978 et 1993) : Veil (1978), Barrot (1979), Questiaux (1981), Beregovoy (1983), Dufoix (1985), Seguin-Barzach (1986), Seguin (1986), Seguin (1987), Evin (1989), Evin (1990), Evin-Durieux (1991), Bianco (1991), Balladur (1993). Ceci atteste de l'absence d'une réforme structurelle notable. Le point commun de tous ces plans est de supposer une hausse du taux de cotisation ou une baisse du taux de remboursement (Mougeot, 1994). En cinquante ans, le remboursement des soins de ville est passé de 80% à 65% (Johanet, 1998). Or, il est de plus en plus difficile de recourir aux prélèvements obligatoires (sous peine de casser la croissance) ou à des emprunts (sous peine de déroger aux critères de Maastricht) pour financer le déficit de l'assurance maladie.

Un double constat rend donc indispensable une réforme structurelle du système de santé :

- Tout d'abord, il n'est plus possible de faire abstraction des contraintes économiques qui pèsent sur le système de santé. Lorsque plus de la moitié des prélèvements obligatoires consiste en des prélèvements sociaux, il est impossible d'envisager une maîtrise desdits prélèvements sans maîtriser les dépenses de santé. Ainsi, pour un salaire inférieur au plafond de Sécurité Sociale, le taux global de cotisation s'est accru de près de 15 points de 1970 à 1991, 4 points sur les 15 étant imputables aux dépenses de santé (Commissariat Général du Plan, 1993 a et b).

- Par ailleurs, si une augmentation des dépenses était possible, rien ne prouve que les mécanismes d'allocation des ressources garantissent une utilisation optimale de celles-ci. En effet, Evans et Stoddardt (1990) établissent que l'état de santé d'une population n'est pas en rapport direct avec le volume de ressources qu'elle consacre aux soins. Il est en revanche corrélé à un grand nombre de facteurs parmi lesquels figurent les inégalités de revenus individuels.

Ce second élément est d'autant plus discuté que la croissance économique est faible. En effet, une meilleure allocation des ressources laisse augurer une

diminution des dépenses de santé. Celle-ci semble d'autant plus réalisable que l'importance des gaspillages a été mise en évidence par le rapport Béraud (1992) qui l'estime à 120 milliards de dépenses sur 575 milliards engagés. Même si le chiffre est contestable, l'inefficacité du système français est mise en exergue de manière quasi unanime.

C'est pourquoi, l'augmentation constante des dépenses de santé à un rythme supérieur au PIB et cette inefficacité latente ont suscité l'inquiétude des pouvoirs publics et motivé les économistes². Les déséquilibres macroéconomiques ont conduit dans le passé à des politiques de contrôle global de la santé et à une maîtrise comptable. (Mougeot, 1999).

Cette vision macroéconomique de la régulation des dépenses sanitaires a occulté la question essentielle des mécanismes microéconomiques à l'origine de ces déséquilibres, qui ne peuvent être laissés pour compte. La maîtrise comptable des dépenses sanitaires étant mal acceptée par les professionnels de la santé, il convient de proposer des règles et incitations pour réguler le système sanitaire afin que ces professionnels souscrivent à cette régulation.

Cependant, "toute analyse du système de santé est délicate car elle implique une interaction entre des variables médicales et économiques, entre des considérations éthiques, sociales et financières et entre des finalités contradictoires" (Mougeot, 1999).

La complexité des problèmes du système de santé provient de la difficile appréhension de la santé. L'état de santé, "bien final" recherché par le patient est obtenu à partir de services médicaux (biens intermédiaires) produits par le système de santé.

Sans doute convient-il, tout d'abord, de préciser la définition du mot santé retenu dans le cadre de cette thèse. Selon l'OMS, la santé est un "état complet de bien-être physique, mental et social". Elle ne se réduit nullement à l'absence de maladie ou d'infirmité. Une telle acception a l'avantage d'insister sur la pluralité des caractéristiques ; elle a le défaut de ne pas être toujours opérationnelle (Labourdette, 1988). Le bien-être devient donc difficile à appréhender. La santé est moins un état idéal vers lequel elle doit tendre qu'une absence de "morbidité diagnostiquée"³. Elle est diagnostiquée par un généraliste ou par un spécialiste qui recourt à une interprétation scientifique, elle-même dépendant partiellement de l'observateur et de l'observé. Elle est coûteuse. L'économiste intervient donc pour l'arbitrage des dépenses quand pèse une contrainte financière. Même si l'économie de la santé met en jeu la vie humaine, elle ne peut nier les considérations économiques, telles que le rendement, les dépenses. C'est pour cela qu nous supposons, dans le cadre de cette thèse, que la santé est principalement un capital, une richesse non monétaire (Becker, 1964 ; Grossman, 1972) et que l'individu arbitre entre sa richesse monétaire et sa richesse sanitaire. Les personnes héritent d'un capital ou stock de santé initial qui se déprécie avec la maladie mais qui peut être restauré avec des dépenses curatives dans un cadre

²Il est à noter que l'article fondateur de l'économie de la santé (Arrow, 1963) est antérieur à la crise du système de santé. Il tentait d'explorer la spécificité des soins médicaux et de leur appliquer les principes de l'économie normative.

³La morbidité caractérise une gêne, une douleur, la crainte qu'une maladie se déclare.

statique et également avec un investissement dans un cadre dynamique. Aussi, nous postulons que le patient cherche à recouvrer son état de santé initial.

Pour recouvrer son état de santé initial, il fait appel à des services médicaux qui servent à “réparer” le corps humain⁵. Ces services sont spécifiques puisqu’un patient mécontent du service offert ne peut le faire reprendre par le producteur sanitaire. “Par sa nature, ce service implique la participation du corps du patient” (Phelps, 1995). Ils n’impliquent pas une obligation de guérison. Si une erreur médicale est commise, il est plus difficile d’accepter un échange de services que de prouver cette première. Si la plupart des erreurs peut être corrigée par un autre service, il n’en demeure pas moins qu’elles sont plus difficiles à corriger que pour des biens et services plus standards.

Le caractère particulier du secteur médical implique qu’il est alors malséant de croire que peuvent être appliqués les résultats traditionnels de la science économique établis pour des biens et produits dans un cadre d’information parfaite. En effet, le secteur de la santé est particulier en raison de l’incertitude qui le régit, des asymétries d’information qui le caractérisent et des services qu’il produit.

1.2 Incertitude et asymétrie d’information

Le secteur de la santé est, tout d’abord, caractérisé par l’omniprésence des phénomènes d’incertitude. Sous différentes formes, elle est présente. Elle débute chez le patient qui fait appel aux soins médicaux sur la base d’événements incertains mais réalisés : jambe cassée, infarctus. Elle se poursuit par la morbidité ressentie c’est-à-dire la modification de l’état de santé jugée par le patient : à la suite de de gênes, de douleurs, de restrictions de capacités, d’anomalies physiques, il craint ou prend conscience qu’une maladie se développe.

Elle est également présente du côté des pourvoyeurs de soins qui ne semblent pas l’avouer. Le diagnostic juste s’avère quasiment impossible plus spécialement pour le généraliste. “Il réfléchit en court-circuitant l’analyse rigoureuse du cas pathologique pour aller à l’essentiel, à la décision thérapeutique” (Guérin, 1995). Prenons un exemple. Un individu se rend chez son médecin pour un état grippal sans fièvre. Après un examen minutieux, le médecin utilise un raisonnement déductif par négation. Le patient ne souffre ni d’une bronchite, ni d’une pneumonie, et il ne s’agit pas d’une grippe. Est-ce un virus ou une surinfection bactérienne ? L’examen diagnostique ne permet pas de le découvrir. La plupart du temps, des antibiotiques sont prescrits afin de détruire les hypothétiques bactéries. De plus, deux praticiens pourront fournir pour un même ensemble de symptômes, un diagnostic différent. Le savoir médical repose depuis un siècle sur la possibilité de repérer les différentes affections, de les nommer et de les classer. Certains symptômes sont communs à plusieurs pathologies. Malgré l’imprécision du diagnostic, la médecine abute le traitement rationnel, l’anticipation et l’évitement des complications possibles.

⁵Phelps (1995) définissait les soins médicaux comme des séries d’activités conçues spécifiquement pour restaurer ou augmenter le capital santé.

Une fois le traitement établi à l'aide du diagnostic, le patient continue à subir l'incertitude. En effet, il ne peut juger de la qualité de la prestation médicale et de la véracité du diagnostic des médecins. Un traitement qui ne produit pas l'effet escompté - la disparition de certains signes pathologiques - est-il la résultante d'une mauvaise qualité (délibérée ou non) de diagnostic qui établit un traitement erroné ou une complication qui survient impossible à prévoir ? L'incertitude quant à la qualité du produit sanitaire est peut être plus grande que celle des autres biens et services (Arrow, 1963). La guérison d'une maladie est aussi imprévisible que l'incidence de la maladie. Des effets externes au traitement physique proprement dit tels que l'effet placebo ou nocebo⁶ ne peuvent être occultés. Pour la plupart des biens et services, la possibilité d'apprendre de sa propre expérience ou de celle des autres est grande parce qu'il existe un nombre d'essais importants. Pour une maladie grave, à l'incertitude due à l'inexpérience s'ajoute la difficulté de la prévision. Pour cette raison, nous considérons qu'un service sanitaire est un bien de confiance. Le patient, après une consultation médicale, est incapable d'en appréhender la qualité.

Cette incertitude revêt un caractère spécifique comme le souligne Arrow (1963). Elle est très différente pour les deux protagonistes : le consommateur et le pourvoyeur de soins. En raison de la complexité médicale, le médecin détient nécessairement plus d'informations que le patient. Il connaît mieux les conséquences et les possibilités de traitement que le patient. Il y a toujours une différence d'informations entre le consommateur et le pourvoyeur de soins sanitaires.

Le contrat résultant de la négociation de ces deux agents dépendra de l'asymétrie d'information. Un ancien proverbe affirme "Qui détient la connaissance, détient le pouvoir". Pour les services médicaux, l'obligation professionnelle, l'éthique et la responsabilité personnelle contribuent à ce que le médecin peut se comporter honnêtement. Cependant, Domenighetti (1994) montre indirectement que certains médecins utilisent leur pouvoir discrétionnaire dans la prescription de certains actes et il qu'il en résulte une sur-utilisation injustifiée. Le médecin a donc la possibilité de tromper son patient et de lui fournir des soins inutiles ou de mauvaise qualité. Ce dernier n'a aucun moyen de contrôler son médecin. Comme le souligne Phelps (1995) "après tout, si les patients décident de consulter les médecins, c'est pour obtenir leur conseil" avisé ou non.

Cette situation d'asymétrie informationnelle n'est pas propre à l'interaction médecin-patient. Elle est présente dans toutes les relations des acteurs du système de santé. Pour simplifier, nous considérons, ci-après, une organisation tripartite du système de santé : le producteur de soins (le médecin), le consommateur (le patient) et la tutelle ou la compagnie d'assurance.

Nous occultons une quatrième relation d'agence qui doit être prise en compte à l'intérieur de la notion de tutelle. Cette distinction se justifierait par des fonctions objectives différentes du gouvernement et de la Sécurité Sociale (tiers-payant) : les objectifs du tiers-payant sont représentés par une simple contrainte

⁶L'effet nocebo caractérise le développement d'effets secondaires annoncés lors de l'absorption d'un placebo.

d'équilibre budgétaire pour un système d'assurance socialisée ou une maximisation des profits pour des assureurs privés et les objectifs du gouvernement sont un contrôle des conditions d'accès aux soins et la mise en place de politiques économiques (voir Béjean et Gadreau, 1992). La prise en compte de cette relation ne modifie pas de manière fondamentale l'analyse mais permet une meilleure appréhension du système sanitaire français (Rochaix, 1993).

L'interaction patient-médecin (première relation d'agence) dépasse le cadre traditionnel du paradigme principal-agent de la théorie des contrats puisque le bien considéré est un bien de confiance. Dans un cadre standard, le principal n'observe pas l'action mais le résultat. Lors du colloque médecin-patient, le patient peut observer l'action du médecin (le contenu de la consultation) mais il ne peut appréhender le résultat⁷ en raison de ses connaissances médicales lacunaires⁸. C'est pour cela qu'il mandate la tutelle⁹ pour qu'elle le remplace afin de définir les termes du contrat avec le médecin (seconde relation d'agence). Nous supposons que les fonctions de la tutelle sont de servir les intérêts du patient quand elle le représente.

Cette substitution s'explique aussi bien par la solidarité que par l'information médicale dont la tutelle dispose. En France, la Sécurité Sociale peut réunir une commission de médecins¹⁰, qui à la lecture des ordonnances, peut définir imparfaitement, la pathologie dont le patient est atteint. La mandature de la tutelle permet de restaurer le cadre traditionnel de la théorie des contrats¹¹. Contrairement au patient, elle n'observe pas l'action mais elle peut juger la juger à l'aune de l'ordonnance ou des actes réalisés. La rémunération du médecin peut donc devenir contractuelle. La seconde raison d'être d'une telle substitution réside dans le principe de solidarité. Elle représente l'intérêt collectif dans la définition des contrats avec le médecin.

La tutelle, bien qu'elle représente les intérêts du patient, rencontre également un conflit d'intérêts dans un cadre assurantiel (troisième relation d'agence). Il est de deux types. Tout d'abord, le comportement de prévention du patient est inobservable par la tutelle alors même qu'il influe sur la probabilité d'occurrence de la maladie. Cette relation d'agence présente donc un risque moral. Cependant, bien que cette relation d'agence s'établisse avec une tutelle relevant d'un système d'assurance obligatoire, cette relation s'inscrit également dans un cadre de sélection adverse puisque l'assurance obligatoire n'est pas complète¹². En

⁷Voir Rochaix (1989).

⁸Atkinson (1999) est très circonspect quant à l'avantage informationnel que possède le médecin sur le patient. "L'avantage informationnel que la théorie attribue aux médecins est fortement limité par l'état des connaissances médicales"

⁹Dans un cadre d'assurance publique, les acteurs sont le médecin, le patient et la tutelle (Sécurité Sociale). Dans un cadre privé, les acteurs sont le médecin, le patient et la compagnie d'assurance. Dans les deux situations, le patient mandate soit la tutelle, soit choisit un Health Maintenance Organization qui contracte avec le médecin.

¹⁰La sécurité sociale dispose environ de 2300 praticiens-conseils pour expertiser.

¹¹La spécificité de la relation d'agence médecin-tutelle tient au fait que c'est l'action (effort) du médecin et non le résultat qui détermine la rémunération. Nous supposons que le résultat est corrélé avec l'action du médecin.

¹²Une assurance obligatoire universelle fournit une réponse de principe au problème de sélection adverse (Drèze, 1997).

effet, nombre de patients ont recours à une assurance complémentaire. Les entreprises d'assurance offrant cette dernière sont confrontées au phénomène d'antisélection. Un patient dont l'état de santé est bon sera moins tenté de souscrire une assurance complémentaire offrant une couverture complète du dommage sanitaire qu'un individu en mauvaise santé.

Il est très aisé pour une compagnie d'assurance automobile de connaître le type du conducteur puisque l'historique des accidents déclarés, provoqués ou subis sont connaissances communes ainsi que les caractéristiques du véhicule décrits par la photocopie de la carte grise. En revanche, il est délicat pour une compagnie d'assurance sanitaire de définir l'état de santé du patient¹³. L'individu ne possède pas de carte génétique (l'équivalent de la carte grise) ainsi que l'historique des maladies malignes ou bénignes. Toutefois, l'apparition du carnet médical aurait au moins permis d'établir l'historique médical du patient. Sous le prétexte du secret médical, les médecins n'ont pas utilisé ce carnet¹⁴.

Les relations bilatérales des trois acteurs du système de santé que nous avons définies s'inscrivent dans un cadre d'asymétrie d'information. Cette dernière est à l'origine de la délégation de décision. "La hiérarchie des pouvoirs que confère la détention d'information forme un écran entre les intérêts des acteurs, qui délèguent leur pouvoir de décision, et les objectifs de ceux qui sont habilités à décider... Ainsi, la non-révélation de l'information caractérise les comportements stratégiques justifiés par la recherche de l'intérêt individuel" (Béjean, 1994). Il convient donc d'inciter l'agent à révéler l'information qu'il possède au principal. La théorie des contrats permet la caractérisation de contrats incitatifs de second rang malgré l'asymétrie d'information qui régit la relation entre deux acteurs.

1.3 Une application de la théorie des contrats

Un marché de concurrence pure et parfaite, défini par la microéconomie traditionnelle, présente quatre caractéristiques : l'atomicité des participants (le marché comprend un grand nombre d'acheteurs et de vendeurs, et les actions d'un seul agent n'influencent pas le prix de marché), l'homogénéité des produits (les produits sont identiques de sorte que les acheteurs sont indifférents à l'identité du vendeur), la libre-entrée et libre sortie (les agents entrent sur un marché tant qu'ils ont un intérêt à y être actifs), et la transparence (les agents sont parfaitement informés du prix et de la qualité du produit). Ces quatre hypothèses excluent les relations stratégiques qui lient les agents. Dans le modèle de l'équilibre général walrassien, seuls les prix coordonnent les actions des agents. L'absence des interactions stratégiques des agents, au sein du cadre de l'équilibre général, est à l'origine de l'émergence et du développement de l'économie de l'information, de la théorie des contrats. Ces "modèles prennent en compte, sur des exemples forcément partiels, toute la complexité des comporte-

¹³Dans le cas des handicaps, le problème est différent puisque la prise en charge des frais médicaux peut être totale. Le problème de sélection adverse ne se pose plus.

¹⁴Le carnet médical était perçu par les professionnels de santé comme un instrument de contrôle des soins.

ments stratégiques d'un petit nombre d'agents au sein des liens institutionnels qui définissent la possibilité des actions... Ils résument les propriétés du cadre institutionnel prévalant à travers un contrat" (Salanié, 1997).

Ils peuvent être classés en trois grandes familles selon deux critères (l'information privée et la forme stratégique du jeu) (Salanié, 1997) :

- Les modèles d'autosélection : la partie non informée des caractéristiques de la partie informée a l'initiative ;
- les modèles de signaux : la partie informée a l'initiative ;
- les modèles d'aléa moral, la partie non informée qui a l'initiative du contrat observe imparfaitement les actions de la partie informée.

Ces modèles utilisent souvent le paradigme Principal-Agent qui met en relation deux agents : la partie informée qui détient une information pertinente pour le marchandage avec une partie non informée. Par simplification, un seul des deux agents possède le pouvoir de marchandage, c'est le principal, l'autre étant l'agent.

En conséquence, les asymétries d'information conduisent à des solutions de second rang engendrées par les intérêts conflictuels des deux agents en relation. Le contrat optimal défini par le principal doit inciter l'agent à révéler l'information dont il dispose. Pour ce faire, le principal peut être conduit à offrir une rente d'information qui diminue son profit (solution de second rang).

Cette thèse s'inscrit dans cette lignée. Son objectif est d'étudier les relations bilatérales qui lient les trois acteurs du système de santé : le médecin, le patient et la tutelle-assurance.

Bien qu'Arrow en 1963 place l'asymétrie d'information au coeur des problèmes de santé, ce n'est qu'assez récemment que les premières modélisations des comportements de production de soins en information imparfaite ont été réalisées. Rochaix (1997) explique ce constat par l'adage selon lequel : "la santé n'aurait pas de prix". L'application récente de la théorie des contrats permet toutefois de mieux prendre en compte les comportements stratégiques des acteurs et la définition de leurs incitations. "La politique de santé est en effet insuffisamment guidée par une analyse des fondements microéconomiques qui président au fonctionnement de ce marché... En France, c'est une politique par défaut et la croissance des dépenses est en conséquence plus subie que choisie" (Rochaix, 1997).

A travers l'application des modèles issus de la théorie des contrats, les avancées des économistes de la santé en matière de détermination de modes de rémunération sont plus probantes pour les établissements de santé que pour la médecine ambulatoire ou l'assurance santé (Rochaix, 1997). Cette thèse se propose donc d'étudier deux relations d'agence bilatérales centrées sur la médecine ambulatoire et l'assurance santé. Un intérêt tout particulier est porté au secteur 2 de la médecine ambulatoire et aux phénomènes d'aléa moral et d'antisélection propres à l'assurance santé.

1.3.1 La médecine ambulatoire

La question de la médecine ambulatoire est très complexe. En France, la tarification à l'acte fixée par l'autorité de tutelle n'incite pas le médecin à limiter le nombre d'actes. En présence d'honoraires fixes, les médecins ont intérêt à accroître leur production et à se différencier de leurs collègues en proposant différents services (arrêts de travail, examens médicaux, etc...) (Mougeot, 1994). Les consultations et visites ont été multipliées par deux entre 1975 et 1990 sans qu'une explication démographique ne puisse justifier une telle évolution.

Ces stratégies de surconsommation sont d'autant plus aisées que l'incertitude médicale régit la relation médecin-patient¹⁵. Ces comportements opportunistes sont pris en compte puisque l'autorité de tutelle définit chaque année, avec les médecins, une enveloppe globale de soins avec primes ou versements si celle-ci n'est pas respectée.

Cependant un autre problème est à considérer, celui de la qualité de la prestation médicale. L'introduction de la liberté tarifaire par la convention médicale de 1980 a-t-elle incité les médecins dont les honoraires sont libres à fournir une meilleure qualité de la prestation médicale ? Un prix plus élevé signale-t-il une qualité plus élevée ?

Les médecins appartiennent soit au secteur 1 et perçoivent des honoraires fixés par la tutelle, soit au secteur 2 et s'affranchissent de cette dernière pour la détermination du prix des actes. Cette liberté tarifaire est amendée depuis 1990 par une nouvelle convention qui gèle l'accès au secteur 2 (sauf pour les anciens chefs de clinique) puisque 20% de médecins en moyenne optaient chaque année pour le secteur 2.

Nous étudions le secteur 2, dans le chapitre 2 et 3, car il convient de s'interroger sur la relation positive entre tarification et qualité¹⁶. Dans le cadre d'un bien d'expérience, la relation est ambiguë. Il semble donc qu'avec un bien de confiance, cette ambiguïté soit renforcée. Nous limitons notre analyse aux médecins généralistes. Quant aux spécialistes qui effectuent de la recherche, ils sont peut-être plus à même de proposer des soins de meilleure qualité. Quant aux médecins généralistes, la littérature montre qu'un généraliste sur deux, en secteur 2, propose une spécialité non reconnue par l'ordre des médecins (la médecine du sport, par exemple). Celle-ci est-elle d'une utilité dans le traitement de maladies bénignes ?

Si les médecins agissent de façon opportuniste dans le secteur 2, c'est-à-dire s'ils proposent une qualité identique à celle du secteur 1 à un prix plus élevé, il convient alors de s'interroger sur la justification de l'existence de ce secteur et

¹⁵Le bien santé est un bien spécifique car c'est un bien de confiance. Le patient ne sait pas si la guérison est la résultante du traitement ; il le croit.

¹⁶Morniche (1995) a montré que le secteur 2 est également générateur d'inégalités sociales. "L'insuffisance d'information sur les prix effectivement pratiqués ne permet pas de faire jouer un véritable mécanisme de marché" (Mougeot, 1999). Ces options tarifaires peuvent conduire à des inégalités sociales d'accès et à l'exclusion de certaines catégories de soins les plus performants (de plus grande qualité) si les patients aux revenus les plus faibles n'ont accès qu'au secteur 1.

de la réforme à lui appliquer pour qu'une augmentation du prix s'accompagne d'une augmentation de la qualité.

1.3.2 L'assurance santé

L'assurance joue un rôle principal en économie de la santé. Aux Etats-Unis, plus de 80% des dépenses sont supportées par les assureurs. Ces derniers, lors de la détermination des polices d'assurance doivent prendre en compte les problèmes d'agence rencontrés dans les autres secteurs de l'assurance.

Toutefois, l'assurance santé se différencie, par exemple, de l'assurance automobile en ce que la survenance d'une maladie affecte aussi bien l'état de santé que la richesse financière de l'individu, puisque celui-ci recourt à des soins curatifs supposés restaurer sa santé. Dans les modèles d'économie de la santé, il est usuel de considérer que lesdits soins restaurent parfaitement l'état de santé. Ainsi seul un risque bénin est donc considéré ce qui ne modifie pas les résultats traditionnels de l'économie de l'assurance. La prise en compte d'un risque malin rend l'analyse plus complexe. Cuttler et Zeckhauser (1999) soulignaient : "The situation is more complex when medical spending fails to restore the person to perfect health and the marginal utility of income is affected by health status... it is not clear how much insurance the person will want"¹⁷. Une seconde difficulté apparaît quand la nature dote les individus d'états de santé différents. L'utilité marginale du revenu se trouve également modifiée.

Les asymétries d'information rencontrées par les assureurs ne sont pas sans conséquence pour la détermination des contrats d'assurance optimaux. En effet, si l'assurance complète était proposée aux agents, les contrats de premier rang pourraient devenir déficitaires car ils n'incitent pas les individus à révéler leur type ou à effectuer un effort préventif¹⁸.

Dans un contexte d'aléa moral standard, l'assurance complète décourage les individus à effectuer un effort préventif ex-ante puisqu'ils sont remboursés intégralement de leurs dépenses curatives (dépenses ex-post en cas de maladie). L'assureur ne peut rendre caduque le contrat car il ne peut prouver la négligence de l'assuré en n'observant pas son niveau d'effort. Les différents modèles d'aléa moral développés dans la littérature de théorie des contrats (effort discret ou continu) considèrent rarement des individus différents. Dans un cadre sanitaire, ceux-ci diffèrent cependant par l'état de santé que la nature leur attribue. L'hétérogénéité des caractéristiques individuelles modifie les choix optimaux des assurés. Un individu en bonne santé est-il plus ou moins incité à effectuer un effort préventif? Pour un risque sanitaire malin, les assurés arbitrent dorénavant entre le gain d'utilité lié à la santé (en espérance, l'utilité liée à la santé est supérieure pour la réalisation d'un effort préventif puisque celui-ci "évite" une détérioration de l'état de santé) et le coût de l'effort préventif.

¹⁷Selden (1998) montre que si les individus ont le même état de santé, un risque malin ne modifie pas les résultats traditionnels de l'assurance.

¹⁸Lors de la caractérisation des contrats d'assurance, nous ne distinguerons pas assureurs privés et publics.

Dans un contexte d'antisélection à la Rothschild et Stiglitz (1976) [noté par la suite RS], les individus les plus risqués (dont la probabilité d'occurrence de la maladie est la plus élevée et dont le coût espéré du dommage est le plus élevé) choisissent le contrat d'assurance complète destiné aux individus les moins risqués (pour la même couverture du dommage, ils s'acquittent d'une prime inférieure). L'introduction d'un état de santé différent selon les types de risque modifie-t-elle les résultats de RS ? L'individu en bonne santé obtient-il une couverture du dommage complète ou partielle¹⁹ ?

Nous supposons que l'état de santé d'un individu détermine sa classe de risque. Un individu en mauvaise santé (faible état de santé) engendre en espérance des dépenses sanitaires curatives supérieures à celles de l'individu en bonne santé (état de santé "plus grand"). Toutefois, rien ne justifie que les dépenses soient identiques pour les individus²⁰. La restriction établie par RS, c'est-à-dire supposer des dommages identiques pour des individus de risque différent, ne semble plus valide. C'est pour cela qu'il convient d'étudier différentes situations d'antisélection pour lesquelles est vérifiée la corrélation négative entre état de santé et espérance des dépenses de santé (l'individu à haut risque est caractérisé par une plus grande espérance des dépenses sanitaires et par un faible capital sanitaire). Si dans un tel cadre, les résultats de RS sont infirmés, il est donc nécessaire de préciser les critères qui permettent de vérifier empiriquement le phénomène d'antisélection. Dans un cadre à la RS, Chiappori et Salanié (1998) établissent une conclusion testable et robuste : dans un menu de contrats, les agents qui ont une probabilité d'occurrence d'une maladie plus élevée choisissent les contrats les plus couvrants vendus à un prix unitaire plus élevé. Si les valeurs des probabilités et des dommages sont hétérogènes, sont-ce les individus à la probabilité la plus élevée qui choisissent le contrat le plus couvrant (résultat RS) ou sont-ce les individus à l'espérance de dommage la plus élevée (donc le plus risqué²¹) ?

1.3.3 Organisation de la thèse

La thèse est organisée de la manière suivante. Les chapitres 2 et 3 traitent du secteur 2 de la médecine ambulatoire. Ce secteur est spécifique en raison de la liberté tarifaire instaurée par la convention de 1980. Les chapitres 4, 5, 6 et 7 sont consacrés à l'étude de l'assurance santé en présence des phénomènes d'aléa moral et d'antisélection. Nous nous appuyons, dans ces quatre chapitres, sur une représentation des préférences des individus par une fonction d'utilité à deux variables : le revenu et l'état de santé. Nous considérons un risque malin dans les chapitres 4 et 5 ; et un risque bénin dans les deux derniers chapitres. Les contributions principales de cette thèse sont détaillées dans les chapitres 3, 5 et 7. Le contenu des différents chapitres ainsi que nos résultats sont exposés ci-dessous.

¹⁹ Quand l'équilibre existe.

²⁰ Par exemple, un individu âgé (faible état de santé) possède la même probabilité de se fracturer la jambe qu'un jeune enfant (meilleur état de santé) (Couffinal, 1999).

²¹ Pour RS, le risque de l'individu était la probabilité d'occurrence.

Le chapitre 2 présente une analyse du secteur 2 de la médecine ambulatoire. Ces travaux proviennent en partie de Batifoulier, Bien et Biencourt (1999), et de Bien, Batifoulier et Biencourt (1998). Nous nous interrogeons sur le contenu informationnel des honoraires. Nous mettons en évidence, à l'aide d'une revue de la littérature non exhaustive, qu'il n'existe pas de signal parfait de qualité pour ce bien de confiance. La caractéristique même du service médical encourage les pratiques opportunistes de la part de certains médecins puisque le patient est incapable d'appréhender parfaitement la qualité de la prestation médicale. Nous posons que la probabilité de découvrir que le médecin a fourni de la mauvaise qualité est une fonction décroissante du nombre de fois que la haute qualité a été délivrée. Plus cette dernière est offerte au patient, moins celui-ci est vigilant. La confiance se traduit, ici, par un contrôle moins minutieux de la qualité. Nous montrons, dans ce cadre d'achats répétés, que le médecin opportuniste se construit une réputation en fournissant la haute qualité. Une fois sa réputation établie, il fournit la mauvaise qualité. Nous prouvons également que les individus les plus "favorisés socialement" se voient délivrer plus longtemps des services médicaux de haute qualité.

Le chapitre 3 s'intéresse à la réforme tarifaire que la tutelle doit appliquer au secteur 2 pour que les pratiques opportunistes des médecins soient annihilées. Nous exposons tout d'abord les différentes rémunérations envisagées dans la littérature pour améliorer l'efficacité des services médicaux. Ces travaux considèrent soit un bien de confiance, soit le cas d'asymétries d'information (inobservabilité de l'effort, de la productivité des médecins). Nous développons un modèle qui considère un bien de confiance produit par un médecin dont le type de productivité est inobservable par la tutelle (asymétrie d'information). Nous supposons que la tutelle ne peut pas saisir la qualité d'un service pour toutes les valeurs. Au dessus d'une valeur seuil, elle ne peut déterminer la qualité de la prestation médicale. Si tous les médecins sont opportunistes, nous montrons qu'en fonction de cette valeur seuil, les contrats sont mélangeants ou séparateurs. En effet, si la tutelle ne peut appréhender la qualité fournie par les médecins les moins productifs, elle propose un contrat mélangeant et demande la fourniture d'une qualité égale à celle de la valeur seuil²². Si cette dernière appartient à l'intervalle défini par les qualités de premier rang (déterminées par l'égalisation du bénéfice marginal au coût marginal), le contrat est séparateur et de troisième rang. La qualité demandée aux médecins les plus productifs est la "qualité seuil". Cette modélisation paraît restrictive car tous les médecins ne peuvent être considérés comme opportunistes. La plupart des médecins sont honnêtes, leur honnêteté étant fondée par le serment d'Hippocrate. Nous intégrons donc une seconde asymétrie d'information : la moralité du médecin. Ainsi, quatre types de médecin sont dénombrées. Nous mettons en évidence que les caractéristiques des contrats dépendent du nombre de médecins honnêtes et de la valeur de l'information à propos de la qualité observée par la tutelle. Si cette dernière ne peut appréhender la qualité fournie par les médecins, le contrat sera mélangeant si le nombre de médecins honnêtes est faible et semi-séparateur s'il

²²Tous les médecins fournissent la même qualité quel que soit leur type de productivité.

est élevé. Dans ce dernier cas, le médecin opportuniste et peu productif ne révèle jamais sa productivité.

Le chapitre 4 propose une analyse d'un risque sanitaire²³. Nous proposons tout d'abord une définition de la prime de risque d'un aléa sanitaire assurable qui vérifie l'équivalence de Pratt (1964) dans un cadre bivarié. Ensuite, nous généralisons la notion de prudence de Kimball (1990) à un risque sanitaire assurable et non assurable. Enfin, nous montrons, sous certaines conditions - la prime de risque d'un aléa sanitaire est inférieure à la prime de prudence par rapport à la santé -, que la prime de risque de cet aléa sanitaire, qui peut être malin ou bénin, est une fonction décroissante de l'état de santé.

Le chapitre 5 caractérise les contrats d'assurance exclusifs en présence d'un risque sanitaire malin, d'aléa moral et d'antisélection (états de santé différents). Dans un cadre discret, nous supposons que les individus peuvent modifier la probabilité d'occurrence du risque sanitaire malin en effectuant un effort préventif. La survenance de la maladie affecte aussi bien le niveau de santé que la richesse de l'individu. Seul le risque financier induit par la maladie est assurable. Nous montrons que l'individu réalise un effort d'autoprotection quand son coût est faible et avec un contrat de pleine assurance. En effet, pour un risque sanitaire malin, les assurés arbitrent entre le gain d'utilité lié à la non réalisation de l'effort et à la diminution de l'espérance d'utilité provoquée par la détérioration de l'état de santé. En fonction du coût de l'effort préventif, les individus en mauvaise santé sont plus enclins à effectuer un effort de prévention que les individus en bonne santé. Dans certains cas, nous établissons que les individus en bonne santé obtiennent une couverture complète contrairement aux patients en mauvaise santé. Enfin, nous mettons en évidence que des individus aux caractéristiques sanitaires différentes obtiennent le même contrat.

Dans le chapitre 6, nous présentons un état des lieux de la sélection adverse en assurance. Nous présentons les résultats obtenus pour une double asymétrie d'information (probabilité d'occurrence du dommage et aversion au risque différentes) : contrats à profits positifs, contrats de type RS et contrat mélangeants. Nous obtenons, pour une double asymétrie d'information (l'assuré de type bas risque est très averse au risque), que le contrat d'équilibre est mélangeant avec une couverture complète. Dans ce cas il n'existe aucun contrat profitable réalisable (absence de surassurance et de sous-assurance) attirant l'individu de type bas risque. Par ailleurs, nous démontrons la coexistence de deux types d'entreprises d'assurance (mutuelles et compagnies d'assurance), condition nécessaire à l'existence d'un équilibre séparateur²⁴ avec profits positifs. Dans les modèles d'économie d'assurance, chaque compagnie offre un unique contrat. Si le contrat des individus de type bas risque génère des profits positifs, toutes ces compagnies

²³Un risque sanitaire est un risque qui détériore l'état de santé. Deux voies sont explorées. Tout d'abord, nous supposons que le risque est assurable en tant que tel, c'est-à-dire qu'en contrepartie d'une prime financière l'individu peut obtenir dans chaque état de la nature l'état de santé moyen. Ensuite, nous considérons un risque sanitaire non assurable. Il n'existe aucune technologie qui permette d'obtenir dans chaque état de la nature un état de santé moyen. L'individu s'assure simplement contre les dépenses de santé occasionnées par ce risque.

²⁴S'il existe.

n'offrent que ce contrat. Si aucun contrat actuariel n'est offert aux individus de type haut risque, l'équilibre n'est plus tenable car ceux-ci choisissent le contrat propre aux bas risque. Ce contrat devient donc déficitaire. Il est donc nécessaire de supposer²⁵ que les entreprises d'assurance sont de deux types juridiques : des compagnies d'assurance qui maximisent leurs profits et des mutuelles à but non lucratif. Ces dernières assurent les individus de type haut risque et permettent l'émergence d'un équilibre du marché d'assurance.

Dans le chapitre 7, nous introduisons une dimension sanitaire au modèle de sélection adverse de RS. Nous posons deux hypothèses. Premièrement, les individus sont dotés d'un capital santé inobservable par les compagnies d'assurance. Deuxièmement, un individu en bonne santé engendre en espérance des dépenses de santé inférieures à celles de l'individu en mauvaise santé. Deux extensions du modèle de RS sont proposées. Tout d'abord, nous supposons que la perception du risque dépend de l'état de santé. Un individu en mauvaise santé cherche à réduire la différence de richesse entre les différents états du monde : il est donc infiniment averse au risque financier. L'individu en bonne santé, quant à lui, maximise son espérance d'utilité. Nous mettons en évidence l'existence de contrats avec subventions croisées à l'équilibre. L'individu en mauvaise santé et de type bas risque (faible probabilité d'occurrence d'une maladie bénigne) subventionne les individus de type haut risque. Ensuite, nous postulons que la classification des risques s'effectue à l'aide de l'espérance mathématique du dommage. Nous levons l'hypothèse de dommage unique posée par RS. Dans ce cadre, nous montrons que les résultats de RS ne sont pas toujours corroborés : existence de contrats à profits positifs, couvertures complètes et inversion des résultats RS. En information imparfaite, la contrainte d'incitation de l'individu dont la probabilité d'occurrence de la maladie est la plus élevée est toujours saturée qu'il soit de type bas risque ou haut risque. Le contrat de pleine assurance est donc offert à cet individu. Pour l'autre individu, si le contrat est actuariel, il peut recevoir une couverture partielle. Ainsi l'individu de type bas risque peut contracter une police d'assurance avec une couverture complète contrairement à l'autre individu. Dans cette situation, les résultats de RS sont inversés. A l'équilibre, si le contrat d'assurance, des individus dont la probabilité d'occurrence du dommage est la plus faible, génère des profits positifs, alors la couverture peut être complète (le montant de la couverture dépend de l'état de santé). Ainsi, en information parfaite, les deux individus obtiennent des couvertures complètes aussi bien pour des contrats actuariels²⁶ qu'à profits positifs.

²⁵Pour obtenir un équilibre, Smart (2000) postule un coût fixe.

²⁶Nous généralisons les résultats de Doherty et Jung (1993) pour un cadre sanitaire.

Chapitre 2

Une analyse des honoraires libres

“L’art pratiqué en médecine est un art à part, un art qui ne correspond pas exactement à ce que les Grecs appelaient *techne*, ni à ce que nous appelons, nous, artisanat ou science”, Gadamer (1998).

2.1 Introduction

Le système de santé français est caractérisé par une forte socialisation. L’autorité de tutelle est omniprésente, ou du moins ressentie comme telle par les praticiens. Historiquement, elle était compétente en matière tarifaire. D’autorité, elle fixait le prix de la consultation. Ce dernier était remboursé au patient par les organismes de Sécurité Sociale. En ce sens, on pouvait dire que le coût, pour le consommateur, d’acquisition du bien “santé” était nul ou faible en cas de remboursement partiel. Dès lors, même si l’expression “prix de la consultation” était souvent usitée, il est évident qu’il ne s’agissait pas d’un “prix de marché” au sens de la théorie économique.

En 1980, une réforme a été entreprise introduisant une relative liberté des prix pour la médecine libérale¹. En d’autres termes, à partir de cette date, l’alternative suivante s’offrait à chaque médecin :

- continuer à percevoir le “prix conventionné”, fixé par l’autorité de tutelle et entièrement remboursé au patient (aujourd’hui 115 francs par consultation) ; les médecins sont alors considérés comme appartenant au secteur 1.

¹A l’exception des prothèses dentaires et de certains produits pharmaceutiques, la pratique des honoraires libres est le seul domaine public où il est a priori possible de raisonner en termes de prix de marché.

- choisir de tarifier au-dessus du “prix conventionné”, c’est-à-dire de percevoir un dépassement d’honoraire non remboursé au patient par la Sécurité Sociale ; les praticiens sont considérés comme membres du secteur 2, ou secteur à honoraires libres.

Durant la période 1985 à 1989, nous avons pu observer que 20 % des médecins, en moyenne chaque année, optaient pour le secteur 2. La part des médecins en honoraires libres est évaluée aujourd’hui à 25 %, dont 20% de généralistes et 33% de spécialistes, sur l’ensemble du territoire national.

En 1990, cette (relative) liberté tarifaire est amendée par une nouvelle convention médicale. Les médecins qui sont déjà en secteur 2 peuvent y demeurer mais, mis à part les chefs de clinique, aucun nouveau praticien ne peut désormais opter pour ce système. En d’autres termes, l’attrait du secteur 2 aux yeux des médecins était tel que la puissance publique a ressenti la nécessité² de le réglementer en gelant le “secteur à honoraires libres”. Coexistent donc aujourd’hui encore deux secteurs dont l’un (le secteur 2) institue une relative liberté des prix.

En conséquence, il convient de s’interroger sur le rôle du “prix” dans un domaine d’activité jusque-là administrativement régulé. Plus précisément, comment les médecins, qui bénéficient de la liberté tarifaire, en usent-ils ?

Cette question ne peut être traitée sans faire référence à la subjectivité de la qualité de la pratique médicale. Arrow (1963) soulignait que l’incertitude quant à la qualité du produit est peut-être plus grande ici que pour n’importe quel bien et service. Le rapport médecin-patient est caractérisé par une double asymétrie d’information : la qualité de la prestation médicale n’est souvent pas connue par le patient qui ne connaît pas sa maladie. Même si le titre de “docteur en médecine” assure d’une compétence minimale garantie par le Conseil de l’Ordre, il ne peut être, à lui seul, un signal de qualité.

La parfaite ou l’imparfaite appréhension de la qualité fournit une classification simple des différents biens³ :

- un bien dont la qualité est identifiée avant consommation moyennant un certain coût (par exemple le temps) est un bien de recherche ;
- un bien dont la qualité est identifiée après consommation est un bien d’expérience (Nelson, 1970) ;
- et un bien dont la qualité reste imparfaitement identifiée même après consommation est un bien de confiance (Darby et Karni, 1973).

Le bien santé semble donc s’apparenter à un bien de confiance⁴. En effet, après une consultation, le patient ne peut affirmer quelle est la qualité de la prestation médicale délivrée par le médecin. Nonobstant son diplôme de médecine, ce dernier peut-il révéler la qualité de sa prestation médicale ? Il dispose

²Si tous les médecins du secteur 1 choisissaient le secteur 2, la création de ce secteur aurait conduit à une hausse du prix de la consultation sans nécessairement aboutir à une hausse de la qualité.

³La plupart des biens ne peuvent être classés aussi simplement car leurs caractéristiques appartiennent à différents types de biens.

⁴Nous omettrons dans cette analyse les prestations médicales simples telle que la vaccination, etc... En effet, ces pratiques médicales ne posent pas le problème de la qualité.

de plusieurs signaux possibles pour informer le patient : la publicité, le prix, la durée de la consultation et le traitement ou examens médicaux.

Si la qualité du bien médical est difficilement appréhendable et si le prix, la publicité et la durée de la consultation ne peuvent être des signaux parfaits de qualité, il est bienséant de douter de la légitimité de l'existence du secteur 2 de la médecine ambulatoire. En effet, l'absence de signal parfait de la qualité de la prestation médicale permet de s'interroger sur les conditions nous garantissant qu'une tarification au-dessus du "prix conventionné" va de pair avec une bonne qualité de la prestation médicale. Les déclarations des médecins affirment que l'augmentation du prix de la consultation s'accompagne d'une hausse du niveau de la qualité qui se traduit soit par un allongement de la durée de la consultation, soit par l'offre d'une spécialisation thérapeutique particulière (du type acupuncture ou homéopathie) (Borrel, 1995 ; Ruelland, 1995 et Sandier, 1993).

Le chapitre est organisé comme suit. Dans une première section, nous tentons d'établir l'absence d'un signal de qualité de la prestation médicale. Tout d'abord, à l'aide d'une revue de la littérature qui ne prétend pas à l'exhaustivité, nous établissons qu'il n'existe pas de corrélation positive entre prix et qualité pour un bien d'expérience. Si pour ce bien, le consommateur ne peut à l'aide du prix connaître la qualité, il semble en être de même pour le bien médical. Ensuite, à l'aide de quelques faits stylisés, nous décrivons le secteur 2 de la médecine ambulatoire. De ces faits, nous en déduisons que le prix n'est pas un signal de qualité. Enfin, nous montrons que la durée de consultation ne peut prétendre au statut de signal de qualité.

L'absence de signal de qualité de la prestation médicale conduit à définir, dans une seconde section, la prestation médicale comme un bien de confiance. Dans un premier temps, nous traitons de la spécificité du bien médical qui annihile toute possibilité de contrôle de la qualité par le patient. Dans un deuxième temps, nous exposons deux modélisations propres à un bien de confiance, c'est-à-dire un service de réparation. Nous obtenons comme résultats, dans un cadre statique, que la basse qualité peut être fournie en lieu et place de la haute qualité en présence d'individus opportunistes. Dans un troisième temps, nous établissons à l'aide d'exemples l'opportunisme de certains médecins malgré l'énonciation du serment d'Hippocrate. Enfin, nous montrons, dans un cadre d'achats répétés, que le médecin opportuniste se construit une réputation en fournissant la haute qualité. Puis sa réputation établie, il fournira la mauvaise qualité. Pour ce faire, nous nous inspirons d'une modélisation liant réputation et corruption (Tirole, 1996) dans laquelle le médecin peut donc pratiquer soit un prix p_L (associé à un certain niveau de qualité q_L et à un coût c_L) qui représente le "tarif conventionné", soit un prix p_H , plus élevé (associé à un certain niveau de qualité q_H et à un coût c_H).

2.2 L'absence d'un signal de qualité

Dans le cadre du colloque patient-médecin, nous pouvons considérer deux canaux de transmission de signaux pour le bien médical : le prix et la publi-

cité⁵. Nous ne considérons à ce propos que les résultats ayant trait aux biens d'expérience produit par un monopole puisqu'ils présentent le plus de similitude avec les biens de confiance. La littérature traitant desdits biens est très peu développée.

2.2.1 Une revue de la littérature

Bien que la publicité ait longtemps été perçue comme un gaspillage et une manipulation, elle permet néanmoins la transmission de deux types d'information, soit directe soit indirecte. Cependant, la plupart des écrans publicitaires apportent peu d'informations sur la qualité du produit (Tirole, 1988, Chapter 2). Ainsi, un montant exagéré de dépenses publicitaires serait un signal de la haute qualité du produit (Nelson, 1974). Il est supposé que la haute qualité engendre des achats répétés et que les consommateurs associent à un niveau de dépenses élevées ladite qualité. Par conséquent, la constitution d'une clientèle avec un produit de haute qualité contribue à augmenter le profit par rapport à la basse qualité. Le monopoleur produisant la haute qualité renonce à son bénéfice présent en engageant des dépenses publicitaires élevées pour se signaler et ainsi augmenter son profit futur contrairement au monopoleur fournissant de la basse qualité.

Schmalensee (1978) s'interroge sur la relation positive existante entre dépenses publicitaires et haute qualité. Bien que le gain futur du monopoleur de la haute qualité soit supérieur à celui de la mauvaise qualité, la position relative des gains présents est inversée. En effet, pour un prix donné du produit, la marge bénéficiaire par client est supérieure pour un monopoleur produisant de la mauvaise qualité si on suppose que le coût de production est une fonction croissante de la qualité. Le monopoleur qui fournit une mauvaise qualité est également incité, à la première période, à effectuer des dépenses publicitaires. Par conséquent, les dépenses publicitaires ne peuvent être considérées comme un signal de qualité⁶ puisque les deux types de monopoleurs peuvent les effectuer. La publicité n'est donc pas un signal parfait révélant la qualité. Les règles internes à la profession médicale interdisant le recours à la publicité sont donc légitimées.

Le second signal usité est le prix. Présentons les principaux résultats obtenus pour les biens d'expérience sans garantie⁷. La durée de la relation entre acheteurs et vendeurs permet de distinguer deux cadres d'analyse : celui des achats non répétés et celui des achats répétés⁸.

Dans le premier cas, quand un monopole vend un bien d'expérience une fois pour toutes à des consommateurs, il est incité à réduire la qualité (Tirole, 1988,

⁵Nous présentons uniquement les modèles dont les signaux sont indépendants. Toutefois, en raison de l'interdiction de publicité établie à l'égard des médecins, nous traitons très succinctement les modèles dont le signal de qualité est la publicité.

⁶Pour une analyse plus détaillée, voir Kihlstrom et Riordan (1984).

⁷Nous nous limitons aux biens sans garantie, car la garantie lève l'incertitude sur les biens dont la qualité n'est pas connue avant consommation. Le consommateur se prémunit contre une possible fourniture de mauvaise qualité en achetant un bien garanti.

⁸Cette classification des résultats est propre à Tirole (1988, chapter 2).

chapter 2). Intuitivement, il fournit la mauvaise qualité ce qui génère un accroissement du profit. En effet, en produisant la mauvaise qualité le monopoleur ne réduit pas la demande puisque le consommateur ne connaît pas la qualité. Ce faisant, il économise le différentiel de coût de production entre la haute et la basse qualité.

Un raffinement susceptible de faire émerger un équilibre de haute qualité est de supposer que les consommateurs entrent séquentiellement sur le marché de sorte que pour un point donné du temps, une fraction α est informée de la qualité du produit (Bagwell et Riordan, 1986). Ainsi, seule l'hypothèse d'information parfaite de la qualité du produit par une fraction de la population permet l'émergence de la haute qualité. Le monopoleur vend à un prix p donné soit la mauvaise qualité au coût de production c_0 , soit la haute qualité au coût c_1 . Le bénéfice du monopole est $p - c_1$ quand il produit la haute qualité, soit $(1 - \alpha)(p - c_0)$ quand il produit la basse qualité puisque seuls les agents non informés ont acquis le produit. La haute qualité est produite si $\alpha p \geq c_1 - (1 - \alpha)c_0$. De cette inégalité, nous déduisons deux résultats. Tout d'abord, la haute qualité est fournie si le prix de vente est suffisamment élevé. Ensuite, plus la fraction de la population informée est importante plus cette inégalité est vraisemblablement vérifiée. Ainsi, un prix élevé signale une haute qualité.

Toutefois, la répétition des achats permet en l'absence de garantie le contrôle de la qualité produite par le monopole, indépendamment d'une règle tarifaire. La temporalité de la qualité du bien caractérise deux cas. Premièrement, si la qualité est inaltérable deux règles sont à envisager :

- Un faible prix (prix de lancement) signale une qualité haute. Toutefois, les consommateurs sont réticents à acheter un produit à bas prix bas car ledit prix peut être un signal de mauvaise qualité. Nous supposons que seule la haute qualité permet la constitution d'une clientèle. Comme dans le cadre statique, deux effets s'opposent. D'une part, un produit de haute qualité génère des achats répétés. Le monopoleur renonce à son bénéfice présent pour que le consommateur expérimente son produit (Nelson, 1974). D'autre part, un produit de faible qualité pour un prix donné procure des bénéfices présents élevés et l'absence de bénéfices futurs (Schmalensee, 1978). Le monopoleur de basse qualité n'est pas incité à vendre à faible prix car cette action ne génère qu'une baisse du profit.

- A contrario, si le producteur de la mauvaise qualité conserve sa clientèle, le prix n'est plus usité pour la constitution d'une clientèle. Un prix élevé est donc un signal de qualité (Bagwell et Riordan, 1986). Le monopoleur de basse qualité n'est pas incité à pratiquer un prix élevé sinon il réduirait la demande qui s'adresse à lui ce qui entraînerait une baisse de son profit.

Deuxièmement, le producteur peut modifier sa qualité au cours du temps. Intuitivement, la bonne qualité présente va construire la réputation future du producteur. Deux modélisations sont à retenir.

- Le modèle de primes de qualité est basé sur l'idée que les consommateurs réagissent à la baisse de la qualité en ne réitérant pas leur achat (Klein et Laffer, 1981 ; Shapiro, 1983). L'équilibre de ce jeu est tel que :

- la réputation du monopole dépend de la qualité antérieure ;
- et le monopole produit toujours la haute qualité si le prix est suffisamment

élevé et supérieur au coût de production.

En effet, si le monopole produit la mauvaise qualité à un prix p , il obtient comme bénéfice $p - c_0$ dans la période présente et 0 après. S'il produit la bonne qualité il obtient $\frac{1+r}{r}(p - c_1)$ comme bénéfice actualisé avec r le taux d'intérêt. Ainsi, la haute qualité n'est produite que si $p - c_1 > r(c_1 - c_0)$. Le terme de gauche est la rente de réputation (prime de qualité minimale) que le monopoleur doit recevoir pour produire la haute qualité. Nous dénotons que plus le différentiel de coût de production est faible, plus le monopoleur est incité à produire la haute qualité. Cet équilibre dépend des anticipations des consommateurs. Si ceux-ci croient que la qualité de demain ne dépend pas de celle d'aujourd'hui, alors le monopoleur n'est pas incité à fournir la haute qualité.

- La seconde modélisation intègre l'hypothèse selon laquelle les consommateurs ne sont pas certains que le monopole est incité à fournir la mauvaise qualité. Par exemple, il existe des monopoleurs honnêtes qui répugnent à fournir de la basse qualité alors qu'ils annoncent de la haute qualité (Kreps et Wilson, 1982; Milgrom et Roberts, 1982). Ils considèrent un modèle à deux périodes et postulent que la basse qualité ne génère pas d'achats répétés. Dans ce cadre, le monopoleur honnête fournit toujours la haute qualité tandis que le monopoleur de basse qualité réalise un arbitrage entre les différentes actions possibles. Son action de production dépendra du différentiel de coûts de production. Si $(c_1 - c_0)$ est grand, il révèle son type⁹ puisque fournir la haute qualité à la première période n'est pas profitable. Même en perdant sa clientèle à la seconde période, il maximise son bénéfice avec cette action. Si $(c_1 - c_0)$ est faible, il fournit la haute qualité à la première période ce qui permet la constitution d'une clientèle qu'il floue à la seconde période en leur vendant la mauvaise qualité en raison de la temporalité du modèle.

Ainsi, pour les biens d'expérience il n'existe pas de corrélation entre prix et qualité. Le prix ne peut être un parfait signal de qualité. Si de plus, le consommateur ne connaît qu'imparfaitement la qualité même après consommation (biens de confiance), alors le prix ne peut être un signal de qualité. En effet, si à cette absence de relation entre prix et qualité est jointe une imperfection de perception de la qualité par le consommateur, alors la relation entre prix et qualité ne saurait être rétablie.

2.2.2 Les faits stylisés

La vérification de l'hypothèse d'induction de la demande (Evans, 1974), qui remet en cause l'application de la théorie néo-classique au marché des soins, a engendré de nombreux travaux empiriques alors que la pratique des honoraires libres ne fait jamais l'objet d'études statistiques spécifiques. Les principaux résultats¹⁰ que nous détaillons ci-après résultent d'analyses traitant soit l'activité médicale globale, soit l'effet d'induction.

⁹En fournissant de la mauvaise qualité à la première période, il est caractérisé par les consommateurs comme un producteur mauvaise qualité.

¹⁰Ce paragraphe et les deux suivants proviennent en partie de Batifoulier, Bien et Biencourt (1999), et Batifoulier et Bien (2000).

Un premier groupe de travaux rassemble des analyses statistiques qui rendent compte de la diversité de la pratique médicale. Elles mettent l'accent notamment sur les déterminants des revenus des médecins libéraux. Les auteurs sont, dans ce cadre, amenés à accorder une attention particulière à la corrélation prix de l'acte-nombre d'actes. De telles études, Borrel (1995), Carrere (1988), Lancry (1989), Mizrahi (1994), Ruelland (1995) et Sandier (1993), apportent des résultats très intéressants concernant la variété des motivations des médecins qui passent du secteur 1 au secteur 2.

Le second groupe d'études analyse la nouvelle pratique tarifaire au regard d'un problème fondamental des économistes de la santé : l'effet d'induction. En opposition avec les conclusions de la microéconomie traditionnelle, ce dernier prédit, en effet, une hausse des prix à la suite d'une augmentation de l'offre, celle-ci générant sa propre demande. Le nombre d'articles consacrés à la vérification de l'hypothèse de la demande induite (liaison positive significative entre activité médicale et densité médicale) dans le cas français est relativement faible (par rapport à ce qu'il est outre-Atlantique). Ces études, Béjean (1997), Béjean et Gadreau (1992), Berault et Jacobzone (1994) et Delattre et Dormont (2000), ne permettent pas toujours de conclure sur la véracité de cette hypothèse mais elles apportent indirectement des éléments explicatifs sur le comportement des médecins qui optent pour le secteur 2.

Bien que ces travaux ne choisissent généralement pas le thème des honoraires libres comme sujet principal d'études, ils mettent en exergue deux faits fondamentaux.

- Bien que les honoraires libres soient les mêmes¹¹ à l'intérieur d'une zone géographique donnée (Batifoulier et Biencourt, 1996), il faut noter que ce sont dans les régions où les médecins sont les plus nombreux - par exemple en Ile-de-France ou en Provence-Alpes-Côte d'Azur - que le secteur 2 est le plus développé (Lancry, 1997). Ce n'est donc pas la rareté qui entraîne la hausse des prix.
- Parmi les généralistes en secteur 2, les médecins en "mode d'exercice particulier" (acupuncture, homoéopathie, etc.) sont massivement représentés (les dépassements d'honoraires proviennent pour 40% de ce type d'exercice (Ruelland, 1995)). Ils se considèrent comme des spécialistes et, en conséquence, ils cherchent avant tout à faire reconnaître la spécificité de leurs pratiques médicales qu'ils estiment de qualité supérieure.

En cela, le passage en secteur 2 peut être perçu comme un indice de qualité. La hausse du tarif est la réponse à un choix d'une qualité supposée supérieure de médecine. Ces pratiques se traduisent en moyenne par des durées de consultation plus grandes et des visites au domicile plus rares (ce dernier point est justement une des caractéristiques des généralistes du secteur 2 (Sandier, 1993)).

Or le revenu des médecins est négativement corrélé avec la durée de la consultation puisque ces médecins sont payés à l'acte. Le choix d'une qualité de soins

¹¹Le secteur à honoraires libres n'est bien sûr pas homogène, mais il est postulé que l'hétérogénéité inter-secteur est plus forte que l'hétérogénéité intra-secteur. Ce travail a moins pour but d'expliquer le choix d'un montant d'honoraire libre que la relation honoraire libre-qualité.

différente implique donc potentiellement une baisse de revenu horaire que vient limiter ou contrecarrer le passage en secteur 2. Quoi qu'il en soit, le dépassement d'honoraires ne se traduit pas toujours par un revenu annuel plus élevé. En conséquence, l'installation en secteur 2 ne découlerait pas d'un comportement maximisateur du revenu. Il témoignerait plutôt de la recherche d'un revenu-cible¹² : le médecin se fixerait un niveau de revenu et ajusterait son activité quantitativement comme qualitativement.

Ainsi, avec l'instauration du secteur à honoraires libres, la notion de prix comme signal de qualité peut être pertinente car elle introduit les deux éléments suivants :

- Un élément de segmentation. Jusqu'en 1990, les médecins libéraux ont pu choisir entre les deux secteurs. Il existe dès lors une segmentation tarifaire des médecins de l'ordre (3/4, 1/4).
- Un élément de marché. En se substituant au droit permanent à dépassement qui était déjà attribué à certains médecins, la convention médicale de 1980 permet *a priori* aux médecins d'appliquer le principe déjà mentionné par Chamberlin (1933) selon lequel prix élevé et qualité vont souvent de pair, et de signaler ainsi la qualité particulière de leurs prestations.

2.2.3 Le prix, un signal de qualité

Le patient désire de la qualité en recherchant un médecin. Il exerce un contrôle sur le médecin en interprétant, comme signal de qualité, la proposition de traitement (Dranove, 1988). Le choix du médecin s'explique par la menace crédible qu'exerce le patient : si ce dernier réalise que la qualité offerte n'est pas conforme à celle espérée, il ne renouvellera pas sa confiance à ce médecin et ira ultérieurement consulter un de ses confrères. En effet, la relation médecin-patient s'inscrit dans le long terme et la notion d'achats répétés est déterminante. Le gain immédiat résultant d'un comportement opportuniste du médecin serait donc circonscrit par une perte de clientèle ou par acquisition d'une mauvaise réputation. Le patient développe une contrainte de marché qui pousse le médecin ayant investi dans la qualité à signaler cet investissement par une hausse de tarif et à faire coïncider la qualité espérée à celle effectivement réalisée.

Ce raisonnement qui fait du dépassement d'honoraires un signal de qualité se heurte à quatre insuffisances :

- Tout d'abord, comme il l'est rappelé, la majorité des médecins libéraux et particulièrement les généralistes (80%) choisit de rester en secteur 1. Ceux-là n'utilisent pas le prix (élevé) comme un signal de qualité.
- En ce qui concerne les médecins en secteur 2, le prix reste une inconnue pour le patient et ne sera dévoilé qu'à la fin de l'acte¹³. Il ne peut donc servir de signal de qualité dans le cas d'une première prise de contact.

¹²Evans (1974) suppose que le revenu-cible modifie la fonction d'utilité du médecin. Il cherche à atteindre un niveau d'utilité.

¹³Contrairement à une proposition du Conseil Economique et Social pour lequel "l'information préalable sur les honoraires et les modalités de remboursement constitue un élément de du rapport de confiance avec le médecin, le Conseil souhaite en particulier que le secteur

De plus, même en présence d'une relation durable, le patient n'a jamais l'assurance que le prix sera le même¹⁴. Il est au total difficile de prétendre que le patient va acquérir de l'information sur la qualité à partir du signal, le montant des honoraires, qu'il ne peut ni observer ni connaître *ex ante*.

- Dans certaines configurations, le patient n'a pas le temps de chercher de l'information sur la qualité, soit parce qu'il fait face à une situation d'urgence, soit parce qu'il y a trop de médecins à prospecter. Le patient se contentera alors du premier venu, quel que soit le tarif pratiqué. Il ne s'intéresse alors pas au prix.
- Dans une logique purement marchande, un prix élevé conduit à un revenu plus important. Dans le système de santé, où les médecins libéraux sont payés à l'acte, une hausse du prix de la consultation conduit mécaniquement à une hausse de revenu. Or, les travaux cités précédemment indiquent que le dépassement d'honoraires se traduit moins par un revenu plus élevé que par un revenu-cible. Le signal qualité se déplacerait vers la durée de la consultation.

2.2.4 La durée de consultation, un signal de qualité ?

La qualité s'apprécie en situation, c'est-à-dire au cours du colloque singulier patient-médecin. La vitesse de réalisation de l'acte joue alors un rôle majeur dans l'évaluation de la qualité. En effet, un choix de qualité opérée par le médecin se traduit par des durées de consultations plus longues qui débouchent inévitablement sur une baisse de revenu horaire¹⁵. Le passage au secteur 2 limite alors les effets de la réduction du nombre d'actes et permet de cibler un revenu. Les études statistiques semblent corroborer ce point de vue. Le choix du secteur 2, en compensant une moindre activité voulue, signalerait une augmentation de la qualité.

L'observation de durées de consultations plus longues est particulièrement caractéristique de la stratégie des jeunes médecins¹⁶. Cela peut s'apparenter

d'appartenance conventionnelle du médecin figure sur sa plaque extérieure ; cette information soit rappelée dans la salle d'attente et ses conséquences sur le niveau de remboursement et de la consultation formulées de manière explicite ; le tarif des consultations et des suppléments éventuels (visite à domicile, le dimanche, etc.) soient indiqués" (Journal Officiel, 1996). Cette suggestion a soulevé de vives critiques de la part des organisations professionnelles des médecins.

¹⁴Un médecin du secteur 2 peut faire varier le tarif suivant le motif de la consultation et, pour un même patient, ne prendre que le tarif conventionné pour quelque chose qu'il juge bénin alors qu'il percevra un dépassement important pour des causes plus graves. Or, le patient n'est pas toujours en mesure d'évaluer lui-même la gravité des symptômes qui le conduisent à consulter.

¹⁵D'autres éléments interviennent dans l'évaluation de la qualité de la prestation médicale, notamment l'accueil au cabinet, mais nous supposons que la durée de la consultation est un facteur déterminant dans l'opinion du malade sur le médecin (parce qu'il atteste, notamment, de sa capacité d'écoute, etc).

¹⁶Si on considère que la durée de consultation apparaît, pour certains patients, comme un indicateur de qualité, alors le jeune médecin, qui cherche à se constituer une clientèle, aura tendance à pratiquer des consultations plus longues.

à la pratique du prix de lancement en économie industrielle où il s'agit de se forger une réputation. Bien sûr, il ne s'agit pas à proprement parler d'un prix de lancement puisqu'ici le jeune médecin ne propose pas un prix inférieur, mais la logique est cependant la même. Une qualité de lancement signifie ici une durée de consultation plus importante et une perte de revenus. Cela ne peut se faire que si ses pertes présentes seront contrebalancées par des gains futurs, c'est-à-dire si la qualité offerte au patient assure au producteur d'être encore présent sur le marché à la période suivante. Une fois sa réputation établie, le médecin pourra se permettre de réduire sa durée de consultation (Berault et Jacobzone, 1994).

Toutefois, si la durée de consultation est considérée comme un indicateur de qualité, elle ne peut prétendre au statut de signal en raison de quatre insuffisances :

- Comme l'a déjà été signalé, la grande majorité des médecins libéraux choisissent de rester en secteur 1. De plus, parmi les médecins en honoraires libres, tous n'ont pas des durées de consultations plus longues. L'association prix élevé-durée de consultation plus longue n'est donc pas pertinente pour signaler la qualité de la prestation médicale.
- La durée de la consultation, à l'instar du prix, n'est pas une variable observable *ex ante* par le patient.
- Il existe plusieurs types de biens médicaux dont certains sont sans réputation (Rochaix, 1987). Sur ce type de prestation (vaccination, etc.), la durée de consultation peut difficilement s'allonger. A la limite, une longue consultation peut être qualitativement mal jugée par le patient.
- Dans le cas de biens avec réputation, le médecin a une grande liberté de choix quant à la durée de sa prestation médicale. Quelle sera la durée qui signalera le mieux au patient la qualité de la prestation ?

La liberté tarifaire peut être utilisée par les médecins pour signaler une hausse de la qualité de la prestation médicale. Cependant, ni le tarif, ni la durée de la consultation ne peuvent appréhender la qualité et être ainsi des signaux de la qualité. Dans le cadre d'achats répétés qui caractérise la relation médecin-patient, l'absence d'un signal tangible de qualité affaiblit la contrainte de marché que le patient fait subir au médecin pour un bien d'expérience. Pour un bien de confiance, les comportements opportunistes des médecins s'en trouvent renforcés. Avant de caractériser le comportement opportuniste des médecins, nous devons tout d'abord prouver que le service médical peut être considéré comme un bien de confiance.

2.3 La prestation médicale : un bien de confiance

Le médecin généraliste occupe une place essentielle dans la cité moderne. Il est le médecin de famille : il soigne, il écoute et tient la main de ceux qui seront bientôt délivrés de la vie. Cependant, le matériel qu'il utilise est des moins sophistiqués : un stylo, un ordonnancier et un stéthoscope. Cette description,

comme le souligne Guérin (1995), met en exergue l'activité prescriptrice du médecin¹⁷, stade final de la consultation.

2.3.1 Un bien spécifique

A l'origine de cette prescription se trouve le diagnostic. Il est si rapide¹⁸ qu'il s'apparente à "une sorte de réflexe qui prend en compte la somme de connaissances engrangées depuis la Faculté et l'habitude quasi instinctive de la pathologie fondée sur des notions de haute probabilité" (Guérin, 1995).

Le diagnostic est en effet fondé sur la triade signes-diagnostic-traitement. Pourtant en médecine générale cette dialectique est erronée. Le diagnostic juste s'avère quasiment impossible¹⁹. Braun (1979), lors d'un relevé épidémiologique en 1978, établit que sur 8000 consultations, seules 11% de celles-ci aboutissent à un diagnostic précis. La médecine ambulatoire abuse le traitement rationnel, l'anticipation et l'évitement des complications possibles. Malgré l'imprécision du diagnostic, l'impuissance médicale face à certaines maladies et la moindre docilité du patient face au pouvoir médical, le généraliste reste le confident du malade.

Il possède, en outre, un pouvoir guérisseur démontré par Balint (1966). Cette efficacité thérapeutique dépend de la relation médecin-malade. Cette dernière doit s'inscrire dans un climat de confiance. Malgré l'imprécision du diagnostic et l'imparfaite information sanitaire du patient, ce dernier doit accorder sa confiance au médecin sous peine que le traitement ne se révèle partiellement inefficace. En effet, même dans la prise de placebo, des études ont montré que les patients développaient les mêmes effets secondaires que lors de la prise des principes actifs (Shapiro, 1960)²⁰. La confiance qu'accorde le patient au médecin est primordiale pour l'efficacité thérapeutique du traitement. Pourtant, elle est de plus en plus difficile à établir puisque l'attitude critique et la non-coopération se sont développées au cours du temps. Des médecins français ayant déjà pris leur retraite ont déclaré que seuls 2% des patients possédaient des informations médicales au début de leur carrière contre 93% à la fin de celle-ci. Par ailleurs, la relation devient plus conflictuelle (moins propice à l'instauration de la confiance) puisque les exigences des patients augmentent (de 8.8% à 80.3%) (Herzlich et al., 1993) et que le recours en justice contre les médecins s'accroît. La révolution scientifique qu'a connue la médecine et sa médiatisation ont suscité bien des espoirs ainsi que des déceptions. Cette médecine moderne est tout aussi capable de réussites spectaculaires ainsi que d'échecs. En même temps que se transformait la médecine, la perception sociétale de celle-ci évoluait. Les revues médicales sont nombreuses et bien documentées. Elles accordent une place importante

¹⁷Molière faisait déclarer à Sganarelle dans le *Médecin malgré lui* : "Il ne faut pas qu'elle meure sans l'ordonnance du médecin".

¹⁸Hippocrate, dans l'Art Médical, déclarait que le diagnostic devait être rapide et la maladie lente.

¹⁹Selon Guérin (1995), cette imprécision résulte d'une attitude intentionnelle des médecins. Prescrire un antibiotique aux pouvoirs thérapeutiques élevés est moins coûteux que déterminer, à l'aide d'analyses approfondies, l'agent causal.

²⁰Effet nocebo.

aux exploits scientifiques, médicaux créant ainsi l'image d'une médecine toute puissante aux pouvoirs guérisseurs infinis. Il n'est donc pas étonnant que l'attitude du patient soit exigeante et ses espoirs déçus (Nicolas, 1996). L'absence de satisfaction l'amène à ester en justice le patient. Pour évaluer le nombre de recours en justice, la meilleure source est celle qui provient des assurances médicales (Nicolas, 1996). Les données fournies par Jacques Pouletty (président du *Sou Médical*) montrent que l'augmentation du nombre de mise en cause des médecins, en France, qui a progressé de 62,5% entre 1980 et 1990 n'est guère différente de la variation du nombre de médecins +68%. Cependant, une étude de la Sofrès en 1994 laisse augurer une évolution différente. 71% de l'échantillon interrogé, en cas d'erreur médicale, ferait un procès au médecin. La réponse pour la tranche d'âge 15-24 ans est de 89%. Le médecin ne se trouve plus face à la justice dans une position privilégiée²¹.

Pour décrire l'interaction médecin-patient, au modèle paternaliste (le médecin prend la décision de traitement seul) se substitue un modèle de décision commune (Beisecker et Besisecker, 1990; Siminoff et Fetting, 1991; Brock et Wartman, 1990)²². Cette modification de la représentation du colloque médecin-patient permet de mieux informer les patients.

Même si les patients sont mieux informés²³, il n'en reste pas moins que leur savoir est un savoir profane (Freidson, 1984). Le médecin perçoit la maladie et prescrit un traitement à l'aide d'un savoir technique tandis que le patient perçoit sa maladie et la consultation médicale selon les exigences de sa vie quotidienne et l'environnement culturel qui lui est propre (Adam et Herzlich, 1996). A l'aune de son savoir profane, le patient juge la qualité de la prestation médicale. Cette difficile appréhension de la qualité de la prestation médicale se trouve renforcée par la variabilité des pratiques médicales. Le traitement que prescrit le médecin ne sera jamais la copie conforme de celui qu'il a proposé au patient précédent²⁴ (Guérin, 1995).

Le diagnostic imprécis du généraliste, la variabilité des pratiques médicales et le savoir profane rendent inopérant un jugement objectif de la qualité médicale. Le jugement de la qualité semble plus découler de critères subjectifs (les attentes

²¹ "En cent cinquante ans, on est passé du niveau de l'irresponsabilité du médecin et de l'irrecevabilité des plaintes à celui de la condamnation pénale et de la prise en charge de l'accident non fautif dans le cadre de la responsabilité administrative" Nicolas (1995).

En 1835, le procureur Dupin affirmait que "ne pouvaient donner lieu à responsabilité les faits reprochés aux médecins, sortant de la classe de ceux qui, par leur nature, sont exclusivement réservés aux doutes et aux discussions de la science".

En 1993, l'arrêt Bianchi du Conseil d'Etat admet le principe de l'indemnisation en l'absence de toute faute, sur le simple fait qu'un risque existait.

²² Cette affirmation n'est plus vérifiée pour les pathologies graves. Pour un cancer, Blanchard et al. (1988) obtiennent comme résultats que 92% des patients veulent connaître les différents traitements possibles mais que seuls 69% désirent être associés à la prise de décisions du traitement.

²³ Ce sont les milieux les plus favorisés socialement (en général aux hauts revenus) qui sont les plus aptes à acquérir une compétence médicale. Boltanski (1971) montre que le rapport à un corps (savoir médical) est modelé par la proximité d'un groupe social avec le discours scientifique dominant.

²⁴ Puisque le médecin peut répondre en partie ou en totalité aux attentes de ses clients.

des patients) que de critères normatifs²⁵. Il est à noter que les pays où l'indice de satisfaction est le plus élevé sont ceux où le médecin leur consacre le plus d'attention. Le sentiment d'être bien soigné dépend de la qualité conviviale de l'écoute²⁶. (Pour la France, le temps de consultation est de 14 minutes, et pour la Grande-Bretagne de 8.2 minutes). Il semble que la qualité des soins ne soit pas différente entre ces pays (Guérin, 1995).

Nous pouvons en conclure que la prestation médicale est un bien de confiance puisque le patient est dans l'impossibilité de déceler la qualité réelle de la consultation et que douter de la qualité de la prestation médicale génère une diminution de l'efficacité thérapeutique du traitement. Ainsi, le patient doit faire confiance au médecin par absence d'information et par souci d'efficacité thérapeutique. Avant de montrer que le médecin peut abuser de la confiance des patients, nous présentons quelques modélisations d'un bien de confiance.

2.3.2 Les différentes modélisations

Très peu d'articles traitent du bien de confiance. L'article fondateur est celui de Darby et Karni (1973) qui a introduit la notion de bien de confiance. Ils ont décrit l'effet de la réputation sur la fraude. Nous pouvons dans la littérature restreinte existante distinguer deux types de modélisation. Une première voie de recherche caractérise un bien alimentaire²⁷ dont les caractéristiques sont inconnues des consommateurs. Cette modélisation est inadaptée au cadre sanitaire.

La seconde voie de recherche traite d'un bien de confiance comme d'un bien

²⁵Ces critères resteront difficiles à obtenir. Même si les médecins se réunissent pour tenter d'harmoniser leurs pratiques, ces réunions ne conduisent pas à l'adoption d'une pratique unique. Elles servent à éliminer les traitements les moins efficaces. Les médecins préservent leur liberté d'action ce qui leur permet de se différencier entre eux.

²⁶L'écoute accordée aux patients dépend des modes de rémunération. Le paiement à l'acte induit une plus grande écoute que la capitation.

²⁷Bureau, Marette et Shivinia (1998) considèrent un modèle statique avec un bien alimentaire produit différemment par deux pays (domestique et étranger). Ce bien alimentaire diffère par l'utilisation d'hormones de croissance qui abaisse le coût de production. Les effets négatifs de l'utilisation des hormones de croissance ne sont pas connus. Toutefois, les consommateurs croient que l'utilisation des hormones de croissance réduit la qualité du bien. La demande des consommateurs dépend de la qualité subjective et relative des deux biens et de leur préférence pour la qualité. Les biens n'étant pas différenciables, un prix unique caractérise l'équilibre du marché.

Dans un cadre autarcique, seuls les individus qui apprécient la haute qualité consomment le bien en question. L'ouverture des frontières permet aux consommateurs qui goûtent peu la qualité de consommer car le prix du bien diminue en raison du coût de production inférieur du pays étranger. Toutefois, l'effet sur le bien être du pays domestique dépendra de la perception de la qualité relative des biens.

Si celle-ci est faible, les deux biens sont considérés comme équivalents. Le prix du marché est abaissé. La qualité du bien diminue peu. Ainsi, les agents qui apprécient la qualité réduisent peu leur demande, tandis que les agents qui goûtent peu la qualité augmentent leur demande en raison de la baisse du prix de vente du produit. Il est évident que le surplus collectif du pays domestique augmente.

Si la différence de la qualité perçue est importante, l'introduction du bien étranger fait diminuer le prix et la qualité du bien. Cette baisse de la qualité modifie la demande qui se déplace vers l'origine. Le surplus collectif diminue car la baisse du surplus engendrée par la baisse de la qualité est supérieure à la hausse du surplus liée à la baisse du prix.

se détériorant et dont le diagnostic ne peut être effectué que par une tierce personne (un expert). Un consommateur dont le bien ne fonctionne plus se rend chez un réparateur-expert. Ce dernier établit un diagnostic qui détermine le montant de la réparation. Deux types de dysfonctionnement sont possibles (faible ou grave panne), ce qui nécessite deux types de réparation (dispendieuse ou à bas prix). Si la panne est grave, une réparation onéreuse est nécessaire pour la remise en état du bien. Par contre si la panne est faible, les deux réparations remettent le bien en service.

La première modélisation de bien de confiance considère un cadre stratégique sans coopération. L'expert propose deux réparations et le consommateur accepte ou rejette le diagnostic de l'expert. Pitchik et Schotter (1987) obtiennent un équilibre de stratégies mixtes. Les consommateurs et l'expert décident aléatoirement de leur action.

Deux autres types de modélisations sont à retenir. Les auteurs postulent soit un coût de recherche pour le consommateur, soit une contrainte d'activité pour l'expert-réparateur. Ces modélisations sont bien adaptées pour saisir l'activité médicale. Les auteurs considèrent deux individus : un agent qui expertise et répare, ce qui est propre au généraliste, et un consommateur dont le bien ne fonctionne plus. Il se rend chez l'agent-expertiseur. Il diagnostique (expertise) et prescrit un traitement (réparation).

Coût de recherche

Wolinski (1993) montre que selon les valeurs du coût de recherche, à l'équilibre, les consommateurs requièrent plus d'un diagnostic, ce qui inhibe le comportement frauduleux des experts. Ces derniers proposent deux types de réparation : une dispendieuse et une à bas prix. Si le coût de recherche d'une seconde expertise est faible, les consommateurs, quand on leur recommande une réparation onéreuse, consultent un second expert ce qui désincite les experts-réparateurs à frauder. Ainsi, à l'équilibre, les experts-réparateurs se spécialisent dans une seule réparation. Les experts-réparateurs qui effectuent la réparation peu coûteuse expertisent tous les biens tandis que les autres réparateurs n'effectuent plus d'expertise. En effet, ce sont les seuls incités à frauder et à diagnostiquer une réparation très coûteuse²⁸. Si le coût de recherche d'un second avis est élevé, les experts-réparateurs sont incités à frauder. En effet, diagnostiquer une réparation dispendieuse est toujours bénéfique puisque les consommateurs ne peuvent consulter un autre expert. L'équilibre de ce jeu, est donc un équilibre de fraude où tous les experts-réparateurs diagnostiquent la réparation coûteuse et l'effectuent.

Ce modèle peut justifier la réforme sanitaire française dont la mise en oeuvre est difficile. Un médecin référent (un médecin généraliste) soigne les maladies bénignes ou usuelles et oriente chez un spécialiste pour les autres pathologies. Cependant, il serait plus approprié de considérer qu'il existe trois types de productions sanitaires ; le médecin généraliste ne pouvant effectuer que la moins

²⁸Car il est supposé qu'un bien qui subit une réparation peu coûteuse au lieu d'une réparation onéreuse ne peut être remis en état.

technique²⁹. Le patient devrait donc faire confiance au spécialiste puisque, pour des pathologies complexes ou peu usuelles, son information médicale semble être nulle. Une telle réforme semble restreindre les pratiques opportunistes du généraliste contrairement à celles du spécialiste.

Contrainte d'activité

Emons (1997) et Emons (2000) considèrent un bien durable qui, au bout d'une certaine usure, nécessite une révision et éventuellement une réparation. Il est supposé que le consommateur ne peut établir le diagnostic. Si le produit est en bon état, il nécessite une réparation avec une probabilité inférieure à celle du produit en mauvais état. Le consommateur consulte un expert qui est contraint par le nombre d'heures de travail. Il s'ensuit que le comportement de l'expert dépendra de la taille de sa clientèle. La nature de l'équilibre dépendra de l'observation de cette variable³⁰.

La taille de la clientèle est observable (Emons, 1997) En observant la taille de la clientèle et le prix, le consommateur en déduit le comportement de l'expert ce qui incite ce dernier à être honnête. Ledit comportement dépendra de la taille de la clientèle. Si la taille de la clientèle est importante, l'expert-réparateur ne peut satisfaire toute sa clientèle. Il sera honnête si, "économiquement", il est indifférent entre effectuer les deux tâches (expertise ou réparation)³¹. Le consommateur connaissant la taille de la clientèle et le prix ne se rendra chez l'expert-réparateur que si les prix des deux services sont égaux. Si la taille de la clientèle est faible, l'expert ne sera honnête que si le coût de la réparation est nul³². Le consommateur ne se rendra chez l'expert que si le prix de la réparation est nul. Si la taille de la clientèle est égale à la capacité de l'expert, l'expert ne sera honnête que si le prix de l'expertise est supérieur au prix de la réparation. Si tel n'est pas le cas, l'expert réparera tous les produits. Le consommateur, en observant le prix et la clientèle, en déduira le comportement frauduleux de l'expert et ne se rendra pas chez lui. Ainsi, à l'équilibre, les experts se comportent honnêtement car le consommateur connaît le prix et la taille de la clientèle. En fonction de cette dernière, les prix d'équilibre sont différents.

²⁹Ces types de productions se différencient par leur technicité. Nous supposons que le généraliste peut effectuer les deux moins techniques et le spécialiste les deux plus techniques. Par ailleurs, nous supposons que la production la plus technique est difficilement contrôlable car innovante.

³⁰Les articles de Emons (1997) et (2000) sont complémentaires. Dans l'article de 1997, il suppose que le diagnostic est inobservable ainsi que la taille de la clientèle et que les experts sont en concurrence. En 2000, il considère un expert en monopole. Il étudie quatre situations : diagnostic observable ou non et taille observable ou non. Nous ne présenterons que les résultats pour lesquels le diagnostic est inconnu du consommateur. Cette absence d'information caractérise le cadre médical.

³¹Si le prix de la réparation est supérieur au prix de l'expertise, tous les produits seront réparés.

³²Il reste du temps pour que l'expert effectue des réparations inutiles.

Appliqué à la médecine ambulatoire, ce modèle prédit qu'un médecin avec une grande clientèle sera honnête puisque le prix de la prestation médicale n'est pas différencié selon les actes : prescription d'un examen, analyse, traitement ou vaccination. Cependant, le médecin avec une faible clientèle prescrit des examens ou des services inutiles afin d'augmenter son chiffre d'affaires. Si le patient observe la taille de la clientèle, alors la clientèle des médecins honnêtes croîtra a contrario de celle des médecins malhonnêtes ce qui augmentera leur incitation à surprescrire. Il faut donc limiter la taille de la clientèle des médecins ou le nombre d'actes par médecin. La réforme qui préconisait une limitation du nombre de clients ou du nombre d'actes par médecin revient donc à limiter les comportements frauduleux des médecins à faible clientèle. En augmentant la taille de la clientèle des médecins à faible activité, la tutelle les incite à être moins opportunistes.

La taille de la clientèle est inobservable (Emons, 2000) Si la taille de la clientèle n'est pas observable, le consommateur ne peut aucunement déterminer le comportement de l'expert. L'observation des prix ne le renseigne que si la taille de la clientèle est connue. Anticipant le comportement malhonnête de l'expert, le consommateur ne consulte pas l'expert-réparateur. Si la taille de la clientèle est inobservable, la fraude ne peut être proscrite.

Les modèles qui caractérisent un bien de confiance comme un bien nécessitant une expertise sont plutôt bien adaptés au cadre médical. Nous obtenons que le comportement frauduleux des agents ne peut toujours être proscrit. Cependant, ces comportements sont mis en exergue avec des agents opportunistes. Les médecins agissent-ils avec opportunisme ou honnêteté ?

2.3.3 L'opportunisme des médecins

Selon de Pouvoirville (1997), il est difficile de contester que l'hétérogénéité des pratiques médicales résulte de différences d'efficacité. Domenighetti (1994) montre qu'il y a réellement un pouvoir discrétionnaire du médecin dans la prescription de certains actes et dont le résultat est une sur-utilisation injustifiée. L'observation des taux de recours des familles des médecins et des avocats à un petit nombre d'intervention à fort pouvoir discrétionnaire met en évidence des différences de l'ordre de 10 à 25% avec le reste de la population.

Cet opportuniste résulte d'un conflit d'intérêt en situation d'asymétrie informationnelle. De Kervasdoué (1998) rappelle qu'en France la question de l'agent est absente des débats traitant de l'éthique médicale. Aux Etats-Unis, 23 Etats ont interdit aux médecins d'être propriétaires d'actions dans les cliniques privées en raison d'un conflit d'intérêt potentiel entre leur rôle de médecins et d'actionnaires. Le médecin généraliste propriétaire de son cabinet est soumis à ce conflit d'intérêt qui consiste à soigner les malades efficacement ou augmenter son chiffre d'affaires par des soins inefficaces.

Quelques faits stylisés relevés par Emons (2000) corroborent également les pratiques opportunistes des médecins. Des patients suisses avec un faible niveau d'éducation ont deux fois plus de chance de subir une intervention sur la vésicule

biliaire que des patients diplômés de l'université (le rapport est de 1,5 pour l'opération de l'articulation coxo-fémorale). De même, les enfants des médecins ont 1,25 fois plus de chance de ne pas subir l'ectomie des amygdales que le reste de la population (KTIP, 22/05/1996). Par ailleurs, un tiers des dépenses courantes de santé aux Etats-Unis résulterait d'actes inutiles (The Economist, 13/02/1999). Même en Chine, la sur-prescription est usuelle : les hôpitaux augmentent leurs profits en effectuant des actes inutiles (The Economist, 07/11/1998). En outre, des études empiriques traitant du marché ambulatoire montrent que le paiement à l'acte tend à augmenter la sur-prescription par rapport à la capitation (Gaynor, 1994).

La France n'est pas exempte de ces comportements opportunistes³³. Le revenu des cardiologues varie de 1 à 3 selon les départements, sans que rien ne justifie, ni le nombre de malades, ni le nombre de cardiologues. De plus, pour les prothèses totales de la hanche, la durée de séjour varie de onze à vingt-six jours selon les établissements. En outre, des exemples de fraude avérée sont à noter. En 1994, la Caisse Primaire d'Assurance Maladie récupère 2,547 millions de francs en raison d'une cotation de 48 chirurgiens en Seine et Marne qui exerçaient dans 14 cliniques privées. Egalement en 1994, pour l'ensemble des actes de coelio-chirurgie, 2/3 sont surcotés entraînant ainsi une surcotation des actes anesthésiques et une majoration des frais de salle d'opération.

Si cette variabilité des pratiques était induite par des considérations médicales, le nombre d'actes effectués ne dépendrait pas de la rémunération de ceux-ci. Contandriopoulos (1980) a montré que dans un système de paiement à l'acte, la modification du prix de l'acte provoque un effet très rapide sur sa pratique et qu'il produit un effet de report sur d'autres actes difficiles à prédire et à prévenir.

Ces comportements opportunistes existaient dès l'Antiquité. L'image d'un médecin âpre au gain apparaissait dans la littérature grecque dès avant l'époque d'Hippocrate³⁴.

Les différentes réformes de la Caisse Nationale d'Assurance Maladie corroborent la suspicion de comportements opportunistes des omnipraticiens de la médecine ambulatoire. L'instauration d'une enveloppe globale des soins de ville en 1996 pour limiter la hausse des dépenses non justifiées atteste, de la part de la tutelle, de la prise en compte de comportements opportunistes des médecins. Ce mécanisme présente l'avantage d'une certaine cohérence globale et d'annonces d'objectifs. La fixation du taux d'évolution des dépenses de l'année n est problématique puisqu'il dépend de l'année précédente³⁵. Il serait efficace à court terme si les sanctions étaient réellement appliquées : remboursement d'une partie des honoraires en cas de dépassement de l'enveloppe pour une zone

³³ Les exemples, ci-dessous, proviennent de Johanet (1998).

³⁴ Le lien entre l'appât des honoraires et l'exercice de la profession médicale est souligné ironiquement par Aristophane. A la question "quel médecin y-a-t-il maintenant dans la ville?" un vieillard athénien répondait "Il n'y a ni trace de salaire, ni art médical".

³⁵ Un tel mécanisme permet "la convergence dans le temps vers une valeur moyenne. Il confère à la moyenne une valeur normative... Il est un obstacle à une évolution cohérente de l'offre de soins tant du point de vue de la qualité des soins que de celui des coûts" Mougeot (1999).

donnée. La tutelle a préféré diminuer le montant des honoraires, pour l'année suivante, de la catégorie concernée. Un changement de règle altère la crédibilité d'une politique sanitaire. L'opposition de l'Ordre des médecins aux politiques efficaces de dépenses de santé en 1992, aux sanctions individuelles en 1993, au dossier médical en 1994, au codage des actes en 1995 et aux sanctions collectives en 1996 ne semble pas être justifiée par l'indépendance que les médecins revendiquent. L'absence de tout contrôle demandé par l'Ordre des médecins tend plutôt à laisser perdurer une certaine opacité qui encourage les comportements opportunistes.

2.3.4 Réputation individuelle et bien de confiance

Nous considérons, dans cette économie, deux agents : un médecin et un patient³⁶. Cette considération n'est pas usuelle en économie de la santé. En effet, la relation médecin-patient a longtemps été analysée sous l'angle d'un effet d'induction. Le médecin, en tant qu'expert, non seulement produit les biens médicaux mais formule également la demande du patient. Selon cette analyse, seul le médecin peut être considéré comme un véritable décideur car le patient ne fait qu'appliquer ses décisions. L'interaction médicale est réduite à la décision d'un seul agent.

Même si notre sujet d'étude - les pratiques tarifaires - nous conduit à nous intéresser naturellement au médecin, nous ne considérerons pas, dans la suite du chapitre et contrairement à une certaine tradition en économie de la santé³⁷, que le patient n'est que la "courroie de transmission" des décisions du médecin.

En effet, le patient peut exercer un contrôle "profane" sur le médecin (Freidson, 1960). Il peut mobiliser un certain nombre d'informations (générales ou privées) pour intervenir dans le choix du médecin et la prescription médicale. Cependant, le patient ne peut être considéré comme un décideur à part entière (Rochaix, 1989).

Toutefois, le patient n'est pas non plus appréhendé comme un principal sous l'angle de la théorie des contrats où le médecin est l'agent de ce premier afin de réaliser en son nom une tâche définie a priori, le soigner³⁸. Dans le modèle, développé ci-après, le patient n'est ni passif comme dans l'induction médicale, ni actif comme dans l'application de la théorie des contrats à l'interaction médicale.

La notion de confiance est introduite dans le modèle utilisé ci-après. En effet, un principal (le patient) contracte pour la première fois avec un agent (le médecin). Le bien qui fait l'objet du contrat est un bien santé (une consultation) dont les caractéristiques sont celles d'un bien de confiance en raison de sa subjectivité. Même après la consultation³⁹, le patient n'est pas capable de discerner

³⁶Cette section reprend en partie un travail de Bien, Batifoulier et Biencourt (1998). Je reste cependant responsable de toute erreur de modélisation.

³⁷Pour un état des lieux récents sur le problème de l'induction médicale voir (Rochaix et Jacobzone, 1997).

³⁸Pour une analyse de l'application de la théorie des contrats à l'interaction médecin-patient voir (Reberioux et Bien, à paraître).

³⁹La consultation regroupe le diagnostic et le traitement. Chez le généraliste, il y a rarement dissociation des deux tâches.

parfaitement la qualité de celle-ci⁴⁰. Cette asymétrie d'information permet au médecin de retirer un bénéfice de la non délivrance de la qualité souhaitée par le patient. Si après la consommation de ce bien, il n'a pas découvert que le médecin lui a fourni une qualité moindre, il retourne consulter le même médecin⁴¹. Notre analyse considère des achats répétés.

Le cadre

On considère une économie stationnaire, indicée par $t \in [0, +\infty]$ dans laquelle chaque agent et principal vivent t périodes⁴². Ce modèle se différencie des modèles de matching usuels car le principal et l'agent se rencontrent plusieurs fois⁴³. Présentons le déroulement du jeu. Le patient se rend chez un médecin du secteur 1 ou du secteur 2. Tant que le patient est satisfait par la qualité délivrée par le médecin, ces deux individus se rencontrent de nouveau.

Notre analyse concerne uniquement les médecins en secteur 2, c'est à dire ceux qui disposent de la "liberté" des honoraires. En effet, les médecins du secteur 1 pratiquent, par définition, des tarifs faibles (le tarif conventionnel). Dans le secteur 1, le prix conventionnel est fixé de telle sorte que les médecins délivrent la qualité minimale permettant de soigner. Ainsi, en secteur 2, se pose le problème de la qualité et du prix. Parler de prix élevé n'est cohérent que pour une population de médecins qui ont opté pour le secteur 2 de la médecine libérale. Nous étudions donc les liaisons entre ce prix élevé et la qualité des soins.

L'agent (médecin) Trois types de médecins sont présents dans le secteur non conventionné de l'économie :

- honnêtes en proportion α ;
 - opportunistes en proportion β ;
 - malhonnêtes en proportion γ ;
- avec $\alpha + \beta + \gamma = 1$.

Ces proportions restent inchangées pour chaque cohorte et donc stables dans la population. Nous n'étudierons dans le reste de cette section que le cas des médecins opportunistes car les médecins honnêtes (respectivement malhonnêtes) offrent toujours la haute (basse) qualité.

Les médecins sont répartis entre deux secteurs :

- Le secteur conventionné. A un prix p_L , équivalent au tarif conventionné, il offre toujours le bien de qualité q_L (à un coût c_L). On suppose que dans ce

⁴⁰Cette hypothèse de bien de croyance altère, voire infirme le contrôle profane du patient sur le médecin (Friedson, 1960).

⁴¹En France, le nomadisme médical est faible : environ 5% des patients changent régulièrement de médecins.

⁴²Pour simplifier les résultats, nous supposons que seuls les agents d'une même cohorte s'apparient.

⁴³Dans le modèle de Tirole (1996), dont notre modélisation s'inspire, le principal et l'agent se rencontrent une seule fois. A contrario de notre modélisation qui accentue la réputation individuelle, son modèle était une première modélisation de la réputation collective.

secteur, il n'est jamais possible d'offrir une qualité moindre⁴⁴.

- Le secteur non conventionné ou secteur des honoraires libres. A un prix p_H , supérieur au tarif conventionné, le médecin peut offrir soit un bien de haute qualité q_H à un coût c_H , soit un bien de basse qualité q_L à un coût c_L . Il décidera de son action en comparant les différents profits actualisés qu'il obtiendra.

Nous supposons que le tarif conventionné est fixé de telle sorte que chaque médecin de ce secteur retire un profit nul :

$$p_L - c_L = 0$$

De plus, nous supposons que produire la haute qualité est plus coûteux que fournir la basse qualité. Soit :

$$\begin{aligned} q_H &\geq q_L \\ c_H &\geq c_L \end{aligned}$$

Nous en déduisons que la qualité q_H ne peut être offerte à un prix conventionné, puisque $p_L - c_H \leq 0$.

Le principal (patient) Il connaît la proportion des types de médecins dans l'économie (α, β, γ) . Il observe imparfaitement la qualité du bien santé.

Nous supposons toutefois que, lorsque le patient pense que le médecin lui a délivré la haute qualité, il ne change pas de médecin puisqu'il existe dans l'économie des médecins honnêtes.

Hypothèse : *Tant que le patient ne découvre pas que son médecin lui a délivré de la mauvaise qualité il le croit honnête.*

Toutefois, le patient exerce une certaine surveillance puisqu'il sait que des médecins sont opportunistes⁴⁵. Nous posons que sa probabilité x_t de découvrir que la qualité est basse alors qu'il pensait avoir acheté la haute qualité est définie par :

$$x_t = \lambda^{T+1}$$

avec $0 < \lambda < 1$ ⁴⁶, T le nombre de fois où la haute qualité a été délivrée pendant la relation, t la durée de la relation médecin-patient et $T \leq t$ ⁴⁷.

Cette définition de la probabilité prend en compte deux effets :

⁴⁴On suppose que la qualité q_L est la qualité minimale nécessaire au soin du patient.

⁴⁵Sa technologie de surveillance est exogène. L'endogénéisation de celle-ci est complexe puisqu'il faudrait caractériser une fonction objectif du patient qui dépendrait du prix et de la qualité de la prestation, et du coût de la surveillance. Cependant, quelle qualité doit-être prise en compte dans ce calcul? La qualité effectivement reçue ou la qualité que le patient croit avoir reçue.

⁴⁶L'exposant est égal à $T + 1$ car le patient ne connaît jamais réellement la qualité du bien santé. Si l'exposant était T , alors, à la première période de la relation, le patient saurait avec exactitude quelle est la caractéristique du bien qu'il a acquis ce qui dénaturerait le bien (il perdrait ses caractéristiques de bien de confiance).

⁴⁷Nous supposons t grand et non infini car il existe des médecins honnêtes qui, à la dernière période, offrent la haute qualité.

- Le premier effet. Le patient n'est pas capable, après consommation du service médical, de déceler la qualité du bien. Sa probabilité de découvrir la qualité médicale (notée λ^{48}) est inférieure à 1 mais n'est pas égale à 0 car les revues médicales et les nouvelles technologies ne laissent pas le patient sans informations. Ces revues créent chez le patient l'espérance de nouveaux traitements largement diffusés, bien plus qu'elles ne l'informent (Atkinson, 1998). Cette probabilité exprime le savoir profane du patient.

- Le second effet est un effet dynamique qui résulte de la confiance. Nous supposons que plus le médecin délivre la haute qualité au patient, plus la vigilance de ce dernier diminue. Le patient accorde une confiance plus grande au médecin et le contrôle moins. Plus T est grand, plus x_t diminue. Cette confiance qui s'instaure est vérifiée par l'absence de nomadisme médical. Selon Breuil-Genier (2000), le pourcentage de patients nomades serait compris entre 5% et 9% maximum.

Par exemple, si on suppose que la durée de consultation est un indicateur de qualité, la baisse de cette dernière peut être perçue par le patient comme une meilleure connaissance de son cas médical par le médecin de famille. Ceci se traduit par une baisse concomitante des durées et qualité d'écoute. Le patient n'a pas nécessairement conscience de ce changement d'attitude du médecin. Si tel est le cas, il découvrirait que le médecin lui fait payer un bien de qualité q_L à un prix p_H . La rupture serait alors consommée. Le patient "voterait avec ses pieds" et irait consulter un autre médecin (Rochaix, 1997).

Offrir la haute qualité est une stratégie dominée

Tout d'abord, nous allons définir les conditions de "viabilité du marché de la haute qualité" pour les médecins opportunistes. Nous montrons, dans un premier temps, que les médecins opportunistes offrent toujours la basse qualité si le prix de la haute qualité est trop faible. Nous qualifions ensuite le temps T^* jusqu'auquel il est optimal d'offrir de la haute qualité. Enfin, nous étayons l'affirmation selon laquelle pour un médecin opportuniste, le choix d'offrir la qualité L à un prix élevé est confirmé à chaque période dès lors qu'il a été effectué une fois.

Le prix permettant la fourniture de la haute qualité Soit $B(q_H, p_H)$ le flux de revenus actualisés du médecin qui produit toujours une qualité haute

⁴⁸ Celle-ci ne dépend pas des différents états d'équilibre possibles. $\lambda \neq h_q$ avec h_q la proportion de la fourniture de la haute qualité dans le secteur 2. Celle-ci caractérise les trois états d'équilibre possible :

- $h_q = \alpha$. L'état d'équilibre est celui de faible qualité. Tous les agents opportunistes délivrent la faible qualité à un prix élevé.

- $h_q = \alpha + \gamma$. L'état d'équilibre est celui de haute qualité. Tous les agents opportunistes délivrent la haute qualité à un prix élevé.

- $h_q = \alpha + f(\gamma)$ avec $0 < f(\gamma) < \gamma$. Les agents opportunistes délivrent aussi bien la basse qualité que la haute qualité.

à un prix haut, nous avons :

$$B(q_H, p_H) = \sum_{i=0}^t [(p_H - c_H) \delta^i]$$

avec soit δ le facteur d'actualisation.

$$\begin{aligned} B(q_H, p_H) &= (p_H - c_H) \sum_{i=0}^t [\delta^i] \\ B(q_H, p_H) &= \frac{(p_H - c_H)}{1 - \delta} \end{aligned}$$

De même, $B(q_L, p_H)$ désigne le flux de revenus actualisés du médecin qui produit toujours une qualité basse à un prix haut :

$$B(q_L, p_H) = \sum_{i=0}^t [(p_H - c_L) (1 - \lambda)^i \delta^i]$$

Explicitons cette formule. Si le médecin offre toujours de la basse qualité à un prix élevé, à la première période le patient est obligé de payer p_H pour la prestation fournie. A la deuxième période, le médecin revoit son patient avec une probabilité $(1 - \lambda)$ car il lui a offert de la basse qualité. A la troisième période, la probabilité pour que le médecin revoit son patient est $(1 - \lambda)^2$.

Cette expression est donc égale à :

$$B(q_L, p_H) = (p_H - c_L) \left\{ \frac{1}{1 - (1 - \lambda)\delta} \right\}$$

Nous voulons caractériser sous quelle condition les médecins opportunistes produisent la haute qualité. Il est aisé de montrer qu'une déviation de la qualité par rapport à la fourniture pérenne de la haute qualité est profitable. Si le médecin offre la qualité q_L à la dernière période, son bénéfice augmente de $(c_H - c_L)\delta^t$.

Fournir toujours la haute qualité rapporte au médecin opportuniste le bénéfice minimal pour que le marché de la haute qualité soit actif (ici la haute qualité est fournie au moins une fois). Nous déterminons donc le prix minimal de la haute qualité pour que le marché de la basse qualité ne soit jamais actif à chaque période.

Le prix minimal de la haute qualité vérifie que :

$$B(q_H, p_H) \geq B(q_L, p_H)$$

C'est-à-dire,

$$\frac{(p_H - c_H)}{1 - \delta} \geq \left\{ \frac{(p_H - c_L)}{1 - (1 - \lambda)\delta} \right\}$$

Nous avons donc :

Proposition 1 *Les médecins opportunistes n'offrent jamais à chaque fois la basse qualité si :*

$$p_H \geq p_H^* \geq p_L$$

$$\text{avec } p_H^* = c_H + \left[\frac{1-\delta}{\delta\lambda} \right] (c_H - c_L) \geq c_H$$

Un prix trop faible dans le secteur 2 est associé à la fourniture d'une faible qualité par les médecins opportunistes. Ce résultat est cohérent avec la théorie du signal (Spence, 1973 et 1974) selon laquelle les médecins vont signaler la qualité de leurs prestations en s'écartant du tarif standard, appliquant ainsi le principe selon lequel prix élevé et qualité vont souvent de pair (Chamberlin, 1933). Le prix est un signal de qualité. Les médecins vont utiliser la nouvelle liberté qui leur est donnée pour signaler la qualité particulière de leur prestation.

Effet de réputation individuelle Afin de montrer qu'il est optimal de toujours offrir une qualité basse, une fois celle-ci déjà offerte, nous allons montrer qu'il est impossible d'offrir, à une période donnée, une qualité haute si le médecin rationnel a offert, à la période antérieure, une qualité basse.

Nous supposons qu'il existe une date $T + 1$ à partir de laquelle le médecin change de qualité (haute à basse). Pour montrer que c'est une stratégie dominante de toujours offrir une qualité basse, en $t' > T + 1$, nous allons raisonner en deux temps.

Tout d'abord, nous considérons un médecin qui offre de la qualité basse à la date $T + 1$. Si ce médecin offre de la basse qualité à cette période, son revenu actualisé noté $B(q_L^{T+1}, p_H)$ doit être supérieur à son revenu actualisé s'il avait toujours offert de la haute qualité, noté $B(p_H)$.

Calculons $B(q_L^{T+1}, p_H)$ ⁴⁹.

$$B(q_L^{T+1}, p_H) = (p_H - c_H) \sum_{i=0}^T [\delta^i] + (p_H - c_L) \delta^{T+1} +$$

$$(p_H - c_H) (1 - \lambda^{T+2}) \sum_{i=T+2}^t [\delta^i]$$

et

$$B(p_H) = (p_H - c_H) \sum_{i=0}^t \delta^i$$

Si le médecin offre une qualité basse à la période $T + 1$, alors ses revenus actualisés sont supérieurs à ceux engendrés par une fourniture permanente de la haute qualité.

Soit $B(q_L^{T+1}, p_H) \geq B(p_H)$

⁴⁹ Comme on a supposé que le patient ne changeait pas de médecin quand celui-ci lui offre de la haute qualité alors à partir de $T + 2$ son bénéfice actualisé est donc égal à la probabilité de ne pas être découvert que multiplie le bénéfice actualisé de n'offrir que de la haute qualité de $T + 2$ à t . La probabilité d'être découvert s'énonce λ^{T+2} car le médecin a offert de la qualité pendant $T + 1$ périodes.

Après simplifications, nous obtenons :

$$(p_H - c_H) + B_{T+2} \leq (p_H - c_L) + B_{T+2}(1 - \lambda^{T+2})$$

$$\text{avec } B_{T+2} = (p_H - c_H) \sum_{i=T+2}^t \delta^{i-(T+1)}$$

Soit,

$$\lambda^{T+2} \leq \frac{c_H - c_L}{B_{T+2}}$$

Comme λ^T et B_T sont des fonctions décroissantes de T , si cette expression est vérifiée pour T , elle l'est également pour $T' > T$.

Ensuite, nous considérons ce même médecin qui offre de la qualité basse à la date $T+1$ et qui préfère offrir à la période suivante de la qualité haute. Si ce médecin n'offre pas de la qualité basse à la date $T+2$ alors son revenu actualisé noté $B(q_L^{T+1}, p_H)$ doit être supérieur à son revenu actualisé quand il offre de la qualité basse à $T+1$ et à la période suivante ($T+2$) noté $B(q_L^{T+1}, q_L^{T+2}, p_H)$

Calculons cette expression :

$$\begin{aligned} B(q_L^{T+1}, q_L^{T+2}, p_H) &= (p_H - c_H) \sum_{i=0}^T [\delta^i] + (p_H - c_L) \delta^{T+1} \\ &\quad + (p_H - c_L) (1 - \lambda^{T+2}) \delta^{T+2} \\ &\quad + (p_H - c_H) (1 - \lambda^{T+2})^2 \sum_{i=T+3}^t [\delta^i] \end{aligned}$$

Pour tout $T' \geq T+2$, la probabilité que le patient découvre que le médecin lui a fourni la mauvaise qualité est λ^{T+2} puisque le médecin n'a délivré que $T+1$ fois la haute qualité.

Si le médecin n'offre pas une qualité basse à la période $T+2$, alors ses revenus actualisés sont supérieurs à ceux engendrés par une fourniture de la mauvaise qualité durant deux périodes.

$$\text{Soit } B(q_L^{T+1}, q_L^{T+2}, p_H) \leq B(q_L^{T+1}, p_H)$$

Après simplification, nous obtenons :

$$(p_H - c_H) + B_{T+3} \geq (p_H - c_L) + B_{T+3}(1 - \lambda^{T+2})$$

$$\text{avec } B_{T+3} = (p_H - c_H) \sum_{i=T+3}^t \delta^{i-(T+1)}$$

$$\lambda^{T+2} \geq \frac{c_H - c_L}{B_{T+3}}$$

Si $\lambda^{T+2} \leq \frac{c_H - c_L}{B_{T+2}}$ alors $\lambda^{T+2} \leq \frac{c_H - c_L}{B_{T+3}}$ car $B_{T+3} \leq B_{T+2}$. Nous en déduisons que si un médecin offre de la qualité basse à un prix élevé à une date

T , alors il offre également cette qualité standard à ce même prix à la période suivante. Par récurrence, nous montrons donc qu'une fois que la qualité standard est offerte à un prix élevé, il est optimal de continuer d'offrir ladite qualité au même prix⁵⁰.

Proposition 2 *Il est optimal pour un médecin opportuniste de continuer à offrir la basse qualité dès qu'il en a délivrée.*

Le basculement d'une qualité haute à une qualité basse est définitif. Calculons T^* , la date de ce basculement, c'est-à-dire la date du changement de qualité.

Un médecin qui offre toujours de la qualité haute a comme revenus actualisés :

$$B(q_H, p_H) = \sum_{i=0}^{T^*-1} [(p_H - c_H) \delta^i] + \sum_{i=T^*}^t [(p_H - c_H) \delta^i]$$

Un médecin qui offre la qualité haute jusqu'à $T^* - 1$ et ultérieurement la qualité basse obtient comme revenus actualisés :

$$B(q_H, q_L, p_H) = \sum_{i=0}^{T^*-1} [(p_H - c_H) \delta^i] + \sum_{i=T^*}^t [(p_H - c_L) (1 - \lambda^{T^*+1})^{t-T^*} \delta^i]$$

Si le médecin offre de la qualité haute puis de la qualité basse, il en retire un revenu actualisé supérieur à la situation où il n'offre que de la haute qualité.

Soit $B(q_H, p_H) \leq B(q_H, q_L, p_H)$.

$$\sum_{i=T^*}^t [(p_H - c_H) \delta^i] \leq \sum_{i=T^*}^t [(p_H - c_L) (1 - \lambda^{T^*+1})^{t-T^*} \delta^i]$$

On peut écrire $\sum_{i=T}^t m^{t-T} = \frac{1}{1-m}$ car t est grand.

Après simplifications, nous obtenons :

$$\lambda^{T+1} \leq \left(\frac{c_H - c_L}{p_H - c_L} \right) \left(\frac{1 - \delta}{\delta} \right)$$

Proposition 3 *Il existe une date T^* , $T^* > 0$, tel qu'il est optimal pour un médecin opportuniste d'offrir*

- la qualité q_H pour tout $i \in [0, T^* - 1]$
- la qualité q_L pour tout $i \in [T^*, t]$.

⁵⁰Bien que les calculs soient différents de ceux de Tirole (1996), on retrouve la même conclusion. Chez Tirole, l'individu, une fois corrompu, est enfermé dans la corruption. Ici, le médecin une fois détourné de la délivrance d'une haute qualité, s'enferme dans la médiocrité.

Preuve :

$$\text{soit } \ln \left[\left(\frac{c_H - c_L}{p_H - c_L} \right) \left(\frac{1 - \delta}{\delta} \right) \right] \geq T + 1 \ln \lambda$$

$\ln \lambda < 0$ car $\lambda < 1$, ce qui implique que

$$T \geq T^* = \frac{\ln \left[\left(\frac{c_H - c_L}{p_H - c_L} \right) \left(\frac{1 - \delta}{\delta} \right) \right]}{\ln \lambda} - 1$$

$$T^* > 0 \text{ ce qui implique que } \left(\frac{c_H - c_L}{p_H - c_L} \right) \left(\frac{1 - \delta}{\delta \lambda} \right) < 1.$$

Cette inégalité est vérifiée si $p_H > p_H^*$ c'est-à-dire qu'offrir la qualité haute, au moins une fois, est un comportement possible pour le médecin opportuniste.

Si $p_H < p_H^*$ le médecin opportuniste offre toujours de la basse qualité. Par conséquent $T^* = 0$. CQFD.

Pour tout $T > T^*$, le médecin doit offrir la qualité basse car il en retire des revenus actualisés supérieurs. Pendant T^* périodes, le médecin construit sa réputation individuelle en offrant une qualité haute. Après, il offre de la qualité basse car le patient est moins vigilant par rapport à la qualité en raison de la délivrance de la haute qualité pendant les T^* premières périodes.

Ce résultat établit que le comportement opportuniste du médecin est possible et que le patient n'est pas en mesure de contraindre le médecin à lui offrir la qualité haute. L'hétérogénéité de la population des médecins ne permet pas à l'individu de connaître la qualité délivrée par son médecin. Les repères standards comme le titre de docteur en médecine ou la spécialité affichée sont des signaux insuffisants. Ils permettent de définir la qualité minimale d'un service médical. Mais ceci est insuffisant pour distinguer une qualité haute d'une qualité basse.

Rechercher de l'information sur la qualité du médecin, c'est chercher à connaître sa réputation (Pauly et Sattertherwaite, 1981). Deux conditions doivent alors être réunies :

- Le patient doit avoir la possibilité de rechercher de l'information. Il doit donc disposer du temps nécessaire, ce qui exclut les demandes en urgence et les patients étrangers au tissu local. En effet, un consommateur local est mieux renseigné qu'un consommateur étranger (de passage par exemple) et la recherche de réputation est plus facile.
- Le bien offert par le médecin doit être un bien à réputation (Pauly 1988). Il est difficile pour le médecin de révéler la haute qualité de son service sur une prestation standard comme les vaccinations, les soins des affections banales (par exemple une angine). C'est au moment où le médecin dispense une prestation spécifique, dédiée au patient, qu'il peut faire émerger une réputation.

Ces réserves étant faites, il est de l'intérêt bien compris du médecin de se construire une réputation individuelle flatteuse en pratiquant une qualité élevée. Ceci lui permettra de fidéliser une clientèle qui le consultera par habitude, indépendamment de la qualité du service offert, et le recommandera aux autres patients. Le médecin a donc besoin d'une haute qualité pour commencer la relation avec le patient. Il peut ensuite revenir à une qualité standard.

Statique comparative

Etudions comment varie T^* lorsque chacune des variables qui composent cette fonction varie. Présentons les dérivés partielles sous forme de tableau :

X	$\frac{\delta T^*}{\delta X}$	signe $\left(\frac{\delta T^*}{\delta X}\right)$
c_H	$\frac{1}{(\ln \lambda)(c_H - c_L)}$	< 0
c_L	$-\frac{(p_H - c_H)}{(\ln \lambda)(p_H - c_L)(c_H - c_L)}$	> 0
p_H	$-\frac{1}{(\ln \lambda)(p_H - c_L)}$	> 0
δ	$-\frac{1}{(\ln \lambda)\delta(1-\delta)}$	> 0
λ	$-\frac{1}{\lambda(\ln \lambda)^2} \ln \left(\frac{(c_H - c_L)(1-\delta)}{(p_H - c_L)\delta} \right)$	> 0

Tableau 2-1 : Statique comparative

Seule l'augmentation du coût de la haute qualité entraîne, toutes choses égales par ailleurs, la diminution de la date T^* , à partir de laquelle le médecin offre de la mauvaise qualité. Ainsi, toute augmentation d'une des quatre autres variables provoque la hausse de la durée de la fourniture de la haute qualité et donc, augmente la qualité moyenne du secteur 2. Or, en France, la convention médicale de 1993 a augmenté les cotisations sociales des médecins en secteur 2⁵¹. Celle-ci a donc engendré une hausse du coût de la haute qualité et, à la vue de nos résultats, on peut en déduire qu'il y a une diminution de l'intervalle pendant lequel va être offerte une prestation de qualité.

Par ailleurs, si l'information que possèdent les patients dépend de leur position sociale (Boltanski, 1971), alors nous pouvons en conclure que les individus les plus favorisés socialement se verront offrir la haute qualité plus longtemps que les autres individus, puisque T^* est une fonction croissante de λ . Cette dernière qui représente la "technologie de surveillance" dépend du savoir médical et non du savoir profane. Pour améliorer ce premier, une politique d'éducation sanitaire est nécessaire et est du ressort de l'Etat ou des organismes sociaux. Les revues médicales, quant à elles, contribuent à l'augmentation du savoir profane, c'est-à-dire à la connaissance des techniques médicales. Les membres des milieux les plus favorisés socialement intériorisent le mieux les catégories conceptuelles du savoir médical. Leur niveau d'éducation plus élevé rend plus aisé la transmission de connaissances du médecin au malade. Le rapport au savoir médical est modelé par la proximité d'un groupe social avec le discours du corps médical. En outre, Friedson (1984) soutient que les configurations de l'interaction médecin-patient dépendent des activités médicales et du statut social du patient. La pédiatrie et la chirurgie reposent sur un modèle de relation activité-passivité. Par ailleurs, cette configuration a également plus de chance de s'imposer en médecine ambulatoire quand le statut social du patient est faible. Nous retrouvons

⁵¹ On suppose que c_H est la somme des coûts totaux auxquels le médecin doit faire face pour son activité médicale. Ainsi, l'augmentation de ses cotisations sociales engendre une hausse de c_H .

donc ici une inégalité sanitaire mise en exergue par les études statistiques. Les individus les plus aisés sont les mieux soignés.

La deuxième réglementation du secteur 2, c'est-à-dire la fermeture de ce secteur, implique à la vue de nos résultats une baisse de la qualité délivrée par les médecins en secteur 2. En effet, en ne laissant entrer que les anciens chefs de clinique, elle diminue le nombre de jeunes médecins présents dans ce secteur. Or ce sont eux, qui délivrent la haute qualité pour fidéliser la clientèle. Ne laisser dans ce secteur, que des vieux médecins (en termes de relation avec le patient) revient à faire diminuer la qualité du secteur 2. Cette réforme se révèle donc inefficace.

2.4 Conclusion

La médecine ambulatoire est caractérisée par la coexistence de deux secteurs dont la fixation des honoraires est différente. Pour le premier, le tarif est déterminé par les autorités tutellaires. Quant au second, les médecins dudit secteur ont libre arbitre pour établir l'honoraire de la consultation. Les honoraires du secteur 2 sont plus élevés qu'en secteur 1. Un tarif plus élevé ne signale par pour autant une qualité plus grande.

La particularité du bien médical est d'être un bien dont la qualité, même après consultation, n'est pas connue avec certitude par le patient. Pour un bien d'expérience, la relation prix-qualité ne peut être déterminée pour des producteurs opportunistes. Qu'en est-il pour un bien de confiance ?

Bien que les médecins prononcent le serment d'Hippocrate, les faits tendent à prouver que certains d'entre eux agissent comme des *homo economicus*. Dans un tel cadre, nous montrons qu'une haute qualité peut être utilisée par les médecins pour amorcer une relation suivie avec le patient. Cette fidélisation étant effectuée, le médecin peut revenir à une qualité standard. Cette affirmation fournit un schéma théorique aux pratiques d'honoraires libres. Les médecins justifient des prix élevés par une haute qualité de leur service, mais continuent à pratiquer ces mêmes tarifs élevés quand ils sont revenus à une qualité standard.

Pour limiter ces pratiques opportunistes, une réglementation favorisant la pérennité d'une haute qualité est nécessaire. Mais celle-ci doit aussi combattre un effet pervers : le caractère éphémère de la hausse de la qualité. Pour concilier ces deux objectifs, la réglementation ne doit pas, selon le mode actuel, instituer la fermeture du secteur 2 car elle entraîne la baisse de la qualité moyenne du secteur 2. Elle doit, au contraire, dans la réglementation actuelle permettre le remplacement d'un sortant par un entrant. C'est en entrant dans le secteur que l'on se signale par une haute qualité. Sans réforme de la rémunération des généralistes du secteur conventionné, il serait souhaitable que le nombre de généralistes en secteur 2 soit constant pour que la qualité moyenne offerte dans ce secteur demeure stable. Ainsi, une sortie du secteur 2 (un départ en retraite, par exemple) doit être compensée par une nouvelle arrivée.

Néanmoins, dans un tel cadre institutionnel, le secteur 2 n'a pas lieu d'exister car il engendre des inégalités sociales : les individus les plus aisés sont les mieux

soignés. Nous avons montré qu'en fonction du statut social, les individus aisés ont plus de chance de se voir offrir la haute qualité que les autres individus car leur technologie de surveillance dépend de leur savoir médical. De plus, le secteur 2 a instauré une discrimination par le revenu qui n'a pas de fondement en termes d'efficacité (Mougeot, 1999). Lorsque tous les spécialistes d'une même ville sont en secteur 2, l'individu dont la disposition à payer est faible n'a plus accès à des soins techniques ou spécialisés en secteur 1.

Le secteur 2 n'a de raison d'exister que si "l'assurance maladie retient des procédés discriminatoires de rémunération des médecins pour inciter ceux-ci à accepter des actes spécifiques (comme tout ce qui relève de l'urgence)" (Mougeot, 1999) ou des actes de haute qualité. C'est pour cela qu'il est nécessaire de déterminer pour les médecins du secteur 2 des contrats inhibant les comportements opportunistes, c'est-à-dire un couple rémunération-qualité qui maximise l'utilité des patients tout en s'assurant de la participation des médecins.

Chapitre 3

La réforme du secteur 2

“L’art médical trouve son accomplissement dans le retranchement de soi-même et dans la restitution à autrui de sa liberté”, Gadamer (1998).

3.1 Introduction

L’existence du secteur 2 se justifie uniquement si le fait de tarifier plus cher s’accompagne d’une meilleure qualité des soins. Mais, du fait de l’imparfaite observabilité de la qualité du bien santé par le patient, nous avons montré dans le chapitre 2 que le médecin opportuniste est incité à tricher, c’est-à-dire à fournir une basse qualité malgré l’élévation du prix. Ces comportements déviants mis en exergue justifient le gel du secteur 2 établi par la convention médicale de 1993 qui ne restreint pas les pratiques opportunistes. C’est pour cela qu’il est nécessaire de s’interroger sur la rémunération des médecins inscrits en secteur 2. La liberté tarifaire n’incite pas les médecins à fournir la haute qualité.

La rémunération des médecins est au coeur de nombreuses réformes sanitaires. Le gouvernement veut réguler le secteur de l’offre de soins en utilisant la rémunération des médecins puisque ceux-ci sont prescripteurs de soins. Afin de limiter le nombre d’actes produits par le médecin, Mougeot (1999) propose un prix non linéaire des actes. A partir d’un certain seuil, les honoraires diminuent¹. La tutelle doit donc résoudre l’arbitrage entre qualité et efficacité. L’augmentation de la qualité semble générer une hausse des dépenses de santé tandis que l’efficacité résulterait d’une meilleure prescription de la part des médecins pour diminuer les dépenses de santé.

Dans ce chapitre, nous ne considérons que le problème de la qualité. Nous cherchons à déterminer la rémunération incitant les médecins à fournir la qualité qu’ils sont supposés fournir. Seule la tutelle peut proposer des contrats incitatifs aux médecins en raison de l’information médicale dont elle dispose. En France,

¹Le Québec utilise cette règle tarifaire depuis 1976.

la Sécurité Sociale peut réunir une commission de médecins² qui, à la lecture des ordonnances, peut définir, peut-être imparfaitement, la pathologie dont le patient est atteint. La mandature de la tutelle permet de restaurer le cadre de la théorie des contrats. Contrairement au patient, elle n'observe pas l'action mais elle peut juger la juger à l'aulne de l'ordonnance ou des actes réalisés. La rémunération du médecin peut donc devenir contractuelle.

Le chapitre est organisé comme suit. Dans une première section, nous exposons les différentes rémunérations envisagées dans la littérature pour améliorer l'efficacité des services médicaux. Dans ces travaux, il est toujours postulé que les médecins sont opportunistes³. Cette hypothèse doit être reconsidérée puisque les médecins prêtent serment même si le serment d'Hippocrate est plus une tradition qu'un acte de foi. Ce serment fonde l'honnêteté des médecins. Dans une deuxième section, nous présentons ce serment ainsi que ses implications. Il est supposé régir le comportement du médecin. Ceux qui respectent ce serment sont honnêtes puisque leur fonction objectif est le bien-être du patient et non leur bien-être propre. Toutefois, nous avons montré dans le chapitre 2 que certains médecins agissaient avec opportunisme. C'est pourquoi, dans une troisième section, nous présentons une modélisation permettant de caractériser la rémunération optimale des médecins tout en supposant qu'ils diffèrent par deux caractéristiques : leur honnêteté et leur productivité⁴ et ce, pour un bien de confiance (le bien médical). Nous montrons que, si le nombre de médecins honnêtes est grand et si la qualité appréhendable par le régulateur est grande, les comportements malhonnêtes des médecins peu productifs sont annihilés. A contrario, si la qualité appréhendable par la tutelle est faible et si le nombre de médecins honnêtes est grand, les comportements opportunistes sont admis à l'équilibre.

3.2 La rémunération des médecins

Deux voies de recherches se distinguent selon que les dépenses de santé résultent d'un comportement de demande ou d'offre⁵. Cette distinction est criti-

²La Sécurité Sociale dispose environ de 2300 praticiens-conseil pour expertiser.

³Il est à noter que Rochaix (1989) introduit dans la fonction objectif du médecin une composante éthique.

⁴Nous reprenons ici la terminologie de Allard, Cremer et Marchand (1999). Nous devrions plutôt évoquer la dextérité ou l'habileté d'un médecin.

⁵Nous ne présentons ici que les modélisations liant qualité et rémunération pour la médecine ambulatoire. Le paiement des médecins a été envisagé dans d'autres cadres quand il existe une asymétrie d'information.

Le modèle de recherche de marché de Rochaix (1989) considère que le patient exerce une contrainte de marché sur le patient. S'il n'est pas satisfait, il consulte un autre médecin. Les résultats du modèle montrent que plus le patient est informé moins le médecin est incité à sur-prescrire. Par ailleurs, un paiement à l'acte incite le médecin à sur-traiter son patient. Dans ce modèle, la variable est l'intensité du traitement. Il me semble que l'intensité diffère de la qualité puisqu'il s'agit d'un problème de prescription et non de la qualité de la prestation médicale.

Fermon (1991) spécifie l'utilité du médecin qui dépend de son revenu, de son effort pour satisfaire le patient et de la pression de demande du patient. Une détermination des honoraires

cable puisque l'interdépendance entre l'offre et la demande ne peut être occultée. Toutefois, cette présentation a le mérite de dissocier les différents phénomènes.

3.2.1 Comportement de demande

Breuil-Genier et al. (1997) montrent qu'une couverture complémentaire de santé augmente la probabilité d'avoir recours à des soins sanitaires. Ils obtiennent également comme résultat que la densité des spécialistes influence positivement la probabilité de recours à un spécialiste. Breuil-Genier (1999) met en évidence que les phénomènes de demande induite sont limités. Chiappori-Durand et Geoffard (1997) établissent que la consommation accrue des traitements résulte d'une meilleure couverture de soins de santé. Par ailleurs, Delattre et Dormont (2000) montrent que l'hypothèse de demande induite est vérifiée pour le secteur 1 et infirmée pour le secteur 2.

La vérification de la demande induite consiste à valider l'hypothèse de l'existence d'une corrélation positive entre densité médicale et consommation de soins (Rochaix et Jacobzone, 1997). Ce lien peut simplement refléter des effets de substitution entre différents inputs du système de santé. Si les comportements stratégiques des acteurs de la santé ne peuvent être éludés, leur importance potentielle est difficile à saisir (Breuil-Genier et Rupprecht, 2000).

3.2.2 Comportement d'offre

D'autres auteurs considèrent que la consommation de soins est la résultante d'une décision du prescripteur (Ma, 1994). Nous retenons cette hypothèse pour la suite de la section.

La médecine ambulatoire est étudiée par Ma et Mc Guire (1997). Dans une relation d'agence tripartite (assureur, médecin et patient), l'effort de diagnostic du médecin est inobservable par les autres agents et la quantité consommée n'est pas observable par les assureurs. Ces derniers se basent sur le report des traitements établis par les médecins en accord avec les patients. Il en résulte que celui-ci dépend de la rémunération des médecins. Un paiement par capitation avec une pénalité sur chaque unité de traitement n'incite pas les médecins à effectuer le vrai report. Ils sont incités à reporter une quantité inférieure pour éviter la pénalité. Dans ce cadre, le paiement optimal est une rémunération mixte combinant paiement à l'acte et capitation, s'il existe une relation négative entre effort et quantité du traitement. Ils reçoivent une partie fixe pour reporter la

par la tutelle et une demande importante réduisent l'incitation du médecin à satisfaire le patient.

Le modèle de Blomqvist (1991) présente un cadre d'agence de double délégation du médecin : agent du patient et de l'assurance. Lorsqu'il est rémunéré à l'acte, il n'est pas incité à réduire son activité. Il peut donc satisfaire les souhaits de surconsommation de ses patients. En revanche, s'il est salarié d'un HMO (Health Maintenance Organization) il est incité à remplir les objectifs de l'assurance au détriment du patient.

Pour un état des lieux de l'application de la théorie de l'agence à l'économie de la santé, voir Rochaix (1997).

quantité de traitement et ils versent une partie du coût du traitement pour les inciter à réaliser un effort.

Le système par capitation est propre au Royaume-Uni et aux Pays-Bas. Le médecin reçoit des organismes payeurs une allocation par patient. Celui-ci s'abonne et choisit librement son médecin⁶. Cette rémunération qui encourage la prévention et la sélection des patients permet le contrôle des dépenses de santé qui s'effectue au détriment de la qualité des soins.

A contrario, le paiement à l'acte - fixé par négociation entre les organismes tutélaires et les médecins - en vigueur en France et en Allemagne favorise la prescription de soins de qualité. Le patient est libre de choisir son médecin et de consulter. Une telle liberté s'accompagne d'une hausse des dépenses de santé et de visites multiples pour une même pathologie. De plus, le paiement à l'acte n'incite pas à la prévention puisque le revenu des médecins dépend du nombre d'actes réalisés.

Franc (2000) montre qu'il est impossible pour un gouvernement d'instaurer un système de rémunération qui permette à la fois de favoriser la qualité et de diminuer les dépenses de santé. Elle considère quatre agents : patients, gouvernement, médecins et fournisseurs de soins. Le patient peut souffrir de deux pathologies bénignes : une faible et une sévère. Le corps médical est composé de deux types de fournisseurs de soins. Le médecin établit le diagnostic avec une possibilité d'erreur et peut réaliser une partie des soins. Les soins restants sont produits par l'autre fournisseur de soins. L'activité de fourniture de soins du médecin dépend d'un arbitrage financier entre prescrire et produire des soins sanitaires. Pour augmenter son revenu, le médecin produira des soins que l'état de santé du patient ne nécessite pas. Si la maladie est sévère, une consommation de soins inutiles ne réduit pas l'utilité du patient a contrario d'une faible maladie.

L'objectif du gouvernement est de réduire les dépenses inefficaces. Pour ce faire, il maximise le bien-être social. Il a instauré une assurance sociale qui prévoit la prise en charge d'une partie des dépenses sanitaires. Le gouvernement dispose de quatre instruments pour sa politique de réduction de dépenses : le forfait versé au médecin par le patient, la part de remboursement des coûts de production des médecins, la part de remboursement des autres coûts de production et le ticket modérateur.

Sans asymétrie d'information et avec un diagnostic exogène (qui ne conditionne pas l'effort des médecins), le médecin ne réalise aucun acte inefficace s'il n'est pas intégralement remboursé des coûts de production desdits soins. Si le médecin peut améliorer la qualité du diagnostic en effectuant un effort, le système de rémunération dépendra de la possibilité de prescription de soins inefficaces. Si le médecin, ne produit pas ces soins, le système efficace est un système mixte entre paiement à l'acte et capitation. Si sa capacité de production de soins inefficaces est limitée, le gouvernement favorise l'amélioration de la qualité a contrario si les capacités de production de soins du médecin sont

⁶L'abonnement est difficilement résiliable.

importantes⁷.

Allard, Cremer et Marchand (1999) considèrent un cadre principal-agent. Ils étudient l'effet du nombre de patients sur la rémunération du médecin. L'originalité de cette modélisation consiste à supposer que les médecins diffèrent par leur productivité inobservable par l'agence de régulation. Cette dernière a pour objectif la qualité des soins et le nombre de patients par médecin. Intuitivement, un médecin plus productif soignera plus de patients. En information parfaite, le médecin le plus productif reçoit un transfert plus élevé que le médecin peu productif. En outre, il traite plus de patient. En information imparfaite, le médecin le moins productif reçoit un transfert inférieur à celui de premier rang pour limiter la rente informationnelle du médecin le plus productif. Il traite moins de patients puisqu'il diminue sa production sanitaire. Le médecin le plus productif reçoit un transfert supérieur à celui de premier rang en raison de la rente. Toutefois, il traite toujours le même nombre de patients car il ne diminue pas sa production sanitaire.

Tous ces résultats établissent que l'efficacité de premier rang est impossible à mettre en oeuvre. Les actions de ces médecins sont gouvernées par la recherche de leur propre intérêt. Ils maximisent leur propre utilité sans tenir compte de l'intérêt du patient. Ces objectifs financiers qui leur sont assignés semblent contradictoires avec la finalité de la médecine : tenir compte du bien-être du patient.

3.3 Homo economicus versus Homo hippocraticus

L'analyse du comportement économique du médecin introduit quelquefois une contrainte d'altruisme ou d'éthique médicale. Ces contraintes stipulent que le bénéfice en termes sanitaires issu de la production de soins est supérieur à un certain niveau (Ma et Mc Guire, 1997). Gaynor (1994) remarque à juste titre qu'il n'y a pas de consensus à propos de l'altruisme ou de l'éthique médicale. Arrow (1963) suggérait que le médecin pouvait prendre en compte l'intérêt du patient dans la maximisation de son bien-être. Ainsi, le médecin maximise son utilité en tenant compte du patient.

Supposer que la prise en compte du patient n'intervient que par l'intermédiaire d'une contrainte revient à postuler que le médecin est opportuniste mais sans excès. Or, cette affirmation semble contradictoire avec le serment d'Hippocrate que chaque médecin prononce.

3.3.1 Le serment d'Hippocrate

Il établit les règles de bonne conduite du médecin et du patient. L'essentiel est la déontologie des médecins. La célébrité du serment vient du fait que la

⁷Ces résultats sont obtenus en supposant que le coût de production des soins inefficaces est nul.

seconde partie des écrits comprend l'essentiel des devoirs du médecin vis-à-vis du malade et de son entourage. Aussi, n'est-ce pas sans fondement que ce serment est prononcé par les médecins dans des cérémonies rituelles, voire solennelles.

Sous l'Ancien Régime, il existait une distinction très nette entre les branches de l'art médical. Au sommet, la médecine était regroupée dans le cadre de facultés. A l'opposé, la chirurgie, l'apothicairerie et l'art des accouchements étaient considérés comme des métiers manuels. Leur formation reposait sur un apprentissage, tandis que les cours de médecine étaient essentiellement théoriques. Bien que distinctes, ces différentes activités médicales n'étaient pas indépendantes les unes des autres. Les médecins, à cette époque, veillaient à conserver l'usage exclusif des termes de Faculté et doctorat. Ils multipliaient à l'envi, les rites (serment d'Hippocrate) et les insignes (robe, bonnet carré) qui les distinguaient des chirurgiens⁸. C'est moins un souci d'éthique qu'une volonté de distinction qui explique la prononciation du serment Hippocrate par les médecins.

Ce serment fonde la déontologie médicale moderne, codifiée dans la loi. "Le maître mot de cette éthique est l'intérêt du malade" (Jouanna, 1995).

Le serment d'Hippocrate (Extraits)

"J'utiliserai le régime pour l'utilité des malades, suivant mon pouvoir et mon jugement : mais si c'est pour leur perte ou une injustice à leur égard, je jure d'y faire obstacle... C'est dans la pureté et la piété que je passerai ma vie et exercerai mon art... Je n'inciserai pas des malades atteints de lithiase, mais je laisserai cela aux hommes spécialisés dans cette intervention... Dans toutes les maisons où je dois entrer, je pénétrerai pour l'utilité des malades... Tout ce que je verrai ou entendrai au cours du traitement, ou même en dehors du traitement, concernant la vie des gens, si cela ne doit jamais être répété au dehors, je le tairai, considérant que de telles choses sont secrètes..."

Ce serment peut se résumer par une maxime des Epidémies I : "avoir dans les malades deux choses en vues : être utile ou ne pas nuire". Hippocrate affirmait avant Platon et Aristote que la finalité de la médecine était l'intérêt du malade. A défaut d'être utile, le médecin ne doit pas aggraver l'état du malade par des interventions intempestives. Cette dimension humaine constitue l'originalité de la pensée hippocratique (Jouanna, 1995).

3.3.2 Les implications

La règle essentielle de la médecine contenue dans ce serment est que le médecin n'agit que dans l'intérêt du patient. Toutes les règles comportementales inscrites dans ce serment sont en vigueur aujourd'hui, mis à part l'avortement et la piété.

⁸La chirurgie a acquis son indépendance en 1730.

L'avortement est contraire à l'éthique hippocratique car à l'époque le divin est prépondérant. Il n'y a pas, dans les écrits d'Hippocrate, une séparation de la médecine et du divin.

Bien qu'Hippocrate affirmait que le médecin exerçait dans la piété, il ne semble pas qu'il refusait les honoraires (Jouanna, 1995). Cependant, ses tarifs étaient différenciés selon les revenus des malades. Tous les médecins recevaient des honoraires pour les soins qu'ils donnaient. "Les médecins gagnent leurs salaires en guérissant les malades" (Aristote, Politique III).

Ce philosophe considère que la médecine est un modèle de désintéressement car elle a pour fonction de procurer un avantage non à celui qui exerce la médecine mais à celui qui elle s'exerce. C'est pourquoi, pour insister sur le désintéressement auquel doivent s'adonner les médecins, Hippocrate utilise les termes de piété et pureté.

Il apparaît pertinent de retenir comme règle éthique de ce serment que le médecin agit dans l'intérêt du patient, qu'il respecte le secret médical, qu'il oriente les patients chez des spécialistes quand il ne possède pas la dextérité nécessaire à certaines interventions et qu'il ne pratique pas l'acharnement thérapeutique. En effet, évoquer la piété dans nos sociétés industrielles et marchandes semble désuet. Le médecin agit dans l'intérêt du patient (en termes de soins sanitaires) tout en ne négligeant pas ses intérêts financiers. Cependant, il ne néglige pas l'intérêt du patient à des fins mercantilistes. Ces affirmations résumeraient la version moderne du serment d'Hippocrate. En effet, le médecin va négocier sa rémunération avec les organismes payeurs (secteur 1) ou la déterminer (secteur 2). Une fois sa rémunération établie, il fournira la qualité propre aux secteurs sans chercher à la diminuer pour augmenter son bénéfice. En résumé, il respecte les termes du contrat implicite ou explicite. Ce comportement honnête du médecin peut-il être mis en exergue ?

3.3.3 Comportements des médecins

Dans un paragraphe précédent, nous avons établi le comportement opportuniste du médecin. Cependant, il paraît restrictif de considérer tous les médecins comme des fraudeurs. Leur moralité ne peut pas toujours être remise en cause.

Fagon, un des médecins de Louis XIV, dès qu'il fût nommé premier médecin du roi, diminua les revenus de sa charge. De même, il abolit des tributs qui étaient attribués à la nomination des professeurs de médecine aux chaires royales des universités et remboursait sur sa charge aux autres médecins de la Cour ce qu'ils payaient pour leurs serments.

Au XIX^e siècle, dans une société très rurale et majoritairement pauvre, une famille réglait son médecin une ou deux fois par an. Plus ou moins débattue, la somme fournie tenait compte du travail du médecin et des éventuels revers de fortune des clients. Si les malades étaient très pauvres, la consultation était gratuite. En effet, par tradition (serment d'Hippocrate) et par conviction (catholicisme), le médecin a le devoir moral de soigner gratuitement. À côté de cette charité ponctuelle, il existait des formes plus institutionnelles de médecine gratuite (Faure, 1994). Dans les hôpitaux, les médecins exerçaient à titre gratuit.

Dans les villes, ils effectuaient le service médical du bureau de bienfaisance. Par ailleurs, la plupart des estimations convergent pour situer le revenu annuel d'un médecin ni célèbre, ni reprimé, d'une agglomération de taille moyenne entre 4000 et 5000 francs au XIX^e siècle (par exemple, Laval). Ces sommes modestes⁹ étaient acquises au prix d'un effort indéniable¹⁰ (Faure, 1994).

En 1998, le revenu moyen net de charges varie, selon les disciplines, de 310 000 francs pour les pédiatres à 1 million de francs pour les radiologues. La moitié des revenus des médecins est située entre 242 000 francs et 510 000 francs (Audric et al., 2001)¹¹. Cette dispersion des revenus s'est atténuée durant la dernière décennie. Résulte-t-elle de différences de comportements des médecins ? L'augmentation du revenu moyen des généralistes est supérieure à la moyenne (4,8% contre 4,3%). Bien que l'activité des médecins soit très fluctuante, le revenu par tête des médecins est croissant (Audric et al., 2001).

Par ailleurs, les résultats de Delattre et Dormont (2000) révèlent que les omnipraticiens se distinguent par des niveaux de taux de croissance des prescriptions supérieurs à ceux des spécialistes. Ils notent également l'importance des dispersions des taux de croissance de l'activité mesurés par quartile. Un quart des médecins a atteint des rythmes de croissance d'activité qui dépassent nettement les objectifs annuels définis au niveau global. Ces disparités accentuées semblent résulter de comportements différents des médecins. Pour simplifier, certains agiraient comme des *homo economicus* et les autres comme des *homo hippocraticus*¹².

Ces différents faits étayent l'hypothèse que tous les médecins n'agissent pas en tant qu'*homo economicus*. Il semble nécessaire de supposer l'hétérogénéité comportementale des médecins dans la détermination d'une règle de rémunération des services médicaux.

3.4 Honnêteté et bien de confiance

Les services médicaux tels que la consultation médicale correspondent aux caractéristiques des biens de confiance. Le patient est incapable de détecter la qualité de la consultation à la fin de celle-ci. Dans ce cas, la réputation d'un médecin s'établit moins à l'aune de la qualité qu'à la représentation subjective de celle-ci. Les résultats standards de Shapiro (1983), Allen (1984) et Grossman and Shapiro (1988) ne peuvent être appliqués puisqu'ils étudiaient des biens de recherche.

L'objectif de cette partie est d'étudier l'incidence de l'introduction d'un bien

⁹A l'époque, un curé de première classe gagne 1200 francs, un haut-fonctionnaire 4 000 francs et un manoeuvre 400 francs, Faure (1994).

¹⁰Léonard (1978) établit que la clientèle d'un médecin rural réside à plus de 10 kilomètres de la localité du praticien. Ces cheminements à la marche ou à cheval ne sont que rarement rétribués. Le médecin arrive souvent trop tard pour soigner la maladie d'un patient pauvre.

¹¹Les médecins à temps complet déclarent travailler 57,5 heures par semaine. (Source : DREES, enquête emploi, 1999).

¹²Contrairement à Rochaix (1989), nous ne supposons pas que les médecins prennent en compte aussi bien leur bien-être que celui des patients.

de confiance dans un modèle standard d'antisélection (Guesnerie et Laffont, 1984 ; Laffont, 1989 ; Maskin et Riley, 1984 ; Moore, 1988 ; Salanié, 1997 ; Stiglitz et Weiss, 1981). Au problème de révélation de l'information privée de l'agent se greffe une difficile appréhension de la qualité par le principal.

Dans un premier temps, nous présentons le cadre analytique du modèle. Dans un deuxième temps, sont exposés les résultats d'information parfaite ainsi que les cas d'asymétrie d'information unique qui caractérisent les cas standards. Dans un troisième temps, nous montrons que dans un cas de double asymétrie d'information, les résultats standards sont retrouvés. Dans un dernier temps, nous supposons que les médecins se différencient par leur productivité ainsi que par leur opportunisme. Dans un tel cadre, deux équilibres apparaissent : un contrat séparateur sans fraude et un contrat semi-séparateur avec fraude.

3.4.1 Le modèle

Nous supposons ici que le principal est une agence de réglementation sanitaire et que l'agent est un médecin. Cette agence propose un contrat au médecin pour soigner le patient. Le patient est passif car, dans le cas d'un bien de confiance, il ne peut pas déterminer la qualité de la prestation médicale. Il confie donc ses intérêts à une agence sanitaire dont les compétences en termes de santé sont supérieures aux siennes. Dans cette économie, la productivité du médecin est de deux types : une forte et une faible. Un médecin productif, par exemple, est un médecin qui établit un meilleur diagnostic que le médecin peu productif pour un même coût. La prestation médicale du médecin permet la délivrance d'un traitement, matérialisé par une ordonnance, qui soigne le patient. L'analyse du traitement (feuille de soins) ne peut renseigner parfaitement l'agence sur la qualité de la prestation médicale. Seul un audit de qualité coûteux permettrait de connaître l'exacte qualité de la prestation médicale.

L'agence sanitaire

Elle maximise le surplus des patients qui s'écrit comme la différence entre la qualité du service médical et le paiement versé au médecin : $q - t$.

L'agence ne connaît pas la valeur de la productivité du médecin mais connaît la probabilité a priori que le médecin soit peu productif π . En outre, elle ne peut pas appréhender parfaitement la haute qualité d'une prestation médicale. Toutefois, il existe un standard de qualité q_0 tel que pour toute qualité strictement inférieure l'agence peut déterminer avec certitude la qualité produite¹³. Par contre, pour toute qualité supérieure ou égale à ce standard, l'agence ne peut déterminer avec certitude la qualité.

¹³Si, par exemple, un médecin ne fournit que des médicaments dont les effets thérapeutiques sont contestés, l'agence de régulation peut en déduire la mauvaise qualité du traitement. Par contre, nous avons décrit dans un paragraphe précédant que pour une pathologie donnée, il n'existait pas de traitement idéal. Lors de réunions de consensus, les médecins tentent d'établir une trame de diagnostic et de traitement. Le traitement idéal n'existe pas car le traitement dépend des habitudes du médecin et des attentes des patients (Guérin, 1995) et de l'influence des laboratoires pharmaceutiques.

Le médecin

L'utilité du médecin est la différence entre le transfert reçu et le coût de la prestation médicale, $t - C(q, \theta)$ avec θ la productivité du médecin¹⁴. $C(q, \theta)$ est une fonction continue, croissante et convexe de la qualité q et décroissante en θ .

Deux valeurs de θ sont possibles : $\theta_1 < \theta_2$. Une faible valeur de θ signifie que le médecin est peu productif. Ainsi, $C(q, \theta_2) < C(q, \theta_1)$. Pour une même valeur de la qualité, le médecin peu productif aura un coût de production du service médical supérieur à celui productif. Cette inégalité traduit la forme discrète de la condition Spence-Mirrlees. La sécance unique des courbes d'indifférence des médecins permet la caractérisation de contrats séparateurs.

Le médecin connaît les difficultés de la tutelle à déterminer la qualité. La valeur de q_0 est connaissance commune de l'agence et des médecins.

3.4.2 Antisélection ou bien de confiance

Deux asymétries d'information caractérisent ce modèle. Nous étudierons ici trois cas. Tout d'abord, le principal connaît la productivité des médecins et discerne la qualité du service médical. Nous retrouvons la solution du premier rang des modèles d'antisélection standards. Ensuite, le principal discerne la qualité mais ne connaît pas la productivité du médecin. Cette situation caractérise la solution du second rang des modèles d'antisélection standards. Enfin, l'agence connaît le type des médecins mais ne discerne qu'imparfaitement la qualité fournie par les médecins.

La solution du premier rang

L'agence observe parfaitement le type du médecin ainsi que la qualité. Elle maximise le surplus du patient en tenant compte de la contrainte de participation du médecin. Elle résout le programme suivant :

$$\begin{aligned} \max_{q_i, t_i} & q_i - t_i \\ \text{sc } & t_i - C(q_i, \theta_i) \geq 0 \end{aligned}$$

A l'optimum, $t_i = C(q_i, \theta_i)$ et q_i^* est définie par l'égalité du coût marginal et du bénéfice marginal.

$$C_{q_i}(q_i^*, \theta_i) = 1$$

où $C_{q_i}(q_i^*, \theta_i)$ est la dérivée du coût par rapport à la qualité calculée en q_i^* .

¹⁴Nous avons observé précédemment que le diagnostic du médecin était imprécis. Ici, la productivité θ peut signifier la compétence du médecin à établir un diagnostic précis. Plus il est compétent, moins il lui est coûteux d'établir un diagnostic précis et plus il paraît probable que la qualité du traitement est grande.

Par exemple, un médecin peu productif est un médecin qui prescrit des principes actifs à la faible efficacité thérapeutique.

En information parfaite, les médecins obtiennent une utilité nulle tandis que le médecin peu productif produit une qualité inférieure à celle du médecin productif. Comme $C(\cdot)$ est une fonction croissante de la qualité et que $\theta_1 < \theta_2$, on en déduit que $q_1^* < q_2^*$. Ces qualités sont dites efficaces.

Définition 4 *Le standard de qualité est défini tel que $q_0 < q_2^*$*

Si $q_0 > q_2^*$, le problème lié à l'inobservabilité de la qualité n'a plus lieu d'être car nous savons que, dans les modèles d'antisélection, les valeurs de la qualité du second rang ne sont jamais supérieures à celle du premier rang.

Inobservabilité de la productivité

L'agence appréhende la qualité du bien santé délivré tandis qu'elle n'apprécie pas la productivité des médecins. Elle connaît seulement la probabilité a priori que le médecin soit peu productif (π) et respectivement productif $(1 - \pi)^{15}$. Si l'agence propose les contrats de premier rang, le médecin productif choisira le contrat peu productif car il en retire une utilité supérieure puisque $C(q_1, \theta_1) - C(q_1, \theta_2) > 0$. Ce différentiel de coût est appelé rente informationnelle. Le coût est une fonction décroissante de la productivité. Les deux médecins choisissent le même contrat.

Le couple de contrats séparateurs sera obtenu en résolvant le programme suivant :

$$\max_{q_1, t_1, q_2, t_2} \pi(q_1 - t_1) + (1 - \pi)(q_2 - t_2)$$

sous les contraintes

$$t_1 - C(q_1, \theta_1) \geq 0 \quad IR1$$

$$t_2 - C(q_2, \theta_2) \geq 0 \quad IR2$$

$$t_1 - C(q_1, \theta_1) \geq t_2 - C(q_2, \theta_1) \quad IC1$$

$$t_2 - C(q_2, \theta_2) \geq t_1 - C(q_1, \theta_2) \quad IC2$$

Les deux premières équations représentent les contraintes de participation. Le médecin accepte le contrat si celui-ci lui procure une utilité supérieure à son utilité de réservation normée à 0. Les deux dernières contraintes, nommées contraintes d'incitation, spécifient que chaque médecin accepte le contrat approprié à son type.

¹⁵Nous supposons que la valeur de π est suffisamment élevée pour qu'il soit optimal de maintenir en activité les deux types de médecin. Si π est faible, le coût marginal de l'individu peu productif est négatif. Un contrat sera seulement offert aux médecins productifs.

A l'optimum, seules deux contraintes sont saturées (*IR1* et *IC2*)¹⁶. L'individu le moins productif reçoit un contrat qui lui procure son utilité de réservation (son coût de production est remboursé) tandis que l'individu le plus productif est économiquement indifférent entre les deux contrats. Ainsi, il obtient une utilité supérieure à 0 en raison du versement d'une rente d'information (gain monétaire nécessaire pour que l'individu le plus productif révèle son type). S'il se déclare de type θ_1 il reçoit un transfert égal à $C(q_1, \theta_1)$ alors que son coût de production est $C(q_1, \theta_2)$. Pour révéler son type, il doit recevoir en sus du remboursement de son coût de production ($C(q_2, \theta_2)$), une rente égale à $C(q_1, \theta_1) - C(q_1, \theta_2)$.

Les conditions du premier ordre sont :

$$\pi(1 - C_{q_1}(q_1, \theta_1)) - (1 - \pi)[C_{q_1}(q_1, \theta_1) - C_{q_1}(q_1, \theta_2)] = 0$$

$$1 - C_{q_2}(q_2, \theta_2) = 0$$

Ainsi (q_1^{**}, q_2^{**}) est solution du système :

$$C_{q_2}(q_2^{**}, \theta_2) = 1$$

$$C_{q_1}(q_1^{**}, \theta_1) = 1 - \frac{(1 - \pi)}{\pi} [C_{q_1}(q_1^{**}, \theta_1) - C_{q_1}(q_1^{**}, \theta_2)] < 1$$

Ainsi, $q_1^{**} < q_1^*$ et $q_2^{**} = q_2^*$

L'individu le plus productif produit la qualité efficace tandis que l'individu peu productif délivre une qualité inférieure à celle efficace. Diminuer la qualité de l'individu peu productif permet de réduire la rente versée à l'individu productif puisque celle-ci est une fonction croissante de q_1 . A l'optimum, le principal effectue un arbitrage entre la qualité q_1 et la rente versée à l'individu productif.

Inobservabilité de la qualité

L'agence appréhende imparfaitement la qualité du bien santé délivré tandis qu'elle connaît la productivité des médecins. Deux cas sont à étudier. Soit q_0 est inférieure à q_1^* (par la suite cette situation aura pour indice q), soit q_0 est supérieure à q_1^* (indice Q).

1^{er} cas : $q_0 < q_1^*$ Si l'agence propose à chaque médecin les contrats de premier rang, ils produiront q_0 puisque pour toute qualité supérieure ou égale à q_0 , l'agence n'est pas capable de déterminer la valeur de la qualité.

L'agence maximise le surplus du consommateur sous contrainte de participation. L'agence tient compte du comportement opportuniste des médecins productifs. Elle résout le programme suivant :

$$\max_{q_1, t_1, q_2, t_2} \pi(q_1 - t_1) + (1 - \pi)\{q_2 - t_2\}$$

¹⁶Si *IR1* est saturée alors *IC2* est saturée.

Si *IC1* n'est pas saturée alors *IC2* est saturée.

IR1 et *IC2* saturées impliquent que *IR2* n'est pas saturée. (Pour la preuve, voir Salanié (1997)).

sous les contraintes

$$t_1 - C(q_1, \theta_1) \geq 0 \quad IR1$$

$$t_2 - C(q_2, \theta_2) \geq 0 \quad IR2$$

$$q_1 \leq q_0 \quad IP1$$

$$q_2 \leq q_0 \quad IP2$$

La contrainte (IP_i) traduit le fait que l'agence en raison de sa déficience informative sur la qualité anticipe que les médecins produiront au mieux q_0 avec $i \in \{1, 2\}$. Soit γ_i le coefficient de Lagrange associé à (IP_i).

A l'optimum, les deux contraintes de participation sont liées puisque l'agence maximise le surplus du patient. Chaque médecin obtient son utilité de réservation.

Les conditions du premier ordre sont :

$$(1 - C_{q_1}(q_1, \theta_1)) - \gamma_1 \geq 0$$

$$(1 - C_{q_2}(q_2, \theta_2)) - \gamma_2 \geq 0$$

Si $\gamma_i = 0$, les contraintes IRi ne sont pas saturées. Nous obtenons que $q_i = q_i^*$. Or, nous avons supposé que $q_0 < q_1^*$. Ainsi, $\gamma_i \neq 0$. Soit $q_i^{*q} = q_0$.

Ils recevront comme transfert le remboursement de leur coût (pour qu'ils acceptent de contracter)¹⁷ puisque l'agence connaît la productivité du médecin.

2nd cas : $q_1^* < q_0 < q_2^*$ Si l'agence propose les contrats de premier rang, les médecins productifs fourniront la qualité q_0 et recevront un transfert $C(q_2^*, \theta_2)$ supérieur aux coûts de production. Seuls les médecins productifs peuvent fauter car q_0 est supérieure à la qualité efficace des agents peu productifs.

L'agence maximise le surplus du consommateur sous contrainte de participation. L'agence tient compte du comportement opportuniste des médecins productifs. Elle résout le programme suivant :

$$\max_{q_1, t_1, q_2, t_2} \pi(q_1 - t_1) + (1 - \pi) \{q_2 - t_2\}$$

sous les contraintes :

$$t_1 - C(q_1, \theta_1) \geq 0 \quad IR1$$

¹⁷Le surplus est une fonction décroissante du transfert.

$$t_2 - C(q_2, \theta_2) \geq 0 \quad IR2$$

$$q_2 \leq q_0 \quad IP$$

La contrainte (*IP*) traduit le fait que l'agence en raison de sa déficience informative sur la qualité anticipe que les médecins productifs fourniront au mieux q_0 . Soit γ le coefficient de Lagrange associé à (*IP*).

À l'optimum, les deux contraintes de participation sont liées puisque l'agence maximise le surplus du patient. Les conditions du premier ordre sont :

$$(1 - C_{q_1}(q_1, \theta_1)) = 0$$

$$(1 - C_{q_2}(q_2, \theta_2)) - \gamma \geq 0$$

Si $\gamma = 0$ alors $q_2 = q_2^*$. Cette solution n'est pas réalisable.

Nous obtenons donc $C_{q_1}(q_1, \theta_1) = 1$ avec $q_1^{*Q} = q_1^*$, et $q_2^{*Q} = q_0$ avec $q_2^{*Q} < q_2^*$.

Proposition 5 *Si le principal observe parfaitement le type de l'agent a contrario de la qualité demandée, les caractéristiques du contrat dépendent de l'information de l'agence*

- si $q_0 < q_1^*$ le contrat est mélangeant : $q_i^{*q} = q_0 < q_1^*$
- si $q_0 \geq q_1^*$, le contrat est séparateur : $q_1^{*Q} = q_1^* < q_2^{*Q} = q_0 < q_2^*$

Quand l'agence observe imparfaitement la qualité, l'agent le plus productif délivre une qualité inférieure à celle du premier rang. Par ailleurs, quand l'agence ne connaît pas le type de l'agent, les agents peu productifs délivrent une qualité inférieure à celle du premier rang. De la combinaison de ces deux asymétries d'information résultera-t-il une production de la part des deux agents inférieure à celle du premier rang ?

3.4.3 Antisélection et bien de confiance

L'agence ne décèle ni la qualité du produit ni le type de l'agent (médecin). En information imparfaite, l'agence sait seulement que les agents peu productifs sont en proportion π et qu'elle ne peut pas reconnaître la valeur de la qualité si elle est supérieure ou égale à q_0 . Les termes du contrat dépendront de l'information de l'agence. Deux cas sont à étudier. La valeur seuil du standard de qualité est q_1^{**} . En effet si $q_0 < q_1^{**}$ tous les agents productifs auront un comportement opportuniste et produiront au mieux q_0 . Si $q_0 > q_1^{**}$, seuls les agents productifs peuvent frauder.

1er cas : $q_0 < q_1^{**}$

Si l'agence propose les contrats séparateurs de premier rang, les agents produiront tous q_0 et recevront comme transfert le coût de production de premier rang. Les individus productifs se déclareront improductifs, car pour une production q' contractée, ils recevront $C(q', \theta_1)$ alors que le coût réel de production est $C(q_0, \theta_2)$. S'ils s'étaient déclarés productifs, leur bénéfice aurait été $C(q', \theta_2) - C(q_0, \theta_2)$. Il faut donner aux individus productifs pour qu'ils se déclarent productifs, une rente égale à $C(q', \theta_1) - C(q', \theta_2)$.

Le contrat proposé aux agents est obtenu en résolvant le programme suivant :

$$\max_{q_1, t_1, q_2, t_2} \pi(q_1 - t_1) + (1 - \pi) \{q_2 - t_2\}$$

sous les contraintes

$$t_1 - C(q_1, \theta_1) \geq 0 \quad IR1$$

$$t_2 - C(q_2, \theta_2) \geq 0 \quad IR2$$

$$t_1 - C(q_1, \theta_1) \geq t_2 - C(q_2, \theta_1) \quad IC1$$

$$t_2 - C(q_2, \theta_2) \geq t_1 - C(q_1, \theta_2) \quad IC2$$

$$q_1 \leq q_0 \quad IP1$$

$$q_2 \leq q_0 \quad IP2$$

Les deux contraintes (IRi) sont les contraintes de participation. Le médecin accepte le contrat si celui-ci lui procure une utilité supérieure à son utilité de réservation normée à 0. Les deux contraintes (IC) spécifient que chaque médecin préfère le contrat approprié à son type à tout autre contrat. L'individu de type θ_1 sélectionnera son contrat si l'utilité qu'il en retire est supérieure à celle obtenue pour tout autre contrat ($IC1$). De même pour ($IC2$). Les contraintes (IPi) stipulent que chaque agent n'offrira pas plus que q_0 . Soient γ_i , les coefficients de Lagrange associés aux contraintes IP_i .

Les conditions du premier ordre sont :

$$\pi(1 - C_{q_1}(q_1, \theta_1)) + (1 - \pi) \{-C_{q_1}(q_1, \theta_1) + C_{q_1}(q_1, \theta_2)\} - \gamma_1 \geq 0$$

$$(1 - C_{q_2}(q_2, \theta_2)) - \gamma_2 \geq 0$$

Si $\gamma_i = 0$ les contraintes IP_i ne sont pas saturées. Ainsi, $q_i = q_i^{**}$. Nous savons que $q_i < q_0$ alors que $q_i^{**} > q_0$.

Nous venons de montrer que $\gamma_i \neq 0$.

Soit $q_i^{**q} = q_0$. L'agence ne connaissant pas le type des agents est obligée d'offrir le même transfert. En effet, pour que les deux agents participent¹⁸, il faut que le transfert soit égal à $C(q_0, \theta_1)$. Si elle proposait $C(q_0, \theta_2)$, seuls les agents de type 2 produiraient.

2nd cas : $q_1^{**} < q_0 < q_2^*$

Si l'agence propose les contrats séparateurs de premier rang, tous les agents hésiteront entre produire q_0 (se déclarer productif) ou se déclarer peu productif. En effet, les agents productifs (respectivement peu productifs) en révélant leur type obtiendraient comme utilité $C(q_2, \theta_2) - C(q_0, \theta_2)$ (respectivement 0) sinon $C(q_1, \theta_2) - C(q_1, \theta_2)$ (respectivement $C(q_2, \theta_2) - C(q_0, \theta_1)$)¹⁹.

L'agence ne proposera pas le contrat $(q_2, t = C(q_2, \theta_2))$ car les agents se déclarant productifs ne produiront que q_0 .

Le contrat proposé aux agents est obtenu en résolvant le programme suivant :

$$\max_{q_1, t_1, q_2, t_2} \pi(q_1 - t_1) + (1 - \pi) \{q_2 - t_2\}$$

sous les contraintes

$$t_1 - C(q_1, \theta_1) \geq 0 \quad IR1$$

$$t_2 - C(q_2, \theta_2) \geq 0 \quad IR2$$

$$t_1 - C(q_1, \theta_1) \geq t_2 - C(q_2, \theta_1) \quad IC1$$

$$t_2 - C(q_2, \theta_2) \geq t_1 - C(q_1, \theta_2) \quad IC2$$

$$q_2 \leq q_0 \quad IP$$

Les deux contraintes (*IR*) sont les contraintes de participation. Le médecin accepte le contrat s'il obtient une utilité non négative. Les deux contraintes (*IC*) spécifient que chaque médecin préfère le contrat approprié à son type à tout autre contrat. La contrainte (*IP*) stipule que l'agence ne demandera jamais aux agents productifs une qualité supérieure à q_0 .

Comme précédemment, seules les contraintes *IC2* et *IR1* sont saturées²⁰. Soit γ le coefficient de Lagrange associé à (*IP*).

¹⁸Nous avons supposé que π est suffisamment grand pour que les deux types d'agent participent.

¹⁹Si $q_2 = q_0$ l'individu improductif ne se déclare pas productif.

²⁰Cette affirmation reste vérifiée pour toute valeur de q_2 .

Les conditions du premier ordre sont :

$$\pi(1 - C_{q_1}(q_1, \theta_1)) + (1 - \pi) \{-C_{q_1}(q_1, \theta_1) + C_{q_1}(q_1, \theta_2)\} = 0$$

$$(1 - C_{q_2}(q_2, \theta_2)) - \gamma \geq 0$$

Si la contrainte IP n'est pas saturée alors $\gamma = 0$. Nous obtenons de la seconde condition que $q_2 = q_2^*$. Or nous savons que $q_2 < q_0 < q_2^*$. Ainsi, nous venons de montrer par l'absurde que $\gamma \neq 0$. Nous obtenons donc :

$$C_{q_1}(q_1, \theta_1) = 1 - \frac{(1 - \pi)}{\pi} [C_{q_1}(q_1, \theta_1) - C_{q_1}(q_1, \theta_2)]$$

$$q_2 = q_0^{21}$$

De cette équation, nous en déduisons que q_1^{**Q} , solution de cette équation, est inférieure à celle de premier rang afin de diminuer la rente versée à l'individu productif. De plus q_2^{**Q} est inférieure à q_2^* en raison de la déficience d'information de l'agence ayant trait à l'évaluation de la qualité.

Proposition 6 *En présence d'une double asymétrie d'information sur la qualité et la productivité, le contrat est*

- si $q_0 < q_1^{**}$ mélangeant : $q_i^{**q} = q_0$
- si $q_1^{**} < q_0 < q_2^*$ séparateur et de troisième rang, avec $q_2^{**Q} < q_2^*$ et $q_1^{**Q} < q_1^*$

Nous obtenons ici un résultat conforme à l'intuition. L'asymétrie d'information sur la qualité engendre une baisse de la qualité. Un résultat standard est également vérifié. L'asymétrie d'information ayant trait à la productivité diminue la qualité demandée à l'agent le moins productif pour diminuer la rente d'information versée à l'autre agent.

3.4.4 Introduction de l'honnêteté des médecins

Nous supposons maintenant que deux caractéristiques distinguent les médecins : leur productivité et leur honnêteté. Ils sont honnêtes ou opportunistes, et productifs ou peu productifs. Un médecin honnête est un médecin qui produit toujours la qualité qu'il a contractée contrairement au médecin opportuniste²².

L'honnêteté des médecins est contenue dans le serment d'Hippocrate²³. Il n'agit en termes sanitaires que dans l'intérêt du patient. Il fournira toujours la qualité contractée. Les différents exemples cités antérieurement montrent que l'opportunisme ou l'honnêteté gouvernent les actions des médecins. Toutefois,

²²La fonction d'utilité des médecins s'énonce : $t_i - C(q_i, \theta_i) + A_j$ où A_j est le paramètre d'honnêteté des médecins. $A_j = 0$ pour un médecin opportuniste. A_j tend vers plus l'infini quand le médecin honnête fournit la qualité contractée et vers moins l'infini autrement.

²³Voir section précédente.

lors de la détermination des honoraires, tous les médecins feront prévaloir leurs intérêts financiers. Ceci ne semble pas contraire à l'éthique médicale. En effet, la Sécurité Sociale ayant garanti aux plus pauvres l'accès gratuit aux soins par l'instauration de la Couverture Maladie Universelle, les honoraires ne sont pas un obstacle aux soins médicaux. De plus, les patients qui rétribuent les médecins sont remboursés en partie par la sécurité sociale, et pour certains d'entre eux par une assurance complémentaire. Le médecin peut donc défendre ses intérêts financiers auprès des organismes sociaux puisque le coût des soins n'est pas prohibitif (nous ne considérons que la médecine ambulatoire).

L'agence connaît seulement la probabilité a priori que le médecin soit opportuniste ($1 - \eta$) (noté O) et respectivement honnête η (noté H). De plus, l'agence ne décèle ni la qualité du produit, ni le type de l'agent. En information imparfaite, l'agence sait seulement que les agents peu productifs sont en proportion π et qu'elle ne peut plus reconnaître la valeur de la qualité la plus élevée si elle est supérieure ou égale à q_0 .

L'honnêteté de certains médecins amène l'agence à réviser ses actions. Si elle propose les contrats de second rang quand la productivité est inobservable, elle sait que seuls les agents opportunistes profiteront de son manque d'information. Deux comportements déviants sont à considérer. Les agents productifs fourniront une qualité moindre mais révéleront leur type. A contrario des agents peu productifs pourraient être incités à ne pas révéler leur type si :

$$\{C(q_2, \theta_2) + C(q_1, \theta_1) - C(q_1, \theta_2)\} - C(q_0, \theta_1) > 0^{24}$$

Nous en déduisons que cette inéquation est vérifiée si q_0 est faible et si q_2 est grand. Les caractéristiques de l'équilibre dépendront de la valeur de l'information de l'agence et du nombre d'agents honnêtes. En effet, plus le nombre d'agents honnêtes sera grand, plus q_2 sera grand et plus la perte de bien-être générée par des comportements opportunistes sera faible. Par ailleurs, plus la valeur de q_0 est élevée, plus le bénéfice lié aux comportements opportunistes est réduit. Nous étudierons deux cas comme précédemment.

1^{er} cas : $q_0 < q_1 < q_2$

Nous supposons que l'agence n'est pas capable de saisir parfaitement la qualité fournie par les médecins.

Si l'agence propose le contrat séparateur de second rang, tous les médecins opportunistes fourniront la qualité q_0 . Les médecins opportunistes peu productifs se déclareront toujours productifs car pour des qualités q_2 et q_1 :

$$\{C(q_2, \theta_2) + C(q_1, \theta_1) - C(q_1, \theta_2)\} - C(q_0, \theta_1) > 0$$

car $q_0 < q_1$.

Les contrats proposés dépendront du nombre de médecins honnêtes. En effet, si tous les agents sont opportunistes, la tutelle proposera comme contrat $(q_0, C(q_0, \theta_1))$. Si tous les agents sont honnêtes, elle proposera les contrats de second rang quand le type des agents n'est pas connu. Par ailleurs, si le nombre

d'agents honnêtes est faible, elle proposera le contrat mélangeant $(q_0, C(q_0, \theta_1))$. Si le nombre de médecins honnêtes augmente, la tutelle proposera un contrat semi-séparateur quand les gains, en termes de surplus liés à la fourniture de la qualité, seront supérieurs aux pertes générées par les comportements opportunistes. Le contrat est dit semi-séparateur car seuls les médecins honnêtes seront séparés.

Ce contrat est obtenu en résolvant le programme suivant :

$$\max_{q_1, t_1, q_2, t_2} \pi(q_1 - t_1)\eta + (1 - \pi)\eta \{q_2 - t_2\} + (1 - \eta) \{q_0 - t_2\}$$

sous les contraintes

$$t_1 - C(q_1, \theta_1) \geq 0 \quad IRH1$$

$$t_2 - C(q_2, \theta_2) \geq 0 \quad IRH2$$

$$t_1 - C(q_1, \theta_1) \geq t_2 - C(q_2, \theta_1) \quad ICH1$$

$$t_2 - C(q_2, \theta_2) \geq t_1 - C(q_1, \theta_2) \quad ICH2$$

$$q_0 \leq q_1 \quad IP1$$

$$q_0 \leq q_2 \quad IP2$$

Nous omettons les contraintes des médecins opportunistes car nous savons qu'elles ne seront jamais saturées.

En effet, si $t_2 - C(q_2, \theta_2) \geq 0$ alors $t_2 - C(q_0, \theta_2) \geq 0$ car $q_2 \geq q_0$. De plus, nous avons montré que leur contrainte d'incitation n'est pas active car ils ne révèlent pas leur type.

A l'optimum, les contraintes (*IRH1* et *ICH2*) sont saturées. Soit :

$$t_1 = C(q_1, \theta_1)$$

$$t_2 = C(q_2, \theta_2) + C(q_1, \theta_1) - C(q_1, \theta_2)$$

Les conditions du premier ordre sont :

$$\pi\eta(1 - C_{q_1}(q_1, \theta_1)) - (1 - \pi)\eta \{C_{q_1}(q_1, \theta_1) - C_{q_1}(q_1, \theta_2)\} - \gamma_1 = 0$$

$$(1 - \pi)\eta - (1 - \pi)\eta C_{q_2}(q_2, \theta_2) - \gamma_2 = 0$$

Si les contraintes IP_i ne sont pas saturées alors $\gamma_i = 0$. Nous obtenons :

$$C_{q_1}(q_1^{**qh}, \theta_1) = 1 - \frac{(1 - \pi\eta)}{\pi\eta} \{C_{q_1}(q_1, \theta_1) - C_{q_1}(q_1, \theta_2)\}$$

$$C_{q_2}(q_2^{**qh}, \theta_2) = \frac{(1 - \pi)\eta}{(1 - \pi\eta)}$$

Ainsi $q_1^{**qh} < q_1^{**}$ et $q_2^{**qh} < q_2^{**}$

Remarque 7 $\frac{\partial q_i^{**qh}}{\partial \eta} > 0$

Plus le nombre de médecins honnêtes est élevé, plus la qualité fournie par les médecins sera grande puisque les comportements opportunistes se raréfient.

Il est aisé de montrer que $q_1^{**qh} < q_2^{**qh}$ ²⁵.

Ainsi, ce contrat semi-séparateur est défini si $q_1^{**qh} \geq q_0$. De la première condition du premier ordre nous déduisons qu'il existe $q_1^{**qh} \geq q_0$ si $\eta \geq \underline{\eta}$. Il est à dénoter que plus q_0 est faible, plus $\underline{\eta}$ est faible.

Si les contraintes IP_i sont saturées, alors $\gamma_i \neq 0$. Nous obtenons que $q_i^{**h} = q_0$. Chaque individu produit la même qualité. Nous avons supposé que π était suffisant important pour que chaque individu produise. Le transfert reçu est égal à $C(q_0, \theta_1)$. Les médecins productifs reçoivent une rente.

Nous pouvons en conclure qu'il existe $\eta > \underline{\eta}$ tel que le contrat semi-séparateur se substitue au contrat mélangeant. Dans ce cas, $q_0 < q_1^{**h} < q_2^{**h}$. Les médecins honnêtes peu productifs (respectivement productifs) produisent la qualité q_1^{**h} (respectivement q_2^{**h}). Les médecins opportunistes choisissent le contrat des agents productifs et fournissent la qualité q_0 .

Proposition 8 si $q_0 < q_1^{**Qh}$ alors

- l'équilibre sera mélangeant sans fraude si $\eta < \underline{\eta}$: tous les médecins produisent la qualité q_0 .
- semi-séparateur avec fraude autrement : les médecins opportunistes choisissent le contrat des médecins honnêtes productifs et produisent q_0 .

Si le nombre de médecins honnêtes est important, la qualité demandée aux agents est supérieure à celle demandée en absence d'honnêteté. Le gain en termes de surplus lié à l'honnêteté est supérieur à la perte générée par les comportements opportunistes. Dans le cadre d'un contrat séparateur d'un bien de confiance, les qualités optimales fournies sont réduites par rapport à la situation de second rang d'un bien de recherche.

²⁵ q_i^{**qh} est une fonction croissante de η .

Nous avons : $q_1^{**qh}(\eta = 0) < q_2^{**qh}(0)$ et $q_1^{**qh}(1) < q_2^{**qh}(1)$. Nous en déduisons que pour tout $\eta \in [0, 1]$, $q_1^{**qh} < q_2^{**qh}$.

2nd cas : $q_1 < q_0 < q_2$

Nous supposons que la valeur limite de la qualité q_0 reconnue par l'agence est moyenne. Les caractéristiques des contrats dépendront du comportement de l'agent peu productif opportuniste. Sa sélection du contrat dépendra de la valeur de q_2 . Si elle est proche de q_0 , sa contrainte d'incitation s'énonce :

$$\{C(q_2, \theta_2) + C(q_1, \theta_1) - C(q_1, \theta_2)\} - C(q_2, \theta_1) < 0$$

Nous avons vu précédemment que la quantité q_2 est une fonction croissante du nombre de médecins honnêtes. Nous supposons dans un premier temps, que le signe de cette contrainte est négatif, c'est-à-dire qu'ils révèlent leur type.

Contrats séparateurs La tutelle maximise le programme suivant :

$$\max_{q_1, t_1, q_2, t_2} \pi(q_1 - t_1) + (1 - \pi)\eta\{q_2 - t_2\} + (1 - \pi)(1 - \eta)\{q_0 - t_2\}$$

sous les contraintes

$$t_1 - C(q_1, \theta_1) \geq 0 \quad IRH1$$

$$t_2 - C(q_2, \theta_2) \geq 0 \quad IRH2$$

$$t_1 - C(q_1, \theta_1) \geq t_2 - C(q_2, \theta_1) \quad ICH1$$

$$t_2 - C(q_2, \theta_2) \geq t_1 - C(q_1, \theta_2) \quad ICH2$$

$$q_0 \leq q_2 \quad IP$$

Soit γ le multiplicateur de Lagrange associé à la contrainte IP .

À l'optimum, les contraintes ($IRH1$ et $ICH2$) sont saturées. Les conditions du premier ordre sont :

$$\pi(1 - C_{q_1}(q_1, \theta_1)) - (1 - \pi)\{C_{q_1}(q_1, \theta_1) - C_{q_1}(q_1, \theta_2)\} = 0$$

$$(1 - \pi)\eta - (1 - \pi)C_{q_2}(q_2, \theta_2) - \gamma \geq 0$$

Nous obtenons :

$$C_{q_1}(q_1, \theta_1) = 1 - \frac{(1 - \pi)}{\pi} \{C_{q_1}(q_1, \theta_1) - C_{q_1}(q_1, \theta_2)\}$$

Si $q_0 < q_2$ alors γ est nul. Soit $C_{q_2}(q_2, \theta_2) = \eta$.

Si $\gamma \neq 0$ alors $q_0 = q_2$. Nous obtenons que $C_{q_2}(q_0, \theta_2) > \eta$ c'est-à-dire que η est faible.

Soit $\underline{\eta}$ défini tel que $C_{q_2}(q_0, \theta_2) = \underline{\eta}$.

Cet équilibre séparateur est de troisième rang car $q_1^{**Qh} = q_1^{**}$ et $q_2^{**Qh} < q_2^*$.

Remarque 9 $\frac{\partial q_2^{**Qh}}{\partial \eta} > 0$

La qualité q_2 est une fonction croissante de η . Le gain de l'individu peu productif opportuniste croît avec le nombre de médecins honnêtes. Quand $\eta = \underline{\eta}$, son gain est négatif car :

$$\{C(q_2, \theta_2) + C(q_1, \theta_1) - C(q_1, \theta_2)\} - C(q_2, \theta_1) < 0$$

et lorsque η tend vers 1, nous pouvons montrer que pour certaines valeurs de q_0 , son gain est positif. En effet, si q_0 tend vers q_1 , son gain est positif car :

$$\{C(q_2, \theta_2) + C(q_0, \theta_1) - C(q_0, \theta_2)\} - C(q_0, \theta_1) > 0$$

Nous pouvons définir $\bar{\eta} > \underline{\eta}$ tel que pour certaines valeur de q_0 (tel que $q_1 < q_0 < q_2$) :

$$\{C(q_2(\bar{\eta}), \theta_2) + C(q_1, \theta_1) - C(q_1, \theta_2)\} - C(q_0 + \varepsilon, \theta_1) = 0$$

Pour tout $\eta > \bar{\eta}$, l'individu peu productif opportuniste ne révèle pas son type, il se déclare productif.

Contrats semi-séparateurs La tutelle maximise le programme suivant :

$$\max_{q_1, t_1, q_2, t_2} \pi \eta (q_1 - t_1) + (1 - \pi) \eta \{q_2 - t_2\} + (1 - \pi) \{q_0 - t_2\}$$

sous les contraintes

$$t_1 - C(q_1, \theta_1) \geq 0 \quad IRH1$$

$$t_2 - C(q_2, \theta_2) \geq 0 \quad IRH2$$

$$t_1 - C(q_1, \theta_1) \geq t_2 - C(q_2, \theta_1) \quad ICH1$$

$$t_2 - C(q_2, \theta_2) \geq t_1 - C(q_1, \theta_2) \quad ICH2$$

$$q_0 \leq q_2 \quad IP$$

Soit γ le multiplicateur de Lagrange associé à la contrainte IP .

A l'optimum, les contraintes ($IRH1$ et $ICH2$) sont saturées. Les conditions du premier ordre sont :

$$\pi \eta (1 - C_{q_1}(q_1, \theta_1)) - (1 - \pi) \eta \{C_{q_1}(q_1, \theta_1) - C_{q_1}(q_1, \theta_2)\} = 0$$

$$(1 - \pi) \eta - (1 - \pi) \eta C_{q_2}(q_2, \theta_2) - \gamma \geq 0$$

Nous obtenons :

$$C_{q_1}(q_1, \theta_1) = 1 - \frac{(1 - \pi\eta)}{\pi\eta} \{C_{q_1}(q_1, \theta_1) - C_{q_1}(q_1, \theta_2)\}$$

Si $q_0 < q_2$ alors γ est nul. Soit

$$C_{q_2}(q_2, \theta_2) = \frac{(1 - \pi)\eta}{(1 - \pi\eta)}$$

q_2 est supérieur à q_0 si $\eta > \tilde{\eta}^{26}$.

Si $\eta \leq \tilde{\eta}$ alors, $q_0 = q_2$.

Toutefois, si $\tilde{\eta} < \bar{\eta}$ alors $q_0 < q_2$ ce qui implique que la contrainte *IP* n'est jamais saturée.

Cet équilibre semi-séparateur est de troisième rang car $q_1^{**Qh} < q_1^{**}$ et $q_2^{**Qh} < q_2^*$.

Proposition 10 si $q_1^{**Qh} < q_0 < q_2^{**Qh}$ l'équilibre de troisième rang sera

- séparateur sans fraude si $\eta < \bar{\eta}$: les médecins productifs fournissent q_2^{**Qh} et les peu productifs q_1^{**Qh} .
- semi-séparateur avec fraude autrement : les médecins opportunistes choisissent le contrat des médecins honnêtes productifs et produisent q_0 .

Dans les deux cas, la qualité demandée à l'agent productif est inférieure à celle de premier rang en raison du choix du contrat par les agents opportunistes. La qualité basse est toujours inférieure à celle de premier rang afin de diminuer la rente versée aux agents peu productifs opportunistes pour qu'ils soient désincités à ne pas révéler leur type. Toutefois, quand un grand nombre d'agents est honnête, la qualité demandée aux médecins productifs n'est pas observable par la tutelle.

3.5 Conclusion

Pour limiter, voire proscrire les pratiques opportunistes des médecins, nous proposons que les honoraires du secteur 2 ne soient plus fixés librement, mais déterminés par la tutelle (le paiement du médecin reste un paiement à l'acte). Nous avons considéré que les caractéristiques des médecins sont de deux types et inobservables par la tutelle (la productivité et l'honnêteté). Il semble raisonnable de postuler que les médecins ne sont pas tous mus par leurs intérêts personnels.

Dans ce cadre, les caractéristiques des contrats dépendent de la valeur de l'information de la tutelle sur la qualité ainsi que du nombre de médecins honnêtes. Si la tutelle ne peut définir la qualité fournie par les médecins, le contrat sera mélangeant si les agents opportunistes sont en grande proportion. Cependant, si le nombre de médecins honnêtes est important, le contrat sera semi-séparateur puisque les médecins peu productifs et opportunistes ne révéleront jamais leur

²⁶ $\tilde{\eta}$ est défini par $C_{q_2}(q_0, \theta_2) = \frac{(1 - \pi)\eta}{(1 - \pi\eta)}$.

type. Les qualités fournies sont inférieures à celle du second rang quand la tutelle n'observe pas la productivité. La qualité la plus élevée diminue car les agents opportunistes font augmenter le coût de production unitaire de la haute qualité. La qualité la plus basse décroît pour diminuer l'incitation des agents peu productifs et opportunistes à ne pas révéler leur type.

Si la tutelle ne peut appréhender que la basse qualité, les caractéristiques du contrat dépendront du nombre de médecins opportunistes. Si celui-ci est faible, le contrat sera séparateur. Les agents, selon leur productivité, ne fourniront pas la même qualité. Toutefois, quand le nombre de médecins honnêtes augmente, le contrat de haute qualité devient plus rentable pour le médecin peu productif opportuniste car la qualité contractée s'éloigne de la qualité observable. C'est pour cela qu'ils ne révèlent pas leur type quand le nombre de médecins honnêtes est élevé. La hausse du bien-être générée par l'augmentation de la qualité fournie par les médecins honnêtes compense la perte de surplus liée aux pratiques opportunistes.

Cette modélisation qui propose une régulation d'un bien de confiance gagnerait à être améliorée. Il faudrait étudier un cas continu pour la productivité des médecins. Intuitivement, le médecin le moins productif ne recevra pas de rente a contrario des autres médecins. Pour un nombre de médecins honnêtes donné, en fonction de leur productivité et de la valeur de l'information ayant trait à la qualité, certains médecins opportunistes productifs ne révéleront pas leur type a contrario de médecins moins productifs. Un second prolongement serait de supposer que la vérification de l'information est coûteuse²⁷. La valeur du standard de qualité serait déterminée par un calcul économique. Elle dépendrait de la technologie de contrôle.

²⁷Nous serions ramenés à un bien de recherche.

Chapitre 4

Une analyse du risque sanitaire

“La santé est un état en creux qui échappe au discours de l’homme”, Gadamer (1998).

4.1 Introduction

Les premiers travaux traitant de la demande d’assurance ont considéré que la décision d’un consommateur de contracter une police d’assurance dépendait d’un unique risque financier assurable. Cette hypothèse a été récusée par Doherty et Schlesinger (1983a, b)¹ qui ont introduit un second risque financier non assurable dans un modèle usuel d’assurance². La décision d’assurance n’est pas isolée. Elle se prend dans des situations d’incertitude pour lesquelles il n’existe pas toujours un moyen de couverture. Toutefois, ces résultats n’ont pu s’appréhender qu’en utilisant la notion de prudence énoncée par Kimball (1990) et en supposant que la prudence était une fonction décroissante de la richesse des individus. Eeckhoudt et Kimball (1991)³ ont caractérisé l’effet du *background risk* sur le niveau de la couverture quand les risques financiers sont dépendants ou indépendants⁴ et quand l’aversion absolue au risque est une fonction décroissante de la richesse.

En économie de la santé, le risque n’est pas exclusivement monétaire. Nous utilisons une fonction d’utilité à deux variables : une monétaire (la richesse) et une sanitaire (le capital santé). Des travaux précédents ont considéré des risques multiples, ou des risques endogènes ou encore des risques multivariés (Briys, Eeckhoudt et Loubergé, 1989; Chalfant et Finkelshtain, 1993; Demers et De-

¹ Ainsi que par Mayers et Smith (1983), Turnbull (1983) et Briys, Kahane et Kroll (1988).

² Ils n’ont discuté que de l’optimalité de la pleine assurance en présence de plusieurs risques.

³ Ils sont les premiers à examiner l’effet du Background risk sur le niveau optimal de la couverture.

⁴ Seul l’un des deux est assurable.

mers, 1991; Doherty, Loubergé et Schleinsinger, 1987; Duncan, 1997; Karni, 1979; Kihlstrom et Mirman, 1974; Levy et Levy, 1991; Pratt, 1988; Pratt et Zeckhauser, 1987; Kihlstrom, Romer et Williams, 1981; Ross, 1981). La littérature traitant du risque sanitaire est peu développée. Seul Courbage (1999) a appliqué l'analyse d'un risque financier exogène, unique et univarié à un risque sanitaire.

L'objectif de ce chapitre est d'appliquer l'analyse des risques bivariés (Courbage, 2001; Demers et Demers, 1991; Finkelshtain et Chalfant, 1993; Karni, 1979; Pratt, 1988). Nous définissons donc la prime de risque d'un aléa sanitaire assurable et non assurable. Puis nous généralisons la notion de prudence d'un risque financier à un risque sanitaire pour étudier l'effet d'une variation de l'état de santé sur la prime de risque.

Le chapitre est organisé comme suit. Dans la section 2, nous proposons tout d'abord une définition de la prime de risque d'un aléa sanitaire assurable et non assurable. Nous montrons que la définition de la prime de risque d'un aléa sanitaire assurable de Courbage (1999) ne vérifie pas l'équivalence de Pratt (1964). La prime de risque définie par Arrow (1965) et Pratt (1964) permet d'évaluer le coût d'un risque financier auquel est soumis un individu. Elle définit le montant que l'individu est prêt à payer pour se départir de ce risque financier. Elle est égale à la différence entre l'espérance de gain et l'équivalent certain. Pratt (1964) a énoncé un théorème prouvant l'équivalence entre la positivité de la prime de risque et l'aversion au risque de l'individu. Nous appliquons cette définition à plusieurs situations sanitaires.

Dans la section 3, nous généralisons la notion de prudence de Kimball (1990) à un risque sanitaire assurable et non assurable⁵. Nous montrons, sous certaines conditions, que la prime de risque d'un risque sanitaire est une fonction décroissante de l'état de santé. Dans un premier temps, nous présentons brièvement le concept de prudence mis en exergue par Kimball (1990). Dans un second temps, nous adaptons cette notion à un risque sanitaire assurable. La prime de prudence est le prix que l'individu est prêt à payer pour obtenir une unité de capital sanitaire supplémentaire (Eeckhoudt, 2002). Enfin, pour un risque sanitaire non assurable bénin puis malin⁶, nous étudions le sens de variation de la prime de

⁵Dans un cadre unidimensionnel, Gollier et Pratt (1996) ont étudié l'effet d'un *background risk* sur des autres risques indépendants, et Eeckhoudt, Gollier et Schlesinger (1996) un *background risk* pour des risques dépendants. L'addition d'un second risque financier de moyenne nulle augmente-t-elle l'aversion au risque par rapport aux autres risques financiers ?

Les premiers ont montré que la vulnérabilité au risque est une condition nécessaire et suffisante pour que l'addition d'un autre risque financier indépendant (dont la condition locale est la suivante : $r'' > r'r$, avec r l'aversion au risque) augmente l'aversion au risque des individus par rapport aux autres risques. La vulnérabilité au risque implique que $u''' > 0$ et $u'''' > 0$. Les seconds ont montré que l'ajout d'un risque financier dépendant augmentait l'aversion au risque des individus par rapport aux autres risques si l'aversion au risque au sens de Ross (1981) est une fonction décroissante de la richesse.

Dans un cadre bivarié, nous cherchons à déterminer comment varie la prime de risque financière quand varie l'état de santé. Notre approche est analogue à celle d'une analyse *background risk* appliquée à un risque bivariée. Un travail de recherche ultérieur serait d'appliquer l'analyse de vulnérabilité d'un risque à un cadre bivarié.

⁶Nous définissons un risque malin comme un risque qui, après traitement, détériore le

risque en fonction de l'état de santé de l'individu.

4.2 Primes de risque

Soit un individu dont les préférences sont représentées par une fonction d'utilité *vNM*, de la forme $u(w, H)$ avec w le niveau de richesse et H l'état de santé assimilé à un "stock de santé". Cette fonction est strictement croissante par rapport aux deux arguments.

Cet individu est doté d'une richesse initiale w_0 et d'un état de santé initial H_2 , l'état de bonne santé. Il est soumis à un risque sanitaire, c'est-à-dire qu'avec une probabilité p_0 son stock de santé se déprécie de $(H_2 - H_0) > 0$ reflétant la gravité de la maladie⁷.

4.2.1 Définition de la prime de risque sanitaire

Pour proposer une définition de la prime de risque d'un aléa sanitaire, nous appliquons la définition d'une prime de risque bivariée partielle (Finkelshtain et Chalfant, 1993) et modifions celle de Courbage (1999)⁸. La prime d'un risque sanitaire assurable, notée r_H , est le montant monétaire que l'individu serait

capital sanitaire contrairement à un risque bénin.

⁷Nous supposons tout au long de ce chapitre que l'individu est confronté à un petit risque santé. Cette hypothèse permet l'utilisation d'un développement limité de Taylor.

⁸Nous n'appliquons pas la définition de Courbage (1999) car elle ne corrobore pas l'équivalence d'Arrow-Pratt.

La définition de la prime de risque d'un aléa sanitaire retenue par Courbage (1999) est la suivante :

"Nous pouvons définir la prime de risque R comme le montant que serait prêt à payer l'agent pour se débarrasser du risque santé et obtenir avec certitude l'état de bonne santé... Cependant, face à un risque de santé, ce n'est pas le niveau de santé moyen que l'individu cherche à atteindre, mais le niveau de bonne santé."

Formellement, la prime R est donc définie par la relation :

$$p_0 u(w_0, H_0) + (1 - p_0) u(w_0, H_2) = u(w_0 - R, H_2)$$

Celle-ci se réécrit :

$$p_0 \{u(w_0, H_2) - u(w_0, H_0)\} = u(w_0, H_2) - u(w_0 - R, H_2)$$

Nous déduisons de cette relation que, quelle que soit l'aversion au risque de l'individu, la prime R est positive car la fonction d'utilité est croissante. En effet, $H_2 - H_0 >$ implique que $R > 0$. R est le prix que l'individu est prêt à payer pour obtenir avec certitude l'état de bonne santé et non pas une prime de risque correspondant à la définition d'Arrow-Pratt. Tout individu est prêt à compenser financièrement la dépréciation de son stock de santé, c'est-à-dire payer R définie ci-dessus puisque nous supposons que les deux biens sont substituables. Cette définition du prix de la bonne santé est intéressante car elle permet de comparer différentes situations sanitaires entre elles puisque les individus ont tous, après le risque sanitaire, le même état de bonne santé (Courbage, 1999).

prêt à payer pour obtenir avec certitude l'état de santé moyen⁹ $E\tilde{H}$ avec $E\tilde{H} = p_0H_0 + (1 - p_0)H_2$. Formellement, nous obtenons :

$$p_0u(w_0, H_0) + (1 - p_0)u(w_0, H_2) = u(w_0 - r_H, E\tilde{H})$$

En utilisant un développement limité de Taylor autour du point $(w_0, E\tilde{H})$, la relation précédente conduit à l'approximation suivante :

$$\begin{aligned} -r_H u_1 &\simeq p_0(H_0 - E\tilde{H})u_2 + \frac{p_0}{2}(H_0 - E\tilde{H})^2 u_{22} \\ &\quad + (1 - p_0)(H_2 - E\tilde{H})u_2 + \frac{(1 - p_0)}{2}(H_2 - E\tilde{H})^2 u_{22} \end{aligned}$$

où u_1 et u_2 représentent respectivement les dérivées partielles premières par rapport au premier et au second argument et u_{22} la dérivée partielle seconde par rapport au stock de santé.

Nous posons $Var(H) = p_0(H_0 - E\tilde{H})^2 + (1 - p_0)(H_2 - E\tilde{H})^2$. Soit :

$$r_H \simeq -\frac{Var(H)}{2} \frac{u_{22}(w_0, E\tilde{H})}{u_1(w_0, E\tilde{H})}$$

Définition 11 La prime de risque d'un aléa sanitaire s'écrit :

$$r_H \simeq -\frac{V(H)}{2} \frac{u_{22}(w_0, E\tilde{H})}{u_2(w_0, E\tilde{H})} \frac{u_2(w_0, E\tilde{H})}{u_1(w_0, E\tilde{H})}$$

Cette prime de risque dépend de la variance du risque santé ($Var(H)$) multipliée par l'indice d'aversion absolue au risque sanitaire ($A(H)$) et par le taux marginal de substitution entre richesse et santé ($\frac{u_2}{u_1}$).

Cette définition de la prime d'un risque sanitaire vérifie bien l'équivalence de Arrow-Pratt en ce sens que l'aversion au risque sanitaire ($u_{22} < 0$) implique la positivité de r_H contrairement à la négativité de r_H si l'individu aime le risque ($u_{22} > 0$).

4.2.2 Primes de risques et soins sanitaires

Dans ce paragraphe, nous définissons les primes de risque pour différentes situations sanitaires¹⁰ : l'individu a recours aux soins curatifs, aux soins préventifs secondaires et primaires. Nous supposons que les soins sanitaires sont efficaces et leurs conséquences déterministes, c'est-à-dire que l'individu obtient une meilleure utilité en recourant à ces soins.

Primes de risque et soins curatifs

Pour limiter la dépréciation de son stock de santé, l'individu a recours à des soins curatifs, ce qui affecte sa richesse. En cas de maladie, l'individu dépensera

⁹Nous supposons qu'il existe une technologie médicale qui permet de lisser l'état de santé de l'individu. Cette hypothèse sera levée dans la section suivante.

¹⁰Nous recommençons l'analyse de Courbage (1999) avec une définition de la prime de risque qui corrobore l'équivalence de Pratt (1964).

sy en soins de maladie où s représente le prix unitaire du traitement sanitaire et y le nombre d'unités de traitement. Une dépense sy augmente de $m(y)$ le niveau de santé¹¹. Comme un individu malade et traité ne peut être sanitaire mieux qu'un individu en bonne santé on a nécessairement $H_1 \leq H_2$ ¹² avec $H_1 = H_0 + m(y)$. L'individu détermine y en maximisant son espérance d'utilité. Soit y_1^* cette solution.

Le recours aux soins curatifs diminue en termes de variance le risque de santé mais crée un risque de richesse. L'individu va devoir arbitrer entre sa santé et sa richesse.

Nous pouvons donc définir une nouvelle prime de risque conditionnelle aux dépenses optimales de soins de santé curatifs, notée r_1 , comme le montant que l'individu est prêt à payer pour obtenir avec certitude l'état de santé moyen $E\tilde{H}_1$ avec $E\tilde{H}_1 = p_0(H_1) + (1 - p_0)H_2$. Soit,

$$p_0u(w_0 - sy_1^*, H_1) + (1 - p_0)u(w_0, H_2) = u(w_0 - r_1, E\tilde{H}_1)$$

En utilisant un développement limité de Taylor autour du point $(w_0 - p_0sy_1^*, E\tilde{H}_1)$, la relation précédente conduit à l'approximation :

$$\begin{aligned} r_1u_1 &\simeq p_0sy_1^*u_1 - \frac{1}{2}p_0(1 - p_0)(H_2 - H_1)^2u_{22} \\ &\quad - \frac{1}{2}p_0(1 - p_0)(sy_1^*)^2u_{12} - p_0(1 - p_0)sy_1^*(H_2 - H_1)u_{11} \end{aligned}$$

Nous savons que $Var(H) = p_0(1 - p_0)(H_2 - H_1)^2$, $Var(w) = p_0(1 - p_0)(sy_1^*)^2$ et $Cov(H, w) = p_0(1 - p_0)sy_1^*(H_2 - H_1)$

Définition 12 La prime de risque d'un aléa sanitaire conditionnelle aux dépenses de soins de santé curatifs s'écrit :

$$r_1 \simeq p_0sy_1^* - \frac{1}{2} \sum_{i=1}^2 \sum_{j=1}^2 \sigma_{ij} \frac{u_{ij}(w_0 - p_0sy_1^*, E\tilde{H}_1)}{u_1(w_0 - p_0sy_1^*, E\tilde{H}_1)}$$

avec $[\sigma_{ij}]$ la matrice de variance-covariance des risques de richesse et de santé.

La prime r_1 dépend de la mesure matricielle d'aversion au risque multivariée $H_R = \begin{bmatrix} -u_{11} \\ -u_{12} \\ u_1 \end{bmatrix}$ définie par Duncan (1977) et de la perte de richesse provoquée par le recours aux soins $p_0sy_1^*$.

Ainsi r_1 supérieur à $p_0sy_1^*$ implique que H_R est semi-définie positive.

H_R est semi-définie positive si $-u_{11} \geq 0$ et si $u_{11}u_{22} - (u_{12})^2 \geq 0$. Une condition nécessaire pour que cette expression soit positive est que $u_{22} < 0$. Nous en déduisons que H_R est semi-définie positive si $-u_{ii} > 0$. Ainsi, la prime de risque d'un aléa sanitaire est positive si les indices d'aversion absolue aux risques santé et financier sont positifs (car $\frac{1}{u_i} > 0$)¹³.

¹¹ $m(y)$ est une fonction non décroissante et concave.

¹² Si avec traitement, l'individu malade retrouvait l'état de bonne santé, il ne ferait face qu'à un risque financier pur.

¹³ H_R est semi-définie négative si $u_{ii} > 0$.

Si u_{11} et u_{22} sont de signes opposés, H_R ne peut être définie.

En effet, les soins étant efficaces chaque individu est prêt, quelle que soit son aversion au risque, à payer en espérance $p_0sy_1^*$. Pour un individu risquophile, sa prime d'assurance est inférieure à $p_0sy_1^*$ car il ne veut s'assurer ni contre le risque financier lié au paiement des soins, ni contre le risque sanitaire. L'équivalence de Pratt est donc vérifiée : si l'individu est risquophile (respectivement risquophile) aux deux risques alors $r_1 - p_0sy_1^* > (<)0$.

Primes de risque et soins préventifs secondaires

En médecine, deux types de soins préventifs sont à distinguer. Les soins primaires aboutent la réduction de la probabilité d'occurrence du dommage. Les soins préventifs secondaires visent à diminuer la gravité de la maladie.

L'individu engage des soins préventifs secondaires à chaque état de la nature notés sy_2^* avec y_2^* la quantité optimale de soins préventifs secondaires qui maximise l'espérance d'utilité de l'agent.

Nous pouvons donc définir une prime de risque conditionnelle aux dépenses optimales de soins préventifs secondaires sy_2^* , notée r_2 , comme le montant que l'individu est prêt à payer pour obtenir avec certitude l'état de santé moyen $E\tilde{H}_2$ avec $E\tilde{H}_2 = p_0H_1' + (1 - p_0)H_2$ avec $H_1' = H_0 + m(y_2^*)$. Soit,

$$p_0u(w_0 - sy_2^*, H_1') + (1 - p_0)u(w_0 - sy_2^*, H_2) = u(w_0 - r_2, E\tilde{H}_2)$$

En utilisant un développement limité de Taylor autour du point $(w_0 - sy_2^*, E\tilde{H}_2)$, nous obtenons :

$$r_2 \simeq sy_2^* - \frac{Var(H)}{2} \frac{u_{22}(w_0 - sy_2^*, E\tilde{H}_2)}{u_1(w_0 - sy_2^*, E\tilde{H}_2)}$$

Définition 13 La prime de risque conditionnelle aux dépenses de soins préventifs secondaires s'écrit :

$$r_2 \simeq sy_2^* - \frac{Var(H)}{2} \frac{u_{22}(w_0 - sy_2^*, E\tilde{H}_2)}{u_2(w_0 - sy_2^*, E\tilde{H}_2)} \frac{u_2(w_0 - sy_2^*, E\tilde{H}_2)}{u_1(w_0 - sy_2^*, E\tilde{H}_2)}$$

La prime de risque r_2 dépend de la perte de richesse liée aux soins préventifs secondaires, de la variance du risque santé multipliée par l'indice d'aversion absolue au risque de santé et par le taux marginal de substitution entre richesse et santé. Elle est indépendante de l'aversion au risque de richesse car ce dernier est nul.

Ainsi r_2 supérieur (respectivement inférieur) à sy_2^* implique que l'individu est risquophile (risquophile) au risque sanitaire. En effet, les soins étant efficaces chaque individu et ce, quelle que soit son aversion au risque, paye py_2^* . Pour un individu risquophile ($u_{22} < 0$), la prime d'assurance est supérieure à py_2^* car il veut s'assurer contre le risque sanitaire. L'équivalence de Pratt est donc vérifiée : si l'individu est risquophile (respectivement risquophile) au risque sanitaire alors $\pi_2 - py_2^* > (<)0$.

Primes de risque et soins préventifs primaires

Nous supposons que l'individu peut effectuer un effort de prévention primaire d'un coût c qui réduit la probabilité d'occurrence du risque de $(p_0 - p_1)$ et que cet effort préventif est efficace.

Nous pouvons donc définir une prime de risque conditionnelle aux dépenses de soins préventifs primaires, notée r_3 , comme le montant que l'individu est prêt à payer pour obtenir avec certitude l'état de santé moyen $E\tilde{H}_3$ avec $E\tilde{H}_3 = p_1 H_0 + (1 - p_1) H_2$. Soit,

$$p_1 u(w_0 - c, H_0) + (1 - p_1) u(w_0 - c, H_2) = u(w_0 - r_3, E\tilde{H}_3)$$

En utilisant un développement limité de Taylor autour du point $(w_0 - c, E\tilde{H}_3)$, nous obtenons :

$$r_3 \simeq c - \frac{\text{Var}(H)}{2} \frac{u_{22}(w_0 - c, E\tilde{H}_3)}{u_1(w_0 - c, E\tilde{H}_3)}$$

Définition 14 La prime de risque conditionnelle aux dépenses de soins préventifs primaires s'écrit :

$$r_3 \simeq c - \frac{\text{Var}(H)}{2} \frac{u_{22}(w_0 - c, E\tilde{H}_3)}{u_2(w_0 - c, E\tilde{H}_3)} \frac{u_2(w_0 - c, E\tilde{H}_3)}{u_1(w_0 - c, E\tilde{H}_3)}$$

Nous retrouvons les mêmes résultats que dans le cas de soins préventifs secondaires. Ainsi, l'équivalence de Pratt est corroborée, car pour un individu risquophobe, $r_3 - c > 0$.

Dans cette section, nous avons appliqué le concept d'une prime de risque bivariée partielle au cas où l'individu est confronté à un risque sanitaire. L'équivalence de Pratt (1964) entre l'aversion au risque et le signe de la prime de risque reste vérifiée. La prime de risque est positive quand l'individu est averse à un risque sanitaire assurable.

4.3 Prudence

Dans cette section, nous étudions l'effet de la variation du capital sanitaire ou financier sur le montant de la prime de risque que l'individu est prêt à payer pour un aléa sanitaire assurable ou non assurable¹⁴.

¹⁴La plupart des travaux considèrent deux risques financiers indépendants dont l'un d'entre eux est non-assurable. Gollier (1996) étudie dans quelle mesure la présence d'un risque non assurable modifie la demande d'un risque assurable. Gollier et Pratt (1996) déterminent une condition nécessaire et suffisante pour laquelle un risque inassurable d'espérance non positive réduit le taux de rétention pour tout autre risque indépendant.

Dans le cas de deux risques assurables séparément, Gollier et Schlesinger (1995) montrent qu'un contrat universel, avec une franchise unique sur le risque agrégé, est meilleur qu'un ensemble de polices d'assurance propre à chaque risque.

Comme le soulignent Gollier et Haritchabalet (2000), peu de résultats théoriques sont à notre disposition pour décrire "l'interaction entre les demandes d'assurance d'un assuré faisant face à plusieurs sources indépendantes de risques non assurables". Seuls Eeckhoudt, Gollier et Levasseur (1994) étudient ce problème dans un modèle d'espérance d'utilité.

4.3.1 Définition

Un individu possède un capital financier w et fait face à un unique risque financier : une perte d avec une probabilité p_0 . La prime de risque noté r_f est définie par :

$$u(w - p_0d - r_f) = p_0u(w - d) + (1 - p_0)u(w)$$

La fonction u est croissante et deux fois dérivable. En utilisant un développement limité de Taylor (Pratt, 1964) autour du point $(w - p_0d)$, nous obtenons que :

$$r_f \simeq \frac{Var(w)}{2} A(w)$$

avec $V(w)$ la variance de la richesse et $A(w)$ l'aversion absolue au risque financier.

Bien que cette notion soit très usitée, elle ne permet pas de caractériser l'effet du risque financier sur l'utilité marginale de la richesse financière (Eeckhoudt et Kimball, 1991). Il est nécessaire de s'intéresser, lorsque plusieurs risques sont en jeu, à la prime de prudence proposée par Kimball (1990) dans un article consacré aux choix d'épargne, appliquée au problème de risques financiers multiples par Eeckhoudt et Kimball (1991).

La prime de prudence¹⁵ ψ_f permet de définir la compensation monétaire dont un individu risquophobe est prêt à se départir afin qu'il ne change pas ses décisions lorsque son environnement se modifie (Eeckhoudt et Gollier, 1992). Formellement, elle s'énonce :

$$u'(w - p_0d - \psi_f) = p_0u'(w - d) + (1 - p_0)u'(w)$$

En utilisant un développement limité autour du point $(w - p_0d)$, nous obtenons :

$$\psi_f \simeq \frac{Var(w)}{2} \eta(w)$$

avec $\eta(w)$ la prudence absolue au risque financier (par rapport à la richesse) définie par Kimball (1990) :

$$\eta(w) = -\frac{u'''(w - p_0d)}{u''(w - p_0d)}$$

Cette notion représente la propension à se prémunir contre le risque.

Pour illustrer cette notion, nous considérons un modèle d'épargne avec incertitude. La prudence absolue doit être positive pour que l'augmentation de

¹⁵Dans le cadre de l'espérance d'utilité, la notion de prime de prudence est interchangeable avec celle de prime de précaution. Cependant, dans le cadre de la théorie duale, ces deux primes sont différentes (Demers et Demers, 1997). La prime de précaution est la somme de la prime de risque et d'un terme qui représente la prime de prudence.

l'incertitude (diminution de la richesse en espérance) augmente l'épargne de prudence. Dans un cadre d'assurance, la prudence absolue au risque doit être positive et l'aversion au risque doit être une fonction décroissante de la richesse pour que la demande d'assurance s'accroisse lorsque la richesse financière diminue. Eeckhoudt et Kimball (1991) considèrent ces hypothèses comme satisfaisantes¹⁶. Nous posons donc que $\psi_f > 0$ et $u''' > 0$.

Pour conclure, la prime de risque est une fonction décroissante de la richesse si et seulement si l'aversion au risque est une fonction décroissante de la richesse¹⁷.

Pour caractériser le sens de variation de la prime de risque par rapport à la richesse, nous dérivons par rapport à la richesse l'équation qui définit la prime de risque. Soit,

$$\left(1 - \frac{\partial r_f}{\partial w}\right)u'(w - p_0d - r_f) = p_0u'(w - d) + (1 - p_0)u'(w)$$

Cette équation se réécrit en utilisant la définition de la prime de prudence :

$$\left(1 - \frac{\partial r_f}{\partial w}\right)u'(w - p_0d - r_f) = u'(w - p_0d - \psi_f)$$

Nous avons supposé la décroissance de l'aversion absolue au risque ce qui implique, en utilisant les définitions des primes de risque et de prudence, que

$$A(w) < \eta(w) \Rightarrow r_f < \psi_f$$

En raison de la concavité de la fonction d'utilité, nous déduisons que

$$u'(w - p_0d - r_f) < u'(w - p_0d - \psi_f)$$

Soit $\left(1 - \frac{\partial r_f}{\partial w}\right) > 1$. Nous en concluons que $\frac{\partial r_f}{\partial w} < 0$. Nous pouvons énoncer le résultat suivant :

Proposition 15 (Kimball, 1990)

$$\frac{\partial A(w)}{\partial w} < 0 \text{ implique que } \frac{\partial r_f}{\partial w} < 0$$

Si l'aversion absolue au risque financier est une fonction décroissante de la richesse alors la prime de risque décroît.

Nous supposons dans la suite de ce chapitre que l'aversion absolue au risque financier (sanitaire) est une fonction décroissante de la richesse (de la santé).

¹⁶Eeckhoudt et Kimball (1992) et Kimball (1993) ont obtenu une condition suffisante pour définir "l'aversion au risque standard". L'aversion au risque est dite standard si l'aversion absolue au risque financier et la prudence absolue au risque financier sont des fonctions décroissantes de la richesse financière.

¹⁷La décroissance de l'aversion au risque implique que : $A(w) < \eta(w)$.

4.3.2 Risque sanitaire assurable

Nous étudions le sens de variation de la prime de risque en fonction de l'état de santé, puis de la richesse. Intuitivement, nous devons retrouver les résultats usuels d'un risque financier. Un individu en bonne santé ou plus riche sera prêt à verser une prime de risque inférieure à celui en mauvaise santé.

Pour ce faire, nous réécrivons la définition de la prime de risque d'un risque sanitaire assurable. La prime de risque est définie par :

$$r_H = \frac{Var(H)}{2} A(H) \frac{u_2(w_0, E\tilde{H})}{u_1(w_0, E\tilde{H})}$$

Nous posons que $H_0 = H - \Delta H$ et $H_2 = H$

Prudence sanitaire par rapport à l'état de santé

Par analogie avec un risque financier, nous pouvons définir la prime de prudence sanitaire par rapport à la santé :

$$u_2(w - \psi_H, E\tilde{H}) = p_0 u_2(w, H - \Delta H) + (1 - p_0) u_2(w, H)$$

où u_2 est l'utilité marginale par rapport au stock sanitaire. Pour un risque sanitaire, la prudence mesure l'effet de ce risque sur l'utilité marginale par rapport à la santé.

En utilisant un développement limité autour du point $(w, E\tilde{H})$, nous obtenons que :

$$\begin{aligned} -\psi_H u_{12} &\simeq p_0 (H - \Delta H - E\tilde{H}) u_{22} + \frac{p_0}{2} (H - \Delta H - E\tilde{H})^2 u_{222} \\ &\quad (1 - p_0) (H - E\tilde{H}) u_{22} + \frac{(1 - p_0)}{2} (H - E\tilde{H})^2 u_{222} \end{aligned}$$

Soit,

$$\psi_H \simeq - \frac{Var(H)}{2} \frac{u_{222}(w_0, E\tilde{H})}{u_{12}(w_0, E\tilde{H})}$$

Définition 16 La prime de prudence sanitaire par rapport à la santé s'énonce¹⁸ :

$$\psi_H \simeq \frac{Var(H)}{2} \eta(H) \frac{u_{22}(w_0, E\tilde{H})}{u_{12}(w_0, E\tilde{H})}$$

$$\text{avec } \eta(H) = - \frac{u_{222}(w_0, E\tilde{H})}{u_{22}(w_0, E\tilde{H})}$$

¹⁸Kimball (1990) montre pour un problème intertemporel de consommation-épargne que la prudence absolue s'énonce pour une fonction d'utilité non séparable :

$$\eta(w, c) = \frac{U_{122}(c, w - c) - U_{222}(c, w - c)}{U_{12}(c, w - c) - U_{22}(c, w - c)}$$

La prime de prudence sanitaire dépend de la variance de l'état de santé multipliée par l'indice de prudence absolue au risque sanitaire par rapport à la santé¹⁹ notée $\eta(H)$ et le taux marginal de substitution de l'utilité marginale par rapport à la santé entre richesse et santé.

Nous posons que $u_{12}(\cdot, \cdot) < 0$ ²⁰. Nous supposons donc que l'utilité marginale de la richesse est une fonction décroissante du stock de santé²¹.

Ainsi, la prime de prudence est positive si $u_{222}(\cdot, \cdot) > 0$. Nous retrouvons ici un résultat de Kimball (1990) transposé à un risque sanitaire assurable.

Proposition 17 *Si $-\frac{u_{22}(w_0, E\tilde{H})}{u_1(w_0, E\tilde{H})}$ est une fonction décroissante de H alors $\frac{\partial r_H}{\partial H} < 0$*

Preuve :
 $\frac{\partial \left(-\frac{u_{22}(w_0, E\tilde{H})}{u_1(w_0, E\tilde{H})} \right)}{\partial H} < 0$ implique que $-\frac{u_{222}(w_0, E\tilde{H})}{u_{12}(w_0, E\tilde{H})} > -\frac{u_{22}(w_0, E\tilde{H})}{u_1(w_0, E\tilde{H})}$. En multipliant ces deux termes par la variance du risque sanitaire, nous obtenons que $\psi_H > r_H$.

Différencions maintenant l'équation qui définit la prime de risque par rapport à H . Nous obtenons (en utilisant l'équation qui définit la prime de prudence sanitaire par rapport à la santé) :

$$-u_1(w - r_H, E\tilde{H}) \frac{\partial r_H}{\partial H} + u_2(w - r_H, E\tilde{H}) = u_2(w - \psi_H, E\tilde{H})$$

$\psi_H > r_H$ implique que $u_2(w - r_H, E\tilde{H}) < u_2(w - \psi_H, E\tilde{H})$ car la fonction d'utilité est une fonction concave en H . Nous en déduisons que $-u_1(w - r_H, E\tilde{H}) \frac{\partial r_H}{\partial H} > 0$. Nous en concluons donc que $\frac{\partial r_H}{\partial H} < 0$ car $u_1(\cdot, \cdot) > 0$. CQFD.

La prime de risque d'un aléa sanitaire assurable est une fonction décroissante de la santé si le produit de $\eta(H)$ et du taux marginal de substitution de l'utilité marginale par rapport à la santé entre richesse et santé est une fonction décroissante de la santé.

Prudence sanitaire par rapport à la richesse

Par analogie à un risque financier, nous pouvons définir la prime de prudence sanitaire par rapport à la richesse notée $\psi_{H/w}$:

$$u_1(w - \psi_{H/w}, E\tilde{H}) = p_0 u_1(w, H - \Delta H) + (1 - p_0) u_1(w, H)$$

où u_1 est l'utilité marginale par rapport à la richesse.

¹⁹ La prudence absolue du risque sanitaire par rapport à la richesse s'énoncerait :

$$\eta'_H = -\frac{u_{221}}{u_{21}}$$

²⁰ Courbage (2001) montre que la prime de risque de risques bivariés assurables peut être négative si $u_{12} > 0$.

²¹ Graphiquement, cette hypothèse traduit le fait que pour un même niveau de richesse financière, la pente de la courbe d'utilité dans un repère (richesse, utilité) est supérieure pour un individu en mauvaise santé par rapport à celle de l'individu en bonne santé.

En utilisant un développement limité autour du point $(w, E\tilde{H})$, nous obtenons que :

$$\begin{aligned} -\psi_{H/w} u_{11} &\simeq p_0(H - \Delta H - E\tilde{H})u_{12} + \frac{p_0}{2}(H - \Delta H - E\tilde{H})^2 u_{122} \\ &\quad + (1 - p_0)(H - E\tilde{H})u_{12} + \frac{(1 - p_0)}{2}(H - E\tilde{H})^2 u_{122} \end{aligned}$$

Soit,

$$\psi_{H/w} \simeq -\frac{Var(H)}{2} \frac{u_{122}(w_0, E\tilde{H})}{u_{11}(w_0, E\tilde{H})}$$

Définition 18 La prime de prudence sanitaire par rapport à la richesse est définie par :

$$\psi_{H/w} \simeq \frac{Var(H)}{2} \eta_w(H) \frac{u_{12}(w_0, E\tilde{H})}{u_{11}(w_0, E\tilde{H})}$$

avec $\eta_w(H) = -\frac{u_{122}(w_0, E\tilde{H})}{u_{12}(w_0, E\tilde{H})}$ la prudence absolue du risque sanitaire par rapport à la richesse.

La prime de prudence sanitaire par rapport à la richesse dépend de la variance de l'état de santé multipliée par l'indice de prudence absolue au risque sanitaire par rapport à la richesse notée $\eta_w(H)$ et par le taux marginal de substitution de l'utilité marginale de la richesse à la santé. Ainsi la prime de prudence est positive si $u_{122}(\cdot, \cdot) > 0$. Nous retrouvons ici un résultat de Kimball (1990) transposé à un risque sanitaire assurable.

Proposition 19 Si $-\frac{u_{22}(w_0, E\tilde{H})}{u_{11}(w_0, E\tilde{H})}$ est une fonction décroissante de w alors $\frac{\partial r_H}{\partial w} < 0$

Preuve :

$\frac{\partial \left(-\frac{u_{22}(w_0, E\tilde{H})}{u_{11}(w_0, E\tilde{H})} \right)}{\partial w} < 0$ implique que $-\frac{u_{122}(w_0, E\tilde{H})}{u_{11}(w_0, E\tilde{H})} > -\frac{u_{22}(w_0, E\tilde{H})}{u_{11}(w_0, E\tilde{H})}$. En multipliant ces deux termes par la variance du risque sanitaire, nous obtenons que $\psi_{H/w} > r_H$.

Différencions maintenant l'équation qui définit la prime de risque par rapport à w , nous obtenons que (en utilisant l'équation qui définit la prime de prudence par rapport à la richesse) :

$$\left(1 - \frac{\partial r_H}{\partial w}\right) u_1(w - r_H, E\tilde{H}) = u_1(w - \psi_{H/w}, E\tilde{H})$$

$\psi_{H/w} > r_H$ implique que $u_1(w - r_H, E\tilde{H}) < u_1(w - \psi_{H/w}, E\tilde{H})$ car la fonction d'utilité est une fonction concave en w . Nous en déduisons que $\frac{\partial r_H}{\partial w} < 0$. CQFD.

4.3.3 Risque sanitaire non assurable

Nous allons étudier deux cas : un risque bénin qui ne détériore pas l'état de santé après traitement curatif et un risque malin qui altère l'état de santé même avec traitement curatif²². Nous caractérisons l'effet de l'état de santé sur le montant de la prime de risque dont l'individu est prêt à se départir pour éliminer le risque financier occasionné par le traitement curatif²³.

Risque sanitaire bénin La prime de risque r_{fb} pour un risque bénin est définie par :

$$u(w - p_0sy^* - r_{fb}, H) = p_0u(w - sy^*, H) + (1 - p_0)u(w, H)$$

En utilisant un développement limité autour du point $(w - p_0sy^*, H)$, nous obtenons :

$$r_{fb} \simeq -\frac{Var(w)}{2}A(w)$$

$$\text{avec } A(w) = \frac{u_{11}(w - p_0sy^*, H)}{u_1(w - p_0sy^*, H)}.$$

Nous étudions maintenant quel est l'effet du capital sanitaire sur la prime de risque financière. Avec une fonction à deux variables, deux primes de prudence peuvent être caractérisées :

- ψ_{fb} par rapport à l'utilité marginale de la richesse
- ψ'_{fb} par rapport à l'utilité marginale de la santé.

La première (appelée prime de prudence par rapport à la richesse) permet de caractériser l'effet de la richesse sur la prime de risque monétaire tandis que la seconde (appelée prime de prudence par rapport à la santé) définit l'effet de la santé sur la prime de risque monétaire. ψ_{fb} et ψ'_{fb} sont définies respectivement par les deux relations suivantes :

$$u_1(w - p_0sy^* - \psi_{fb}, H) = p_0u_1(w - sy^*, H) + (1 - p_0)u_1(w, H)$$

$$u_2(w - p_0sy^* - \psi'_{fb}, H) = p_0u_2(w - sy^*, H) + (1 - p_0)u_2(w, H)$$

En utilisant un développement limité autour du point $(w - p_0sy^*, H)$, nous obtenons pour la seconde relation :

$$\begin{aligned} -\psi'_{fb}u_{12} &\simeq p_0(-sy^* + p_0sy^*)u_{12} + (1 - p_0)(p_0sy^*)u_{12} \\ &\quad + \frac{p_0}{2}(-sy^* + p_0sy^*)^2u_{112} + \frac{(1 - p_0)}{2}(p_0sy^*)^2u_{112} \end{aligned}$$

Soit,

$$\psi'_{fb} \simeq -\frac{Var(w)}{2} \frac{u_{112}(w - p_0sy^*, H)}{u_{12}(w - p_0sy^*, H)}$$

²²Nous supposons dans les deux cas que le traitement curatif est efficace.

²³Soit sy^* le montant des soins curatifs que l'individu doit déboursier en cas de maladie.

Définition 20 La prime de prudence par rapport à la santé est définie par :

$$\psi'_{fb} \simeq \frac{\text{Var}(w)}{2} \eta_H(w)$$

avec $\eta_H(w) = -\frac{u_{112}(w_0 - p_0 sy^*, E\tilde{H})}{u_{12}(w_0 - p_0 sy^*, E\tilde{H})}$ la prudence absolue du risque financier par rapport à la santé.

Nous n'étudions que ψ'_{fb} car l'analyse de ψ_{fb} est effectuée au paragraphe précédent. La prime de prudence par rapport à la santé est positive si $u_{112}(.,.) > 0$ car $u_{12}(.,.) < 0$.

Proposition 21 Si $\frac{\partial A(w)}{\partial H} < 0$ alors $\frac{\partial r_{fb}}{\partial H} < 0$

Corollaire 22 Si $\eta_H(w) > A(w)$ alors $\frac{\partial r_{fb}}{\partial H} < 0$

Preuve :

$\frac{\partial A(w)}{\partial H} < 0$ implique que $-\frac{u_{112}(w, H)}{u_{12}(w, H)} > -\frac{u_{11}(w, H)}{u_1(w, H)}$. Soit $\eta_H(w) > A(w)$. Ainsi $\psi'_{fb} > r_{fb}$.

En dérivant l'équation qui définit la prime de risque par rapport à la santé, nous obtenons :

$$\begin{aligned} & u_2(w - \psi'_{fb} - p_0 sy^*, H) \\ &= -u_1(w - r_{fb} - p_0 sy^*, H) \frac{\partial r_f}{\partial H} + u_2(w - r_{fb} - p_0 sy^*, H) \end{aligned}$$

Il est aisé de montrer que $\psi'_{fb} > r_{fb}$ implique que $\frac{\partial r_{fb}}{\partial H} < 0$ ²⁴. CQFD

Risque sanitaire malin

Dans ce cas, l'individu ne peut pas retrouver son état de santé initial²⁵. La prime de risque r_{fm} pour un risque sanitaire malin est définie par ²⁶ :

$$\begin{aligned} & p_0 u(w - p_0 sy^* - r_{fm}, H - \Delta H) + (1 - p_0) u(w - p_0 sy^* - r_{fm}, H) \\ &= p_0 u(w - sy^*, H - \Delta H) + (1 - p_0) u(w, H) \end{aligned}$$

En utilisant un développement limité autour du point $(w - p_0 sy^*, E\tilde{H})$, nous obtenons après simplifications²⁷ :

$$\begin{aligned} -r_{fm} u_1 &\simeq \frac{p_0}{2} (-sy^* + p_0 sy^*)^2 u_{11} + \frac{(1 - p_0)}{2} (p_0 sy^*)^2 u_{11} \\ &\quad + p_0 (-sy^* + p_0 sy^*) (H - \Delta H - E\tilde{H}) u_{12} \\ &\quad + (1 - p_0) (p_0 sy^*) (H - E\tilde{H}) u_{12} \end{aligned}$$

²⁴La démonstration est similaire à celle de la section précédente.

²⁵L'individu après traitement subit une perte de capital sanitaire de ΔH .

²⁶Kihlstrom et Mirman (1974) ont défini une prime de risque bivariee π_1 tel que :

$E[u(w_1 + \tilde{z}_1, w_2 + \tilde{z}_2)] = E[u(w_1 - \pi_1, w_2)]$

Chalfant et Finkelshtain (1993) ont défini une prime de risque partielle multivariee π_1 définie par :

$E[u(w_1 + \tilde{z}_1, w_2 + \tilde{z}_2)] = E[u(w_1 - \pi_1, w_2 + \tilde{z}_2)]$.

Ici, nous appliquons la définition d'une prime de risque partielle multivariee.

²⁷Comme Chalfant et Finkelshtain (1993), nous ignorons les termes qui contiennent l'expression r_{fm}^2 .

Soit,

$$r_{fm} \simeq -\frac{Var(w)}{2} \frac{u_{11}}{u_1} - Cov(w, H) \frac{u_{11}}{u_1} \frac{u_{12}}{u_{11}}$$

Définition 23 La prime de risque d'un aléa sanitaire malin s'énonce :

$$r_{fm} \simeq A(w) \left[\frac{Var(w)}{2} + Cov(w, H) \frac{u_{12}(w-p_0sy^*, E\tilde{H})}{u_{11}(w-p_0sy^*, E\tilde{H})} \right]$$

$$\text{avec } A(w) = \frac{u_{11}(w-p_0sy^*, E\tilde{H})}{u_1(w-p_0sy^*, E\tilde{H})}.$$

La prime de risque d'un aléa sanitaire dépend de l'aversion absolue au risque financier multipliée par la somme de la variance du risque richesse et de la covariance des risques santé et richesse. Elle est positive car $Cov(w, H) > 0$ et $\frac{u_{12}}{u_{22}} > 0$.

Proposition 24 La prime de risque d'un aléa sanitaire malin est supérieure à celle d'un aléa sanitaire bénin.

Preuve :

La prime de risque d'un aléa sanitaire malin s'énonce :

$$r_{fm} = A(w) \left[\frac{Var(w)}{2} + Cov(w, H) \frac{u_{12}}{u_{11}} \right]$$

et $r_{bm} = \frac{1}{2}Var(w)A(w)$. Ainsi $r_{fm} > r_{fb}$ si $Cov(w, H)A(w) \frac{u_{12}}{u_{11}} > 0$.

Or $A(w) > 0$, $Cov(w, H) = 2(1-p_0)p_0sy^*(H_2 - H_1) > 0$ et $\frac{u_{12}}{u_{22}} > 0$.

Ainsi $r_{fm} > r_{fb}$. CQFD.

La prime de risque d'un risque malin est supérieure à celle d'un risque bénin car l'individu est prêt à payer une somme d'argent plus importante puisque son état de santé se détériore.

Nous pouvons donc étudier deux effets : celui de la richesse sur la prime de risque et celui de la santé sur la prime de risque.

La prudence par rapport à la richesse La prime de prudence du risque financier d'un aléa sanitaire malin par rapport à la richesse notée ψ_{fm} est définie par :

$$\begin{aligned} & p_0u_1(w - p_0sy^* - \psi_{fm}, H - \Delta H) + (1 - p_0)u_1(w - p_0sy^* - \psi_{fm}, H) \\ = & p_0u_1(w - sy^*, H - \Delta H) + (1 - p_0)u_1(w, H) \end{aligned}$$

En utilisant un développement limité autour du point $(w - p_0sy^*, E\tilde{H})$, nous obtenons après simplifications²⁸ :

$$\begin{aligned} -\psi_{fm}u_{11} & \simeq \frac{p_0}{2}(-sy^* + p_0sy^*)^2u_{111} + \frac{(1-p_0)}{2}(p_0sy^*)^2u_{111} \\ & + p_0(-sy^* + p_0sy^*)(H - \Delta H - E\tilde{H})u_{112} \\ & + (1-p_0)(p_0sy^*)(H - E\tilde{H})u_{112} \end{aligned}$$

²⁸ Comme Chalfant et Finkelshtain (1993), nous ignorons les termes qui contiennent l'expression ψ_m^2 .

Soit,

$$\psi_{fm} \simeq -\frac{\text{Var}(w)}{2} \frac{u_{111}}{u_{11}} - \text{Cov}(w, H) \frac{u_{111}}{u_{11}} \frac{u_{112}}{u_{111}}$$

Définition 25 La prime de prudence du risque financier d'un aléa sanitaire malin par rapport à la richesse est définie par :

$$\psi_{fm} \simeq \eta(w) \left[\frac{\text{Var}(w)}{2} + \text{Cov}(w, H) \frac{u_{112}}{u_{111}} \right]$$

$$\text{avec } \eta(w) = -\frac{u_{111}(w-p_0sy^*, E\tilde{H})}{u_{11}(w-p_0sy^*, E\tilde{H})}$$

La prime de prudence du risque financier d'un aléa sanitaire malin par rapport à la richesse est égale au produit de l'indice de prudence absolue d'un risque financier et de la somme de la variance du risque richesse et du produit de la covariance que multiplie le taux marginal de substitution de l'utilité marginale seconde de la richesse entre la richesse et la santé. La prime de prudence par rapport à la richesse est donc positive si $u_{111}(.,.) > 0$.

Proposition 26 Si $\frac{\partial A(w)}{\partial w} < 0$ et si $\frac{\partial \left(\frac{u_{12}}{u_{11}}\right)}{\partial w} < 0$ alors $\frac{\partial r_{fm}}{\partial w} < 0$

Preuve :

$\frac{\partial A(w)}{\partial w} < 0$ implique que $-\frac{u_{111}(w-p_0sy^*, E\tilde{H})}{u_{11}(w-p_0sy^*, E\tilde{H})} > -\frac{u_{11}(w-p_0sy^*, E\tilde{H})}{u_{11}(w-p_0sy^*, E\tilde{H})}$. Soit $\eta(w) > A(w)$.

Si $\frac{\partial \frac{u_{12}}{u_{11}}}{\partial w} < 0$ alors $\frac{u_{112}}{u_{111}} > \frac{u_{12}}{u_{11}}$. Nous obtenons donc que $\psi_{fm} > r_{fm}$.

Nous pouvons maintenant dériver l'équation qui définit la prime de risque par rapport à la richesse et nous obtenons :

$$\begin{aligned} & Eu_1/H(w - \psi_{fm} - p_0sy^*, .) \\ &= -Eu_1/H(w - r_{fm} - p_0sy^*, .) \frac{\partial r_{fm}}{\partial w} + Eu_1/H(w - r_{fm} - p_0sy^*, .) \end{aligned}$$

avec $Eu_1/H(W, .)$ l'espérance de l'utilité marginale de la richesse par rapport à la santé pour un niveau de richesse donné W .

$\psi_{fm} > r_{fm}$ implique que $Eu_1/H(w - r_{fm} - p_0sy^*, .) < Eu_1/H(w - \psi_{fm} - p_0sy^*, .)$ car $Eu_1/H(W, .)$ est une fonction concave. Nous en déduisons que $-Eu_1/H(w - \pi - p_0sy^*, .) \frac{\partial r_{fm}}{\partial w} > 0$. En raison de la positivité de $Eu_1/H(W, .)$, nous en déduisons que $\frac{\partial r_{fm}}{\partial w} < 0$. CQFD

La prime de risque d'un aléa sanitaire malin est une fonction décroissante de la richesse si l'aversion absolue au risque financier est une fonction décroissante de la richesse et si le taux marginal de substitution de l'utilité marginale de la richesse entre la richesse et la santé est une fonction décroissante de la richesse.

La prudence par rapport à la santé La prime de prudence du risque financier d'un aléa sanitaire malin par rapport à la santé notée ψ'_{fm} est définie par :

$$\begin{aligned} & p_0u_2(w - p_0sy^* - \psi'_{fm}, H - \Delta H) + (1 - p_0)u_2(w - p_0sy^* - \psi'_{fm}, H) \\ &= p_0u_2(w - sy^*, H - \Delta H) + (1 - p_0)u_2(w, H) \end{aligned}$$

En utilisant un développement limité autour du point $(w - p_0sy^*, E\tilde{H})$, nous obtenons après simplifications²⁹ :

$$\begin{aligned} -\psi'_{fm}u_{12} &\simeq \frac{p_0}{2}(-sy^* + p_0sy^*)^2u_{112} + \frac{(1-p_0)}{2}(p_0sy^*)^2u_{112} \\ &\quad + p_0(-sy^* + p_0sy^*)(H - \Delta H - E\tilde{H})u_{122} \\ &\quad + (1-p_0)(p_0sy^*)(H - E\tilde{H})u_{122} \end{aligned}$$

Soit,

$$\psi'_{fm} \simeq -\frac{Var(w)}{2} \frac{u_{112}}{u_{12}} - Cov(w, H) \frac{u_{112}}{u_{12}} \frac{u_{122}}{u_{112}}$$

Définition 27 La prime de prudence du risque financier d'un aléa sanitaire malin par rapport à la santé est définie par :

$$\psi'_{fm} \simeq \eta_H(w) \left[\frac{Var(w)}{2} + Cov(w, H) \frac{u_{122}}{u_{112}} \right]$$

La prime de prudence du risque financier d'un aléa sanitaire malin par rapport à la richesse est égale au produit de l'indice de prudence absolue d'un risque financier par rapport à la santé et de la somme de la variance du risque richesse et du produit de la covariance que multiplie le taux marginal de substitution de l'utilité marginale croisée de la richesse et de la santé, entre la richesse et la santé. La prime de prudence du risque financier d'un aléa sanitaire malin par rapport à la santé est positive si $u_{112}(\dots) > 0$.

Proposition 28 Si $\frac{\partial A(w)}{\partial H} < 0$ et si $\frac{\partial \frac{u_{12}}{u_{11}}}{\partial H} < 0$ alors $\frac{\partial r_{fm}}{\partial H} < 0$

Preuve :
 $\frac{\partial A(w)}{\partial H} < 0$ implique que $-\frac{u_{112}(w-p_0sy^*, E\tilde{H})}{u_{12}(w-p_0sy^*, E\tilde{H})} > -\frac{u_{11}(w-p_0sy^*, E\tilde{H})}{u_1(w-p_0sy^*, E\tilde{H})}$. Soit $\eta_H(w) > A(w)$.

Si $\frac{\partial \frac{u_{12}}{u_{11}}}{\partial w} < 0$ alors $\frac{u_{122}}{u_{112}} > \frac{u_{12}}{u_{11}}$. Nous en déduisons que $\psi'_{fm} > r_{fm}$.

Nous pouvons maintenant dériver l'équation qui définit la prime de risque par rapport à la santé, nous obtenons :

$$\begin{aligned} &Eu_2/H(w - \psi'_{fm} - p_0sy^*, \cdot) \\ &= -Eu_2/H(w - r_{fm} - p_0sy^*, \cdot) \frac{\partial r_{fm}}{\partial H} + Eu_2/H(w - r_{fm} - p_0sy^*, \cdot) \end{aligned}$$

avec $Eu_2/H(W, \cdot)$ l'espérance de l'utilité marginale de la santé par rapport à la santé pour un niveau de richesse donné W .

$\psi'_{fm} > r_{fm}$ implique que $Eu_2/H(w - r_{fm} - p_0sy^*, \cdot) < Eu_2/H(w - \psi'_{fm} - p_0sy^*, \cdot)$ car $Eu_2/H(W, \cdot) < 0$. Nous en déduisons que $-Eu_2/H(w - \pi - p_0sy^*, \cdot) \frac{\partial r_{fm}}{\partial H} > 0$. En raison de la positivité de $Eu_2/H(W, \cdot)$, nous en déduisons que $\frac{\partial r_{fm}}{\partial H} < 0$. CQFD

²⁹ Comme Chalfant et Finkelshain (1993), nous ignorons les termes qui contiennent l'expression ψ_m^2 .

La prime de risque d'un risque malin est une fonction décroissante de la santé si l'aversion absolue au risque financier par rapport à la santé est une fonction décroissante de la santé et si le taux marginal de substitution de l'utilité marginale de la richesse entre la richesse et la santé est une fonction décroissante de la santé.

Nous retrouvons les résultats usuels d'un risque financier pur de Kimball (1990). Pour un risque sanitaire, sous certaines conditions analogues à celles de Kimball (décroissance de l'aversion absolue au risque financier par rapport à la santé), nous obtenons que la prime de risque d'un aléa monétaire est une fonction décroissante de la santé ou de la richesse. Un individu en bonne santé sera prêt à payer une prime de risque inférieure à celle de l'individu en mauvaise santé, que le risque sanitaire soit assurable ou non.

4.4 Conclusion

Dans ce chapitre, nous avons appliqué l'analyse bivariée partielle (Finkelsh-tain et Chalfant, 1993) à un risque sanitaire. L'équivalence de Pratt (1964) est donc vérifiée pour un risque sanitaire. La prime de risque est positive quand l'individu est averse au risque sanitaire.

De plus, nous avons généralisé l'analyse de la prudence d'un risque financier (Kimball, 1990) à un risque sanitaire. Sous certaines conditions (décroissance de l'aversion absolue au risque financier par rapport à la santé) présentant une ressemblance à celles d'un risque financier, la prime de risque est une fonction décroissante de la santé ou de la richesse. Un individu en bonne santé sera prêt à payer une prime de risque inférieure à celle de l'individu en mauvaise santé que le risque sanitaire soit assurable ou non³⁰.

L'analyse de la prime de risque effectuée, nous caractérisons, dans le chapitre suivant, les contrats d'assurances d'individus qui diffèrent par leurs états de santé et qui peuvent effectuer un acte préventif visant à la réduction de la probabilité d'occurrence du dommage. De l'analyse *background risk*, nous déduisons que l'individu en mauvaise santé sera plus enclin à effectuer un acte préventif qu'un individu en bonne santé et ce, pour une même valeur du coût de l'effort.

³⁰Un travail de recherche futur serait de généraliser, à des risques bivariés ou multivariés, la notion de risque propre, de tolérance au risque, etc...

Chapitre 5

Santé, prévention et assurance

“Mais ce lointain où se situent les possibilités est, dans le même temps, pour l’homme, extrêmement proche : il y vit”, Gadamer (1998).

5.1 Introduction

La prévention¹ a d’abord été étudiée dans un cadre où l’individu est soumis à un risque financier. L’article pionnier de Ehrlich et Becker (1972) - noté par la suite EB - marque l’essor de l’étude de la prévention dont les principales extensions et discussions sont les travaux de Dionne et Eeckhoudt (1985), Briys et Schlesinger (1991), Konrad et Skaperdas (1993) et Chiu (2000). Toutefois, ces articles ont seulement considéré des risques financiers dans le cadre d’une fonction d’utilité unidimensionnelle.

Or, cette approche est inadaptée à l’étude du risque santé qui ne peut être uniquement appréhendé par une dimension de pertes financières. Eeckhoudt, Godfroid et Marchand (1998) ont été les premiers à présenter une approche bidimensionnelle : santé et richesse. En effet, la survenance de la maladie affecte aussi bien le niveau de santé que la richesse de l’individu. Ils récusent également l’absence de périodicité du modèle EB dans lequel toutes les décisions sont effectuées au début de la période. Leur modélisation - notée par la suite EGM - considère que les individus effectuent deux actions : une avant la survenance de la maladie (appelée médecine préventive) et une après la déclaration de la maladie (médecine curative).

Ils adoptent une fonction d’utilité séparable avec neutralité vis-à-vis du risque financier et aversion vis-à-vis du risque santé. Cette distinction est justifiée par la plus grande diversification des risques financiers par rapport aux

¹Les agents peuvent modifier la probabilité d’occurrence d’un risque en effectuant un effort coûteux.

risques sanitaires. Dans ce cadre, ils étudient les propriétés de la demande de prévention primaire² et de prévention secondaire³ en relation avec la médecine curative. Ils montrent que la demande de soins curatifs est indépendante de la demande de prévention primaire tandis que la demande de soins curatifs est corrélée négativement à la demande de prévention secondaire. Ces résultats sont étendus au cadre de la Théorie duale du risque⁴ par Courbage (2000).

Contrairement à ces auteurs, nous ne considérons qu'un cadre statique et d'assurance. Nous supposons que l'individu averse au risque n'effectue qu'une action de prévention au début de la période et contrairement au modèle EB, que l'apparition de la maladie affecte aussi bien la richesse de l'individu que son niveau de santé. Seul le risque financier engendré par les dépenses curatives de l'individu est assurable. Ainsi, ce cadre bidimensionnel caractérise une double asymétrie d'information : aléa moral (prévention) et sélection adverse (les individus ont un état de santé initial différent déterminé par la nature au début du jeu). L'objectif de ce chapitre est, dans un cadre de double asymétrie d'information, propre au risque sanitaire, de caractériser les contrats d'assurance d'équilibre.

Très peu d'articles traitent conjointement de cette double asymétrie d'information. Deux axes de recherche sont néanmoins à mentionner.

- Tout d'abord, l'antisélection porte sur l'aversion au risque. Dans ce modèle, le coût de prévention est financier. Jullien, Salanié et Salanié (2000) démontrent que l'incitation à l'effort de l'agent croît à la fois avec son aversion au risque et avec le caractère incitatif du contrat de réservation⁵. Ils mettent en évidence que l'individu le moins averse au risque reçoit une couverture du dommage moindre.
- Ensuite, l'antisélection concerne le coût pour effectuer l'effort de prévention, coût psychologique et non monétaire. Chiappori et Chassagnon (1995) et Chopin-Kambia (2000) caractérisent les contrats optimaux d'assurance quand la variable d'effort est dichotomique, pour un marché d'assurance en concurrence. Les principaux résultats établissent que les individus les plus risqués (ils présentent un coût de prévention important) préfèrent les contrats des individus moins risqués. L'équilibre sera séparateur si la valeur du coût d'autoprotection est faible. Les individus ne sont incités à effectuer un effort que si le dommage n'est pas intégralement remboursé. Si la valeur du coût est élevée, aucun individu n'effectue d'effort. Les individus présentent le même risque financier. Ils reçoivent un contrat d'assurance complète.

²La prévention primaire est une action médicale qui agit sur la probabilité d'apparition de la maladie.

³La prévention secondaire est une action médicale qui affecte le degré de sévérité de la maladie.

⁴Contrairement à l'espérance de l'utilité, elle considère la linéarité des préférences en la richesse et la non-linéarité en les probabilités (par une transformation subjective de ces dernières) Yaari (1987).

⁵L'effort est une variable continue.

Ce chapitre est organisé comme suit. Dans la seconde section, nous décrivons le modèle et discutons ses propriétés. Nous présentons les contrats d'assurance de premier rang quand aucune asymétrie d'information ne subsiste. Dans une troisième section, nous caractérisons les contrats d'assurance quand une seule asymétrie d'information est considérée. La quatrième section présente les propriétés des contrats d'assurance avec inobservabilité de l'état de santé et en présence d'aléa moral. Nous montrons que les individus en mauvaise santé obtiennent un contrat d'assurance avec couverture partielle contrairement aux individus en bonne santé.

5.2 Le modèle en information parfaite

Nous considérons un individu dont l'utilité dépend du niveau de la richesse (noté W), du stock santé (noté H) et de la désutilité à effectuer un effort de prévention (notée c). Au début de la période, l'individu ne sait pas s'il sera malade. En cas de maladie (un événement de probabilité p dont la valeur est modifiée par une action préventive e_i avec $i \in [0, 1]$), l'individu subit une perte monétaire d qui correspond au coût du traitement. Cette maladie maligne, après traitement, entraîne une dépréciation de son capital santé (notée ΔH) qui reflète la gravité de celle-ci⁶. L'individu possède une richesse initiale notée w et dispose d'un stock de santé initial, noté H_j avec $j \in \{1, 2\}$. Nous posons $H_1 < H_2$. L'individu de type 1 sera appelé l'individu en mauvaise santé et l'individu de type 2 celui en bonne santé. L'individu a le choix entre deux actions préventives primaires : soit $e_0 = 0$ (absence de prévention), soit $e_1 = 1$ (prévention). Pour simplifier, nous noterons $p(e_1) = p_1$ et $p(e_0) = p_0$ avec $p_1 < p_0$. L'individu pour effectuer son action de prévention, arbitre entre le gain lié à la prévention (c'est-à-dire la somme de son utilité liée à sa richesse et son état de santé) et son coût.

Le problème, ici évoqué, peut caractériser un régime alimentaire que l'individu doit effectuer pour diminuer le risque sanitaire qu'il encourt (cholestérol ou diabète). Nous supposons que la modification des habitudes alimentaires génère un effet de richesse financière nulle. L'individu réalloue son budget à la consommation d'autres biens alimentaires. Respecter un régime alimentaire diminue la survenance de la maladie qui détériore l'état de santé du patient.

Dans ce cadre, les individus diffèrent par leur état de santé initial. Ils encourrent un risque sanitaire dont les conséquences financières sont assurables. Les individus de type 1 sont ceux qui supportent le risque sanitaire le plus important car pour ces individus en mauvaise santé, une perte de santé ΔH entraîne une diminution de l'utilité liée à la santé supérieure à celle des individus en bonne santé (type 2).

⁶Si la maladie était bénigne, $\Delta H = 0$. Nous supposons par la suite $\Delta H \neq 0$. Si l'individu n'était pas soigné sa perte de capital santé serait supérieure à ΔH . $\Delta H = \Delta H_M - \Delta H_T$ avec ΔH_M la perte de capital santé générée par la maladie et ΔH_T le gain en capital santé lié au traitement.

Cette hypothèse est usuelle en économie de la santé. Nous supposons qu'un individu malade et traité ne peut obtenir un meilleur état de santé qu'un individu non malade (EGM).

Pour faciliter l'analyse, nous supposons que la fonction d'utilité est additive et séparable. Contrairement au modèle EGM, nous supposons que les individus manifestent de l'aversion vis-à-vis du risque financier et du risque sanitaire. Formellement, la fonction d'utilité s'écrit :

$$U(W, H, e) = u(W) + \delta v(H) - ce$$

Les fonctions u et v vérifient les propriétés usuelles de croissance et de concavité. Le paramètre δ indique l'importance relative que l'individu accorde à sa santé par rapport à sa richesse. Nous supposons que δ et c sont connaissance commune.

Les individus peuvent s'assurer contre le risque financier qu'ils encourent en souscrivant une assurance auprès d'une compagnie d'assurance sur un marché en concurrence.

L'assureur est neutre au risque et couvrira partiellement ou totalement ses assurés en cas de survenance de la maladie. Un contrat d'assurance (C) est défini par la couple (π, q) avec $\pi \geq 0$ la prime d'assurance et $q \geq 0$ le niveau de couverture. La concurrence entre compagnies d'assurance implique que la prime, en l'absence de coûts de transaction, est actuariellement juste : $\pi = pq$.

Le jeu considéré comporte quatre étapes :

- 1 - Chaque assureur offre un contrat,
- 2 - Les assurés choisissent un contrat,
- 3 - En fonction du contrat choisi, ils effectuent ou non un effort de prévention,
- 4 - Les aléas de santé se réalisent et les assureurs effectuent les prélèvements prévus par les contrats choisis.

En information parfaite, l'assureur observe l'effort de l'assuré et connaît son type sanitaire. Lorsque le niveau de prévention e_i est observable, la prime est fonction de e_i . Elle s'écrit donc : $\pi(e_i) = p(e_i)q$. Il détermine le montant de la couverture pour un effort donné.

Il résout donc le programme suivant :

$$\max_q p(e_i)u(w - p(e_i)q + q - d) + (1 - p(e_i))u(w - p(e_i)q) \\ + \delta((1 - p(e_i))v(H_j) + p(e_i)v(H_j - \Delta H)) - ce_i$$

avec $i \in \{0, 1\}$ et $j \in \{1, 2\}$

Nous obtenons comme condition d'optimalité⁷ :

$$p(e_i)(1 - p(e_i))(u'(w - p(e_i)q + q - d) - u'(w - p(e_i)q)) = 0$$

avec $u'(\cdot)$ la dérivée de la fonction par rapport à l'argument.

A l'optimum, la couverture d'assurance du risque financier est complète. En effet, comme l'effort de prévention est parfaitement observable, l'individu est

⁷Si nous supposons une fonction d'utilité non séparable entre la richesse et la santé, la condition du premier ordre implique que $q = D$ quel que soit le risque sanitaire (bénin ou malin).

contraint d'effectuer l'effort qu'il s'est engagé à réaliser dans sa police d'assurance. Toutefois, la valeur de l'effort dépendra ici de la valeur du coût de la désutilité de l'effort et de l'état de santé du patient.

L'effort de prévention sera fourni si :

$$u(w - p_1d) - u(w - p_0d) + \delta((p_0 - p_1)(v(H_j) - v(H_j - \Delta H))) \geq c$$

Cette condition implique que l'effort de prévention sera effectué si le gain d'utilité lié à la richesse financière et à la santé est supérieur au coût de la désutilité de l'effort. Nous pouvons donc distinguer la valeur seuil de désutilité (noté c_j^{IP}) pour laquelle l'individu est indifférent entre réaliser ou non un effort préventif⁸.

$$c_j^{IP} = u(w - p_1d) - u(w - p_0d) + \delta((p_0 - p_1)(v(H_j) - v(H_j - \Delta H)))$$

avec IP information parfaite.

Proposition 29 *En information parfaite, l'individu en mauvaise santé réalise un effort de prévention pour des valeurs de la désutilité de l'effort plus grandes que celles de l'individu en bonne santé.*

Preuve : $H_1 < H_2$ et v concave impliquent que $(v(H_1) - v(H_1 - \Delta H)) > (v(H_2) - v(H_2 - \Delta H))$. Ainsi $c_2^{IP} < c_1^{IP}$. CQFD.

Nous remarquons que plus la valeur du stock de santé initial est faible, plus l'individu est incité à effectuer un effort de prévention. En termes sanitaires, l'individu le plus risqué⁹ (type 1) effectue un effort de prévention pour des valeurs de c plus grandes que celles du type 2. L'individu en mauvaise santé est incité à effectuer un effort de prévention pour des valeurs plus grandes de la désutilité de l'effort car son gain en capital santé est plus important que celui d'un individu en bonne santé. Il arbitre entre le coût de l'effort et le gain en utilité (aussi bien lié à la richesse qu'à la santé) généré par l'effort.

Nous pouvons maintenant caractériser les contrats d'assurance d'information parfaite. La prime d'assurance diffère selon la valeur de l'effort préventif. Les différentes situations sont résumées dans le tableau suivant :

Coût	0	c_2^{IP}	c_1^{IP}
Contrat du type 1	$C_1^{IP} = (p_1d, d)$	$C_1^{IP} = (p_1d, d)$	$C_1^{IP} = (p_0d, d)$
Contrat du type 2	$C_2^{IP} = (p_1d, d)$	$C_2^{IP} = (p_0d, d)$	$C_2^{IP} = (p_0d, d)$

Tableau 5-1 : Les contrats d'assurance en information parfaite

⁸Pour un risque bénin et une utilité non séparable entre richesse et santé, $c_j^{IP} = u(w - p_1d, H_j) - u(w - p_0d, H_j)$.

Comme nous supposons $u_{RH} < 0$ (voir chapitre 4), nous obtenons $c_1^{IP} > c_2^{IP}$.

⁹L'individu de type 1, en termes sanitaires, est le plus risqué en ce sens que la survenance de la maladie entraîne pour lui une perte d'utilité plus importante que pour l'individu de type 2.

Pour des coûts d'effort moyen ($c \in [c_2^{IP}, c_1^{IP}]$), les individus en bonne santé (type 2) préfèrent les contrats d'assurance de premier rang des individus en mauvaise santé, puisque ce sont des contrats à primes plus faibles avec une couverture totale. Ici, nous retrouvons la situation usuelle des modèles de marché d'assurance avec antisélection. En termes de risque financier, les agents de type haut risque (ceux qui réalisent e_0 , c'est-à-dire les individus en bonne santé) veulent souscrire le contrat de type bas risque (ceux qui effectuent un effort, individus en mauvaise santé). Étudions maintenant, l'introduction d'une seule asymétrie d'information.

5.3 Le modèle avec une seule asymétrie d'information

Deux asymétries d'information sont à considérer. La première porte sur l'état de santé de chaque individu. La seconde a trait à l'effort préventif que l'individu peut fournir. Nous montrons que seule l'inobservabilité de l'effort préventif génère des contrats d'assurance de second rang.

5.3.1 L'état de santé de l'individu est inobservable

L'assureur ne connaît pas l'état de santé de ses assurés (le risque sanitaire est ici de subir l'altération de son capital sanitaire). Il connaît par contre, dans la population, la proportion λ de chacun des deux types¹⁰. Par rapport à la situation d'information parfaite, seul un cas semble problématique. En effet, quand le coût de l'effort est faible ou important, les deux agents obtiennent le même contrat. Quel que soit leur état de santé, ils effectuent le même effort. Ce dernier étant observable, il peut figurer dans les contrats d'assurances. Ils obtiennent donc, dans ces deux cas, les contrats d'assurance d'information parfaite. Si les agents n'effectuent pas d'effort, l'assureur peut donc prouver le non-respect du contrat d'assurance puisque l'effort est observable. Nous ne considérons pas la législation applicable en cas de non-respect du contrat. Nous supposons seulement que la caducité du contrat entraîne le non-remboursement du dommage financier par la compagnie d'assurance.

Par rapport à la situation d'information parfaite, pour $c \in [c_2^{IP}, c_1^{IP}]$, les deux individus n'obtiennent pas les mêmes contrats d'assurance. Les individus en bonne santé ne fournissent pas d'effort préventif. Ils préféreraient choisir le contrat des individus en mauvaise santé. Toutefois, le coût de l'effort est trop important pour qu'ils réalisent un effort préventif. Ils ne peuvent donc choisir le contrat des individus en mauvaise santé car l'effort est observable.

Proposition 30 *Pour la compagnie d'assurance, ne pas connaître l'état de santé n'est pas problématique lorsque l'effort est observable. L'équilibre concu-*

¹⁰Soit λ la proportion des individus en bonne santé. Nous supposons que λ est suffisamment grand pour que l'équilibre séparateur existe. Ainsi, dans la population, le nombre d'individus en bonne santé est supérieur à celui en mauvaise santé.

rentiel du marché de l'assurance conduit les individus à choisir les mêmes contrats que dans une situation où l'état de santé est observé par les assureurs.

Ne pas connaître l'état de santé des assurés ne pose pas de problème puisqu'observer l'effort revient à déterminer le risque financier de l'individu, seul risque pertinent pour la compagnie d'assurance. En fonction de leur état de santé et du contrat d'assurance, les individus déterminent la valeur de leur effort préventif.

5.3.2 L'effort d'autoprotection est inobservable

Si l'effort est inobservable, les termes du contrat ne peuvent dépendre de l'effort préventif. La prime d'assurance sera une prime actuarielle en raison de la concurrence du marché d'assurance.

Définition de l'équilibre avec aléa moral

Les propriétés d'un équilibre compétitif avec risque moral ont été caractérisées par Arnott et Stiglitz (1991). L'équilibre compétitif est défini comme un équilibre de Nash avec libre entrée et sortie sur le marché de l'assurance. La nature de l'équilibre dépend de la définition des contrats existants à l'équilibre. Deux types d'équilibre sont à distinguer.

L'équilibre en quantité La compagnie d'assurance rationne la quantité d'assurance contractualisable par l'assuré. Pour des contrats déterministes, Pauly (1974) montre que si la quantité totale d'assurance contractée par l'assuré est observable et que la fonction d'utilité est séparable, les firmes d'assurance offrent des contrats d'assurance exclusifs et à profits nuls¹¹. Ce résultat a été étendu à des contrats aléatoires par Arnott et Stiglitz (1991).

L'équilibre en prix Dans un tel cadre, la quantité d'assurance contractée par les agents n'est pas observable par la compagnie d'assurance. Plus les individus achètent des polices d'assurance, moins ils sont incités à se prémunir contre le risque assuré. La quantité d'assurance modifie le comportement préventif de l'agent (Arnott et Stiglitz, 1991).

Trois types d'équilibre sont possibles¹² (Pauly, 1974) :

- un équilibre avec pleine assurance et prime actuarielle,
- un équilibre avec profits positifs,
- un équilibre sans assurance.

L'absence d'assurance à l'équilibre n'est pas contre-intuitive. Si le prix de la police est trop élevé, les individus ne s'assurent pas. Si le prix de la police est

¹¹Un contrat d'assurance est dit exclusif si l'individu ne peut s'assurer qu'une seule fois contre un même risque financier.

¹²En raison de la non-convexité de la contrainte d'aléa moral des assurés (Arnott, 1991) si la fonction d'utilité est non séparable. Pour une fonction d'utilité séparable, Chopin-Kambia (2001) établit la convexité de la contrainte d'aléa moral.

trop faible, les individus s'assurent plusieurs fois et n'effectuent pas d'effort de protection. Les contrats deviennent donc déficitaires.

Dans le reste de ce chapitre, nous supposons que les contrats sont déterministes et que la quantité totale d'assurance achetée par l'assuré est observable par les compagnies d'assurance.

En assurance santé, l'individu, pour se faire rembourser les soins médicaux par sa mutuelle ou une compagnie d'assurance privée, leur envoie sa feuille de remboursement adressée par la sécurité sociale. Il ne peut donc se faire rembourser qu'une fois. Le contrat d'assurance santé est ici exclusif.

L'équilibre que nous retenons est un équilibre en quantité. Le jeu décrit précédemment se trouve modifié. Le point 2 (décrit dans la section 2) devient : Les assurés choisissent un seul contrat. A l'équilibre, aucun contrat ne peut être déficitaire.

Détermination de l'équilibre

En présence d'aléa moral, les termes du contrat ne peuvent plus être indexés sur le niveau de l'effort. Pour un type j donné, l'assureur caractérise le contrat d'assurance en résolvant le programme suivant :

$$\begin{aligned}
 & \max_{q_j} p(e_j)u(w - \pi_j + q_j - d) + (1 - p(e_j))u(w - \pi_j) \\
 & - ce_j + \delta((1 - p(e_j))v(H_j) + p(e_j)v(H_j - \Delta H)) \\
 & e_j \in \arg \max_{x_i} p(x_i)u(w - \pi_j + q_j - d) + (1 - p(x_i))u(w - \pi_j) \\
 & - cx_i + \delta((1 - p(x_i))v(H_j) + p(x_i)v(H_j - \Delta H)) \quad \text{sc1} \\
 \pi_j & \geq p(e_j)q_j \quad \text{sc2} \\
 \text{avec } x_i & = e_i, i \in \{0, 1\} \text{ et } j \in \{1, 2\}
 \end{aligned}$$

La contrainte (sc1) est posée pour garantir la réalisation de l'effort de prévention maximal, par les assurés, pour un contrat d'assurance donné. La contrainte (sc2) implique que les contrats ne sont pas déficitaires.

En raison de l'exclusivité des contrats et de la concurrence des compagnies d'assurance, les profits espérés sur les contrats sont nuls (Arnott, 1991). La contrainte (sc2) est saturée.

L'effort de prévention n'étant pas une variable continue, nous ne pouvons pas appliquer les conditions du premier ordre¹³.

De la contrainte (sc1), nous déduisons que l'individu de type j effectuera, pour un contrat d'assurance donné, un effort de prévention si :

$$\begin{aligned}
 c & \leq (p_0 - p_1)(u(w - \pi_j) - u(w - \pi_j + q_j - d)) \\
 & + \delta((p_0 - p_1)(v(H_j) - v(H_j - \Delta H)))
 \end{aligned}$$

Le gain lié à la réalisation d'un effort est une fonction décroissante de q , le niveau de la couverture. Toutefois, nous remarquons, dans ce contexte d'aléa

¹³Ces conditions du premier ordre sont énoncées dans Winter (1991).

5.3. LE MODÈLE AVEC UNE SEULE ASYMÉTRIE D'INFORMATION 101

moral, que l'individu peut être amené à fournir un effort même si le contrat implique une assurance complète¹⁴.

Proposition 31 *En présence d'un risque sanitaire et d'aléa moral, le contrat d'assurance optimal offre une couverture complète du risque financier quand le coût de l'effort est faible ($c < \underline{c}_j^{AM}$) et qu'un effort préventif est réalisé.*

Preuve : Un contrat d'assurance complète avec effort stipule que $q_j = d$ et $\pi_j = p_1 d$. Avec un tel contrat, l'individu effectue un effort si :

$$\delta((p_0 - p_1)(v(H_j) - v(H_j - \Delta H))) \geq c \quad j \in \{1, 2\}$$

A l'égalité, un individu de type j est indifférent entre effectuer ou non un effort préventif (avec AM faisant référence au contexte aléa moral). Notons \underline{c}_j^{AM} , cette valeur seuil.

$$\underline{c}_j^{AM} = \delta(p_0 - p_1)(v(H_j) - v(H_j - \Delta H)) \quad j \in \{1, 2\}$$

Si $c < \underline{c}_j^{AM}$, l'individu de type j fournit un effort pour un contrat d'assurance complète. CQFD.

Bien que le risque financier soit totalement couvert pour de faibles valeurs du coût de l'effort, il n'en demeure pas moins que l'individu supporte toujours le risque sanitaire pour lequel la couverture est nulle. Contrairement au modèle d'aléa moral standard, ne pas effectuer d'effort avec un contrat d'assurance de premier rang entraîne en termes d'utilité, une économie de la valeur du coût de l'effort mais implique également, en espérance, une perte liée à la dépréciation du stock sanitaire. Ne pas effectuer d'effort ne conduit donc pas toujours à un gain d'utilité comme c'est le cas dans les modèles d'aléa moral usuels.

Dans ces modèles, la présence d'aléa moral implique une couverture partielle du risque financier pour de faibles coûts d'effort (Shavell, 1979). Cette conclusion reste valide uniquement pour des valeurs moyennes du coût ($\underline{c}_j^{AM} < c < \bar{c}_j^{AM}$) de l'effort que nous allons définir avec \bar{c}_j^{AM} , la valeur seuil du coût de l'effort pour laquelle l'individu est indifférent entre réaliser un effort préventif avec une couverture partielle du dommage et ne pas réaliser d'effort pour une couverture complète du dommage assuré.

L'individu reçoit une couverture partielle du dommage pour l'amener à réaliser un effort préventif si $\underline{c}_j^{AM} < c < \bar{c}_j^{AM}$. Dans ce cas, la contrainte ($sc1$) est saturée. Soit,

$$c = (p_0 - p_1)(u(w - \pi_j) - u(w - \pi_j + q_j - d)) + \delta((p_0 - p_1)(v(H_j) - v(H_j - \Delta H))) \quad (A)$$

¹⁴Pour une utilité non séparable entre la richesse et la santé, deux cas sont possibles :

Dans le cas d'un risque bénin, nous obtenons qu'un individu est incité à faire un effort uniquement si la couverture n'est pas totale. En effet, $(p_0 - p_1)(u(w - \pi, H_j) - u(w - \pi, H_j))$ ne peut être supérieur à c .

Dans le cas d'un risque malin, un contrat d'assurance complète incite l'individu à effectuer un effort si $(p_0 - p_1)(u(w - \pi, H_j) - u(w - \pi, H_j - \Delta H)) > c$.

Nous obtenons $\underline{c}_1^{AM} > \underline{c}_2^{AM}$ car $u_{RH} < 0$.

En utilisant la définition de \underline{c}_j^{AM} , cette équation devient

$$c = (p_0 - p_1)(u(w - \pi_j) - u(w - \pi_j + q_j - d)) + \underline{c}_j^{AM}$$

$c > \underline{c}_j^{AM}$ implique que $u(w - \pi_j) - u(w - \pi_j + q_j - d) > 0$. Soit $q_j < d$ car le contrat d'assurance est actuariel. Ainsi, quand le coût de la prévention n'est pas trop faible ($\underline{c}_j^{AM} < c < \bar{c}_j^{AM}$), la couverture n'est pas complète pour inciter les individus à fournir un effort.

Nous devons déterminer la valeur \bar{c}_j^{AM} pour laquelle l'individu est indifférent entre ce contrat de couverture partielle avec effort et un contrat d'assurance complète sans effort. En effet, quand le coût de l'effort est trop important, une couverture d'assurance complète du risque financier est offerte aux agents puisqu'il n'est plus possible de les inciter à réaliser un effort (Shavell, 1979). Pour des couvertures du dommage partielles trop faibles, un contrat de pleine assurance, sans la réalisation d'un effort, procure une utilité plus grande à l'individu qu'un contrat d'assurance partielle avec la fourniture d'un effort.

Un assuré est indifférent entre un contrat de couverture partielle avec effort et un contrat d'assurance complète sans effort si :

$$\begin{aligned} & p_1 u(w - p_1 q_j + q_j - d) + (1 - p_1) u(w - p_1 q_j) \quad (\text{B}) \\ & + \delta(p_1 v(H_j - \Delta H) + (1 - p_1) v(H_j)) - c \\ = & u(w - p_0 d) + \delta(p_0 v(H_j - \Delta H) + (1 - p_0) v(H_j)) \end{aligned}$$

En réécrivant l'équation (B) et en remplaçant c par sa valeur tirée de l'équation (A), nous obtenons l'équation (C) :

$$p_0 u(w - p_1 q_j + q_j - d) + (1 - p_0) u(w - p_1 q_j) = u(w - p_0 d) \quad (\text{C})$$

Le terme de gauche de l'équation (C) noté Γ est une fonction croissante de q_j . En effet,

$$\frac{\partial \Gamma}{\partial q_j} = p_0(1 - p_1) u_{q_j}(w - p_1 q_j + q_j - d) - p_1(1 - p_0) u_{q_j}(w - p_1 q_j)$$

$$\frac{\partial \Gamma}{\partial q_j} > 0 \text{ car } \frac{(1-p_1)}{p_1} > \frac{(1-p_0)}{p_0} \text{ et } u_{q_j}(w - p_1 q_j + q_j - d) > u_{q_j}(w - p_1 q_j)^{15}.$$

Quand la couverture est complète ($q_j = d$), nous obtenons que $\Gamma > u(w - p_0 d)$ tandis que pour une couverture nulle ($q_j = 0$), $\Gamma < u(w - p_0 d)$. De la croissance stricte de Γ nous déduisons qu'il existe $0 < \underline{q}_j < d$ tel que

$$p_0 u(w - p_1 \underline{q}_j + \underline{q}_j - d) + (1 - p_0) u(w - p_1 \underline{q}_j) = u(w - p_0 d)$$

Nous remarquons que \underline{q}_j ne dépend pas des caractéristiques individuelles des individus. Nous omettons donc l'indice j pour ce qui suit. Nous calculons, en

¹⁵car $q_j < d$.

5.3. LE MODÈLE AVEC UNE SEULE ASYMÉTRIE D'INFORMATION¹⁰³

reportant la valeur de \underline{q} dans l'équation (A), la valeur seuil (\bar{c}_j^{AM}) pour laquelle l'individu de type j est indifférent entre effectuer un effort ou non. Soit,

$$\begin{aligned}\bar{c}_j^{AM} &= (p_0 - p_1)(u(w - p_1\underline{q}) - u(w - p_1\underline{q} + \underline{q} - d)) \\ &\quad + \delta((p_0 - p_1)(v(H_j) - v(H_j - \Delta H)))\end{aligned}$$

Plus la valeur du stock initial sanitaire est faible, plus la valeur \bar{c}_j^{AM} est grande. En effet, $v(H_2) - v(H_2 - \Delta H) < v(H_1) - v(H_1 - \Delta H)$ implique que $\bar{c}_2^{AM} < \bar{c}_1^{AM}$ ¹⁶.

Quand $\underline{c}_j^{AM} < c < \bar{c}_j^{AM}$, l'individu choisit un contrat d'assurance partielle avec la réalisation d'un effort, car l'utilité obtenue avec ce contrat est supérieure à celle obtenue avec un contrat d'assurance complète sans effort.

Quand $c > \bar{c}_j^{AM}$, l'individu de type j obtient une utilité supérieure avec un contrat d'assurance complète sans la réalisation d'un effort à celle obtenue avec un contrat d'assurance partielle l'incitant à faire un effort. Les compagnies d'assurance qui maximisent l'utilité des assurés, leur offrent donc un contrat d'assurance avec une couverture complète.

Avant de résumer ces différents résultats (tableau 2), nous devons déterminer si \bar{c}_2^{AM} est inférieur ou supérieur à \underline{c}_1^{AM} .

$\bar{c}_2^{AM} > \underline{c}_1^{AM}$ implique que

$$\begin{aligned}&(u(w - p_1\underline{q}) - u(w - p_1\underline{q} + \underline{q} - d)) \\ &> \delta([v(H_1) - v(H_1 - \Delta H)] - [v(H_2) - v(H_2 - \Delta H)])\end{aligned}$$

Si la dépréciation du stock de santé n'est pas trop grande ou si la différence de l'état sanitaire n'est pas trop grande, alors cette inégalité est vérifiée¹⁷ si $\Phi H < \overline{\Phi H}$ avec $[v(H_1) - v(H_1 - \Delta H)] - [v(H_2) - v(H_2 - \Delta H)] = \Phi H$ et $\overline{\Phi H} = \frac{1}{\delta}(u(w - p_1\underline{q}) - u(w - p_1\underline{q} + \underline{q} - d))$

$\Phi H < \overline{\Phi H}$					
Coût	0	\underline{c}_2^{AM}	\underline{c}_1^{AM}	\bar{c}_2^{AM}	\bar{c}_1^{AM}
Bonne santé	C_2^{AM} ($p_1 d, d$)	C_2^{AM} ($p_1 q_2, q_2$)	C_2^{AM} ($p_1 q'_2, q'_2$)	C_2^{AM} ($p_0 d, d$)	C_2^{AM} ($p_0 d, d$)
Mauvaise Santé	C_1^{AM} ($p_1 d, d$)	C_1^{AM} ($p_1 d, d$)	C_1^{AM} ($p_1 q_1, q_1$)	C_1^{AM} ($p_1 q'_1, q'_1$)	C_1^{AM} ($p_0 d, d$)
$\Phi H > \overline{\Phi H}$					
Coût	0	\underline{c}_2^{AM}	\bar{c}_2^{AM}	\underline{c}_1^{AM}	\bar{c}_1^{AM}
Bonne santé	C_2^{AM} ($p_1 d, d$)	C_2^{AM} ($p_1 q_2, q_2$)	C_2^{AM} ($p_0 d, d$)	C_2^{AM} ($p_0 d, d$)	C_2^{AM} ($p_0 d, d$)
Mauvaise santé	C_1^{AM} ($p_1 d, d$)	C_1^{AM} ($p_1 d, d$)	C_1^{AM} ($p_1 d, d$)	C_1^{AM} ($p_1 q_1, q_1$)	C_1^{AM} ($p_0 d, d$)

¹⁶Il est trivial de montrer que $\bar{c}_j^{AM} > \underline{c}_j^{AM}$ car $\bar{c}_j^{AM} = (p_0 - p_1)(u(w - p_1\underline{q}) - u(w - p_1\underline{q} + \underline{q} - d)) + \underline{c}_j^{AM}$.

¹⁷Ou si l'aversion au risque sanitaire est très grande pour des états de santé donnés.

Tableau 5-2 : Les contrats d'assurance avec aléa moral

avec $0 < q'_2 < q_2 < d$ et $0 < q'_1 < q_1 < d$.

Ce tableau nous indique que, quelle que soit la différence d'état de santé des individus, un contrat de pleine assurance ne désincite pas l'assuré à réaliser un effort préventif quand son coût est faible. En termes d'utilité, le gain à ne pas réaliser un effort est inférieur à l'espérance de la perte liée à la dépréciation du capital sanitaire. Inversement, pour des valeurs élevées du coût de l'effort, l'individu ne réalise pas d'effort préventif et reçoit également un contrat d'assurance complète. Pour des valeurs moyennes du coût de l'effort ($\underline{c}_j^{AM} < c < \bar{c}_j^{AM}$), un contrat d'assurance partielle est choisi par l'assuré et l'incite à réaliser un effort préventif.

5.4 Inobservabilité de l'effort et de l'état de santé

Dans ce cadre, les contrats d'assurance optimaux doivent vérifier aussi bien les contraintes d'incitation que les contraintes d'autosélection des agents. Le tableau 2 nous indique que, si l'effort et l'état de santé des individus ne sont pas observables, il faut tenir compte du comportement opportuniste des individus. Par exemple, si $\Phi H > \overline{\Phi H}$ et si $c \in [\underline{c}_2^{AM}, \underline{c}_1^{AM}]$, l'individu en bonne santé voudra prendre le contrat des individus en mauvaise santé car, pour une couverture totale du dommage, il s'acquittera d'une prime d'assurance inférieure.

L'inobservabilité de l'effort (aléa moral) et de l'état de santé (antisélection) nous amènent à modifier notre définition de l'équilibre du marché d'assurance.

5.4.1 Définition de l'équilibre

Les propriétés de l'équilibre dépendront de la façon dont les compagnies d'assurance réagissent aux contrats proposés par leurs concurrents. Dans un environnement compétitif (libre-entrée et sortie), il est raisonnable de supposer que chaque entreprise considère le comportement de ses concurrents comme donné. Elles adoptent un comportement purement Nashien (Rothschild et Stiglitz, 1976)¹⁸. Nous supposons également que les compagnies d'assurance n'offrent qu'un seul contrat. Le point 1 du jeu (décrit à la section 3) devient : chaque assureur offre un seul contrat. Ainsi, les contrats à subventions croisées ne peuvent être des contrats d'équilibre¹⁹.

Un menu de contrats sera donc un équilibre si :

- a) aucun contrat à l'équilibre n'est déficitaire,
- b) aucun contrat supplémentaire ne peut réaliser des profits positifs.

¹⁸Pour une présentation exhaustive du phénomène de l'antisélection sur les marchés d'assurance, voir Dionne, Doherty et Fombaron (2001).

¹⁹Crocker et Snow (1985) ont montré que des contrats à subventions croisées dominent au sens de Pareto les contrats séparateurs de Rothschild et Stiglitz (1976).

L'équilibre défini, nous allons caractériser les contrats d'équilibre. Nous présentons tout d'abord, le programme que résolvent les compagnies d'assurance. Nous montrons ensuite, à l'aide des contraintes d'autosélection, que les contrats de second rang d'aléa moral ne peuvent pas toujours être offerts. Dans les problèmes standards d'antisélection, si le contrat de premier rang d'un individu de type j ne vérifie pas la contrainte d'autosélection dudit individu, alors ce contrat est proposé à l'équilibre et le contrat du type $-j$ est caractérisé à partir de la contrainte d'autosélection du type j ²⁰. Nous déterminons donc quelle contrainte d'autosélection est saturée. Nous caractérisons pour finir les contrats d'équilibre qui vérifient les points a) et b) de la définition de l'équilibre retenue.

5.4.2 Le programme

L'assureur détermine les contrats optimaux d'assurance exclusifs en ajoutant au programme de la section précédente, la contrainte d'antisélection (*sc3*). Celle-ci indique que chaque individu préfère le contrat lié à son type à tout autre, pour un effort optimal et un contrat donné.

$$U_j(C_j, H_j, e_j(C_j)) \geq U_j(C_{-j}, H_j, e_{-j}(C_{-j})) \quad \text{sc3}$$

avec x et $i \in \{0, 1\}$, $(j, -j) \in \{1, 2\}^2$, $j \neq -j$ et $U_j(C_h, H_h, e_h(C_h))$ l'utilité de l'individu de type j qui a souscrit un contrat d'assurance C_h et qui effectue l'effort optimal e_h compte tenu du contrat d'assurance h choisi.

Pour un type j donné, le contrat optimal d'assurance exclusif est déterminé par le programme suivant :

$$\begin{aligned} & \max_{q_j} p(e_j(C_j))u(w - \pi_j + q_j - d) + (1 - p(e_j(C_j)))u(w - \pi_j) \\ & - ce_j(C_j) + \delta((1 - p(e_j(C_j)))v(H_j) + p(e_j(C_j))v(H_j - \Delta H)) \\ & e_h(C_h) \in \arg \max_{x_i} p(x_i)u(w - \pi_h + q_h - d) + (1 - p(x_i))u(w - \pi_h) \\ & - cx_i + \delta((1 - p(x_i))v(H_j) + p(x_i)v(H_j - \Delta H)) \quad \text{sc1} \\ \pi_j & \geq p(e_j(C_j))q_j \quad \text{sc2} \\ & p(e_j(C_j))u(w - \pi_j + q_j - d) + (1 - p(e_j(C_j)))u(w - \pi_j) \\ & - ce_j + \delta((1 - p(e_j(C_j)))v(H_j) + p(e_j(C_j))v(H_j - \Delta H)) \\ & \geq p(e_{-j}(C_{-j}))u(w - \pi_{-j} + q_{-j} - d) + (1 - p(e_{-j}(C_{-j})))u(w - \pi_{-j}) \quad \text{sc3} \\ & - ce_{-j} + \delta((1 - p(e_{-j}(C_{-j})))v(H_j) + p(e_{-j}(C_{-j}))v(H_j - \Delta H)) \\ \text{avec } x_i & = e_i, i \in \{0, 1\}, (j, -j) \in \{1, 2\}^2, j \neq -j \text{ et } h \in \{j, -j\} \end{aligned}$$

Pour caractériser les contrats d'équilibre, nous raisonnons en deux temps. Tout d'abord, nous déterminons si les contrats de second rang d'aléa moral

²⁰Dans le modèle de Rothschild et Stiglitz (1976), l'individu de type H préfère le contrat de premier rang de l'individu de type L . A l'équilibre de second rang, le contrat de l'individu de type H est son contrat de premier rang. Quant à celui du type L , il vérifie la contrainte d'autosélection de l'individu de type H .

vérifient les contraintes d'autosélection. Ensuite, nous caractérisons les contrats d'assurance respectant la définition de l'équilibre retenue.

5.4.3 Etude des contraintes

En raison de l'impossibilité de contrats à subventions croisées et d'une offre unique de contrat par les compagnies d'assurance, les profits espérés sur les contrats ne peuvent être déficitaires. Par ailleurs, la séparabilité de la fonction d'utilité et la concurrence entre les compagnies d'assurance impliquent que les profits espérés sur les contrats ne peuvent être positifs. Par conséquent, les profits espérés sont nuls. La contrainte (*sc2*) est donc saturée²¹.

Les contrats de second rang d'aléa moral sont des contrats actuariels qui vérifient pour chaque agent la contrainte (*sc1*). Pour être des contrats d'équilibre, ils doivent également satisfaire la contrainte (*sc3*).

En fonction de la valeur du coût de l'effort, six cas sont à distinguer :

- **cas a** : $c < \underline{c}_2^{AM}$. Les individus reçoivent un contrat d'assurance complète et effectuent un effort de prévention.

- **cas b** : $c > \bar{c}_1^{AM}$. Les individus reçoivent un contrat d'assurance complète et n'effectuent pas d'effort de prévention.

- **cas c** : Soit $\Phi H < \overline{\Phi H}$ et $\underline{c}_2^{AM} < c < \underline{c}_1^{AM}$, soit $\Phi H > \overline{\Phi H}$ et $\underline{c}_2^{AM} < c < \bar{c}_2^{AM}$. Les deux types d'individus réalisent un effort de prévention. L'individu en bonne santé reçoit un contrat d'assurance partielle et l'individu en mauvaise santé une couverture complète.

- **cas d** : $\Phi H > \overline{\Phi H}$ et $\bar{c}_2^{AM} < c < \underline{c}_1^{AM}$. Les deux types d'individus reçoivent une couverture complète. L'individu en mauvaise santé fournit un effort préventif contrairement à l'individu en bonne santé.

- **cas e** : Soit $\Phi H < \overline{\Phi H}$ et $\bar{c}_2^{AM} < c < \bar{c}_1^{AM}$, soit $\Phi H > \overline{\Phi H}$ et $\underline{c}_1^{AM} < c < \bar{c}_1^{AM}$. L'individu en bonne santé ne réalise pas d'effort préventif et reçoit une couverture complète. l'individu en mauvaise santé fournit un effort préventif et reçoit une couverture partielle.

- **cas f** : $\Phi H < \overline{\Phi H}$ et $\underline{c}_1^{AM} < c < \bar{c}_2^{AM}$. Les deux individus réalisent un effort préventif et reçoivent une couverture partielle.

Cas a

Les contrats d'aléa moral de second rang vérifient les contraintes d'autosélection car les individus contractent la même police d'assurance malgré leurs états de santé différents.

Cas b

Avec les contrats d'aléa moral de second rang, la contrainte d'autosélection est vérifiée pour chaque type d'assuré.

²¹Lors de la caractérisation des contrats d'assurance d'équilibre, nous vérifierons qu'aucun contrat à profits positifs n'est préféré à un contrat actuariel.

Cas c

Bien que les deux types d'individus réalisent un effort de prévention, les contrats de second rang d'aléa moral sont différents. L'individu en mauvaise santé reçoit une couverture complète contrairement à l'individu en bonne santé. L'utilité d'un assuré croît avec le niveau de la couverture. Ainsi, pour des contrats actuariels, l'individu en bonne santé voudra prendre le contrat de l'individu en mauvaise santé. La contrainte d'autosélection de l'individu en bonne santé est saturée, contrairement à celle de l'individu en mauvaise santé. Le contrat des individus en mauvaise santé ne peut être proposé à l'équilibre. En effet, une couverture complète n'incite pas les individus en bonne santé à réaliser un effort, ce qui rend le contrat déficitaire s'il est proposé aux deux agents.

Cas d

Les deux individus reçoivent une couverture complète. Seul l'individu en mauvaise santé fournit un effort préventif. L'individu en bonne santé voudra prendre le contrat de celui en mauvaise santé car, pour une même couverture du dommage, il paiera une prime d'assurance inférieure²². La contrainte d'autosélection de l'individu en bonne santé est donc saturée contrairement à celle de l'individu en mauvaise santé. Le contrat de l'individu en mauvaise santé ne peut être proposé à l'équilibre. En effet, il est déficitaire s'il est proposé aux deux agents, car la valeur du coût de l'effort est telle que l'individu en bonne santé ne peut être incité à réaliser un effort.

Cas e

L'individu en bonne santé ne réalise pas d'effort préventif et reçoit une couverture complète. L'individu en mauvaise santé fournit un effort préventif et reçoit une couverture partielle. Avec ce contrat, il obtient une utilité supérieure à celle obtenue avec le contrat de l'individu en bonne santé. De la contrainte d'incitation de l'individu en mauvaise santé, nous obtenons que la couverture du dommage, notée \tilde{q} ²³, est supérieure à \underline{q} car $c < \overline{c_1^{AM}}$.

Or, comme $c > \overline{c_2^{AM}}$, l'individu en bonne santé n'effectuera un effort que si son niveau de couverture est inférieur à \underline{q} . Si ledit individu choisit le contrat de l'autre, il n'effectuera pas d'effort de prévention. Cet individu ne choisit pas le

²²L'individu en mauvaise santé ne prendra jamais le contrat de celui en bonne santé, car réaliser un effort préventif est optimal. Il s'acquitterait, pour un même montant de couverture, d'une prime d'assurance supérieure.

²³ \tilde{q} est définie par :

$$\begin{aligned} & p_1 u(w - p_1 \tilde{q} + \tilde{q} - d) + (1 - p_1) u(w - p_1 \tilde{q}) \\ & + \delta(p_1 v(H_1 - \Delta H) + (1 - p_1) v(H_1)) - c \\ = & u(w - p_0 d) + \delta(p_0 v(H_1 - \Delta H) + (1 - p_0) v(H_1)) \end{aligned}$$

contrat approprié à son type si :

$$p_0 u(w - p_1 \tilde{q} + \tilde{q} - d) + (1 - p_0) u(w - p_1 \tilde{q}) > u(w - p_0 d)$$

En utilisant la contrainte d'incitation à l'effort de cet individu, nous obtenons que :

$$p_0 u(w - p_1 \underline{q} + \underline{q} - d) + (1 - p_0) u(w - p_1 \underline{q}) = u(w - p_0 d)$$

Soit,

$$\begin{aligned} & p_0 u(w - p_1 \tilde{q} + \tilde{q} - d) + (1 - p_0) u(w - p_1 \tilde{q}) \\ > & p_0 u(w - p_1 \underline{q} + \underline{q} - d) + (1 - p_0) u(w - p_1 \underline{q}) \end{aligned}$$

Cette inéquation est toujours vérifiée car $\tilde{q} > \underline{q}$. L'individu en bonne santé ne choisira pas son contrat d'aléa moral de second rang. Si le contrat de second rang d'aléa moral des individus en mauvaise santé est proposé à l'équilibre, il sera déficitaire. La contrainte d'autosélection de l'individu en bonne santé, est donc saturée contrairement à celle de l'individu en mauvaise santé.

Cas f

Les deux individus reçoivent une couverture partielle différente (notée q_1 et q_2) pour les inciter à faire un effort. Pour une valeur de c donnée, nous devons déterminer la position relative de q_2 par rapport à q_1 .

Des contraintes d'incitations à l'effort des individus, nous obtenons que :

$$\begin{aligned} c = & (p_0 - p_1)(u(w - p_1 q_1) - u(w - p_1 q_1 + q_1 - d)) \\ & + \delta((p_0 - p_1)(v(H_1) - v(H_1 - \Delta H))) \end{aligned}$$

$$\begin{aligned} c = & (p_0 - p_1)(u(w - p_1 q_2) - u(w - p_1 q_2 + q_2 - d)) \\ & + \delta((p_0 - p_1)(v(H_2) - v(H_2 - \Delta H))) \end{aligned}$$

De ces deux équations, nous déduisons que :

$$\begin{aligned} & u(w - p_1 q_2) - u(w - p_1 q_2 + q_2 - d) \\ > & u(w - p_1 q_1) - u(w - p_1 q_1 + q_1 - d) \end{aligned}$$

Cette différence d'utilité est une fonction décroissante de la couverture du dommage. Ainsi $q_1 > q_2$.

L'utilité des assurés croît avec le niveau de la couverture (q). L'individu en bonne santé voudra donc prendre le contrat de l'agent en mauvaise santé. Or pour ce contrat, il n'effectuera plus d'effort, ce qui rend ce contrat déficitaire s'il est proposé aux deux agents. La contrainte d'autosélection de l'individu en bonne santé est donc saturée, contrairement à celle de l'individu en mauvaise santé.

Nous remarquons que si les contrats de risque moral de second rang sont proposés aux deux agents, l'agent en bonne santé veut contracter la police des agents en mauvaise santé. Seule la contrainte (*sc3*) de l'individu en bonne santé est toujours saturée. Si l'on considère uniquement le risque financier, nous retrouvons les résultats usuels des modèles d'antisélection²⁴. Les agents qui jouent le rôle des hauts risques du modèle de Rothschild et Stiglitz (1976) sont ici les individus en bonne santé.

5.4.4 Les contrats d'équilibre

L'équilibre du marché d'assurance santé dépend de la valeur du coût de l'effort. Six cas sont à considérer. Les contrats d'équilibre doivent vérifier les points *a* et *b* de la définition de l'équilibre et sont caractérisés par la notation *AI* pour Aléa moral et Inobservabilité de l'état de santé.

Le point *a* de la définition est toujours vérifié. En effet, en supposant que les compagnies d'assurance n'offrent qu'un seul contrat et que les subventions croisées entre contrats sont proscrites, le contrat ne peut être déficitaire.

Cas a

Pour de faibles valeurs du coût de l'effort, c'est-à-dire $c \leq \overline{c_2^{AM}}$, les deux individus choisissent le même contrat d'assurance complète du risque financier et effectuent un effort de prévention et ce, quelle que soit la dépréciation du capital santé. Quand $c < \overline{c_2^{AM}}$, les contraintes d'incitations à l'effort (*sc1*) des individus ne sont pas saturées pour une couverture complète du risque financier. L'effort des individus étant identique (ils présentent un même risque financier), leur contrat d'assurance est donc identique. Pour de faibles valeurs de l'effort, la non réalisation de celui-ci entraîne une dépréciation du capital sanitaire plus importante que l'économie du coût de l'effort d'autoprotection :

$$c \leq \overline{c_2^{AM}} = \delta((p_0 - p_1)(v(H_j) - v(H_j - \Delta H)))$$

Ce résultat fait l'objet de la proposition suivante :

Proposition 32 *En présence d'aléa moral et d'inobservabilité de l'état de santé, un contrat d'assurance complète du risque financier (avec réalisation d'un effort préventif) est offert aux individus quand $c \leq \overline{c_2^{AM}}$.*

Preuve :

Deux cas sont à étudier²⁵.

²⁴ Quand la probabilité d'occurrence du dommage diffère pour les deux individus.

²⁵ Un troisième cas ne sera pas étudié car il regroupe les deux autres cas étudiés.

Soit un contrat à profits positifs qui ne modifie que l'effort réalisé par un assuré. Des valeurs seuils des coûts de l'effort, nous déduisons qu'il existe un intervalle pour lequel l'individu en bonne santé modifie son effort contrairement à l'individu en mauvaise santé.

Pour l'individu en bonne santé, voir le second cas. Pour l'individu en mauvaise, voir le premier cas.

- Soit un contrat à profits positifs $C_a^\pi(\pi + p_1q_a, q_a)$ qui ne modifie pas la valeur de l'effort réalisé par les deux types d'individu. Ce contrat est préféré à un contrat d'assurance complète et actuarielle si :

$$\begin{aligned} & p_1u(w - \pi - p_1q_a + q_a - d) + (1 - p_1)u(w - \pi - p_1q_a) \\ > & u(w - p_1d) \end{aligned} \quad (\text{D})$$

Si $q_a = d$, l'inégalité (D) n'est jamais vérifiée car $u(w - p_1d) > u(w - \pi - p_1d)$.

Si $q_a < d$, nous savons que : $p_1u(w - \pi - p_1q_a + q_a - d) + (1 - p_1)u(w - \pi - p_1q_a) < u(w - \pi - p_1q_a)$. Par conséquent, (D) n'est jamais vérifiée.

- Soit un contrat à profits positifs $C_a^{\pi'}(\pi + p_0q_a, q_a)$ qui modifie la valeur de l'effort réalisé par les deux assurés. Ce contrat est préféré à un contrat d'assurance complète et actuarielle si :

$$\begin{aligned} & p_0u(w - \pi - p_0q_a + q_a - d) + (1 - p_0)u(w - \pi - p_0q_a) \\ > & u(w - p_1d) \end{aligned} \quad (\text{E})$$

Si $q_a = d$, l'inégalité (E) n'est jamais vérifiée car $u(w - p_1d) > u(w - \pi - p_0d)$ car p .

Si $q_a < d$, nous savons que : $p_0u(w - \pi - p_0q_a + q_a - d) + (1 - p_0)u(w - \pi - p_0q_a) < u(w - \pi - p_0q_a)$. Par conséquent, (E) n'est jamais vérifiée.

Le contrat actuariel $C_a^{AI}(p_1d, d)$ est un contrat d'équilibre car aucun contrat à profits positifs ne lui est préféré. CQFD.

Cas b

A contrario, quand le coût de l'effort est très grand, c'est-à-dire $c > \bar{c}_1^{AM}$, nous retrouvons un résultat usuel de l'aléa moral (Shavell, 1979) et de la combinaison de l'aléa moral et de l'antisélection (Chassagnon et Chiappori, 1995 ; Chopin-Kambia, 2000). Les agents, quel que soit leur type, reçoivent un contrat d'assurance complète sans effectuer d'effort de prévention.

Proposition 33 *En présence d'aléa moral et d'inobservabilité de l'état de santé, un contrat d'assurance complète du risque financier (sans réalisation d'un effort préventif) est offert aux individus quand $c > \bar{c}_1^{AM}$.*

Preuve :

- Soit un contrat à profits positifs $C_b^\pi(\pi + p_0q_b, q_b)$. Ce contrat est préféré à un contrat d'assurance complète et actuarielle si

$$\begin{aligned} & p_0u(w - \pi - p_0q_b + q_b - d) + (1 - p_0)u(w - \pi - p_0q_b) \\ > & u(w - p_0d) \end{aligned} \quad (\text{F})$$

Si $q_b = d$, l'inégalité (F) n'est jamais vérifiée car $u(w - p_0d) > u(w - \pi - p_0d)$ car $\pi > 0$.

Si $q_b < d$, nous savons que :

$$p_0u(w - \pi - p_0q_b + q_b - d) + (1 - p_0)u(w - \pi - p_0q_b) < u(w - \pi - p_0q_b)$$

Par conséquent, (F) n'est jamais vérifiée.

Le contrat actuariel $C_b^{AI}(p_0d, d)$ est donc un contrat d'équilibre. CQFD.

Cas c

Les deux individus font un effort préventif et reçoivent une couverture du dommage différente. Toutefois, l'individu en bonne santé reçoit une couverture partielle du risque financier pour l'inciter à faire un effort. Cet individu désire le contrat de l'agent en mauvaise santé²⁶. Le contrat de ce dernier individu doit vérifier la contrainte d'autosélection des individus en bonne santé et la contrainte d'incitation à l'effort des individus en mauvaise santé. Soit,

$$\begin{aligned} & p_1u(w - p_1q_{c2} + q_{c2} - d) + (1 - p_1)u(w - p_1q_{c2}) \\ = & p_1u(w - p_1q_{c1} + q_{c1} - d) + (1 - p_1)u(w - p_1q_{c1}) \end{aligned} \quad (G)$$

et

$$\begin{aligned} & p_1u(w - p_1q_{c1} + q_{c1} - d) + (1 - p_1)u(w - p_1q_{c1}) \\ & - c + \delta((1 - p_1)v(H_1) + p_1v(H_1 - \Delta H)) \\ > & p_0u(w - p_1q_{c1} + q_{c1} - d) + (1 - p_0)u(w - p_1q_{c1}) \\ & + \delta((1 - p_0)v(H_1) + p_0v(H_1 - \Delta H)) \end{aligned} \quad (H)$$

avec q_{c_j} la couverture du contrat de l'individu de type j dans le cas c.

De l'équation (G), nous déduisons que $q_{c1} = q_{c2} < d$. L'équation (H) est vérifiée car $\underline{c}_1^{AM} - c > 0$ et $q_{c1} < d$. En effet,

$$(p_0 - p_1)(u(w - p_1q_{c1}) - u(w - p_1q_{c1} + q_{c1} - d)) + \underline{c}_1^{AM} - c > 0$$

Proposition 34 *En présence d'aléa moral et d'inobservabilité de l'état de santé, un contrat mélangeant d'assurance partielle du risque financier (avec réalisation d'un effort préventif) est offert quand $\Phi H < \bar{\Phi} \bar{H}$ et $\underline{c}_2^{AM} < c < \underline{c}_1^{AM}$, ou $\Phi H > \bar{\Phi} \bar{H}$ et $\underline{c}_2^{AM} < c < \bar{c}_2^{AM}$*

Preuve :

Deux cas sont à étudier.

- Soit un contrat à profits positifs $C_c^\pi(\pi + p_1q_c, q_c)$ qui ne modifie pas la valeur de l'effort réalisé par les deux assurés²⁷. Ce contrat est préféré à un contrat d'assurance complète et actuarielle si :

$$\begin{aligned} & p_1u(w - \pi - p_1q_c + q_c - d) + (1 - p_1)u(w - \pi - p_1q_c) \\ > & p_1u(w - p_1q_{c1} + q_{c1} - d) + (1 - p_1)u(w - p_1q_{c1}) \end{aligned} \quad (I)$$

Si $q_c = q_{c1}$, l'inégalité (I) n'est jamais vérifiée car $\pi > 0$.

²⁶ L'individu de type 1 préfère son contrat car il obtient une couverture complète du risque financier.

²⁷ $q_c \leq q_{1c}$ sinon l'individu en bonne santé n'effectuerait pas d'effort.

Si $q_c < q_{c1}$, nous savons que :

$$\begin{aligned} & p_1 u(w - \pi - p_1 q_c + q_c - d) + (1 - p_1) u(w - \pi - p_1 q_c) \\ < & p_1 u(w - \pi - p_1 q_{1c} + q_{1c} - d) + (1 - p_1) u(w - \pi - p_1 q_{1c}) \\ < & p_1 u(w - p_1 q_{c1} + q_{c1} - d) + (1 - p_1) u(w - p_1 q_{c1}) \end{aligned}$$

Par conséquent, (I) n'est jamais vérifiée.

- Soit un contrat à profits positifs $C_c^{\pi'}(\pi + p_0 q_c, q_c)$ qui modifie la valeur de l'effort réalisé par les deux individus. Ce contrat est préféré à un contrat d'assurance partielle si :

$$\begin{aligned} & p_0 u(w - \pi - p_0 q_c + q_c - d) + (1 - p_0) u(w - \pi - p_0 q_c) \\ > & p_1 u(w - p_1 q_{c1} + q_{c1} - d) + (1 - p_1) u(w - p_1 q_{c1}) \quad (\text{J}) \end{aligned}$$

Les assurés préfèrent un contrat d'assurance partielle (avec la réalisation d'un effort) à un contrat d'assurance complète si :

$$\begin{aligned} & p_1 u(w - p_1 q_{c1} + q_{c1} - d) + (1 - p_1) u(w - p_1 q_{c1}) \\ & - c + \delta((p_0 - p_1)(v(H_j) - v(H_j - \Delta H))) \\ > & u(w - p_0 d) \end{aligned}$$

Or, nous savons que

$$\begin{aligned} & u(w - p_0 d) \\ > & p_0 u(w - \pi - p_0 q_c + q_c - d) + (1 - p_0) u(w - \pi - p_0 q_c) \end{aligned}$$

car $\pi > 0$ et $q_c < d$. Par conséquent, (J) n'est jamais vérifiée.

Le contrat actuariel $C_c^{AI}(p_1 q_{c1}, q_{c1})$ est un contrat d'équilibre car aucun contrat à profits positifs ne lui est préféré. CQFD.

Cas d

En présence d'aléa moral, les deux individus reçoivent une couverture complète mais n'effectuent pas le même effort. Seul l'individu en mauvaise santé effectue un effort. Les deux individus reçoivent une couverture totale du risque financier tandis que l'individu en mauvaise santé s'acquitte d'une prime moindre. L'individu en bonne santé ne choisira pas son contrat d'assurance d'aléa moral de second rang. Par conséquent, le contrat des individus en mauvaise santé $C_d^{AI}(p_1 q_{d1}, q_{d1})$ doit vérifier la contrainte d'autosélection des individus en bonne santé. Soit,

$$p_0 u(w - p_1 q_{d1} + q_{d1} - d) + (1 - p_0) u(w - p_1 q_{d1}) = u(w - p_0 d)$$

Nous en déduisons que $q_{d1} = \underline{q}$. Il est aisé de montrer que la contrainte d'incitation à l'effort de l'individu en mauvaise santé est vérifiée pour cette couverture partielle car $c < \bar{c}_1^{AM}$.

Proposition 35 *En présence d'aléa moral et d'inoobservabilité de l'état de santé, un menu de contrats d'assurance séparateurs est offert à l'équilibre quand $\Phi H > \overline{\Phi H}$ et $\overline{c}_2^{AM} < c < \underline{c}_1^{AM} : C_d^{AI}(p_0 d, d)$ et $C_d^{AI}(p_1 \underline{q}, \underline{q})$.*

Preuve :

Deux cas sont à étudier.

- Soit un contrat à profits positifs $C_d^\pi(\pi + p_1 q_d, q_d)$ qui ne modifie pas la valeur de l'effort réalisé par l'individu en mauvaise santé. Deux sous-cas apparaissent :

a) Soit ce contrat est préféré par les deux agents.

Nous obtenons alors :

$$\begin{aligned} & p_1 u(w - \pi - p_1 q_d + q_d - d) + (1 - p_1) u(w - \pi - p_1 q_d) \\ > & p_1 u(w - p_1 \underline{q} + \underline{q} - d) + (1 - p_1) u(w - p_1 \underline{q}) \end{aligned} \quad (\text{K})$$

$$\begin{aligned} & p_0 u(w - \pi - p_1 q_d + q_d - d) + (1 - p_0) u(w - \pi - p_1 q_d) \\ > & u(w - p_0 d) \end{aligned} \quad (\text{L})$$

Pour que ce contrat soit profitable, il faut que $\pi + p_1 q_d - \overline{p} q_d > 0$ (profit = primes reçues – remboursements).

L'équation (K) devient si $\pi + p_1 q_d - \overline{p} q_d = a$

$$\begin{aligned} & p_1 u(w - a - \overline{p} q_d + q_d - d) + (1 - p_1) u(w - a - \overline{p} q_d) \\ > & p_1 u(w - p_1 \underline{q} + \underline{q} - d) + (1 - p_1) u(w - p_1 \underline{q}) \end{aligned}$$

Or nous avons supposé que le contrat mélangeant ne pouvait déstabiliser le menu de contrats d'équilibre. Ainsi, l'individu en mauvaise santé ne choisira jamais le contrat mélangeant actuariel, encore moins si ce contrat est à profits positifs. L'équation (K) n'est jamais vérifiée.

Pour que ce contrat soit à profits positifs, réalisés sur les assurés en bonne santé, il faut $\pi + p_1 q_d - p_0 q_d = b$.

L'équation (L) devient :

$$\begin{aligned} & p_0 u(w - b - p_0 q_d + q_d - d) + (1 - p_0) u(w - b - p_0 q_d) \\ > & u(w - p_0 d) \end{aligned}$$

Or, nous savons que ce contrat à profit positifs n'est jamais choisi par l'individu en bonne santé (voir preuve du cas b).

b) Soit le contrat à profits positifs est préféré uniquement par un assuré.

Si un contrat à profits positifs est préféré par l'individu en mauvaise santé, il l'est forcément par l'individu en bonne santé.

Si l'individu en mauvaise santé préfère le contrat à profits positifs alors

$$\begin{aligned} & p_1 u(w - \pi - p_1 q_d + q_d - d) + (1 - p_1) u(w - \pi - p_1 q_d) \\ > & p_1 u(w - p_1 \underline{q} + \underline{q} - d) + (1 - p_1) u(w - p_1 \underline{q}) \end{aligned}$$

Cette inéquation est vérifiée si $\underline{q} < q_d < d$. Si cette inéquation est vérifiée pour p_1 , elle l'est pour p_0 car $q_d > \underline{q}$. Soit,

$$\begin{aligned} & p_0 u(w - \pi - p_1 q_d + q_d - d) + (1 - p_0) u(w - \pi - p_1 q_d) \\ & > p_0 u(w - p_1 \underline{q} + \underline{q} - d) + (1 - p_0) u(w - p_1 \underline{q}) \end{aligned}$$

De la contrainte d'autosélection de l'individu nous obtenons que

$$p_0 u(w - p_1 \underline{q} + \underline{q} - d) + (1 - p_0) u(w - p_1 \underline{q}) = u(w - p_0 d)$$

Ainsi, nous déduisons que

$$\begin{aligned} & p_0 u(w - \pi - p_1 q_d + q_d - d) + (1 - p_0) u(w - \pi - p_1 q_d) \\ & > u(w - p_0 d) \end{aligned}$$

L'individu en mauvaise santé préfère également le contrat à profits positifs.

Si un contrat est préféré par un seul individu alors il ne s'agit que de l'individu en bonne santé. Pour que ce contrat soit profitable, la prime d'assurance est égale à $\pi + p_1 q_d - p_0 q_d = b$. Nous savons que ce contrat ne sera jamais choisi par les agents en bonne santé. (voir la preuve du cas b).

- Soit un contrat à profits positifs $C_d^{\pi'}(\pi + p_0 q'_d, q'_d)$ qui modifie la valeur de l'effort réalisé par l'individu en mauvaise santé. Il ne choisira jamais le contrat $C_d^{\pi'}(\pi + p_0 q'_d, q'_d)$. (voir la preuve du cas b).

Il n'existe aucun contrat à profits positifs qui déstabilise le menu de contrats séparateurs d'équilibre. CQFD.

Cas e

Les deux individus ne présentent pas le même risque financier. Seul l'individu en mauvaise santé effectue un effort. Comme $q_1 > \underline{q}$, l'individu en bonne santé préfère le contrat des agents en mauvaise santé à son contrat de couverture complète et actuariel. Le contrat des assurés en mauvaise santé $C_e^{AI}(p_1 q_{1e}, q_{1e})$ doit donc vérifier la contrainte d'antisélection des agents en bonne santé. Soit

$$p_0 u(w - p_1 q_{1e} + q_{1e} - d) + (1 - p_0) u(w - p_1 q_{1e}) = u(w - p_0 d)$$

Nous en déduisons que $q_{1e} = \underline{q}$. Nous devons nous assurer que, pour cette couverture, la contrainte d'incitation à l'effort de l'individu en mauvaise santé est vérifiée et que ce contrat est préféré à un contrat d'assurance complète du risque financier avec effort nul. Comme $c \in \left[\underline{c}_1^{AM}, \overline{c}_1^{AM} \right]$, alors pour $q_{1e} = \underline{q}$, ces deux conditions sont vérifiées. Ainsi, les contrats optimaux d'assurance exclusifs sont séparateurs.

Proposition 36 *En présence d'aléa moral et d'inobservabilité de l'état de santé, un menu de contrats d'assurance séparateurs est offert à l'équilibre quand $\Phi H < \overline{\Phi H}$ et $\underline{c}_2^{AM} < c < \overline{c}_1^{AM}$ ou $\Phi H > \overline{\Phi H}$ et $\underline{c}_1^{AM} < c < \overline{c}_1^{AM}$: $C_e^{AI}(p_0 d, d)$ et $C_e^{AI}(p_1 \underline{q}, \underline{q})$.*

Preuve :

(Voir la preuve du cas d). CQFD.

Cas f

Les deux types d'individus font un effort préventif et reçoivent une couverture partielle du risque financier pour les inciter à faire un effort. Pour une valeur de c donnée, nous savons que la couverture de l'individu en mauvaise santé (q_1) est supérieure à celle de l'individu en bonne santé (q_2).

Comme précédemment, l'individu en bonne santé veut prendre le contrat des individus en mauvaise santé. Le contrat des individus en mauvaise santé $C_f^{AI}(p_1q_{1f}, q_{1f})$ doit vérifier la contrainte d'autosélection des agents en bonne santé. Nous en déduisons que $q_{1f} = q_2$. Pour cette valeur q_{1f} , la contrainte d'incitations des agents en mauvaise santé est vérifiée tandis que ce contrat est préféré à un contrat d'assurance complète du risque financier sans effort. Ainsi, les contrats optimaux d'assurance exclusifs sont identiques.

Proposition 37 *En présence d'aléa moral et d'inobservabilité de l'état de santé, un contrat mélangeant d'assurance partielle du risque financier (avec réalisation d'un effort préventif) $C_f^{AI}(p_1q_2, q_2)$ est offert quand $\Phi H < \overline{\Phi H}$ et $\underline{c}_1^{AM} < c < \overline{c}_2^{AM}$.*

Preuve :

(Voir la preuve du cas c). CQFD.

En présence d'aléa moral et d'inobservabilité de l'état de santé, deux types d'équilibre émergent : un équilibre mélangeant et un équilibre séparateur. L'équilibre est dit mélangeant car des assurés aux caractéristiques sanitaires différentes obtiennent le même contrat. Toutefois, en termes de risque financier, il est inapproprié de considérer ce contrat comme mélangeant car ces individus présentent le même risque financier, seul risque pertinent pour les compagnies d'assurance. L'équilibre est séparateur car les individus aux caractéristiques sanitaires différentes obtiennent des contrats différents. Cependant, aux individus de type haut risque sanitaire est offert un contrat d'assurance partielle contrairement aux individus de type bas risque sanitaire. Ces résultats sont contraires à ceux obtenus par RS.

5.5 Conclusion

Nous avons caractérisé les contrats d'assurance en présence d'un risque sanitaire quand l'effort de prévention et l'état de santé de l'individu sont inobservables. Nous avons montré, contrairement aux résultats usuels de Chassagnon et Chiappori (1995) et de Chopin-Kambia (2000), qu'un contrat d'assurance complète du risque financier ne désincite pas à effectuer un effort quand le coût d'autoprotection est faible. En effet, ne pas effectuer d'effort n'entraîne pas uniquement une économie d'utilité. Elle provoque une diminution de l'espérance d'utilité liée à l'état de santé de l'individu. Par ailleurs, pour des contrats séparateurs, les individus en bonne santé reçoivent une couverture complète du risque financier car, en termes financiers, ce sont des mauvais risques. Leur état de santé les incite moins que les individus en mauvaise santé à effectuer un effort

d'autoprotection. Ces derniers reçoivent par conséquent un contrat d'assurance partielle du risque financier.

Une extension possible de notre modélisation consisterait à supposer un effort continu mais cette hypothèse rendrait l'analyse très complexe. Plutôt que d'introduire cette complication technique, il vaudrait mieux considérer que chaque partie contractuelle (assureur ou assuré) possède un avantage informatif par rapport à l'autre. L'assuré ne connaîtrait pas son état de santé tandis que l'assureur n'observerait pas le niveau de l'effort d'autoprotection de l'assuré. Cette voie de recherche suppose de prolonger les travaux de Villeneuve (2000) qui caractérisaient les contrats d'assurance lorsque l'assureur connaît le risque de l'assuré contrairement à ce dernier. L'individu déduirait du contrat offert par la compagnie d'assurance son état de santé. En fonction de celui-ci, il effectuerait ou non un effort préventif.

Chapitre 6

Antisélection et assurance : un état des lieux

“Comment ce constant trop peu se mue et devient ce trop, vide”,
Rilke.

6.1 Introduction

L'antisélection est un phénomène qui survient dans bien des relations d'assurance. Les marchés d'assurance automobile, santé et vie sont caractérisés par ce phénomène. En effet, la méconnaissance du risque encouru par l'assuré est susceptible de menacer l'existence de l'assurance privée.

“Insurance markets differ from most markets because in insurance markets competition can destroy the market rather than make it better”, Rothschild et Stiglitz (1997).

En général, les modèles d'antisélection considèrent que la population se scinde en deux classes de risques : les hauts risques et les bas risques. La différence de risque a trait à une probabilité d'occurrence du sinistre différente. Les autres caractéristiques des individus sont identiques (revenu, montant du dommage, aversion au risque). Les principaux résultats et extensions ont été établis en supposant que les individus possédaient une autre caractéristique différente.

Rothschild et Stiglitz (1976), noté par la suite RS, indiquent quelques extensions possibles à leur modélisation. Ils ont suggéré de reconsidérer leurs résultats dans un cadre où les agents se différencient par leur aversion au risque. Très peu d'articles traitent conjointement d'une double asymétrie d'information. Deux voies de recherches sont à distinguer.

- Tout d'abord, les agents se différencient par leur probabilité d'occurrence et leur aversion au risque. Landsberger et Meilijson (1994) ont caractérisé les contrats d'assurance d'une société d'assurance en situation de mo-

nopole offrant une assurance aléatoire¹. Dans ce cas, les contrats sont toujours séparateurs. L'individu le moins risquophobe accepte l'assurance aléatoire ce qui déserte la contrainte d'autosélection et permet d'augmenter l'utilité de cet individu. Villeneuve (1996) dans un cadre de concurrence où ne subsiste qu'une unique asymétrie d'information (l'individu le plus risquophobe possède la probabilité d'occurrence la moins élevée), montre que des profits positifs peuvent être réalisés à l'équilibre aussi bien avec une assurance déterministe qu'aléatoire. Smart (2000) obtient des résultats similaires pour un contrat déterministe, avec des agents différenciés par la probabilité d'occurrence de l'événement et leur aversion au risque.

- La seconde voie considère que les agents se différencient par la probabilité d'occurrence du dommage et la taille du dommage. Seuls deux travaux ont repris le modèle de RS en introduisant des niveaux de dommage différents parmi les assurés. Tout d'abord, Doherty et Schlesinger (1995) supposent que le montant du dommage est affecté par un "bruit blanc", c'est-à-dire une variable aléatoire d'espérance nulle intervenant de façon additive. Ils considèrent deux classes de risques (hauts et bas risques). Ils montrent que cette variable aléatoire permet aux bas risques de souscrire un niveau de couverture plus élevé. Ensuite, Doherty et Jung (1993) supposent que la fonction de densité d'occurrence du dommage des individus et le dommage observable sont différents. Seul un cas à deux agents est envisagé. L'individu haut risque est caractérisé par une espérance de dommage plus élevée que l'individu bas risque. Sous certaines conditions, l'équilibre de premier rang (assurance complète pour tous les agents) émerge.

Ce chapitre a pour objectif de présenter les résultats de la littérature et de préciser certains résultats lorsque l'assureur est soumis à une double déficience informationnelle dans le cas d'un contrat exclusif². Nous montrons la coexistence, sur le marché de l'assurance, de mutuelles et de compagnies d'assurance quand les contrats génèrent des profits et l'existence d'un contrat mélangeant à couverture complète.

Pour présenter les résultats de la littérature, nous n'étudions qu'un cadre à deux agents. Dans un premier temps, les propriétés de l'équilibre RS sont rappelées. Dans un deuxième temps, nous exposons trois extensions pertinentes pour la suite de notre propos. Les individus sont différenciés par la probabilité d'occurrence d'un sinistre et par l'aversion au risque ou par la taille du dommage. Dans un dernier temps et dans un cadre plus général, les individus se différencient par leur fonction d'utilité et la fonction de densité d'occurrence du dommage. Cette

¹Une assurance aléatoire stipule que le montant du remboursement n'est pas connu ex ante. Il est déterminé aléatoirement ex post.

²D'autres auteurs ont modifié le cadre du modèle. Fluet et Pannequin (1997) supposent que les individus sont affectés par plusieurs sources de risques monétaires. Ils montrent que, si l'assureur d'un risque donné possède une information parfaite concernant la couverture des autres risques, alors la couverture dudit risque n'est pas toujours partielle. Par ailleurs, Doherty et Jung (1995) considèrent que les individus se différencient par la fonction de densité d'occurrence du dommage et non par la distribution des dommages. Si l'équilibre considéré est celui de RS, alors l'introduction de dommages différents modifie le niveau de couverture.

présentation des résultats achevée, nous discutons dans un troisième temps de l'existence du phénomène d'antisélection sur les marchés d'assurance et plus spécifiquement en santé.

6.2 Le modèle de RS (1976)

Le modèle fondateur en économie de l'assurance ayant trait à la sélection adverse est le modèle de RS (1976). Les caractéristiques de l'équilibre proviennent des hypothèses comportementales des compagnies d'assurance (équilibre de Nash). Nous présentons ci-après, les propriétés de ce modèle fondateur.

6.2.1 Le cadre

Le modèle, développé dans les travaux de RS (1976) et Wilson (1977)³, considère des individus averses au risque et soumis à un risque monétaire ainsi que des compagnies d'assurance neutres au risque et concurrentes. Le contrat souscrit par l'individu est exclusif, c'est-à-dire qu'un seul contrat peut être acquis.

Ces agents sont dotés d'un revenu W au début de la période. Ils encourent, au cours de la période considérée, un risque (maladie) avec une probabilité p qui génère une perte de revenu notée D avec $D < W$. L'aversion au risque est caractérisée par une fonction d'utilité de von Neumann-Morgenstern notée $u(R_i)$, strictement croissante, concave et trois fois différentiable où R_i est la richesse aléatoire de l'individu dans les deux états du monde possibles (état 1 : pas de maladie ; état 2 : état de maladie). Ces individus ne diffèrent que par la probabilité d'occurrence de la maladie. Une proportion de la population notée λ est dite à haut risque (H) car sa probabilité d'occurrence de la maladie (p_H) est supérieure à celle (p_L) de la population dite à bas risque (L) avec $p_L < p_H$. Nous pouvons donc définir le risque moyen de la population de subir une maladie : $\bar{p} = \lambda p_H + (1 - \lambda)p_L$. Chaque individu connaît son type H ou L .

Les compagnies d'assurance sont neutres au risque et offrent un contrat unique. Elles offrent, à travers la souscription d'un contrat d'assurance, le partage du risque encouru par les assurés⁴. Le contrat d'assurance noté C est défini par un couple (π, q) où π est le montant de la prime et q le remboursement total ou partiel du dommage monétaire. On suppose toujours que la couverture est positive et n'excède pas le montant du dommage (exclusion de la surassurance et de l'assurance négative) : $0 \leq q \leq D$. Toutefois, bien que les compagnies d'assurance connaissent parfaitement les valeurs des caractéristiques des individus, elles ne peuvent, a priori, identifier le type de chaque individu.

Chaque individu accepte le meilleur contrat d'assurance qui lui est offert si son utilité est supérieure à l'utilité d'un état sans assurance notée V_j^0 (contrainte de participation). Soit j l'indice qui distingue les deux types d'individus, avec

³Les premiers considèrent un équilibre du marché d'assurance en situation de concurrence tandis que le second une situation de monopole.

⁴Le partage du risque reste possible si les compagnies d'assurance sont risquophobes. Il suffit que ces dernières soient moins risquophobes que les assurés.

$j \in \{H, L\}$

$$\begin{aligned} V_j(C) &= p_j u(W - D + q - \pi) + (1 - p_j) u(W - \pi) \\ V_j(C) &\geq V_j^0 = p_j u(W - D) + (1 - p_j) u(W) \end{aligned}$$

Pour un contrat d'assurance C , nous pouvons calculer la pente des courbes d'indifférence des consommateurs qui est égale au taux marginal de substitution entre les niveaux de richesse entre les deux états du monde. Ainsi, $\frac{dR_2}{dR_1} |_C = -\frac{1-p_j}{p_j} \frac{u'(R_1)}{u'(R_2)} |_C$. Pour tout contrat C , la pente d'une courbe d'indifférence d'un individu de type H est donc toujours inférieure à celle d'un agent de type L (single crossing condition). Cette contrainte permet, lorsque l'équilibre du marché d'assurance existe, d'obtenir des contrats séparateurs.

Le profit des compagnies d'assurance réalisé sur la souscription d'un contrat par un agent de type j est :

$$\Pi = p_j(\pi - q) + (1 - p_j)\pi = \pi - p_j q. \text{ Celles-ci offrent un contrat si } \Pi \geq 0.$$

6.2.2 L'équilibre

Les caractéristiques de l'équilibre dépendent de la façon dont réagit une compagnie d'assurance aux actions des ses rivales. Dans un environnement concurrentiel, chaque compagnie adopte une stratégie de Nash, c'est-à-dire qu'elle considère l'action de sa rivale comme donnée. Elle n'anticipe pas la réaction des assureurs rivaux à des modifications de sa propre stratégie.

La notion d'équilibre retenue dans les modèles concurrentiels d'assurance est donc celle de l'équilibre de Nash⁵. Dans RS, le jeu se déroule en deux étapes. A la première étape, les compagnies d'assurance (agents non informés du type du consommateur) offrent simultanément un ou des contrats en anticipant les caractéristiques de sa clientèle et en supposant donnée l'offre des compagnies rivales. A la seconde étape, les consommateurs opèrent un choix parmi ces contrats et révèlent ex-post leur type en choisissant le contrat qui leur est propre.

Tout contrat d'assurance offert à l'équilibre ne sera pas déficitaire. De plus, un contrat ne pourra réaliser de profits positifs. En effet, si un contrat profitable est proposé par une compagnie d'assurance, une autre offrira un contrat moins profitable ce qui augmente l'utilité de l'agent. La concurrence entre compagnies génère des profits nuls à l'équilibre. La définition d'un équilibre RS est donc :

Définition 38 *L'équilibre RS est un ensemble de contrats tels que :*

- a) *chaque contrat est non déficitaire*
- b) *aucun contrat hors équilibre n'est profitable.*

Avec une telle définition, l'équilibre concurrentiel en information parfaite est déterminé par la maximisation de la fonction d'utilité de chaque type sous

⁵Nous ne considérons pas des équilibres plus sophistiqués : type Wilson (1977) qui intègre l'anticipation de la réaction des concurrents ; type Miyasaki (1977) - Spence (1978) - Wilson qui ne nécessite plus des contrats actuariels ; sans oublier Grossman (1979) et Riley (1979). Rothschild et Stiglitz (1997) soulignaient que les assureurs ne pouvaient avoir des comportements élaborés comme ceux susmentionnés.

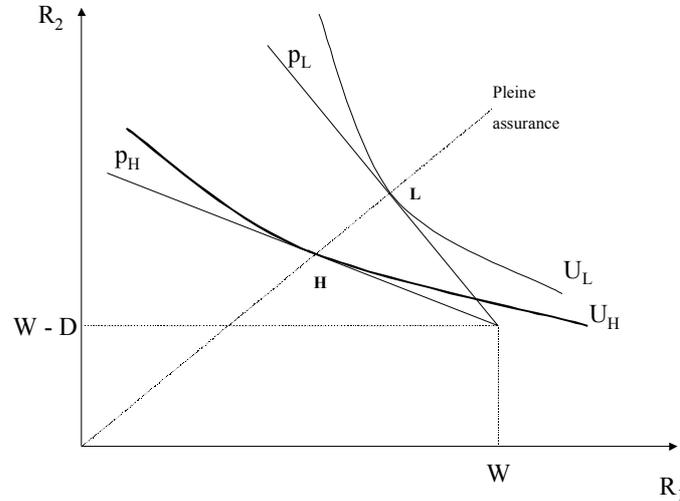


FIG. 6.1: Contrats d'assurance en information parfaite

contrainte de profits nuls. Ce programme s'énonce donc :

$$C_j \in \arg \max_{C_j} V_j(C_j) + \mu(\pi - p_j q)$$

avec μ le multiplicateur de Lagrange associé à la contrainte de profits. La condition du premier ordre conduit pour chaque type à l'égalisation des utilités marginales calculées dans les deux états de la nature. Ainsi, les contrats d'assurance en information parfaite offrent aux deux types une assurance complète au prix actuariel de chaque type. Soit $C_j = (p_j D, D)$, noté j (cf. figure 6-1).

avec p_L, p_H les droites actuarielles des bas risque et des hauts risque.

En présence d'information imparfaite, les contrats d'assurance d'information parfaite ne peuvent être reconduits. En effet, le contrat du type L est toujours choisi par le type H , car il obtient une même couverture du risque pour une prime plus faible⁶. Ce comportement d'antisélection amène les compagnies maintenant ces contrats à réaliser des déficits, ce qui est contraire à la définition de l'équilibre RS retenue.

En information imparfaite, les assureurs qui n'observent pas le type des individus devront tenir compte de ce comportement d'antisélection lors de l'offre des contrats. Dans un tel cadre, deux types de contrats correspondent à la définition de l'équilibre RS :

- contrats séparateurs (les deux types s'autosélectionnent et souscrivent des contrats différenciés),

⁶ A contrario, l'individu de type L ne choisira pas le contrat de type H car, pour une même indemnité, il payerait une prime plus élevée.

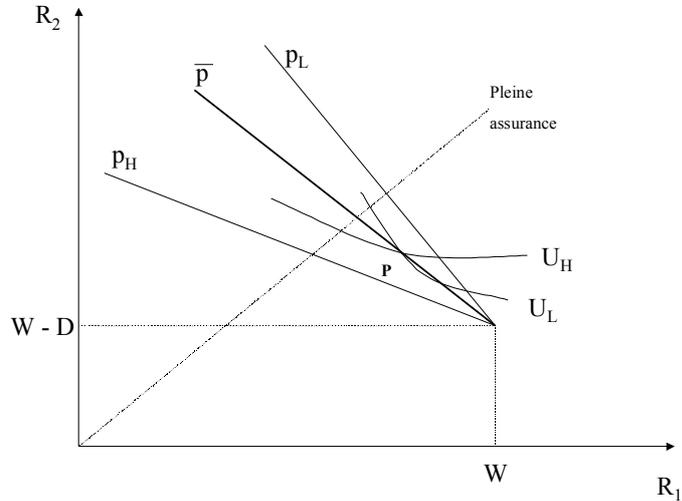


FIG. 6.2: Inexistence d'un contrat mélangeant

- contrat mélangeant (un seul contrat est souscrit par les deux types d'individus).

Seuls les contrats séparateurs sont des contrats d'équilibre. En effet, en utilisant la figure 6-2, il est aisé de voir que, s'il existe un contrat mélangeant non profitable P sur \bar{p} (la droite actuarielle du risque moyen), alors n'importe quelle compagnie peut proposer un contrat d'assurance dans la zone grisée qui augmente l'utilité des individus de type L et diminue celle de type H . Ce contrat est donc profitable car il n'attire que les L et est situé sous la droite actuarielle des individus de type L .

L'individu de type H obtient en assurance imparfaite le contrat d'assurance parfaite car ce contrat ne sera pas choisi par l'individu de type L . Il faut néanmoins s'assurer que l'individu de type H ne choisisse pas le contrat de l'individu de type L .

Nous pouvons donc énoncer le programme qui caractérise les contrats d'assurance séparateurs (Dionne et al., 2000).

$$C_L \in \arg \max V_L(C_L)$$

sous la contrainte d'incitation

$$V_H(C_H) \geq V_H(C_L)$$

sous la contrainte de profits non négatifs

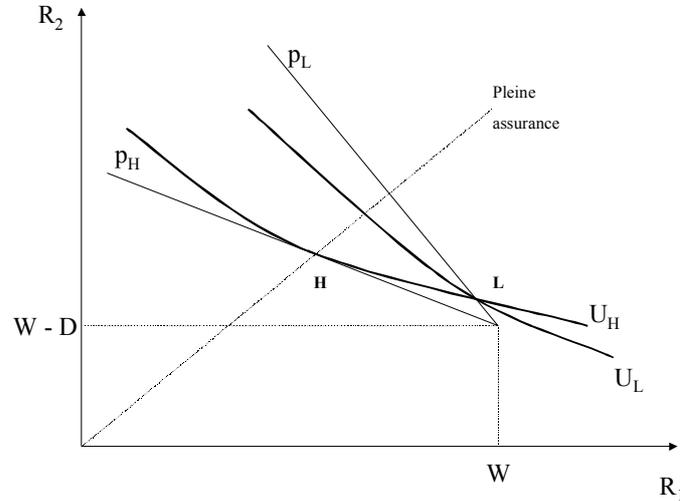


FIG. 6.3: Contrats d'assurance en information imparfaite

$$\pi_L - p_L q_L \geq 0$$

avec $C_H = (p_H D, D)$.

L'individu de type L obtient moins que la pleine assurance. L'ensemble des contrats d'assurance est $C = (C_H, C_L)$ avec $C_H = (p_H D, D)$ et $C_L = (p_L q_L, q_L)$ avec $0 < q_L < D$ où q_L est défini par $V_H(C_H) = V_H(C_L)$ (cf. figure 6-3).

Propriétés de l'équilibre RS

Les propriétés d'un tel équilibre, quand il existe, sont les suivantes :

Propriété 1 : *l'équilibre est non mélangeant.*

Propriété 2 : *les hauts risques obtiennent la pleine assurance.*

Propriété 3 : *la contrainte d'incitation des hauts risques est saturée.*

Propriété 4 : *les bas risques obtiennent moins que la pleine assurance.*

Propriété 5 : *les entreprises réalisent des profits nuls.*

(Pour une démonstration complète de toutes ces propriétés, voir Dionne et al., 2000).

Une fois définies les propriétés de l'équilibre RS, il est naturel de s'intéresser à son existence. Un équilibre RS ne peut être mélangeant. Toutefois, un équilibre mélangeant peut dominer au sens de Pareto l'équilibre séparateur, c'est-à-dire que les deux types d'agents obtiennent une utilité supérieure avec un tel contrat. Dans ce cas, aucun équilibre n'existe. L'équilibre RS existe si la proportion de

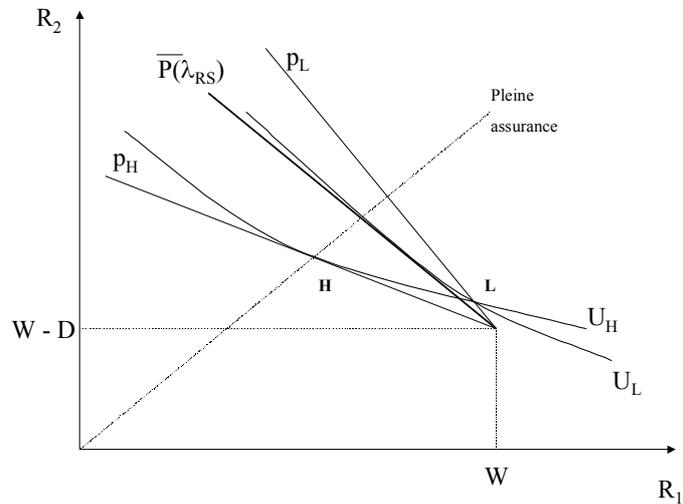


FIG. 6.4: Existence de l'équilibre RS

mauvais risques n'est pas trop faible. En effet, si λ est faible, les agents de types L acceptent de subventionner ceux de type H . A contrario, si λ est grand, les individus de type L refusent de subventionner ceux de type H et préfèrent le contrat séparateur. Quelle que soit la proportion d'individus de types H , ceux-ci obtiennent toujours une utilité supérieure avec un contrat mélangeant. (cf. figure 6-4)

Les individus de type L préfèrent le contrat séparateur si leur utilité est supérieure à l'utilité obtenue avec un contrat mélangeant. Ainsi, les individus de type L sont indifférents entre les deux contrats si leur courbe d'indifférence est tangente à la droite actuarielle de risque moyen.

Nous pouvons donc définir la proportion maximale des hauts risques, notée λ^{RS} , telle que l'individu de type bas risque est indifférent entre un contrat séparateur et un contrat mélangeant, noté C_M , et qu'il n'existe aucun contrat

mélangeant lui procurant une utilité supérieure⁷. Soit,

$$\begin{aligned} V_L(C_L) &= V_L(C_M) \\ &= \frac{1 - p_L}{p_L} \frac{u' \left(W - \left[\lambda^{RS} p_H + (1 - \lambda^{RS}) p_L \right] q \right)}{u' \left(W - \left[\lambda^{RS} p_H + (1 - \lambda^{RS}) p_L \right] q - D + q \right)} \\ &= \frac{1 - \left[\lambda^{RS} p_H + (1 - \lambda^{RS}) p_L \right]}{\left[\lambda^{RS} p_H + (1 - \lambda^{RS}) p_L \right]} \end{aligned}$$

Propriété 6 : a) si $\lambda \geq \lambda^{RS}$, l'équilibre RS existe.
b) si $\lambda < \lambda^{RS}$, l'équilibre RS n'existe pas.

L'équilibre RS existe lorsqu'il n'est pas Pareto-dominé par un contrat mélangeant. Toutefois, si l'on élimine l'hypothèse de contrat unique, il peut être dominé par des contrats séparateurs à subventions croisées (Pour la démonstration, voir Crocker et Snow (1985, 1986)).

Dès lors, il est naturel de se demander si la modification d'une ou plusieurs hypothèses n'altère pas les propriétés de l'équilibre.

6.3 Extensions

Les extensions sont de deux types. Soit est considérée, en sus de la différence de probabilité d'occurrence du sinistre, une différence de taille du dommage, soit une différence d'aversion au risque.

6.3.1 Tailles différentes de dommage

Doherty et Jung (1993) considèrent une relation entre le type de risque (la probabilité) et la taille du sinistre. L'individu le plus risqué est supposé être caractérisé à la fois par une probabilité d'occurrence et une taille de dommage plus élevées. Si la taille du dommage est observable sans coût, nous pouvons penser que les contrats de premier rang peuvent être proposés aux agents. Toutefois, le problème d'asymétrie d'information demeure pour les individus qui ne subissent pas de dommages. Pour que la mutualisation des risques soit possible, il faut que tous les individus de type H choisissent le contrat approprié à leur type (Chiappori, 1996). Or, si aucune réglementation ne contraint les agents à ne pas frauder, les agents de type H ne choisiront pas toujours les contrats propres à leur risque. Doherty et Jung (1993) proposent pour un individu qui est pris à frauder une pénalité importante pour l'inciter à ne pas frauder. Cependant,

⁷Graphiquement, cette seconde condition implique l'égalité entre la pente de la droite actuarielle d'un risque moyen et la pente de la courbe d'indifférence de l'individu L obtenue avec un contrat mélangeant.

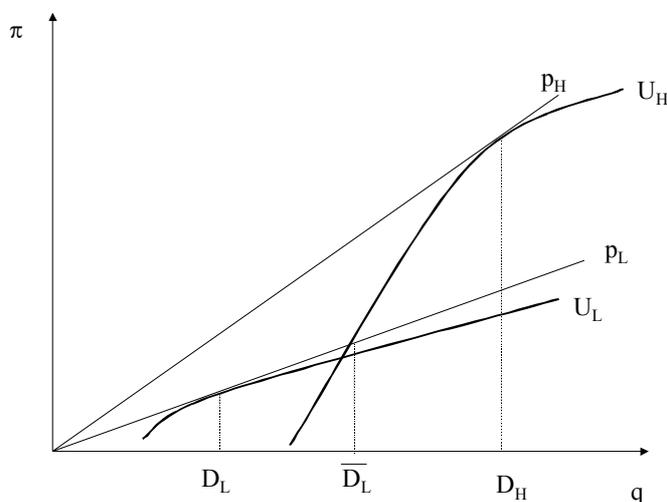


FIG. 6.5: Contrats de premier rang

cette réglementation n'est pas nécessaire dans un cas où les individus, quel que soit leur type choisiront, leur contrat de premier rang.

Tel sera le cas si la différence de dommage est suffisamment élevée, les contraintes d'incitation des individus n'étant alors pas saturées.

Soient D_L et D_H les montants des dommages que subissent les individus. Doherty et Jung (1993)⁸ montrent graphiquement que si $D_L \leq \bar{D}_L$ l'individu de type H n'est jamais incité à choisir le contrat des L (cf. figure 6.5). De plus, si $D_L > \bar{D}_L$, l'individu de type H est toujours incité à choisir le contrat du type L .

Ainsi, les contrats de premier rang peuvent être proposés aux individus si la différence de sinistre entre les individus est grande en présence d'asymétrie d'information.

6.3.2 Aversions au risque différentes

En général, les modèles d'antisélection en assurance considèrent une inobservabilité unidimensionnelle. Dans cette section, nous supposons que le principal fait face à deux asymétries d'information : la probabilité d'occurrence du sinistre

⁸Ils démontrent que si les distributions des probabilités d'occurrence des dommages ainsi que leurs tailles observables sont telles que l'espérance de dommage du haut risque est supérieure à celle du bas risque (avec la valeur extrême de la distribution du dommage des individus de type haut risque supérieure à celle des bas risques), alors, pour un certain montant d'amende, une différence de perte de dommages et une valeur de l'aversion au risque, la contrainte des individus de type haut risque n'est pas saturée.

et l'aversion au risque du consommateur. Cette double asymétrie permet l'émergence de contrats à profits positifs et de contrats mélangeants. L'apparition de ces deux types de contrats à l'équilibre du marché, étrangers à l'analyse de RS, s'explique par le non-respect de la propriété de croisement unique des courbes d'indifférence. Deux auteurs ont étudié cette extension.

Tout d'abord, Villeneuve qui, en 1996, montra l'existence de contrats positifs dans un cadre restreint. Dans un modèle à deux agents, il supposait que l'individu à la probabilité d'occurrence la plus élevée était le moins averse au risque.

Ensuite, Smart (2000) a caractérisé les contrats d'assurance quand l'assureur devait faire face à une double asymétrie d'information⁹. Par rapport aux résultats de RS, un des contrats d'assurance offert à l'équilibre peut être mélangeant, et de plus, un profit positif peut être réalisé et ce, uniquement pour certaines différences d'aversion au risque.

Nous présentons dans le paragraphe suivant, le cas d'un modèle à deux agents où l'individu de type haut risque est le moins averse au risque¹⁰. La propriété de croisement unique des courbes d'indifférence n'est plus respectée. Dans ce cas, les courbes d'indifférence sont deux fois sécantes. Ainsi, il existe toujours un point de tangence entre les deux courbes¹¹. Nous montrons qu'en fonction du point de tangence, l'équilibre est soit séparateur à profits positifs, soit séparateur, soit mélangeant.

Des contrats positifs à l'équilibre ?

Le cadre général de cette modélisation est celui de RS (1976). Toutefois, on suppose que les individus diffèrent par leur aversion au risque. L'utilité d'un agent de type j s'écrit :

$$V_j(C_j) = p_j u_j(W - D + q_j - \pi_j) + (1 - p_j) u_j(W - \pi_j)$$

avec $j \in \{L, H\}$

L'individu de type L est plus averse au risque que l'individu de type H , alors (Pratt, 1964)

$$\text{pour tous } W_1 > W_2 > 0 : \frac{u'_L(W_2)}{u'_L(W_1)} > \frac{u'_H(W_2)}{u'_H(W_1)}$$

Nous reprenons également l'hypothèse simplificatrice que chaque assureur fait une offre unique de contrat d'assurance, spécifiant la prime et le montant de

⁹Dans ce modèle, quatre types d'agents sont rencontrés :

- type haut risque et très averse au risque,
- type haut risque et peu averse au risque,
- type bas risque et très averse au risque,
- type bas risque et peu averse au risque.

¹⁰Si l'individu le plus risqué est le plus averse au risque, la propriété de croisement unique est respectée. Dans ce cas, l'équilibre est de type RS si la proportion des hauts risques est élevée (Villeneuve, 1996).

¹¹Les courbes d'indifférence sont monotones.

l'indemnité. Cet assureur accepte tout agent souscrivant le contrat d'assurance qu'il propose.

En information parfaite, chaque individu est remboursé intégralement du montant du dommage en échange du paiement d'une prime d'assurance actuariellement juste. En information imparfaite, ces contrats séparateurs ne sont plus réalisables car, quelle que soit l'aversion au risque des individus, l'agent de type H veut souscrire le contrat du type L , puisque pour une même indemnité, il verse une prime d'assurance moindre.

L'individu de type H se voit offrir un contrat de pleine assurance, tandis que le contrat du type L est tel que l'agent de type H est indifférent entre les deux contrats. Ainsi, dans le cas de contrats séparateurs, le programme s'énonce :

$$C_L \in \arg \max V_L^{AL}(C_L)$$

sous la contrainte d'incitation

$$V_H^{AH}(X_H) \geq V_H^{AH}(C_L)$$

sous la contrainte de profits non négatifs

$$\pi_L - p_L q_L \geq 0$$

avec $C_H = (p_H D, D)$.

Nous supposons que les contraintes de participation et d'absence de surassurance sont respectées. Ce programme se réécrit :

$$C_L \in \arg \max_{C_L} L = V_L^{AL}(C_L) + \mu (V_H^{AH}(C_L) - V_H^{AH}(C_H)) + \kappa (\pi_L - p_L q_L)$$

avec μ et κ , les multiplicateurs de Lagrange associés respectivement à la contrainte d'antisélection et de profits non négatifs.

Les conditions du premier ordre sont :

$$\frac{\delta L}{\delta \pi_L} = 0 \text{ et } \frac{\delta L}{\delta q_L} = 0$$

Soit,

$$-p_L u'_L(R_2) - (1 - p_L) u'_L(R_1) + \mu \{p_H u'_H(R_2) + (1 - p_H) u'_H(R_1)\} - \kappa = 0$$

$$p_L u'_L(R_2) - \mu p_H u'_H(R_2) + \kappa p_L = 0$$

avec $R_2 = W - D + q_L - \pi_L$ et $R_1 = W - \pi_L$.

Ici, la contrainte d'incitation sera saturée. En effet, maximiser l'utilité de l'individu L revient à offrir un contrat L tel que l'agent de type H soit indifférent entre les deux contrats. Soit $\mu \neq 0$.

La seule contrainte problématique est celle des profits non négatifs. Celle-ci peut être saturée ou non. Chassagnon (1996), Villeneuve (1997) et Smart (2000) ont en effet montré que les profits pouvaient être positifs quand les agents se différenciaient par leur aversion au risque.

Nous devons donc étudier deux cas :

Soit $\pi_L - p_L q_L > 0$ ce qui implique $\kappa = 0$.

Soit : $\pi_L - p_L q_L = 0$ ce qui implique $\kappa > 0$.

Montrons à quelle(s) condition(s) les profits positifs sont possibles à l'équilibre.

Les conditions du premier ordre se réécrivent :

$$0 = -p_L u'_L(R_2) - (1 - p_L)u'_L(R_1) + \mu \{p_H u'_H(R_2) + (1 - p_H)u'_H(R_1)\} \quad (1a)$$

$$0 = p_L u'_L(R_2) - \mu p_H u'_H(R_2) \quad (2a)$$

(2a) devient $\mu = \frac{p_L u'_L(R_2)}{p_H u'_H(R_2)}$.

(1a) devient

$$\frac{p_L u'_L(R_2)}{(1 - p_L)u'_L(R_1)} = \frac{p_H u'_H(R_2)}{(1 - p_H)u'_H(R_1)} = -\frac{dR_1}{dR_2} \quad (1a')$$

(1a') implique l'égalité des taux marginaux de substitution entre les deux états du monde pour les deux individus.

Ainsi, les profits sont positifs si la courbe d'indifférence de l'individu L est tangente à celle de l'individu H .

Nous devons vérifier la positivité des profits et l'absence de surassurance. On doit obtenir $q_L < D$ et $\pi_L - p_L q_L > 0$. Ces deux conditions impliquent que $R_1 > R_2$.

A quelle condition obtenons-nous la positivité des profits réalisés sur les L ?

Par hypothèse, $p_L < p_H$ ce qui implique d'après (1a')

$$\frac{u'_H(R_2)}{u'_H(R_1)} < \frac{u'_L(R_2)}{u'_L(R_1)}$$

Si l'agent de type L est plus risquophobe que le type H (u_L est plus concave que u_H) et pour certaines valeurs des probabilité $p_L < p_H$, alors les deux courbes sont tangentes ce qui conduit à des profits positifs sur les bas risques (cf. figure 6-6)

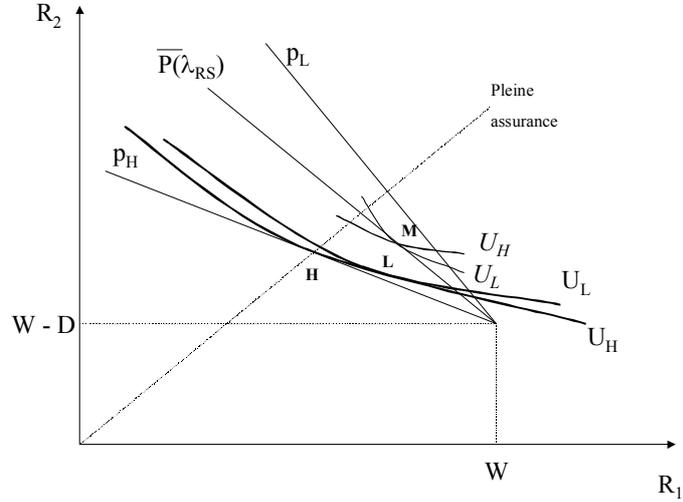


FIG. 6.6: Existence de profits positifs

Proposition 39 *Villeuneuve (1996)* - Si $u_H(\cdot)$ est strictement moins concave que $u_L(\cdot)$, alors quels que soient W , D , et p_H , il existe une probabilité p_L (avec $p_L < p_H$) telle que tout équilibre du marché, s'il existe, implique un profit strictement positif sur l'agent de type bas risque.

Preuve :

$u_H(\cdot)$ est strictement moins concave que $u_L(\cdot)$ implique que

$$\frac{u'_H(R_2)}{u'_H(R_1)} < \frac{u'_L(R_2)}{u'_L(R_1)}$$

De plus, les profits positifs impliquent que les deux courbes d'indifférence sont tangentes, c'est -à-dire que les pentes de ces deux courbes sont égales. Soit

$$\frac{p_L u'_L(R_2)}{(1 - p_L) u'_L(R_1)} = \frac{p_H u'_H(R_2)}{(1 - p_H) u'_H(R_1)}$$

Ainsi p_L est définie par

$$\frac{p_L}{(1 - p_L)} = \frac{p_H u'_H(R_2) u'_L(R_1)}{(1 - p_H) u'_H(R_1) u'_L(R_2)}$$

De plus, on peut vérifier que $p_L < p_H$. CQFD.

Nous avons montré que les courbes d'indifférence des individus peuvent être tangentes, à l'équilibre du marché, si l'aversion au risque des hauts risques est inférieure à celle des bas risques. Le contrat qui maximiserait l'utilité des bas

risques, sous la contrainte de ne pas être choisi par les hauts risques et de ne pas être déficitaire, serait un contrat à profits positifs. Cependant, à ce stade, nous n'avons pas vérifié son existence.

Si le nombre des bas risques est faible, alors ce contrat séparateur à profits positifs est dominé au sens de Pareto par un contrat mélangeant (cf. figure 6-6, point M) instable. En effet, tout contrat dans la zone grisée attire uniquement les bas risques. Comme à l'équilibre RS, il faut que le nombre de hauts risques soit suffisamment élevé pour que l'équilibre séparateur existe. Cependant, cette condition est nécessaire et non suffisante.

En effet, si les entreprises offrent un seul contrat, elles ne voudront offrir que le contrat à profits positifs défini ci-dessus, ce qui rend impossible l'existence de cet équilibre sans hypothèse supplémentaire. Si au contraire, elles anticipent qu'au moins une entreprise offre le contrat à profits nuls (celui de pleine assurance des hauts risques), toutes proposeront le même contrat à profits positifs. Comme le montre Villeneuve (1996), une condition nécessaire à cet équilibre est donc l'existence d'une entreprise qui assure les hauts risques¹². S'il y a une telle entreprise qui assure les hauts risques, alors cet équilibre à profits positifs existe.

Proposition 40 *La coexistence de deux types d'entreprises d'assurance - des mutuelles (à but non lucratif) et des compagnies d'assurance - est nécessaire à l'existence d'un équilibre de contrats à profits positifs.*

Dans cette situation, les mutuelles assurent les hauts risques et les compagnies les bas risques. En effet, les mutuelles ne peuvent proposer les contrats de premier rang. Dans ce cas, seul le contrat des bas risques serait choisi par tous les assurés : il deviendrait donc déficitaire. En ne proposant que des contrats actuariels, elles n'assurent que les hauts risques. Si elles offrent le contrat des bas risques de second rang, il ne sera pas choisi car les bas risques lui préféreront le contrat des compagnies d'assurance.

Une autre solution au problème de l'existence de l'équilibre en présence de profits positifs est proposée par Smart (2000). Celui-ci considère que les entreprises qui offrent un contrat unique supportent un coût fixe qui permet de réguler l'entrée sur le marché. Dans ce cadre, une entreprise unique offre un contrat à profits nuls et toutes les autres entreprises offrent un contrat à profits positifs. Plus le nombre d'entreprises offrant un contrat à profits positifs est élevé, moins leur clientèle est grande et plus leur bénéfice est faible. A l'équilibre, chaque entreprise réalise un profit nul.

Contrats séparateurs RS

Si le point de tangence entre les deux courbes d'indifférence a lieu dans le domaine d'assurance négative ($R_1 > W$) (point P_1 , cf. figure 6-7), alors dans la zone réalisable (absence de sur-assurance et d'assurance négative), nous obtenons que la pente de la courbe d'indifférence des individus H est moins

¹²Cependant, il ne décrit pas le type d'entreprise susceptible d'assurer les hauts risques.

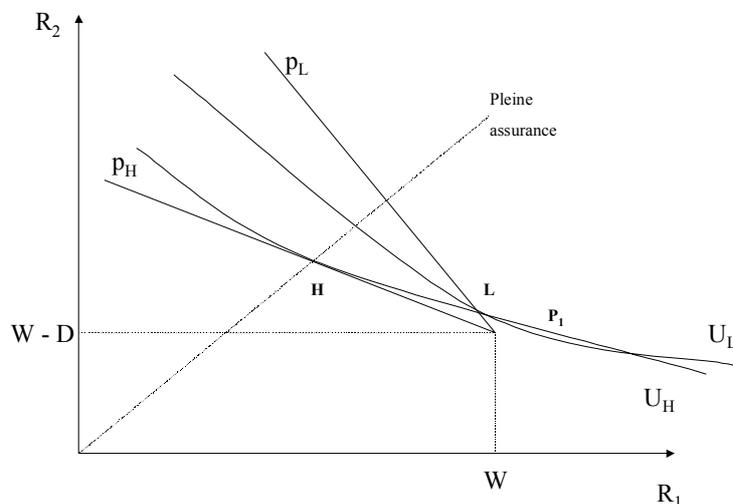


FIG. 6.7: Configuration RS

pentue que celle des individus L pour un même contrat (cf. figure 6-7). On retrouve la configuration proposée par RS.

Proposition 41 *Si $u_H(\cdot)$ est moins concave que $u_L(\cdot)$ et si le point de tangence des courbes d'indifférence est tel que $R_1 > W$, alors l'équilibre, s'il existe, est séparateur*

Preuve :

Si le point de tangence est tel que $R_1 > W$ alors en raison de la stricte décroissance des courbes d'indifférence et de la plus grande concavité de $u_L(\cdot)$ par rapport à $u_H(\cdot)$, pour tout contrat réalisable tel que $R_1 > R_2$, la courbe d'indifférence de l'individu de type L est plus pentue que celle de l'individu de type H . Nous retrouvons la condition de croisement unique qui autorise un équilibre de type RS.

Toutefois, pour s'assurer de son existence, il suffit que la proportion des bas risques ne soit pas trop importante (on pose $\lambda \geq \lambda^{RS}$). CQFD.

Néanmoins, une configuration n'a pas été traitée. Que se passe-t-il lorsque le point de tangence se trouve à gauche de la droite actuarielle juste des hauts risques ?

Un contrat mélangeant

Dans le cas où le point de tangence se trouve dans la zone de sur-assurance, nous obtenons, dans la zone réalisable, que la pente de la courbe d'indifférence

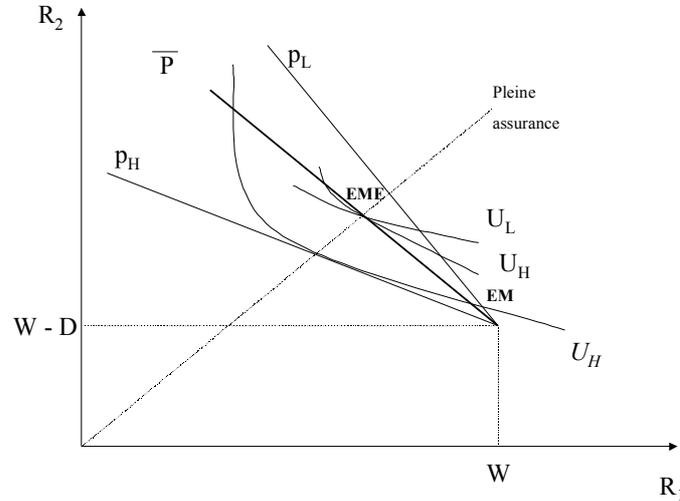


FIG. 6.8: Equilibre mélangeant

des individus L est moins pentue que celle des individus H pour un même contrat (cf. figure 6-8, point EM).

Proposition 42 *Si $u_H(\cdot)$ est moins concave que $u_L(\cdot)$ et si le point de tangence des courbes d'indifférence est tel que $R_1 < R_2$ et si $\lambda^{RS} \geq \lambda$ alors le contrat d'assurance est mélangeant et propose une assurance complète.*

Ce résultat est à rapprocher de Smart (2000) qui montre l'existence d'un contrat mélangeant sans le caractériser (avec quatre agents, ce sont les agents de types croisés qui obtiennent un contrat mélangeant)

Preuve :

Si le point de tangence est tel que $R_1 < R_2$ (point P_2) alors en raison de la stricte décroissance des courbes d'indifférence et de la plus grande concavité de $u_L(\cdot)$ par rapport à $u_H(\cdot)$, pour tout contrat réalisable tel que $R_1 > R_2$, la courbe d'indifférence de l'individu de type L est moins pentue que celle de l'individu de type H . Ainsi, le contrat mélangeant sera celui de pleine assurance car il maximise l'utilité des deux agents (cf. figure 6-8, point EME). Avec ledit contrat, aucun autre contrat réalisable ne peut attirer uniquement les bas risques¹³. Il n'existe donc pas de déviation profitable.

Cet équilibre est atteint si la proportion de bas risques est telle que $\lambda^{RS} \geq \lambda$. Ainsi, l'équilibre mélangeant domine au sens de Pareto l'équilibre séparateur. Cet équilibre n'est pas instable car il n'existe pas de contrat possible attirant uniquement les bas risques. CQFD.

¹³ Les seuls contrats attirant les bas risques sont des contrats de surassurance (zone grise).

Nous venons de montrer que si les individus les plus risqués sont aussi les plus risquophobes, alors les résultats de RS peuvent être généralisés. Toutefois, dans la situation où les individus les plus risqués sont les moins risquophobes alors les résultats de RS peuvent être invalidés. En effet, nous avons montré qu'en fonction de la position du point de tangence des courbes d'indifférence, trois cas sont à distinguer :

- existence de contrats séparateurs à profits nuls,
- existence de contrats séparateurs à profits positifs sur les bas risques,
- existence d'un contrat mélangeant.

6.3.3 Une généralisation

Dans un modèle à quatre agents, Smart (2000) pose que l'utilité d'un agent de type j et d'aversion de type h s'écrit :

$$V_j^h(C_j) = p_j u(W - D + q_j - \pi_j, \theta_h) + (1 - p_j) u(W - \pi_j, \theta_h)$$

avec $j \in \{L, H\}$, $h \in \{L, H\}$ et θ l'aversion absolue au risque.

Il établit le résultat suivant :

Proposition 43 (Smart, 2000)

Il existe des paramètres $\theta_2 > \theta_1 > \theta_L$ tels que pour tout θ_H , l'équilibre s'il existe est :

- a) si $\theta_H \in (\theta_L, \theta_1]$, séparateur à profits nuls,
- b) si $\theta_H \in (\theta_1, \theta_2]$, séparateur à profits positifs,
- c) si $\theta_H > \theta_2$, séparateur et mélangeant.

Preuve : (voir Smart 2000). CQFD.

Les individus de type (H, θ_H) et (L, θ_L) obtiennent des contrats d'assurance de type RS : assurance complète pour le type (H, θ_H) et assurance incomplète pour le type (L, θ_L) . Les contrats des individus de type croisé (L, θ_H) et (H, θ_L) dépendent de la valeur de l'aversion au risque.

Landsberger et Meilijson (1999) caractérisent dans un cadre multidimensionnel les contrats d'assurance optimaux. Les deux types d'agents diffèrent par leur fonction d'utilité croissante et concave, non ordonnée par le critère d'aversion au risque d'Arrow-Pratt et par le risque qu'ils encourent (la distribution de la probabilité ne vérifie pas nécessairement les propriétés usuelles de dominance au risque). La classification des agents ne peut ici s'effectuer qu'à l'aide des équivalents certains, car les deux caractéristiques des agents ne sont pas monétaires. Un résultat fondamental mis en exergue est que les agents dont l'équivalent certain est le plus faible reçoivent la pleine assurance. Ce résultat corrobore celui de RS. Dans le cadre RS, les individus de type H obtenant la pleine assurance sont ceux qui possèdent l'équivalent certain le plus faible ¹⁴.

¹⁴ Soit $U(EC_h)$ l'équivalent certain, dans le cadre RS, de l'individu h défini par : $U(EC_h) = p_h U(w - d) + (1 - p_h) U(w)$. $p_H > p_L \Rightarrow EC_H < EC_L$.

Chiappori, Jullien, Salanié et Salanié (2001) considèrent une classe de modèles qui généralise les modèles de Rothschild et Stiglitz (1976), Arnott et Stiglitz (1988), Chiappori et Chassagnon (1997), De Meza et Webb (2001) et Jullien, Salanié et Salanié (2000). La modélisation de ces derniers ne nécessite pas la vérification de la condition de Spence-Mirlees. Dans ce cadre, un agent de type θ , pour un contrat donné, peut choisir sa probabilité de perte¹⁵ qui dépend du type θ de l'individu, avec θ qui peut représenter le risque, l'aversion au risque, les deux, etc... Ils montrent qu'un individu dont la probabilité d'occurrence du dommage est la plus élevée choisit la couverture la plus élevée.

La vérification de la présence du phénomène d'antisélection sur les marchés d'assurance est assez récente. Bien que les premiers résultats (Puelz et Snow, 1994) mettaient en évidence la présence d'antisélection, les dernières contributions tendent à prouver que la classification élimine toute antisélection résiduelle.

6.4 Validation empirique de l'antisélection

Valider empiriquement le phénomène d'antisélection implique de vérifier les résultats principaux du modèle de RS :

- inexistence de l'équilibre du marché d'assurance quand la population des individus de type haut risque est faible,
- et dans le cas contraire : les individus de type bas risque obtiennent une couverture inférieure à celle des hauts risques ; les individus de type haut risque s'acquittent d'une prime d'assurance supérieure.

Selon Chiappori et Salanié (1997), tester l'hypothèse de l'existence de l'antisélection nécessite de vérifier que l'individu dont la probabilité d'occurrence du sinistre est la plus élevée, choisit la meilleure couverture du dommage. Cependant, Chassagnon et Chiappori (1997) montrent que cette affirmation reste vérifiée même en présence d'aléa moral. En conclusion, si cette affirmation est infirmée, les auteurs en concluent à l'absence d'asymétrie d'information.

Dans ce sens, aucune étude n'a été menée dans le cadre de l'économie de la santé pour la France. Par conséquent, nous étayons notre propos en utilisant les résultats du marché de l'assurance automobile. Dans un premier temps, nous présentons les résultats obtenus sur ce marché. Dans un second temps nous exposons quelques résultats du marché de l'assurance santé.

L'assurance automobile

Les articles récents examinent la relation entre le risque observable ex-post et la couverture du dommage choisie ex-ante. La relation positive entre ces deux variables semblent être récusée.

Bien que les études traitant du risque moral ont commencé dans les années 80 (Boyer et Dionne, 1989 ; Dahly, 1991), l'analyse empirique de l'antisélection a débuté dans les années 90 avec une contribution importante de Puelz et Snow (1994).

¹⁵Dans ce cas, aléa moral. Si la probabilité dépend du type de l'agent, antisélection.

Ils testent la présence d'antisélection pour le marché de l'assurance automobile de la Georgie. A l'aide d'un modèle de prix hédonistes, ils estiment simultanément deux équations. La première représente la prime d'assurance (en fonction du niveau de couverture et des caractéristiques individuelles). La seconde décrit le choix d'une couverture par l'agent (en fonction du nombre d'accident, une probabilité espérée du nombre d'accident et des caractéristiques). Ils mettent en évidence que le montant de la franchise dépend négativement du nombre d'accidents. Ainsi, les hauts risques s'assureraient plus que les bas risques.

Chiappori (1999), Chiappori et Salanié (2000) et Dionne, Gouriéroux et Vanasse (2001) établissent que les résultats obtenus par Puelz et Snow (1994) proviennent d'une mauvaise spécification du modèle.

Dionne, Gouriéroux et Vanasse (2001) montrent, en utilisant les données des assureurs du Québec, qu'une mauvaise spécification du modèle conduit à des résultats erronés. Avec un premier modèle mal spécifié, les agents qui déclarent le plus grand nombre d'accidents choisissent le contrat à la franchise la plus faible. En intégrant dans ce modèle, une estimation du nombre espéré d'accidents par conducteur, le nombre déclaré d'accidents n'est plus une variable significative. En d'autres termes, une classification appropriée revient à éluder le phénomène d'antisélection des marchés d'assurance automobile¹⁶.

Cette absence d'antisélection pour le marché d'assurance automobile était déjà vérifiée pour le marché français. Toutefois, seuls deux contrats étaient pris en compte : un contrat tous risques (noté TR) et un contrat au tiers¹⁷ (noté AT). Richaudeau (1999), à l'aide d'un modèle probit, estime la probabilité de choisir soit une assurance TR ou AT. Il teste ce modèle avec le nombre d'accidents déclarés et effectifs (ses données proviennent de l'Institut National de la Recherche sur les Transports et leur Sécurité). Par ailleurs, il introduit une variable difficilement observable : le nombre de kilomètres parcourus. Un conducteur qui déclare plus d'accidents qu'un autre est-il un conducteur de type haut risque s'il parcourt plus de kilomètres? Il montre que l'estimation qui consiste à ne pas prendre en compte le nombre de kilomètres comme variable, corrobore l'existence d'antisélection. En conclusion, une classification appropriée conduit à résoudre le problème d'antisélection.

Chiappori et Salanié (2000) examinent le marché de l'automobile français. Ils testent la présence d'asymétries d'information (aussi bien aléa moral qu'antisélection). Ils estiment deux variables binaires :

- la quantité d'assurance schématisée par le choix d'un contrat AT ou TR,
- le risque d'accidents ex-post mesuré également par une variable binaire (aucun accident ou au moins un pour lequel la responsabilité du conducteur est engagée).

Dans ce cadre, la présence d'asymétrie d'information implique que les distributions d'erreurs des variables estimées sont corrélées. Ils construisent leurs échantillons selon l'expérience de la population : les jeunes conducteurs et les

¹⁶ "We showed, however, that risk classification is sufficient, in the sense that there is no residual adverse selection on risk types in the automobile insurance portfolio studied". Dionne, Gouriéroux et Vanasse (2001).

¹⁷En pratique, pour ces contrats différents montants de franchise sont proposés.

seniors. Pour un jeune conducteur, ils en concluent à l'absence d'asymétrie d'information. Le jeune conducteur ne connaît pas encore son type. Quant aux seniors, ils ne peuvent rejeter l'hypothèse de corrélation nulle entre les deux distributions d'erreurs.

En conclusion, les études économétriques ayant trait au marché de l'assurance automobile infirment la présence d'antisélection (voire d'asymétries d'information). Les assureurs disposent de données en nombre suffisant pour classer par types les conducteurs.

Cette classification semble moins aisée pour le marché de la santé¹⁸. Les assureurs n'ont pas accès aux informations médicales des patients. Lors d'un accident automobile, la responsabilité d'un conducteur est engagée et l'accident décrit. Pour un risque sanitaire, la responsabilité est difficile à définir ainsi que le risque sanitaire lui-même (de Kervasdoué, 1998)¹⁹. Cependant, un individu qui souffre de troubles cardiaques a une probabilité plus grande d'être atteint d'un infarctus. Cette asymétrie est-elle révélée dans les études traitant du marché de l'assurance santé ?

L'assurance santé

Il est observé, sur le marché de l'assurance santé complémentaire, que les personnes qui présentent un risque faible s'assurent moins systématiquement que celles qui présentent un risque élevé (Couffignal, 2000). Les études ayant trait à la vérification de présence de l'antisélection pour les marchés d'assurance santé montrent que l'hypothèse d'antisélection est validée. Cette thèse est corroborée par Cutler and Zeckhauser (1999) : sur 27 études recensées, 22 confirment l'existence d'antisélection pour le marché d'assurance santé²⁰. Seules quelques études sont présentées ci-après. Des phénomènes d'antisélection sont mis en évidence sans que soient réellement testés les deux derniers résultats de RS.

Sur le marché français, peu d'études testent l'existence du phénomène d'antisélection. Causset et Glaude (1993) confirment l'existence, pour un revenu donné, de l'existence simultanée du risque moral et de l'antisélection tandis que Genier (1998) infirme son existence. Ces résultats sont à nuancer car, en France, le rôle de l'assurance privée est faible.

¹⁸Les assurés qui souscrivent une police d'assurance automobile sont dans l'obligation de délivrer la photocopie de leur carte grise (qui indique les caractéristiques du véhicule) ainsi qu'un historique de leurs accidents (résumé par le coefficient bonus-malus). A partir de ces éléments, l'assureur peut effectuer une classification des risques. A contrario, lors de la souscription d'un contrat d'assurance complémentaire, seules des variables démographiques sont prises en compte pour chaque individu (sexe, âge, profession). Pour éliminer les phénomènes d'antisélection, il faudrait donner une photocopie du carnet de santé (qui décrirait l'historique des dommages sanitaires) ainsi que les résultats d'un test génétique (Chiappori, 1996) qui indiquerait les caractéristiques des individus. Ainsi, les individus pourraient être classés par risques.

¹⁹Supposons que je sois propriétaire d'une centrale nucléaire. Si vous me dites, c'est parce qu'il travaille dans cette centrale qu'il souffre d'une leucémie, je vous rétorquerai qu'il y a des gens qui sont atteints de leucémie et qui ne travaillent pas dans une centrale nucléaire.

²⁰Voir table 9, pp.107-112 (Cutler et Zeckhauser, 1999).

Aux Etats-Unis, Wolfe et Goldeeris (1991) mettent en évidence une corrélation positive entre l'état de santé déclaré et le montant d'assurance sélectionné. Plus un individu déclare un mauvais état de santé, plus il a tendance à s'assurer. Par ailleurs, ils observent que des personnes qui ont du faire face à des dépenses de santé non expliquées par leur état de santé déclaré et leurs caractéristiques individuelles achètent plus d'assurance. Cette étude corrobore la présence d'antisélection puisque des informations auxquelles n'a pas accès l'assureur modifie les choix d'assurance des individus. Ce résultat est à nuancer car l'échantillon testé est constitué de malades du système Medicare.

Seul Browne (1992) teste l'hypothèse selon laquelle les bas risques achètent, sur un marché avec antisélection, moins de couverture d'assurance que sur un marché sans antisélection. Le marché sans antisélection est constitué d'un marché d'assurance subventionné par les employeurs qui proposent des contrats de groupes. Le marché avec antisélection est un marché d'assurance individuelle. Il montre que la variable indicatrice de l'état santé déclaré est positivement corrélée avec la différence de couverture entre les deux marchés²¹. En d'autres termes, une personne se déclarant en bonne santé (faible risque) se couvre moins sur le marché individuel que sur le marché de groupes.

Bien que les études précitées concourent à valider l'existence de l'antisélection, une seule atteste réellement l'existence de l'antisélection selon les critères énoncés par Chiappori et Salanié (2000). Cependant, le risque de l'individu est difficile à définir. Un individu qui est en mauvaise santé^{22,23}, est-ce un individu qui devra engager plus souvent des faibles dépenses pour se soigner, ou moins souvent un montant de dépenses élevées ?

6.5 Conclusion

L'introduction d'une aversion au risque différente invalide partiellement les conclusions de RS. En effet, dans une économie à deux types d'agents, si le bas risque est plus risquophobe que le haut risque, l'équilibre séparateur à profits nuls n'existe plus. Deux nouveaux types d'équilibres apparaissent : un équilibre séparateur à profits positifs sur les bas types et un contrat mélangeant à couverture complète.

L'introduction d'une aversion au risque différente nécessite la présence de deux types d'entreprises sur le marché de l'assurance pour que l'équilibre existe. Dans le cas de contrats à profits positifs réalisés sur les agents de type bas risque par les compagnies d'assurance, une entreprise à but non lucratif doit assurer les hauts risques (une mutuelle) pour lesquels le contrat est actuariellement juste.

Cependant, ces contrats de second rang, qui caractérisent théoriquement l'équilibre d'un marché d'assurance avec antisélection, ne semblent pas être offerts dans le cas de l'assurance automobile. Dionne, Gourrieroux et Vanasse

²¹ Quand l'état de santé augmente, la différence en valeur absolue diminue.

²² Le risque sanitaire est-il corrélé positivement avec le risque financier ?

²³ En espérance, nous supposons qu'un individu en mauvaise santé engage en espérance plus de dépenses qu'un individu en bonne santé.

(2001) montrent que la classification par risque des individus élimine toute anti-sélection résiduelle sur les marchés d'assurance automobile. Il existe suffisamment de variables observables pour définir la classe de risque d'un individu. Sur les marchés d'assurance santé, la classification des risques est plus difficilement réalisable car les variables de classification observables sont plus rares (absence de cartes d'identité génétique, absence d'historique de maladie) et le risque plus difficile à appréhender. ?

Chapitre 7

Antisélection et santé

“On ne peut être contre la nature que si l’on est soi-même nature et si la nature est avec vous”, Gadamer (1998).

7.1 Introduction

Le marché de l’assurance santé se caractérise par la présence du phénomène d’antisélection et par une difficile définition de la classe de risque des individus. Dans les 27 études recensées par Cutler et Zeckhauser (1999), les variables utilisées pour l’analyse sont l’âge, l’état de santé déclaré, les dépenses de l’année précédente, la fréquence des visites sanitaires, une classification démographique. Une variable pertinente pour expliquer les choix de contrats d’assurance est l’état de santé déclaré, non observable par l’assureur. Une autre variable explicative est le montant des dépenses. Un des résultats mis en exergue est qu’un individu en mauvaise santé engendre plus de dépenses qu’un individu en bonne santé. Cependant, nous ne pouvons en conclure que la probabilité d’occurrence du risque est la variable qui permet de classer les individus comme dans le modèle RS. En effet, les individus peuvent faire face à un sinistre avec la même probabilité et supporter un dommage différent.

La plupart des extensions du modèle de RS ont considéré que les assurés appartenant à des classes de risques différentes subissaient le même niveau de dommage. Néanmoins, Doherty et Jung (1993) (en supposant une corrélation positive entre la taille du dommage et la probabilité d’occurrence de celui-ci) montraient que si la différence de perte était suffisamment grande, l’équilibre était séparateur et de premier rang. Le choix de polices d’assurance différentes n’implique pas nécessairement l’existence du phénomène d’antisélection.

Dans ce chapitre, nous généralisons le modèle de RS à un cadre sanitaire puisque les études économétriques ayant trait au marché de l’assurance santé n’infirmant pas la présence du phénomène d’antisélection. Nous supposons qu’un individu est soumis à un risque sanitaire bénin non assurable. Il s’assure contre les conséquences financières du risque sanitaire dont le traitement est efficace.

Nous présentons également deux extensions de ce modèle. Dans celles-ci, l'utilité des individus est bidimensionnelle - la richesse et l'état de santé - et la probabilité d'occurrence du sinistre est différente selon les individus.

Tout d'abord, nous considérons que les individus se différencient aussi par leur état de santé, ce qui modifie leur perception du risque. Ensuite, l'état de santé et le montant du dommage sont également différents selon les individus. Dans ce cas, nous supposons que la classification des individus s'effectue à l'aide de l'espérance mathématique du dommage.

Le chapitre est organisé comme suit. Dans la section 2, nous supposons qu'un individu en bonne santé (notée S_L) a moins de chance d'être malade qu'un individu en moins bonne santé (notée S_H). Nous caractérisons l'équilibre de RS dans ce cadre. Dans la section 3, nous supposons que les individus en mauvaise santé sont infiniment averses au risque. L'équilibre obtenu est mélangeant, à subventions croisées. Enfin, dans la section 4, les individus en bonne santé ont une espérance de dommages inférieure à un individu en mauvaise santé. Supposer l'hétérogénéité de la probabilité d'occurrence et de la taille du dommage revient à invalider la propriété de croisement unique. L'équilibre obtenu ne sera donc plus forcément de type RS. Dans ce cadre, nous obtenons que l'équilibre peut être celui de premier rang, un équilibre à profits positifs ou un équilibre RS inversé.

7.2 Généralisation du modèle RS à un cadre sanitaire

Le cadre d'analyse de cette partie est celui de RS. La fonction d'utilité des individus dépend de deux variables : la richesse monétaire et l'état de santé représenté par un capital santé. Un individu en bonne santé possède un capital santé plus élevé qu'un individu en mauvaise santé¹. Nous posons comme hypothèse que l'individu en mauvaise santé est l'individu dont la probabilité d'occurrence du risque est la plus élevée.

La fonction d'utilité est une fonction croissante, en chacun des deux arguments concave, et trois différentiable. Elle s'écrit $u(R, S)$ avec R la richesse aléatoire de l'individu et S son capital santé.

Nous posons comme hypothèse que la prime d'assurance que l'individu est prêt à payer pour se départir du risque est une fonction décroissante du capital santé et de la richesse monétaire². Ces résultats impliquent que les prudences absolues d'un risque financier par rapport à la richesse ou à la santé sont supérieures à la prime d'assurance d'un risque financier³.

¹Selden (1998) étudie le modèle de RS avec une fonction d'utilité à deux variables mais en supposant que l'état de santé est le même pour les deux types d'individus mais différent selon les états de la nature. Les agents diffèrent uniquement par leur probabilité d'occurrence du risque maladie. Il montre que les contrats à subventions croisées peuvent dominer des contrats séparateurs.

²Voir Chapitre 3, section 2.

³Dans la suite de ce chapitre, nous ne spécifierons plus que le risque est financier.

Une partie de la population, en proportion λ , est en mauvaise santé, c'est-à-dire que son stock de santé est inférieur à celle en bonne santé. Chaque individu connaît son type, c'est-à-dire son état de santé et sa probabilité d'occurrence du sinistre. La demande d'assurance de chaque individu résulte d'un programme de maximisation de sa fonction d'utilité sous la contrainte de participation individuelle. Les entreprises d'assurance ne connaissent pas le type de l'agent. Elles adoptent une stratégie de Nash. Elles n'anticipent pas que ses rivales puissent réagir à des modifications de sa propre stratégie. Le jeu se déroule en deux étapes. A la première étape, chaque entreprise offre un unique contrat. A la seconde étape, les individus opèrent un choix parmi ces contrats et révèlent ex-post leur type en choisissant le contrat qui leur est propre.

En information parfaite, les individus obtiennent la pleine assurance. Cependant, en information parfaite les individus en mauvaise santé désirent le contrat de l'assuré en bonne santé. Le contrat des agents de type L (dont la probabilité est p_L et l'état de santé S_L) est choisi de telle sorte que l'individu de type H est indifférent entre les deux contrats. L'individu en mauvaise santé obtient la pleine assurance, et l'individu en bonne santé une assurance partielle. Nous retrouvons un des résultats mis en exergue par les études économétriques précitées.

Ce contrat séparateur existe si, en valeur absolue, la pente de la courbe d'indifférence des individus en bonne santé est supérieure, pour un contrat donné, à celle des individus en mauvaise santé (condition de croisement unique).

Soit,

$$\frac{(1-p_H) u_R(R_1, S_H)}{p_H u_R(R_2, S_H)} < \frac{(1-p_L) u_R(R_1, S_L)}{p_L u_R(R_2, S_L)}$$

avec u_X la dérivée partielle par rapport à X .

Une condition suffisante pour que cette inégalité soit vérifiée :

$$\frac{u_R(R_1, S_H)}{u_R(R_1, S_L)} < \frac{u_R(R_2, S_H)}{u_R(R_2, S_L)}$$

$R_1 > R_2$ implique que $\frac{u_R(R, S_H)}{u_R(R, S_L)}$ est une fonction décroissante de R . Soit,

$$-\frac{u_{RR}(R, S_L)}{u_R(R, S_L)} < -\frac{u_{RR}(R, S_H)}{u_R(R, S_H)}$$

Cette inégalité est vérifiée si l'aversion absolue au risque financier par rapport à la richesse $\left[-\frac{u_{RR}(R, S)}{u_R(R, S)}\right]$ est une fonction décroissante de la santé, S .

Nous obtenons ainsi :

$$-\frac{u_{RRS}(R, S)}{u_{RS}(R, S)} > -\frac{u_{RR}(R, S)}{u_R(R, S)}$$

Nous déduisons de cette inégalité que la prime de prudence d'un risque financier par rapport à la santé est supérieure à la prime de risque d'un risque financier. Cette inégalité est vérifiée car nous avons supposé que la prime de risque est une fonction décroissante de S (voir chapitre 4).

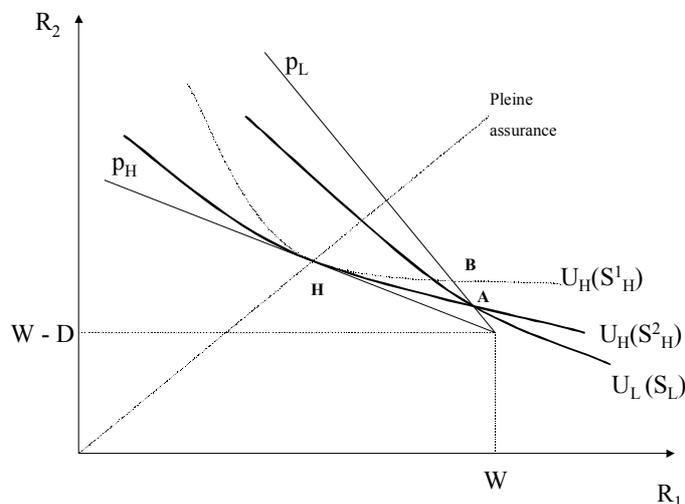


FIG. 7.1: Equilibre séparateur en information parfaite

En un point donné (R_1, R_2) , en valeur absolue, la pente de la courbe d'indifférence des individus en bonne santé est donc supérieure à celle des individus en mauvaise santé. La condition de croisement unique est donc vérifiée.

Proposition 44 *Dans un cadre bidimensionnel (richesse et santé), où l'individu en mauvaise santé supporte la probabilité d'occurrence du sinistre la plus élevée, l'équilibre du marché d'assurance, s'il existe, est de type RS :*

- les individus en mauvaise santé obtiennent la pleine assurance,
- les individus en bonne santé obtiennent une couverture partielle.

Preuve : cf. figure 7-1. CQFD.

Un changement de l'état de santé d'un individu modifie la convexité de la courbe d'indifférence, et donc la valeur de l'aversion absolue au risque par rapport à la richesse. Plus un individu est averse au risque, plus sa courbe d'indifférence est convexe.

L'aversion absolue au risque monétaire est mesurée par le rapport $-\frac{u_{RR}(R,H)}{u_R(R,H)}$. Elle est une fonction décroissante de l'état de santé (voir chapitre 4).

Nous pouvons énoncer le résultat suivant :

Proposition 45 *Plus l'individu de type haut risque est en mauvaise santé, plus la couverture des individus en bonne santé est grande.*

Preuve :

Sur la figure 7-1, nous notons que la courbe d'indifférence en pointillé (notée $U_H(S_H^1)$) est sécante avec la droite actuarielle des bas risques au point B. La

courbe d'indifférence en trait plein $(U_H(S_H^2))^4$ est sécante avec cette droite au point A . Au point B , le revenu des individus en bonne santé est supérieur à celui au point A . Sa couverture est donc améliorée.

En conclusion, plus la courbe d'indifférence de l'individu en mauvaise santé est convexe, plus sa santé s'altère et plus la couverture de l'individu en bonne santé est améliorée. CQFD.

Si l'état de santé est corrélé négativement à la probabilité d'occurrence du sinistre, alors l'introduction d'une dimension au modèle de RS ne modifie pas les propriétés de l'équilibre.

7.3 Hétérogénéité de l'utilité

Le cadre d'analyse de cette section est celui de RS. La fonction d'utilité des individus dépend de deux variables : la richesse monétaire et l'état de santé représenté par un capital santé.

Dans cette section, nous supposons que le niveau de l'état de santé de l'individu modifie sa perception du risque monétaire. L'utilité de l'assuré dépend de son état de santé. Un individu en mauvaise santé, en raison d'une espérance de vie moindre, a une aversion au risque monétaire infinie. Son espérance de vie étant faible, il désire réduire l'écart de revenus entre les différents états du monde. A contrario, un individu en bonne santé, peut supporter des variations importantes de revenu car son horizon temporel est élevé. Il peut donc lisser son revenu sur un nombre important de périodes⁵.

Nous posons \tilde{S} , l'état de santé seuil à partir duquel un individu change sa perception du risque. L'utilité d'un assuré de type j^6 dont l'état de santé est S_i s'écrit :

$$\begin{aligned} - \text{Si } S_i &\geq \tilde{S} & u_j(R, S_i) &= p_j u(W - D + q_j - \pi_j, S_i) + (1 - p_j) u(W - \pi_j, S_i) \\ - \text{Si } S_i &< \tilde{S} & u_j(R, S_i) &= \max \min(u(W - D + q_j - \pi_j), u(W - \pi_j)) \end{aligned}$$

avec j et $i \in \{L, H\}$

Un individu en bonne santé possède un capital santé (S_L) plus élevé qu'un individu en mauvaise santé (S_H) avec $S_H < \tilde{S} < S_L$.

⁴ $S_H^1 < S_H^2 < S_L$.

⁵ Notre hypothèse est, avec prudence, à rapprocher des travaux qui décrivent les plans de consommation des individus en univers incertain (voir Gollier, 2001). Un individu ayant un faible capital santé sait son espérance de vie faible. Il veut augmenter sa consommation dans l'état du monde le plus défavorable. L'incertitude sur la mortalité étant levée, il cherche en moyenne à augmenter sa consommation. Ce résultat est similaire à celui de Eeckhoudt, Gollier et Treich (2001). Ils étudient l'effet d'un changement de temporalité du risque. Ils montrent que la "résolution de l'incertitude" moins tardive augmente la consommation à la période initiale si l'individu est prudent, c'est-à-dire s'il réduit son épargne de précaution (dans un cas particulier : "the rate of pure preference for the present is equal to the risk free rate"). Cette condition est nécessaire mais n'est plus suffisante dans les autres cas.

⁶ Le type d'un individu est défini par la probabilité d'occurrence de la maladie car c'est la variable pertinente de classification pour une compagnie d'assurance.

La fonction d'utilité est une fonction croissante, en chacun des deux arguments concave, et trois différentiable.

Une partie de la population, en proportion λ , est en mauvaise santé. Chaque individu connaît son type (sa probabilité d'occurrence du sinistre) et son état de santé. La demande d'assurance de chaque individu résulte d'un programme de maximisation de sa fonction d'utilité sous la contrainte de participation individuelle. Les entreprises d'assurance ne connaissent pas le type de l'agent, ni l'état de santé des agents. Elles adoptent une stratégie de Nash. Elles n'anticipent pas le fait que ses rivaux puissent réagir à des modifications de sa propre stratégie. Le jeu se déroule en deux étapes. A la première étape, chaque entreprise offre un unique contrat. A la seconde étape, les individus opèrent un choix parmi ces contrats et révèlent ex-post leur type en choisissant le contrat qui leur est propre.

En information parfaite, chaque individu reçoit une couverture complète du montant du dommage en échange du paiement d'une prime d'assurance actuariellement juste.

En information imparfaite, ces contrats séparateurs ne sont plus réalisables car tous les agents de type H veulent souscrire le contrat d'information parfaite des assurés de type L : pour une même couverture, ils versent une prime d'assurance moindre. Le contrat des individus de type L en bonne santé doit donc vérifier la contrainte d'incitation des agents de type H en mauvaise santé⁷. Nous raisonnons d'abord avec les individus en bonne santé. Puis, nous étudions l'effet de l'introduction des individus en mauvaise santé.

Les individus de type H en bonne santé obtiennent leur contrat d'assurance d'information parfaite et les individus de type L en bonne santé un contrat d'assurance avec couverture incomplète. Cependant, les individus de type L en mauvaise santé préfèrent le contrat des agents de type H à celui des agents de type L en bonne santé. En effet, le revenu obtenu en cas de maladie est supérieur avec le premier contrat. Ils ne peuvent être séparés des individus en mauvaise santé. Nous obtenons donc un contrat mélangeant qui réunit les agents (H, S_L) , (H, S_H) et (L, S_H) . Avec ce contrat (cf. figure 7-2, point A), nous pouvons vérifier qu'il n'existe aucun contrat réalisable et profitable n'attirant que les agents (L, S_H) ⁸. Le point A appartient à la droite actuarielle des profits des assureurs dont l'équation est définie par :

$$\pi_H - \{\lambda p_H + (1 - \lambda)\eta p_L\} D = 0$$

Avec ce contrat mélangeant, à subventions croisées, les agents (L, S_H) subventionnent les agents (H, S_H) et (H, S_L) et augmentent indirectement l'utilité des agents (L, S_L) .

⁷La contrainte d'incitation des agents de type H en mauvaise santé ne peut définir le contrat des agents de type L en bonne santé, sinon la contrainte d'incitation des agents de type H en bonne santé n'est pas vérifiée.

Le contrat des agents de type L doit vérifier les contraintes d'incitations des agents de type H et ce, quel que soit leur état de santé.

⁸Seul un contrat de surassurance attire ces agents.

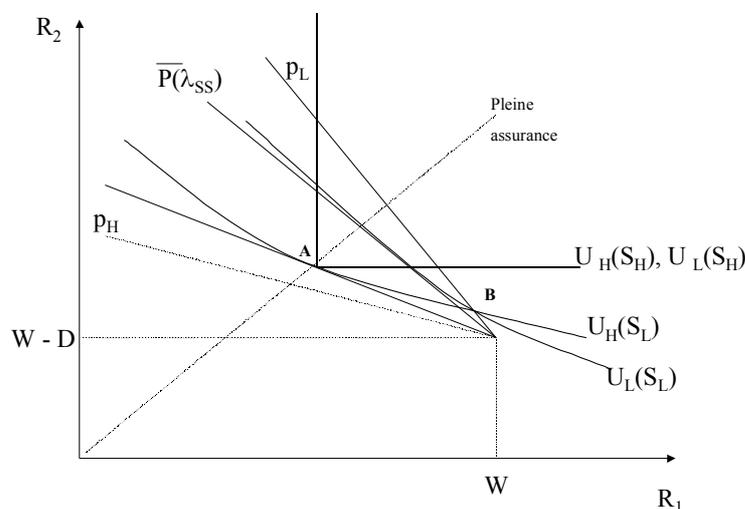


FIG. 7.2: Equilibre semi-séparateur à subventions croisées

Le contrat des individus de type L en bonne santé est quant à lui toujours défini par la contrainte d'incitation des individus en bonne santé et de type H (cf. figure 7-2, point B).

Proposition 46 Dans une économie dans laquelle les agents peuvent être de deux types H et L , et d'état de santé S_L et S_H , l'équilibre du marché est, s'il existe, semi-séparateur à subventions croisées :

- mélangeant pour les individus (H, S_H) , (H, S_L) et (L, S_H)
- séparateur pour les individus (L, S_L) .

La présence d'agents en mauvaise santé de type L augmente l'utilité des agents de type L en bonne santé et de tous les agents de types H . En effet, lesdits agents subventionnent les agents de type H , ce qui "déserre" la contrainte d'incitation. La présence d'individus de type L en mauvaise santé renforce donc l'existence de l'équilibre du marché d'assurance.

L'équilibre du marché d'assurance, dans une économie où sont présents des individus en mauvaise santé, existe si la proportion des individus de type L n'est pas trop importante $\lambda < \lambda^{SS}$ (cf. figure 7-2).

Nous remarquons, dans le cadre d'une double asymétrie informationnelle⁹, bien que les courbes d'indifférence des agents en mauvaise santé de type H

⁹L'article fondateur est celui de Laffont, Maskin et Rochet (1985) qui montrent qu'à l'optimum le contrat peut être mélangeant. Armstrong (1996) montre qu'à l'optimum certains agents sont exclus du marché. Rochet et Choné (1997) étudient plus particulièrement une double asymétrie informationnelle. Ils montrent que la solution peut être stochastique et que l'équilibre mélangeant est une solution générique.

et celles des individus en bonne santé ne sont sécantes qu'une seule fois (la conditions de Spence-Mirlees est respectée) : l'équilibre est mélangeant. Nous en concluons que ce n'est pas la sécance unique des courbes d'indifférence qui fonde la séparation des contrats. Les courbes peuvent être sécantes plusieurs fois. Seule la position relative des courbes d'indifférence dans la zone réalisable prime¹⁰. La condition de Spence-Mirlees dans un cadre de double asymétrie informationnelle doit être redéfinie.

Proposition 47 *Avec une double asymétrie informationnelle, les contrats sont séparateurs si et seulement si, dans la zone réalisable, en valeur absolue, la pente de la courbe d'indifférence des individus de type L est supérieure à celle des individus de type H .*

Nous avons montré que si l'état de santé modifiait la perception du risque par l'individu, alors que la propriété de croisement unique des courbes d'indifférence était respectée, l'équilibre était mélangeant. Nous obtenons que les individus en mauvaise santé de type L subventionnent tous les individus de type H .

7.4 Dommages hétérogènes

Le modèle fondateur de RS stipule un montant de dommages identiques. Cette hypothèse conditionne les résultats de ce modèle. Doherty et Jung (1993) montrent que l'équilibre de premier rang est réalisable avec asymétrie d'information si les montants des dommages sont hétérogènes et corrélés positivement avec la probabilité d'occurrence du dommage (voir chapitre 6).

Supposer des dommages différents revient à s'affranchir d'une classification des risques par probabilité. En effet, pour les assureurs, un individu qualifié de bas risque est un assuré dont l'espérance mathématique de pertes n'est pas trop élevée. Ainsi, la classification doit s'effectuer à l'aide de l'espérance mathématique des dommages.

Postuler des dommages différents rend l'analyse plus réaliste. En assurance automobile, les assureurs dénotent que les femmes ont une probabilité d'accident plus élevée ainsi qu'un montant de dommage inférieur à celui des hommes. En assurance santé, pour une jambe cassée, les personnes âgées et jeunes ont la même probabilité d'occurrence du dommage et un coût de traitement différent, très supérieur pour les premières¹¹. Cependant, ces caractéristiques telles que l'âge ou le sexe sont observables par les compagnies d'assurance. Par ailleurs, pour une même pathologie, des individus développent des complications et génèrent des dépenses supplémentaires. Le coût d'une pathologie est différent selon les individus.

Selon Bakker et Van Viet (1993), il est empiriquement constaté qu'un indi-

¹⁰Dans notre cadre assurantiel, la zone réalisable est définie par :

- profits non négatifs
- absence de surassurance et d'assurance négative.

¹¹Voir Couffinal (1999).

vidu plus risqué¹² sélectionnera le contrat dont la couverture est la plus généreuse. Ce choix n'implique pas nécessairement la présence d'antisélection. Des niveaux de couverture différents peuvent résulter d'une allocation d'équilibre de premier rang (Doherty et Jung, 1993). Chaque assuré souscrit une police d'assurance de couverture complète de dommages différents en échange de primes différentes.

Dans cette section, nous supposons que l'individu en mauvaise santé est l'individu le plus risqué, c'est-à-dire celui qui supporte l'espérance mathématique du dommage la plus élevée que l'individu en bonne santé¹³. Dans une première partie, nous présentons le modèle et les notations. Dans une deuxième partie, dans un cadre à deux types d'agent, nous étudions les contraintes d'incitation des individus et nous montrons que les résultats de RS peuvent être invalidés¹⁴. Dans une troisième partie, nous caractérisons l'équilibre du marché d'assurance avec ces deux types d'agent.

7.4.1 Le modèle

Les agents sont dotés d'un revenu W et d'un capital santé S . Les individus sont dits à bas risque (noté L) lorsque leur capital santé est plus élevé et que leur espérance de perte du dommage est plus faible. Soit $S_H < S_L$ et $p_H D_H > p_L D_L$. Cette distinction entre bas risque et haut risque recouvre en fait trois situations possibles :

- **cas a** : $p_H > p_L$ et $D_H > D_L$
- **cas b** : $p_H > p_L$ et $D_L > D_H > \frac{p_L}{p_H} D_L$
- **cas c** : $p_L > p_H$ et $D_H > \frac{p_L}{p_H} D_H > D_L$

Tout d'abord, le cas le plus intuitif. (**cas a**). Les agents en mauvaise santé sont caractérisés par une probabilité d'occurrence et un montant de dommage plus élevés. Ainsi, $p_L D_L < p_H D_H$ est bien vérifiée.

Ensuite, ces mêmes agents ont une probabilité d'occurrence du sinistre plus élevée et un montant du dommage plus faible que ceux de type L (**cas b**). Si $p_H > p_L$, on en déduit, d'après $p_L D_L < p_H D_H$, que $\frac{p_L}{p_H} D_L < D_H$. $D_L > D_H$ vérifie bien notre classification des risques.

¹²La plupart du temps, les études économétriques vérifiant la présence de l'antisélection sur les marchés d'assurance santé considèrent qu'un individu de type haut risque supporte un montant de dépenses santé supérieur à un individu de type bas risque.

¹³Cette modélisation est différente de celle de Chiappori, Jullien, Salanié et Salanié (2001), car le montant du dommage n'est pas corrélé avec le risque de l'individu.

¹⁴Dans un cadre purement monétaire, Bardey (2000) montrait à l'aide d'une approximation de Taylor d'ordre 1, que ce sont les assurés caractérisés par la probabilité de perte la plus élevée qui sont incités à souscrire le contrat d'information parfaite des individus caractérisés par la probabilité la plus faible. Toutefois, cette approximation d'ordre 1 ne permettait pas de retrouver les résultats de Doherty et Jung (1993). Une approximation d'ordre 2 est nécessaire. De plus, Bardey (2000) se limitait à déterminer si les contrats d'information parfaite vérifiaient les contraintes d'incitation des assurés. L'objectif de son travail n'était pas de caractériser l'équilibre du marché d'assurance mais de récuser la définition de l'antisélection proposée par Dionne (1994).

Enfin, le cas inverse du cas b (**cas c**). Les agents en mauvaise santé ont une probabilité d'occurrence plus faible et un montant plus élevé du dommage que ceux de type L . Si $p_L > p_H$, on en déduit, d'après la classification des risques que, $\frac{p_L}{p_H} D_L < D_H$. Une condition nécessaire (et non suffisante) est que $D_L < D_H$.

La fonction d'utilité est une fonction croissante, en chacun des deux arguments concave et trois fois différentiable.

Nous posons $S_L > S_H > \tilde{S}$. L'utilité d'un assuré de type j dont l'état de santé est S_j s'écrit¹⁵ :

$$u_j(R, S_j) = p_j u(W - D + q_j - \pi_j, S_j) + (1 - p_j) u(W - \pi_j, S_j)$$

avec $j \in \{L, H\}$ ¹⁶.

Une partie de la population, en proportion λ , est en mauvaise santé. Chaque individu connaît son type, c'est-à-dire son état de santé. La demande d'assurance de chaque individu résulte d'un programme de maximisation de sa fonction d'utilité sous la contrainte de participation individuelle.

Les entreprises d'assurance ne connaissent pas le type de l'agent. Elles adoptent une stratégie de Nash. Elles prennent le comportement de leurs concurrentes comme donné.

Supposer l'hétérogénéité de la probabilité d'occurrence et de la taille du dommage revient à invalider la propriété de croisement unique. L'équilibre obtenu ne sera donc plus forcément de type RS. Dans le chapitre précédent, nous avons obtenu comme équilibre, un équilibre mélangeant et un équilibre à profits positifs. Dans ce dernier cas, il est nécessaire de supposer que deux types d'assureurs existent dans cette économie : des compagnies d'assurance qui maximisent leurs profits et des mutuelles qui proposent des contrats actuariellement neutres. En effet, l'existence des mutuelles permet l'existence de l'équilibre du marché de l'assurance dans le cas de profits positifs. Si les mutuelles n'assuraient pas les agents de type haut risque, aucune compagnie n'assureraient ces agents car elles n'assurent que les agents de type bas risque¹⁷ en raison des profits réalisés sur ces agents. L'introduction d'un second agent d'assurance ne modifie pas la définition de l'équilibre au sens de RS mais permet l'existence de l'équilibre dans le cas de profits positifs.

Par ailleurs, nous supposons que les montants du dommage ne sont pas observables par l'assureur sans coût. S'ils sont vérifiables à coût nul dans un cadre à deux agents, connaître le montant du dommage d'un agent revient à déterminer son risque¹⁸. Nous posons également comme hypothèse qu'une politique

¹⁵Nous posons comme hypothèse que la prime d'assurance que l'individu est prêt à payer pour se départir du risque est une fonction décroissante du capital santé et de la richesse monétaire. Elles impliquent que les prudences absolues d'un risque financier par rapport à la richesse ou à la santé sont supérieures à la prime d'assurance d'un risque financier.

¹⁶Les paramètres sont fixés de telle sorte que l'état de santé de l'individu en mauvaise santé est toujours supérieur à \tilde{S} .

¹⁷Chaque entreprise d'assurance n'offre qu'un seul contrat.

¹⁸En présence d'un coût d'audit nul, dans une économie à quatre types de risque (où toutes les configurations sont possibles), déclarer son montant de dommage ne permet pas de résoudre l'asymétrie d'information concernant les probabilités d'occurrence du dommage.

d'audit est très couteuse. Dans ce cas, Picard (2000) montre que l'audit n'est effectué que pour de fortes valeurs de dommage.

Si les assureurs proposent un contrat stipulant que l'assuré est remboursé du montant du dommage s'il est égal à celui couvert par la police d'assurance¹⁹, alors les individus qui n'ont pas souscrit la police d'assurance relative à leur dommage sous-estimeront celui-ci pour être remboursé. Cette fraude est permise car le dommage n'est pas observable sans coût. Il est donc nécessaire de déterminer les contrats optimaux qui incitent les individus à révéler leur type et inhibent les comportements frauduleux. Ainsi, la détermination de contrats incitatifs implique un coût de gestion nul des contrats et entraîne une meilleure efficacité. En effet, si les incitations sont telles que chaque agent choisit le contrat approprié à son type, le problème de l'observabilité du dommage est éludé puisqu'aucun agent n'a intérêt à ne pas déclarer le montant réel de son sinistre.

7.4.2 Etudes des contraintes d'incitation

Dans les paragraphes suivants, nous étudions les trois cas possibles : cas a), cas b) et cas c). Tout d'abord, les agents en mauvaise santé présentent un niveau de dommage et une probabilité d'occurrence du dommage supérieurs à ceux des bas risques (cas a). Ensuite, ils subissent un dommage moindre que ceux en bonne santé (cas b). Enfin, la probabilité d'occurrence du dommage de ces individus est inférieure à celle des bas risques (cas c). Nous supposons dans le reste de ce travail que p_H et p_L sont fixées ainsi que D_H et D_L . Nous allons étudier les conditions pour lesquelles les assurés de chaque type ont intérêt à souscrire leur contrat d'information parfaite.

Chaque assuré a intérêt à souscrire son contrat d'assurance parfaite plutôt que celui de l'autre type si :

- pour les individus en mauvaise santé (de type H)

$$U(W - p_H D_H, S_H) \geq p_H U(W - D_H + D_L - p_L D_L, S_H) + (1 - p_H) U(W - p_L D_L, S_H) \quad (H)$$

- pour les individus en bonne santé (de type L)

$$U(W - p_L D_L, S_L) \geq p_L U(W - D_L + D_H - p_H D_H, S_L) + (1 - p_L) U(W - p_H D_H, S_L) \quad (L)$$

Cas a : $p_H > p_L$ et $D_H > D_L$

Les individus en bonne santé ne choisiront jamais le contrat de ceux en mauvaise santé car, ce faisant, ils prendraient une surassurance avec une prime supérieure à la prime actuarielle, ce qui serait moins bien pour eux que l'assurance complète à prime actuarielle. Nous savons que la pleine assurance maximise l'utilité de chaque individu quel que soit son type. A contrario, les individus

¹⁹Ce type de contrat ne serait pas robuste à des perturbations aléatoires sur la taille du sinistre (bruit blanc de moyenne nulle sur le montant du dommage).

en mauvaise santé peuvent être incités à prendre le contrat de ceux en bonne santé, car même si le remboursement du dommage avec ledit contrat est inférieur à celui stipulé dans le contrat propre à leur type, ils s'acquittent d'une prime d'assurance inférieure. Ils arbitrent entre un gain de prime d'assurance et une diminution du montant du remboursement du dommage. Contrairement au modèle RS, les individus en mauvaise santé ne choisissent pas toujours le contrat de premier rang des individus en bonne santé. Nous retrouvons, dans un cadre sanitaire, le résultat de Doherty et Jung (1993).

L'équation H est vérifiée si, pour des montants de dommage fixé, le capital santé de l'individu en mauvaise santé est faible, ou si pour un capital santé donné, la différence entre les montants des dommages est importante²⁰. Pour le montrer, nous utilisons un développement limité d'ordre 2 au voisinage de $W - p_H D_H$. Nous obtenons après réarrangements²¹ que :

$$0 > (p_H - p_L)D_L U_R(W - p_H D_H, S_H) + \left\{ \begin{array}{l} p_H(1 - p_H)(D_H - D_L)^2 \\ + (p_H - p_L)^2 D_L^2 \end{array} \right\} \frac{U_{RR}(W - p_H D_H, S_H)}{2}$$

Soit $R^A(c, S_H)$ l'indice d'aversion absolue au risque financier, avec $c = W - p_H D_H$. \overline{S}_H est défini par l'expression :

$$0 = (p_H - p_L)D_L U_R(W - p_H D_H, \overline{S}_H) + \left\{ \begin{array}{l} p_H(1 - p_H)(D_H - D_L)^2 \\ + (p_H - p_L)^2 D_L^2 \end{array} \right\} \frac{U_{RR}(W - p_H D_H, \overline{S}_H)}{2}$$

Nous en déduisons que l'individu choisit son propre contrat si :

$$R^A(c, S_H) > R_H^A(c, \overline{S}_H) = \frac{2(p_H - p_L)D_L}{\{p_H(1 - p_H)(D_H - D_L)^2 + (p_H - p_L)^2 D_L^2\}}$$

L'aversion au risque est une fonction décroissante de l'état de santé. En effet, nous avons supposé que la prime de risque d'un aléa monétaire est une fonction décroissante de l'état de santé²². Ainsi, $S_H < \overline{S}_H$.

Nous pouvons formuler la proposition suivante²³ :

Proposition 48 *Si $p_H > p_L$ et si $D_H > D_L$, il existe une valeur seuil \overline{S}_H de l'état de santé telle que tout agent en mauvaise santé est incité à choisir son contrat d'assurance de premier rang lorsque $S_H < \overline{S}_H$.*

Plus l'individu est en mauvaise santé, plus il est averse au risque monétaire et plus il préfère un contrat d'assurance complète à un contrat d'assurance incomplète car il désire réduire la différence de richesse entre les états du monde.

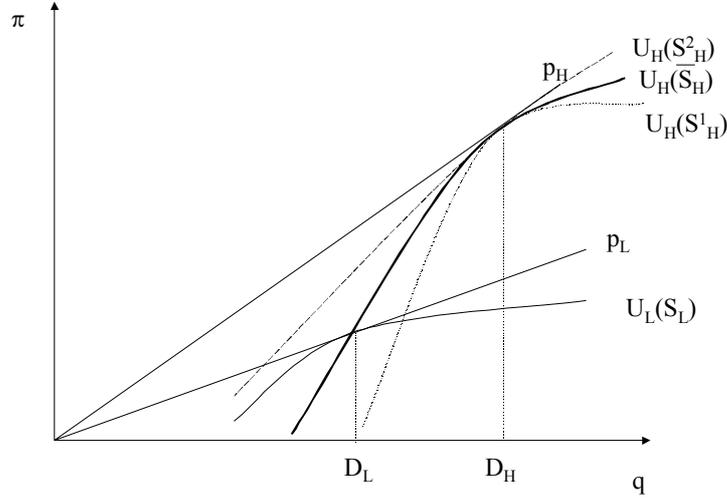
²⁰ Modifier le capital santé des individus en mauvaise santé revient à modifier la convexité de la courbe d'indifférence. L'introduction du capital santé ne modifie pas les résultats obtenus uniquement avec un risque financier.

²¹ Pour obtenir cette expression, il suffit de remarquer que

$$p_H D_H - p_L D_L - D_H + D_L = D_L(1 - p_L) - D_H(1 - p_H).$$

²² Voir première partie.

²³ Nous supposons que S_L est grand, c'est-à-dire que $S_L > \overline{S}_H$.

FIG. 7.3: Contraintes d'incitation : $p_H > p_L$ et $D_H > D_L$

Graphiquement, la convexité de la courbe d'indifférence est une fonction décroissante de l'état de santé (cf. figure 7-3). Soient trois états de santé ordonnés tels que $S_H^1 < \bar{S}_H < S_H^2 < S_L$.

Plus l'individu est en mauvaise santé, plus sa courbe d'utilité est convexe et moins il est incité à choisir le contrat de l'agent en bonne santé (passage de S_H^2 à S_H^1). Plus son aversion au risque financier croît (convexité de la courbe) en raison du changement de son état sanitaire, plus il désire réduire sa différence de richesse entre les différents états du monde.

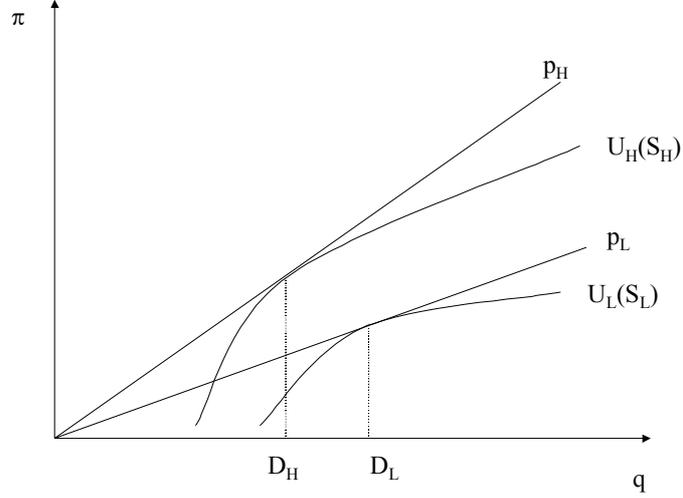
Cas b : $p_H > p_L$ et $D_H < D_L$

Les individus en bonne santé (de type L) choisiront toujours leur contrat du premier rang car le contrat des individus en mauvaise santé leur assure un remboursement du dommage inférieur pour l'acquittement d'une prime supérieure : $D_H < D_L$ et $p_H D_H > p_L D_L$ impliquent que $p_L D_L < D_L - D_H + p_H D_H$.

A contrario, les individus en mauvaise santé préféreront le contrat de ceux en bonne santé car ils obtiendront un remboursement du dommage supérieur pour le versement d'une prime inférieure.

Proposition 49 *Si $p_H > p_L$ et si $D_H < D_L$ alors les individus en mauvaise santé choisiront toujours le contrat de premier rang des individus en bonne santé contrairement aux individus en bonne santé.*

Cette proposition signifie que l'équation (H) n'est jamais vérifiée alors que l'équation (L) l'est toujours.

FIG. 7.4: Contraintes d'incitation : $p_H > p_L$ et $D_H < D_L$

Graphiquement, nous remarquons que, quel que soit l'état sanitaire des individus en mauvaise santé (schématisé par la convexité de leur courbe d'utilité), ils choisiront toujours le contrat de premier rang des agents en bonne santé (cf. figure 7-4).

Cas c : $p_H < p_L$ et $D_H > D_L$

Pour déterminer quel individu préfère le contrat de l'autre agent, nous ré-écrivons l'équation H (respectivement L) en utilisant un développement limité d'ordre 2 au voisinage du point $W - p_H D_H$ ($W - p_L D_L$) pour les individus en mauvaise (bonne) santé.

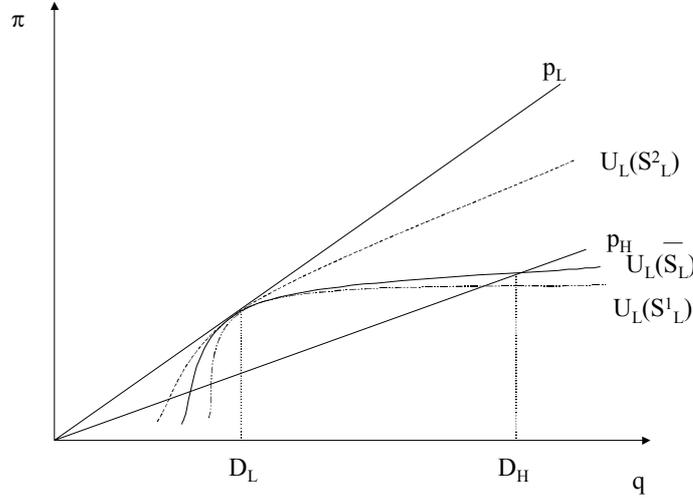
Après réarrangements, nous obtenons pour les individus en mauvaise santé :

$$0 > (p_H - p_L)D_L U_R(W - p_H D_H, S_H) + \{p_H(1 - p_H)(D_H - D_L)^2 + (p_H - p_L)^2 D_L^2\} \frac{U_{RR}(W - p_H D_H, S_H)}{2}$$

Cette expression est toujours vérifiée car $(p_H - p_L) < 0$ et $U_{RR} < 0$. Les individus en mauvaise santé choisissent toujours leur contrat de premier rang.

Pour les individus en bonne santé :

$$0 > (p_L - p_H)D_H U_R(W - p_L D_L, S_L) + \{p_L(1 - p_L)(D_H - D_L)^2 + (p_L - p_H)^2 D_H^2\} \frac{U_{RR}(W - p_L D_L, S_L)}{2}$$

FIG. 7.5: Contraintes d'incitation : $p_L > p_H$ et $D_H > D_L$

Soit $R^A(c, S_L)$ l'indice d'aversion absolue au risque.

$$R^A(c, S_L) > \overline{R^A}(c, \overline{S_L}) = \frac{2(p_L - p_H)D_H}{\{p_L(1 - p_L)(D_H - D_L)^2 + (p_L - p_H)^2 D_H^2\}}$$

avec $\overline{S_L}$ défini par :

$$0 = (p_L - p_H)D_H U_R(W - p_L D_L, \overline{S_L}) + \{p_L(1 - p_L)(D_H - D_L)^2 + (p_L - p_H)^2 D_H^2\} \frac{U_{RR}(W - p_L D_L, \overline{S_L})}{2}$$

L'aversion absolue au risque est une fonction décroissante de l'état sanitaire. Nous en déduisons que $S_L < \overline{S_L}$. Nous pouvons donc énoncer les résultats suivants :

Proposition 50 *Si $p_L > p_H$ et si $D_H > D_L$ alors pour tout état sanitaire tel que $S_H < S_L < \overline{S_L}$, tout agent est incité à choisir son contrat d'assurance de premier rang.*

Proposition 51 *Si $p_L > p_H$ et si $D_H > D_L$ alors pour tout état sanitaire tel que $\overline{S_L} < S_L$, l'agent en bonne santé (de type L) choisira toujours le contrat de premier rang des individus en mauvaise santé (type H).*

Graphiquement, la convexité de la courbe d'indifférence est une fonction décroissante de l'état de santé (cf. figure 7-5). Soient trois états de santé ordonnés tels que $S_H < S_L^1 < \overline{S_L} < S_L^2$. Lorsque l'état sanitaire de l'individu en bonne

santé diminue (passage de S_L^2 à S_L^1), il devient plus averse au risque. Il veut réduire la différence de richesse entre les divers états du monde. Ainsi, il préfère son propre contrat qui lui assure un revenu identique quel que soit l'état du monde.

Lorsque $S_L^2 > \overline{S_L}$, l'individu de type L (en bonne santé) préfère le contrat de premier rang de l'individu de type H (en mauvaise santé). Nous obtenons ici un résultat contraire à celui de RS.

Synthèse

Ces trois cas nous enseignent que, selon la valeur de l'état de santé des agents, le contrat d'assurance d'information parfaite peut être offert et ce, quelle que soit la position relative de p_H et p_L .

Nous montrons que si $D_H > D_L$, alors quelle que soit la position relative des probabilités de maladie ($p_H > (<)p_L$), les contrats de premier rang peuvent être offerts en présence d'asymétries d'information dans un cadre sanitaire. Nous avons donc généralisé le résultat de Doherty et Jung (1993) puisqu'il était établi uniquement pour $p_H > p_L$ et $D_H > D_L$ dans un cadre purement financier²⁴.

Par ailleurs, nous obtenons comme résultat fondamental que ce n'est pas l'individu en mauvaise santé qui préfère toujours le contrat de premier rang de l'individu en bonne santé mais l'individu dont la probabilité de maladie est la plus élevée²⁵. Nous retrouvons donc le résultat de RS pour lesquels la classe de risque et la probabilité d'occurrence du dommage étaient confondus. Ce résultat est résumé dans la proposition suivante²⁶ :

Proposition 52 *Si la classification des individus est fondée sur l'espérance mathématique du dommage, alors ce ne sont pas les individus de type haut risque, mais ce sont ceux qui présentent la probabilité d'occurrence du dommage la plus élevée.*

7.4.3 Equilibre du marché d'assurance

Selon les valeurs de l'état de santé des agents et les configurations mises en évidence précédemment, plusieurs cas retiennent notre attention.

Cas a : $p_H > p_L$ et $D_H > D_L$

D'après l'étude des contraintes d'incitation des individus, deux configurations sont possibles.

$S_H \leq \overline{S_H}$ Dans ce cas, nous savons que tout individu préfère son contrat d'assurance d'information parfaite à tout autre contrat. Le fait que les types des individus ne soient pas observables n'entraîne pas d'inefficacité économique (cf. figure 7-6). Nous retrouvons le résultat de Doherty et Jung (1995) qui le

²⁴Nos résultats restent vérifiés si $S_H = S_L$.

²⁵Nos résultats restent vérifiés si $S_H = S_L$.

²⁶Nous discuterons ce résultat après avoir caractérisé l'équilibre du marché d'assurance.

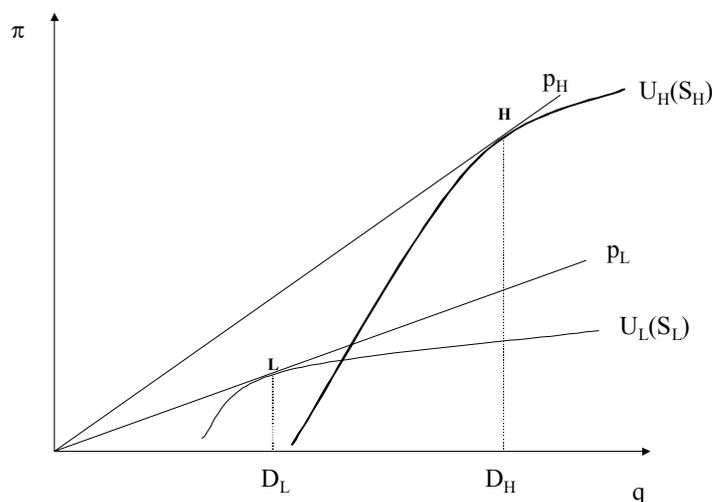


FIG. 7.6: Equilibre de premier rang

montraient pour une perte aléatoire avec une distribution différente pour les individus. Un choix de contrat différent par chaque agent n'implique pas toujours l'existence de phénomènes d'antisélection.

Les entreprises d'assurance offrent des contrats à profits nuls car aucun contrat profitable ne peut attirer les assurés²⁷. Il s'agit donc bien d'un équilibre. Soit H (respectivement L) le contrat des agents en mauvaise (bonne) santé.

Proposition 53 *Si $S_H \leq \overline{S_H}$, chaque type se voit offrir son contrat d'assurance de premier rang.*

$S_H > \overline{S_H}$ Dans cette configuration, nous savons que les individus en mauvaise santé souhaitent prendre le contrat des agents en bonne santé. A l'équilibre, les agents en mauvaise santé seront indifférents entre leur contrat et celui de ceux en bonne santé. Leur contrainte d'incitation est saturée. Nous retrouvons la configuration de RS (1976). Nous savons que si les courbes d'indifférence ne se croisent qu'une fois dans l'espace réalisable (pas de surassurance et absence d'assurance négative), alors aucun contrat à profits positifs n'augmente l'utilité des individus en bonne santé²⁸. Tous les contrats proposés seront actuariellement

²⁷La concurrence entre les compagnies d'assurance n'est plus nécessaire pour assurer la nullité du profit.

Si celles-ci offrent des contrats à profits positifs, les agents souscriront une police d'assurance des mutuelles car leur utilité est plus élevée.

²⁸Si les courbes d'indifférence des agents se coupent deux fois dans cet espace, alors Ville-neuve (1996) et Smart (2000) ont montré qu'un contrat à profits positifs pouvait, en fonction des paramètres, générer une augmentation d'utilité pour les bas risques. Voir le chapitre 6.

équilibrés²⁹. Avant de caractériser l'équilibre, nous devons déterminer la position relative des pentes des courbes d'indifférence.

Pour un contrat d'assurance donné $C(\pi, q)$ avec π la prime d'assurance et q la valeur du remboursement du dommage, la pente de la courbe d'indifférence des agents i est :

$$\frac{\partial \pi}{\partial q} \Big|_i = \frac{1}{1 + \frac{(1-p_i)U_R(W-\pi, S_i)}{p_i U_R(W-\pi+q-D_i, S_i)}}$$

Dans le repère (q, π) ³⁰, la pente de la courbe d'indifférence des individus en mauvaise santé est supérieure à celle des individus en bonne santé si :

$$\frac{(1-p_H)U_R(W-\pi, S_H)}{p_H U_R(W-\pi+q-D_H, S_H)} < \frac{(1-p_L)U_R(W-\pi, S_L)}{p_L U_R(W-\pi+q-D_L, S_L)}$$

La condition $\frac{(1-p_H)}{p_H} < \frac{(1-p_L)}{p_L}$ implique que cette inéquation est vérifiée (condition suffisante), si :

$$\frac{U_R(W-\pi, S_H)}{U_R(W-\pi+q-D_H, S_H)} < \frac{U_R(W-\pi, S_L)}{U_R(W-\pi+q-D_L, S_L)}$$

En effet, comme U est une fonction croissante et en chacun des deux arguments concave alors $U_R(W-\pi+q-D_L, S_H) > U_R(W-\pi+q-D_L, S_L)$

Nous posons $R_1 = W-\pi+q-D_L$ et $R_2 = W-\pi$.

Cette inéquation est vérifiée si $\frac{U_R(R, S_H)}{U_R(R, S_L)}$ est une fonction décroissante de la richesse. Cette condition est réalisée si :

$$-\frac{U_{RR}(R, S_L)}{U_R(R, S_L)} < -\frac{U_{RR}(R, S_H)}{U_R(R, S_H)}$$

Cette inéquation est vérifiée car l'indice d'aversion absolue au risque monétaire est une fonction décroissante de l'état de santé. En effet, nous avons supposé que la prime de risque d'un aléa monétaire est une fonction décroissante de S (voir chapitre 4).

Pour un contrat d'assurance donné $C(\pi, q)$, la pente de la courbe d'indifférence des hauts risques est toujours supérieure à celle des bas risques. La condition de Spence-Mirlees est vérifiée (propriété de croisement unique). L'équilibre séparateur, s'il existe, est tel que l'agent en mauvaise santé obtient une assurance complète (contrat noté H) et l'individu en bonne santé obtient une assurance partielle (contrat noté L). Le contrat des bas risques est déterminé de telle sorte que l'individu haut risque soit indifférent entre les deux contrats (respect de la contrainte d'incitation) (cf. figure 7-7).

Proposition 54 Si $S_H > \overline{S}_H$ l'équilibre du marché d'assurance, s'il existe, est de type RS.

²⁹Nous pouvons supposer que les individus sont répartis uniformément chez les entreprises d'assurance. La personnalité morale de l'entreprise d'assurance (mutuelle ou compagnie) n'influence pas le choix de l'agent.

³⁰Dans ce repère, l'utilité des agents croît quand on se déplace vers le quadrant sud-est.

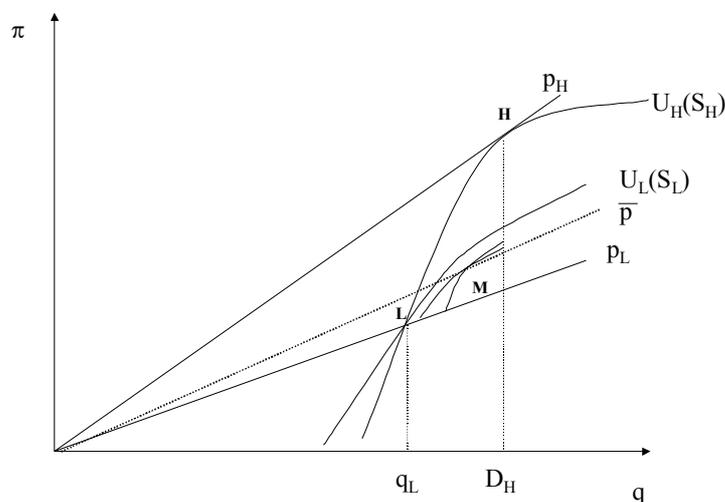


FIG. 7.7: Equilibre de type RS

Nous remarquons graphiquement que si la proportion des individus en bonne santé est trop importante dans l'économie alors tout contrat mélangeant noté M^{31} déstabilise le contrat séparateur (cf. figure 7-7). Cependant ce contrat est instable car les compagnies d'assurance peuvent offrir un contrat séparateur à profit positif attirant uniquement les bas risques (appartenant à la zone grisée). Les mutuelles ne proposent pas de contrat déviant car le contrat M respecte leur objectif de profit nul. Le contrat M devient déficitaire à la suite d'une telle déviation car il n'attire alors plus que les individus en mauvaise santé³². Ainsi, il existe une valeur $\bar{\lambda}_H$ telle que pour tout $\lambda \leq \bar{\lambda}_H$, l'équilibre n'existe pas.

Cas b : $p_H > p_L$ et $D_H < D_L$

Dans cette configuration, nous savons que l'individu en mauvaise santé désire toujours le contrat de premier rang des individus en bonne santé. Sa contrainte d'incitation est donc saturée.

Pour caractériser l'équilibre du marché d'assurance, nous devons déterminer la position relative des pentes des courbes d'indifférence dans le domaine réalisable (pas d'assurance négative ou de surassurance) et dans le plan (π, q) . En fonction de leurs positions relatives (voir chapitre 6), trois types d'équilibre

³¹En absence de surassurance, la couverture de ce contrat est inférieure ou égale à D_L .

³²Notons que le contrat mélangeant peut être un contrat d'équilibre si nous retenons la définition de l'équilibre de Wilson (1977). Cette définition intègre l'anticipation par les assureurs de la réaction de leurs concurrents qui retirent du marché des contrats devenus déficitaires.

peuvent émerger. Nous devons donc comparer, dans le domaine réalisable³³, les pentes des courbes d'indifférence des agents en bonne et mauvaise santé, c'est-à-dire :

$$\frac{1}{1 + \frac{(1-p_L)U_R(W-\pi, S_L)}{p_L U_R(W-\pi+q-D_L, S_L)}} \hat{a} \frac{1}{1 + \frac{(1-p_H)U_R(W-\pi, S_H)}{p_H U_R(W-\pi+q-D_H, S_H)}}$$

Nous savons que $\frac{(1-p_H)}{p_H} < \frac{(1-p_L)}{p_L}$, $U_{RS} < 0$, $p_H < p_L$, $D_H < D_L$, U est une fonction croissante et en chacun des deux arguments concave et que (en reprenant les notations précédentes) $\frac{U_R(R_2, S)}{U_R(R_1, S)}$ est une fonction décroissante de S .

Si pour $S_H = S_L$, la pente de la courbe d'indifférence des individus en mauvaise santé est inférieure à celle de ceux en bonne santé, alors elle le sera pour $S_H < S_L$ car $\frac{\delta \left[\frac{U_R(R_2, S)}{U_R(R_1, S)} \right]}{\delta S} < 0$. Dans ce cas, l'équilibre n'est pas de type RS puisqu'il est nécessaire que la pente des individus en mauvaise santé soit supérieure à celle de ceux en bonne santé.

Si pour $S_H = S_L$, la pente de la courbe d'indifférence des individus en mauvaise santé est inférieure à celle de ceux en bonne santé, alors elle peut ne plus l'être pour $S_H < S_L$ car, pour un contrat d'assurance donné, la pente de la courbe d'indifférence est une fonction décroissante de l'état de santé. Trois configurations sont donc possibles.

L'équilibre, s'il existe, sera de type RS si pour le contrat $(p_L q_L, q_L)$ (cf. figure 7-8, point P_1) défini par la contrainte d'incitation des individus en mauvaise santé, la pente des agents en mauvaise santé est supérieure à celle des agents en bonne santé (cf. figure 7-8)

Nous pouvons définir un cas limite pour lequel les deux courbes sont tangentes pour ce contrat $(p_L q_L, q_L)$ (cf. figure 7-9, point P_1). Plus l'individu est en mauvaise santé, plus la pente de sa courbe d'indifférence augmente.

Nous pouvons donc définir \underline{S}_H tel que :

$$\frac{(1-p_L)U_R(W-p_L q_L, S_L)}{p_L U_R(W-p_L q_L + q_L - D_L, S_L)} = \frac{(1-p_H)U_R(W-p_L q_L, \underline{S}_H)}{p_H U_R(W-p_L q_L + q_L - D_H, \underline{S}_H)}$$

Ainsi pour $S_H \leq \underline{S}_H$, l'équilibre est séparateur (s'il existe³⁴) car la pente de la courbe d'indifférence des individus en mauvaise santé est supérieure à celle des individus en bonne santé.

Si $S_H > \underline{S}_H$, avec un contrat $(p_L q_L, q_L)$ la pente des individus en mauvaise santé est inférieure à celle des agents en bonne santé. Ainsi, graphiquement nous dénotons qu'au point P_2 , les deux courbes d'indifférence sont tangentes et que l'utilité des individus en bonne santé est supérieure avec le contrat schématisé par le point P_2 par rapport au contrat actuariellement équilibré (point P_1). Ils obtiennent un contrat à profits positifs (cf. figure 7-10) si l'équilibre existe. Ici, les mutuelles offrent un contrat actuariel aux individus en mauvaise santé tandis

³³Absence de surassurance et d'assurance négative.

³⁴L'équilibre existe si le nombre d'individu en mauvaise santé est important.

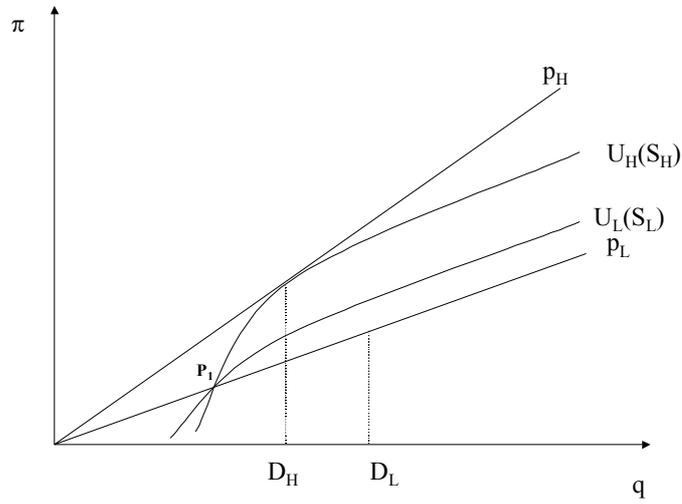


FIG. 7.8: Equilibre séparateur

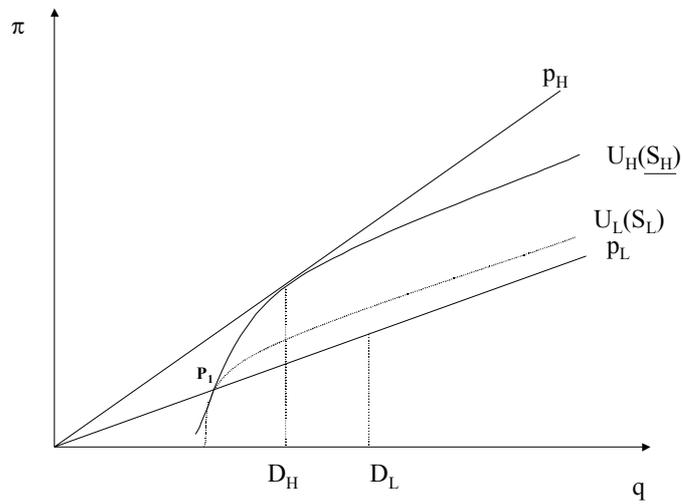
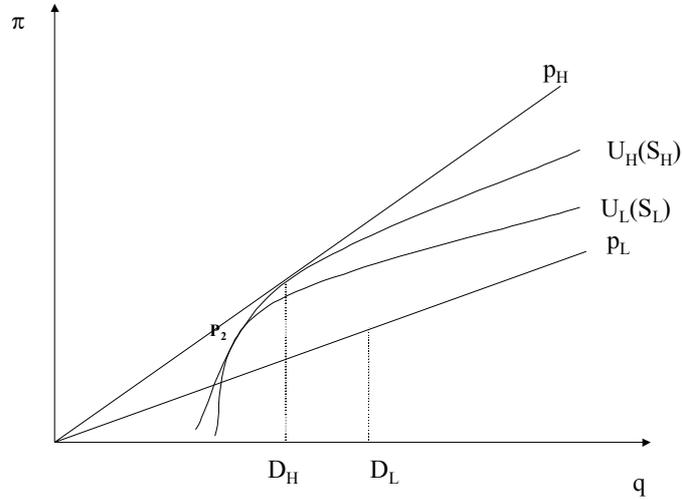


FIG. 7.9: $S_H = \underline{S}_H$

FIG. 7.10: Contrat L à profits positifs et couverture partielle

que les compagnies d'assurance offrent un contrat à profits positifs aux individus en bonne santé.

Plus S_H augmente, moins la courbe d'indifférence de l'individu en mauvaise santé est concave : le point de tangence se déplace de P_1 à P_2 . Il existe un point limite (cf. figure 7-11, point P_3) lorsque les deux courbes d'indifférence sont tangentes pour le contrat à profits positifs et pour une couverture totale du dommage pour les individus en bonne santé.

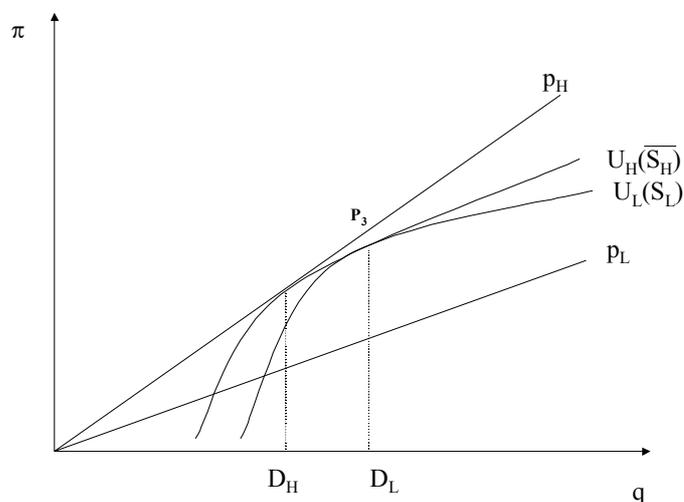
Nous pouvons donc calculer $\overline{S_H}$ tel que

$$\frac{(1 - p_L)}{p_L} = \frac{(1 - p_H)U_R(W - p_L D_L + \pi', \overline{S_H})}{p_H U_R(W - p_L D_L + D_L - D_H + \pi', \overline{S_H})}$$

avec π' les profits réalisés sur les agents en bonne santé.

Pour tout $\overline{S_H} < S_H$, le point de tangence des deux courbes d'indifférence est situé à droite de P_3 . Nous avons montré dans la partie précédente que pour tout point P_3 , le contrat mélangeant émergeait. Ici, le contrat mélangeant ne peut offrir une couverture supérieure à D_H car nous avons postulé l'absence de surassurance.

Quand $\overline{S_H} < S_H$, l'équilibre à profits positifs et couverture totale du risque est réalisé quel que soit le nombre d'individus en mauvaise santé, si et seulement si l'utilité de l'individu en bonne santé obtenue avec le contrat (cf. figure 7-12, point P_3 ,) est supérieure à l'utilité obtenue avec le contrat mélangeant (noté M) avec couverture D_H (cf. figure 7-12). Les mutuelles offrent le contrat des individus en mauvaise santé tandis que les compagnies d'assurance offrent le

FIG. 7.11: Contrat L à profits positifs et couverture complète

contrat des individus en bonne santé.

Si le contrat mélangeant M déstabilise l'équilibre séparateur, il est instable (courbes d'indifférence des H en gras et des L en pointillés). En effet, les compagnies d'assurance peuvent offrir un contrat à profits positifs n'attirant que les individus en bonne santé (cf. figure 7-12, zone grisée), contrat pour lequel la couverture du dommage est supérieure à D_H .

Nous pouvons résumer nos résultats dans la proposition suivante :

Proposition 55 Si $p_H > p_L$ et $D_H < D_L$, la nature de l'équilibre, s'il existe, dépendra de la valeur de S_H :

- si $S_H \leq \underline{S}_H$: séparateur à profits nuls (équilibre RS),
- si $\underline{S}_H < S_H \leq \overline{S}_H$: séparateur avec des profits positifs et couverture partielle (respectivement nuls et couverture totale) réalisés sur les individus en bonne santé (mauvaise santé)
- si $S_H > \overline{S}_H$: séparateur avec une couverture totale pour les deux types et avec des profits positifs (respectivement nuls) réalisés sur les individus en bonne santé (mauvaise santé).

Cas c : $p_H < p_L$ et $D_H > D_L$

D'après l'étude des contraintes d'incitation des individus, deux configurations sont possibles.

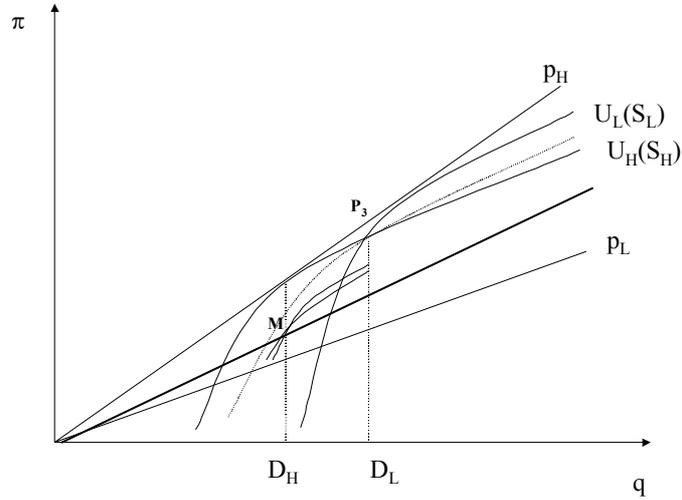


FIG. 7.12: Existence de l'équilibre

$S_H < S_L \leq \overline{S_L}$ Dans ce cas, nous savons que tout individu préfère son contrat d'assurance d'information parfaite à tout autre contrat. La méconnaissance des types des individus n'entraîne pas d'inefficacité économique (cf. figure 7-13). Un choix de contrat différent par chaque agent n'implique pas toujours l'existence de phénomènes d'antisélection.

Les entreprises d'assurance offrent des contrats à profits nuls car tout contrat à profits positifs n'est pas choisi par les assurés³⁵ puisqu'il génère une perte d'utilité. Soit H (respectivement L) le contrat des agents en mauvaise (bonne) santé.

Proposition 56 *Si $S_L \leq \overline{S_L}$, chaque type se voit offrir son contrat d'assurance de premier rang.*

$S_L > \overline{S_L}$ Dans cette configuration, nous savons que les individus en bonne santé souhaitent prendre le contrat des agents en mauvaise santé. A l'équilibre, les agents en bonne santé seront indifférents entre leur contrat et celui de ceux en bonne santé. Leur contrainte d'incitation est saturée. Peut-on retrouver la configuration inverse de RS (1976) ?

Pour caractériser l'équilibre, nous devons déterminer la position relative des pentes des courbes d'indifférence. Pour un contrat donné $C(\pi, q)$ avec π la prime

³⁵La concurrence entre les compagnies d'assurance n'est plus nécessaire pour assurer la nullité du profit.

Si celles-ci offrent des contrats à profits positifs, les agents souscriront une police d'assurance des mutuelles car leur utilité avec un tel contrat est plus élevée.

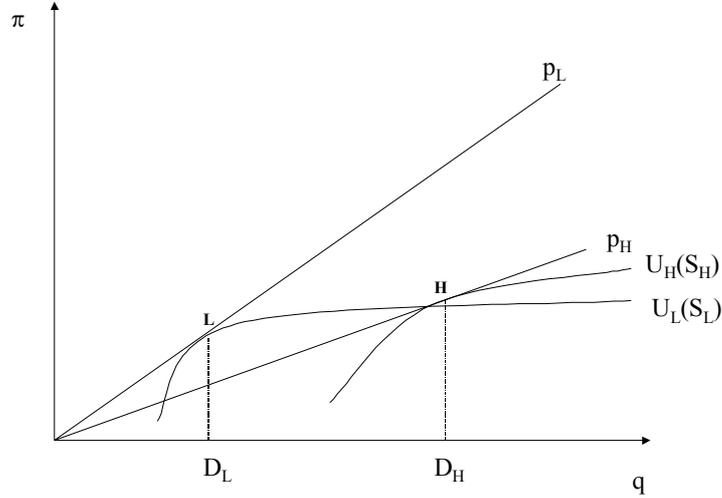


FIG. 7.13: Equilibre de premier rang

d'assurance et q la valeur du remboursement du dommage, la pente de la courbe d'indifférence des agents i est :

$$\frac{\partial \pi}{\partial q} \Big|_i = \frac{1}{1 + \frac{(1-p_i)U_R(W-\pi, S_i)}{p_i U_R(W-\pi+q-D_i, S_i)}}$$

Dans le repère (q, π) ³⁶, la pente de la courbe d'indifférence des individus en bonne santé est plus grande que celle des individus en mauvaise santé si :

$$\frac{(1-p_H)U_R(W-\pi, S_H)}{p_H U_R(W-\pi+q-D_H, S_H)} > \frac{(1-p_L)U_R(W-\pi, S_L)}{p_L U_R(W-\pi+q-D_L, S_L)}$$

Les conditions $\frac{(1-p_H)}{p_H} > \frac{(1-p_L)}{p_L}$ et $\frac{U_R(W-\pi+q-D_H, S_H)}{U_R(W-\pi, S_H)} > \frac{U_R(W-\pi+q-D_L, S_L)}{U_R(W-\pi, S_L)}$ impliquent que cette inéquation est vérifiée pour certaines valeurs de S_L (S_H est fixé).

Nous supposons que les valeurs des paramètres (p_L, p_H, D_H, D_L) sont fixées de telle sorte que l'inégalité est vérifiée pour $S_L > \bar{S}_L$. Le membre de droite de l'inégalité P est une fonction croissante de S_L car $\frac{U_R(R_2, S)}{U_R(R_1, S)}$ est une fonction décroissante de S .

Dans ce cas, l'équilibre est séparateur et actuariellement équilibré avec pleine assurance pour les individus de type bas risque et assurance partielle pour les assurés de type haut risque. Nous obtenons un équilibre RS inversé. Cet équilibre séparateur (contrat H et L) existe si la proportion des individus en bonne santé

³⁶ Dans ce repère, l'utilité des agents croit quand on se déplace vers le quadrant sud-est.

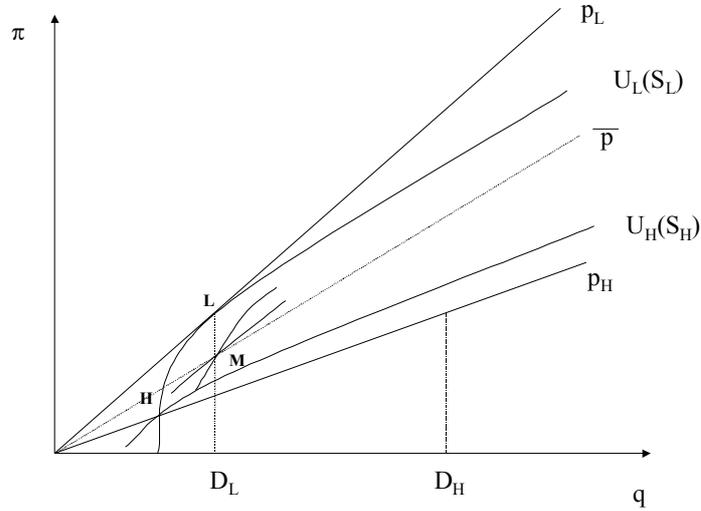


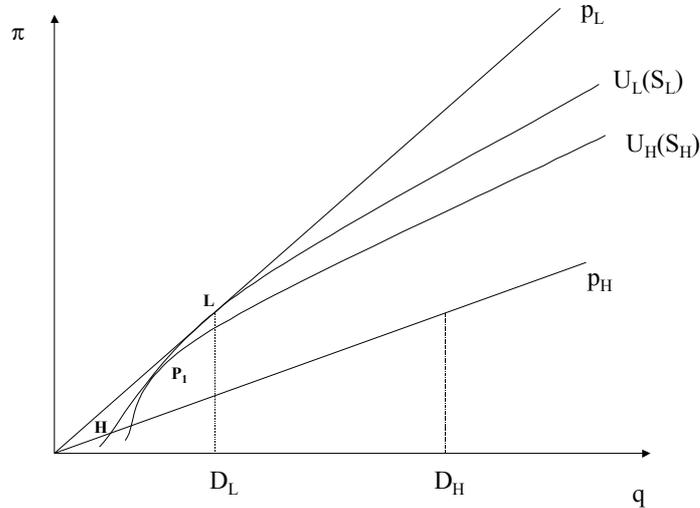
FIG. 7.14: Equilibre RS inversé

n'est pas trop importante. Dans le cas contraire, le contrat mélangeant (noté M) le déstabilise. Cet équilibre est instable pour une couverture D_L , car il existe un contrat à profits positifs (dans la zone grisée) qui attirent uniquement les mauvais risques (cf. figure 7-14).

Comme dans la section précédente, la variation de S_L va modifier la concavité de la courbe d'indifférence des individus en bonne santé. Nous pouvons calculer le niveau \underline{S}_L^1 , tel que pour le contrat H ($p_H q_H, q_H$), les deux courbes d'indifférence sont tangentes. \underline{S}_L^1 est défini par :

$$\frac{(1 - p_H)U_R(W - p_H q_H, S_H)}{p_H U_R(W - p_H q_H + q_H - D_H, S_H)} = \frac{(1 - p_L)U_R(W - p_H q_H, \underline{S}_L^1)}{p_L U_R(W - p_H q_H + q_H - D_L, \underline{S}_L^1)}$$

Si $S_L > \underline{S}_L^1$, avec le contrat H la pente de la courbe d'indifférence des individus en bonne santé est inférieure à celle des agents en mauvaise santé. Ainsi, graphiquement nous notons qu'au point P_1 , les deux courbes d'indifférence sont tangentes et que l'utilité des individus en mauvaise santé est supérieure avec le contrat schématisé par le point P_1 qu'avec le contrat actuariel (point H). Ils obtiennent un contrat à profits positifs (cf figure 7-15) si l'équilibre existe. Ici, les mutuelles offrent un contrat actuariel aux individus en mauvaise santé tandis que les compagnies d'assurance offrent un contrat à profits positifs aux individus en bonne santé. De même, l'existence de l'équilibre dépend du nombre d'individus en mauvaise santé. En effet, nous pouvons montrer que le contrat mélangeant est instable.

FIG. 7.15: Contrat H à profits positifs et couverture partielle

Plus S_L augmente, moins la courbe d'indifférence de l'individu en bonne santé est concave : le point de tangence se déplace de P_1 à P_3 . Il existe un point limite (noté P_3) lorsque les deux courbes d'utilité sont tangentes pour le contrat à profits positifs et couverture totale du dommage pour les individus en mauvaise santé (cf. figure 7-16).

Nous pouvons donc calculer $\overline{S_L^2}$ tel que

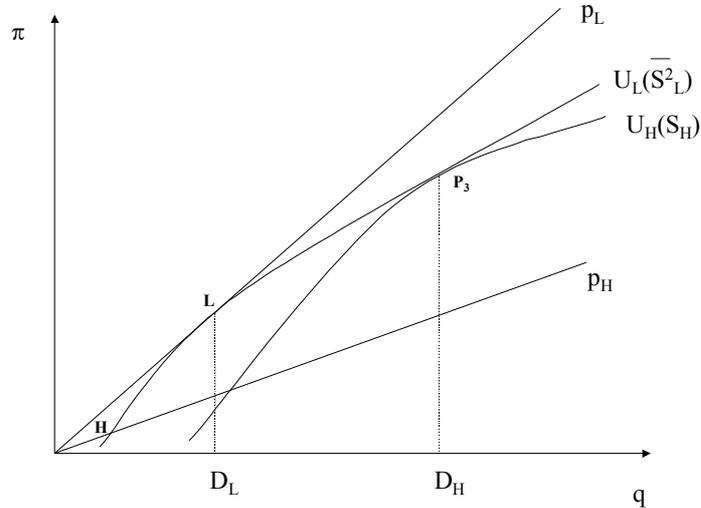
$$\frac{(1 - p_H)}{p_H} = \frac{(1 - p_L)U_R(W - p_H D_H + \pi', \overline{S_L^2})}{p_L U_R(W - p_H D_H + D_H - D_L + \pi', \overline{S_L^2})}$$

avec π' les profits réalisés sur les agents en mauvaise santé.

Pour tout $\overline{S_L^2} < S_L$, le point de tangence des deux courbes d'indifférence est situé à droite de P_3 . Nous avons montré dans la partie précédente que pour tout point P_3 , le contrat mélangeant émergeait. Ici, le contrat mélangeant ne peut offrir une couverture supérieure à D_L car nous avons postulé l'absence de surassurance.

Quand $\overline{S_L^2} < S_L$, l'équilibre à profits positifs et couverture totale du risque est réalisé quel que soit le nombre d'individus en mauvaise santé, si et seulement si l'utilité de l'individu avec le contrat (point P_3) est supérieure à l'utilité obtenue avec le contrat mélangeant avec couverture D_L (au point M) (cf. figure 7-17). Les mutuelles offrent le contrat des individus en bonne santé tandis que les compagnies d'assurance offrent le contrat des individus en mauvaise santé.

Si le contrat mélangeant M déstabilise l'équilibre séparateur, il est instable.

FIG. 7.16: Contrat H à profits positifs et couverture totale

En effet, les compagnies d'assurance peuvent offrir un contrat à profits positifs n'attirant que les individus en mauvaise santé (cf. figure 7-17, zone grisée), contrat pour lequel la couverture du dommage est supérieure à D_L .

Nous pouvons résumer nos résultats dans la proposition suivante :

Proposition 57 *Les caractéristiques de l'équilibre, s'il existe, dépendent de la valeur de S_L :*

- si $\bar{S}_L < S_L \leq \bar{S}_L^1$ l'équilibre est de type RS inversé
- si $\bar{S}_L^1 < S_L \leq \bar{S}_L^2$ l'équilibre est séparable avec profits positifs sur les individus en mauvaise santé
- et si $\bar{S}_L^2 < S_L$ l'équilibre est séparable avec couverture totale et profits positifs sur les individus en mauvaise santé.

Nous avons montré que ce n'est pas la classe de risque qui déterminait les contraintes saturées mais la valeur de la probabilité de la maladie. Il n'est pas surprenant d'obtenir dans une configuration à deux agents un équilibre RS inversé, c'est-à-dire que l'agent bas risque obtient une assurance complète a contrario de l'agent haut risque (assurance partielle). Nous résumons ce résultat par la proposition suivante :

Proposition 58 *Quand l'équilibre est séparable, ce n'est pas l'agent de la classe de risque la plus élevée qui obtient un contrat d'assurance complète et actuarielle mais l'individu dont la probabilité d'occurrence du dommage est la plus élevée.*

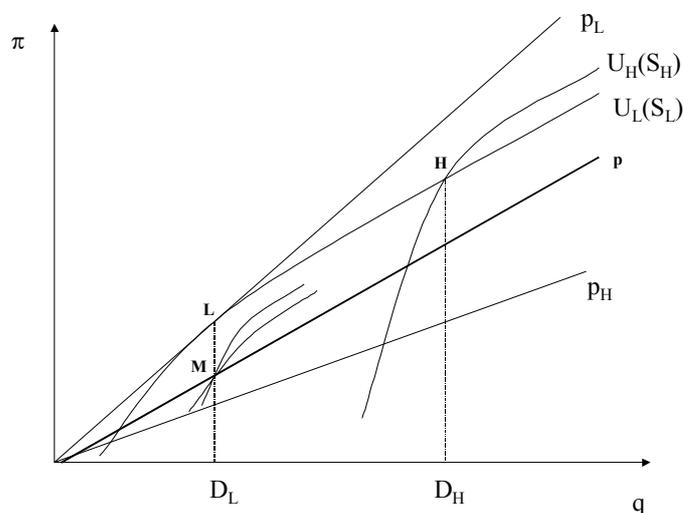


FIG. 7.17: Existence de l'équilibre

7.5 Conclusion

Certaines études empiriques prouvent que la variable expliquant le choix des contrats est l'état de santé déclaré des patients, état inobservable pouvant générer le phénomène d'antisélection. C'est pourquoi, dans le cadre standard RS, nous avons introduit en sus de la richesse, une variable d'état de santé. Nous avons postulé que l'état de santé est corrélé négativement avec le risque de l'individu. Dans le cadre standard de RS, en information parfaite, nous retrouvons les résultats usuels (équilibre séparateur).

Supposer que les dommages ne sont plus identiques nécessite que soit redéfinie la classification des individus. Nous postulons que l'individu en mauvaise santé a l'espérance mathématique des dépenses médicales la plus grande. Dans un tel cadre, nous avons infirmé les résultats de RS. Tout d'abord, l'équilibre peut être de premier rang. De plus, ce ne sont pas les individus les plus risqués qui obtiennent le contrat de premier rang, mais les individus à la probabilité d'occurrence de la maladie la plus grande et ce, quel que soit le montant du dommage. Il n'est donc pas contre-intuitif d'obtenir sous certaines conditions un équilibre RS inversé. Nous montrons également qu'en présence d'antisélection, les profits des compagnies d'assurance peuvent être positifs. Par ailleurs, les individus quel que soit leur type peuvent recevoir une couverture complète du dommage. Dans un tel cadre, l'absence de surassurance (la plus grande couverture offerte est le montant de la perte minimale) empêche la réalisation d'un équilibre mélangeant.

Ces résultats théoriques justifient l'adoption d'une très grande prudence pour

vérifier la présence du phénomène d'antisélection sur les marchés d'assurance santé. Les études économétriques tendent à montrer que les individus dont le risque est le plus élevé choisissent le contrat à la couverture la plus élevée. L'introduction de l'hétérogénéité du montant du dommage infirme cette conclusion. En effet, théoriquement, nous obtenons que l'individu de type bas risque peut obtenir une couverture complète contrairement à l'individu haut risque. Comme le souligne Bardey (2000), la définition de l'antisélection proposée par Dionne (1994) doit donc être modifiée. Nous proposons la définition suivante :

L'antisélection est donc le phénomène selon lequel un type de risque peut se faire passer pour un autre type de risque parce que l'assureur ne peut observer toutes les caractéristiques déterminant le type de risque.

Il conviendrait cependant de généraliser notre travail à un cadre multidimensionnel. Il s'agirait de considérer quatre classes de risque et de supposer que les individus en mauvaise santé sont ceux pour lesquels les espérances de pertes financières sont les plus élevées. Si nous appliquons à un tel cadre les résultats de Smart (2000), nous savons que les deux individus de types extrêmes ne posent pas de problème pour la détermination des contrats d'assurance. Il établit que l'individu le plus risqué reçoit un contrat d'assurance parfaite et le moins risqué un contrat d'assurance partielle.

Chapitre 8

Conclusion Générale

“Nous n’entendons pas bourdonner la litanie de Dieu, nous ne l’entendons que lorsqu’elle s’est tue”, Corossa.

L’objectif de cette thèse était d’étudier plus particulièrement les relations bilatérales qui lient les trois principaux acteurs du système de santé : médecin, patient, tutelle-assurance. Ce système est spécifique par l’incertitude qui y prévaut, l’asymétrie d’information qui le caractérise et le bien qu’il produit. Nous avons étudié deux secteurs du système de santé - la médecine ambulatoire et l’assurance - pour lesquels l’application de la théorie des contrats fournissait des résultats qui semblaient assez peu probants.

8.1 La médecine ambulatoire

En France, la médecine ambulatoire est composée de deux secteurs : un secteur sous tutelle dont le prix est réglementé (secteur 1) et un secteur dont le prix est déréglementé (secteur 2). L’instauration de deux secteurs a-t-elle permis aux médecins de signaler une qualité supérieure par un prix supérieur ? La spécificité des services médicaux nous a conduit à réfuter un tel lien dans le chapitre 2. Il s’agit, en effet, de biens intermédiaires qui n’impliquent pas nécessairement la “réparation” du corps, ce pourquoi ils sont produits. Le patient ne peut définir la qualité du bien médical puisque le résultat ne permet pas de déduire la qualité des services médicaux. Cette incertitude définit donc la prestation médicale comme un bien de confiance. En absence d’information parfaite sur la qualité de la prestation médicale par le patient, nous nous sommes attachés à montrer qu’un comportement rationnel de médecins opportunistes devrait les conduire à établir une réputation, dans la première phase de leur activité, en fournissant une qualité supérieure. Une fois celle-ci établie, il devient ensuite optimal d’offrir une qualité inférieure. Ces comportements opportunistes mis en exergue, il convient de s’interroger sur la légitimité de l’existence du secteur 2.

Depuis la convention médicale de 1993, l'accès au secteur 2 est gelé. Aucun médecin généraliste ne peut y accéder hormis les anciens chefs de clinique (spécialistes). Notre modélisation conduit à penser que cette réforme aboutit à une baisse de la qualité, puisque c'est lors des premiers temps de l'exercice que le médecin délivre une qualité supérieure. Ces comportements opportunistes légitiment donc une régulation de la tutelle du secteur 2, celle-ci étant vraisemblablement plus à même de saisir la qualité que le patient. Le patient, ne pouvant la contrôler, il ne peut proposer un contrat incitatif au médecin en vue d'être mieux soigné, puisque la rémunération d'un tel contrat devrait être liée à l'évaluation du résultat. A contrario, l'appréhension imparfaite de la qualité par la tutelle permet de caractériser les contrats optimaux.

Les résultats du chapitre 3 montre que si les médecins sont classés selon deux critères, productivité et honnêteté, les termes du contrat dépendent du nombre de médecins malhonnêtes. Deux types de comportements opportunistes apparaissent possibles. Tout d'abord, un médecin opportuniste peut être incité à fournir une qualité inférieure si celle-ci ne peut être vérifiée. Par ailleurs, un médecin "peu productif" opportuniste peut être incité à se faire passer plus productif qu'il ne l'est réellement. Si le nombre de médecins honnêtes est important alors la tutelle demandera une qualité élevée aux médecins qui se déclarent productifs puisque la fraude des médecins peu productifs opportunistes génère peu de conséquences en termes de bien être. Par contre, si les médecins honnêtes sont peu nombreux dans l'économie, la tutelle demandera une qualité moins grande qu'elle peut contrôler car, pour une qualité supérieure, en termes de bien-être, les pertes de bien-être générées par la fraude sont supérieures aux gains de bien-être engendrés par la fourniture de cette qualité. Ainsi, nous montrons que si les agents sont majoritairement honnêtes, l'équilibre de l'économie est caractérisé par un certain degré de fraude.

8.2 L'assurance santé

L'assurance santé apparaît spécifique dans le panorama des problèmes d'assurance en ce sens qu'une maladie ou un accident affecte aussi bien l'état de santé que la richesse financière (coût du traitement pour restaurer l'état de santé). Prendre en compte ces deux effets nécessite de reconsidérer les concepts usuels de la théorie du risque : la prime de risque, l'aversion au risque et la prudence. Deux types de risques sanitaires sont à étudier : un risque sanitaire assurable (il existe une technologie qui permet au patient à chaque état de la nature d'obtenir un état de santé moyen) et un risque sanitaire non assurable (l'individu ne peut que restaurer son état de santé). Pour un risque sanitaire assurable, le chapitre 4 a montré que l'équivalence de Pratt est respectée : un individu averse au risque santé s'acquittera d'une prime d'assurance positive et a contrario négative s'il aime le risque santé. Il convient également de déterminer la variation de la prime de risque en fonction de l'état de santé. Il a été établi que la prime de risque d'un aléa sanitaire assurable est une fonction décroissante de l'état de santé si le produit de l'aversion absolue au risque sanitaire et

du taux marginal de substitution de l'utilité marginale par rapport à la santé entre de la richesse à la santé est une fonction décroissante de l'état de santé. Quant au risque sanitaire non assurable, qu'il soit bénin ou malin, sous certaines conditions analogues à celle de Kimball et Eeckhoudt (1991) (décroissance de l'aversion absolue au risque financier par rapport à la santé), nous avons établi que la prime de risque est une fonction décroissante de la santé. Pour un même risque, un individu en bonne santé sera prêt à payer une prime de risque inférieure à celle de l'individu en mauvaise santé.

Si l'état de santé modifie le montant de la prime d'assurance que l'individu est prêt à payer, il est toutefois inobservable ou en tout cas inobservable par les compagnies d'assurance. Ces dernières, pour remédier à l'asymétrie d'information, doivent proposer des contrats d'assurance incitatifs (nous avons étudié uniquement des contrats exclusifs).

Dans un contexte d'aléa moral discret, pour un risque sanitaire malin non assurable, les résultats du chapitre 5 ont montré que la pleine assurance ne décourageait pas la réalisation d'un effort préventif si son coût est faible. Nous avons également établi, pour un contrat séparateur, que les individus en bonne santé obtiennent une couverture d'assurance supérieure à celle des individus en mauvaise santé. En effet, s'ils sont prêts à payer une prime d'assurance plus grande, ils sont de même prêts à effectuer un effort préventif pour des valeurs du coût dudit effort plus grandes que les individus en bonne santé.

Dans un contexte d'antisélection à la Rothschild-Stiglitz, nous avons dressé, dans le chapitre 6, un état des lieux théorique et empirique de ce phénomène. L'introduction d'une seconde asymétrie d'information modifie les résultats standards : apparition d'un contrat mélangeant et d'un contrat à profits positifs sous certaines hypothèses. Nous avons établi que le contrat mélangeant offrait une couverture complète du dommage. Nous avons également démontré la nécessaire coexistence de deux types d'entreprises (mutuelles et compagnies d'assurance) pour soutenir l'existence de contrats à profits positifs à l'équilibre.

Dans le chapitre 7, nous nous sommes intéressés à l'incidence de l'introduction de l'hétérogénéité de l'état sanitaire¹ qui ne modifie pas les résultats standards. Si nous modifions la définition de la classe de risque, en ne la faisant plus reposer sur la probabilité d'occurrence de la maladie mais sur l'espérance du dommage, nous avons obtenu des résultats différents de ceux de RS. Premièrement, l'équilibre peut être de premier rang, c'est-à-dire que chaque type d'agent obtient un contrat actuariel avec couverture complète. Deuxièmement, ce ne sont pas les individus les plus risqués qui obtiennent le contrat de premier rang, mais les individus à la probabilité d'occurrence la plus grande et ce, quel que soit le montant du dommage. Il n'est donc pas contre-intuitif d'obtenir, sous certaines conditions, un équilibre RS inversé.

Ces résultats théoriques justifient une très grande prudence pour vérifier empiriquement la présence du phénomène d'antisélection sur les marchés d'assurance santé. En effet, le choix de contrats différents ne résulte pas toujours du phénomène d'antisélection. De plus, ce ne sont plus les individus les plus

¹Nous supposons que le risque est négativement corrélé à l'état de santé.

risqués (qui subissent les dépenses de santé les plus importantes) qui choisissent toujours une plus grande couverture comme il l'est postulé dans la littérature.

Bibliographie

- [1] Adam, P. et Herzlich, C. (1996), *Sociologie de la Maladie et de la Médecine*, Nathan Université.
- [2] Allard, M., Cremer, H. et Marchand, M. (1999), “Incentive Contracts and the Compensation of Health Care Providers”, *Mimeo*.
- [3] Allen, F. (1984), “Reputation and product quality”, *Rand Journal of Economics*, vol 15, n°3, pp. 311-327.
- [4] Armstrong, M. (1996), “Multiproduct nonlinear pricing”, *Econometrica*, vol 64, pp. 51-75.
- [5] Arnott, J.R. (1991), “Moral hazard in competitive insurance markets”, in Dionne, G. (ed.), *Contributions to Insurance Economics*, Kluwer Academic Publishers, pp. 61-96.
- [6] Arnott, J.R. et Stiglitz, J.E. (1988), “The basic analytics of moral hazard”, *Scandinavian Journal of Economics*, vol 90, pp. 383-413.
- [7] Arnott, R. and Stiglitz, J.E. (1991), “Equilibrium in competitive insurance markets with moral hazard”, *NBER Working Paper*, n° 3588.
- [8] Arrow, J.A. (1963), “Uncertainty and the welfare economics of medical care”, *American Economic Review*, vol 53, pp. 941-973.
- [9] Arrow, K. (1965), *Aspects of the Theory of Risk-Bearing*, Yrjo Jahnsson Saatio, Helsinki.
- [10] Athey, S. (1997), “Comparative statics under uncertainty : single crossing properties and log-supermodularity”, *mimeo*.
- [11] Atkinson, A. (1999), “Commentaire”, in Mougeot, M. (dir.), *Régulation du Système de Santé*, Conseil d’Analyse Economique, La Documentation Française.
- [12] Audric, S., Niel, X., Sicart, D. et Vilain, A. (2001), “Les professions de santé : éléments d’informations statistiques”, *Dossiers Solidarité Santé*, n°1, pp. 115-136.
- [13] Balint, M. (1966), *Le médecin, son Malade et la Maladie*, PUF.
- [14] Bagwell, K. et Riordan, M. (1986), “Equilibrium in competitive insurance markets and the welfare economics and moral hazard”, *Discussion Paper*, n°705, Northwestern University.

- [15] Bakker, F. et van Vliet, R. (1993), "The effect of deductibles on premiums in health insurance. A case study on prescription drugs", *2nd European Workshop on Econometrics and Health Economics*, Center of health economics, University of York.
- [16] Bardey, D. (2000), "Phénomènes d'antisélection et hétérogénéité des niveaux de dommage sur les marchés d'assurance", *mimeo*.
- [17] Batifoulier, P. (1999), "Ethique professionnelle et activité médicale : une analyse en termes de convention", *Finance, Contrôle, Stratégie*, vol 2, pp. 5-24.
- [18] Batifoulier, P. et Bien, F. (2000), "Les honoraires libres en médecine ambulatoire : une lecture par la théorie des conventions", *Sciences Sociales et Santé*, vol 18, pp. 5-22.
- [19] Batifoulier, P., Bien, F. et Biencourt, O. (1999), "La pratique des honoraires libres en médecine ambulatoire : le prix signale-t-il la qualité ?", *Journal d'Economie Médicale*, vol 17, n°9, pp. 1-9.
- [20] Batifoulier, P. et Biencourt, O. (1996), "Localisation des médecins à honoraires libres et conventions de qualité", *Communication aux XVIe journées de l'association d'économie sociale*, Rennes.
- [21] Becker, G. (1964), *Human Capital*, Columbia University Press.
- [22] Beisecker, A. et Beisecker, T.D. (1990), "Patient information-seeking behaviors when communicating with doctors", *Medical Care*, vol 23, n°19.
- [23] Bejean, S. (1994), *Economie du Système de Santé - Du Marché à l'Organisation*, Economica.
- [24] Béjean, S. (1997), "L'induction de la demande par l'offre en médecine ambulatoire : quelques évidences empiriques issues du contexte français", *Cahiers de Sociologie et Démographie Médicales*, vol 37, pp. 311-340.
- [25] Bejean, S. et Gadreau, M. (1992), "Asymétries d'information et régulation en médecine ambulatoire", *Revue d'Economie Politique*, vol 102, pp. 208-227.
- [26] Béraud, G. (1992), *La Sécu c'est Bien, en Abuser ça Craint*, CNAMTS.
- [27] Bérault, N. et Jacobzone, S. (1994), "Dynamique de l'offre médicale en France", *Rapport de Stage*, INSEE.
- [28] Bien, F., Batifoulier, P. et Biencourt, O. (1998), "Quelle réforme pour le secteur 2", *Communication au 38e Congrès de la Société Canadienne de Science Economique*, Canada.
- [29] Blanchard, C.G., Labrecque, M.S., Ruckdeschel, J.C. et Blanchard, E.B. (1988), "Information and decision-making preferences of hospitalized adult cancer patients", *Social Science and Medicine*, vol 27, pp. 1139.
- [30] Blomqvist, A. (1991), "The doctors as a double agent : information asymmetry, health insurance, and medical care", *Journal of Health Economics*, vol 10, pp. 411-432.

- [31] Boltanski, L. (1971), "Les usages sociaux du corps", *Annales, Economies, Sociétés, Civilisation*, n°1, pp. 205-231.
- [32] Borrel, C. (1995), "Le revenu des médecins libéraux et ses déterminants", *Solidarité Santé, Etudes Statistiques*, vol 1, pp. 35-49.
- [33] Boyer, M. et Dionne, G. (1989), "An empirical analysis of moral hazard and experience rating", *Review of Economics and Statistics*, vol 71, pp.128-134.
- [34] Braun, P. (1979), *Pratique, Critique et Enseignement de la Médecine Générale*, Payot.
- [35] Breuil-Genier, P. (2000), "Généraliste puis spécialiste : un parcours peu fréquent", *Insee Première*, n°709.
- [36] Breuil-Genier, P. (1999), "Episodes et filière de soins - une approche micro-économétrique à partir de l'enquête santé 1991-1992", *mimeo*.
- [37] Breuil-Genier, P. et al. (1997), "Analyse empirique de la consommation de soins de ville au niveau micro-économique", *mimeo*.
- [38] Breuil-Genier, P. et Rupprecht, F. (2000), "Comportements opportunistes des patients et des médecins : l'apport d'analyses par épisodes de soins", *Economie et Prévision*, 2000, vol 142, n°1, pp. 163-182.
- [39] Briys, E., Eeckhoudt, L. et Loubergé, L.(1989), "Endogeneous risks and the risk premium", *Theory and Decision*, vol 26, n°1, pp. 37-46.
- [40] Briys, E., Kahane, Y. et Kroll, Y. (1988), "Voluntary insurance coverage, compulsory insurance and risky-riskless portfolio opportunities", *Journal of Risk and Insurance*, vol 4, pp. 713-722.
- [41] Briys, E. et Schlesinger, H. (1991), "Risk aversion and the propensities for self-insurance and self-protection", *Southern Economic Journal*, vol 18, pp. 458-467.
- [42] Brock, D. et Wartman S.A. (1990), "When competent patients make irrational choices", *New England Journal of Medicine*, vol 322, pp. 1595.
- [43] Browne, M.J. (1992), "Evidence of adverse selection in the individual health insurance market", *Journal of Risk and Insurance*, vol 60, n°2, pp. 12-33.
- [44] Bureau, J.C., Marette, S. and Schiavina, A. (forcoming), "Non-tariff trade barriers and consumers'information : the case of EU-US trade dispute on beef", *European Review of Agricultural Economics*.
- [45] Carrere, M.O. (1988), "Les médecins libéraux face à la liberté des prix. Une analyse économique", *Solidarité Santé, Études statistiques*, vol 5, pp. 93-103.
- [46] Caussat, L. et Glaude, M. (1993), "Dépenses médicales et couverture sociale", *Economie et Statistique*, vol 265, n°5, pp. 300-312.
- [47] Chamberlin, E.H. (1933), *La Théorie de la Concurrence Monopolistique. Une Nouvelle Orientation de la Théorie de la Valeur*, traduction française, PUF, 1953.

- [48] Chassagnon, A. and Chiappori, P.A. (1995), "Insurance under moral hazard and adverse selection : the case of a pure competition", *Document de travail*, n°28, Delta.
- [49] Chassagnon, A. (1996), *Sélection Adverse : Modèle Générique et Applications*, PhD dissertation.
- [50] Chiappori, P.A. (1996), *Risque et Assurance*, Dominos Flammarion.
- [51] Chiappori, P.A. (1999), "Asymmetric information in automobile insurance : an overview", *Assurances*, n°4, pp. 629-644.
- [52] Chiappori, P.A., Durand, F. et Geoffard, P.Y. (1998), "Moral hazard and the demand for physicians services", *European Economic Review*, vol 42, n°3-5, pp. 499-513.
- [53] Chiappori, P.A., Jullien, B., Salanié, B. et Salanié, F. (2001), "Two results on a robust prediction of asymmetric information in insurance market", *mimeo*.
- [54] Chiappori, P.A. et Salanié, B. (1997), "Empirical contract theory : the case of insurance data", *European Economic Review*, vol 41, pp. 943-950.
- [55] Chiappori, P.A. et Salanié, B. (2000), "Testing for asymmetric information in insurance markets", *Journal of Political Economics*, vol 108, n°1, pp. 56-79.
- [56] Chiu, W.H. (2000), "On the propensity to self protect", *Journal of Risk and Insurance*, vol 67, n°4
- [57] Chopin-Kambia, B. (2001), "Coûts d'autoprotection et équilibre d'un marché concurrentiel", *mimeo*.
- [58] Coestier, B. (1996), *Asymétrie de l'Information sur la Qualité, Concurrence et Certification*, Thèse de Doctorat, Université Paris X-Nanterre.
- [59] Coestier, B. (1998), "Asymétrie de l'information, réputation et certification", *Annales d'Economie et de Statistique*, vol 51, pp. 49-78.
- [60] Commissariat Général du Plan (1993a), *Santé 2010*, Soubie, R. (dir.), La Documentation Française.
- [61] Commissariat Général du Plan (1993b), *Santé 2010 Equité et Efficacité du Système*, Soubie, R. (dir.), La Documentation Française.
- [62] Commissariat Général du Plan (1994), *Livre Blanc sur le Système de Santé et d'Assurance Maladie*, Soubie, R. (dir.), La Documentation Française.
- [63] Constandriopoulos, A.P. (1980), "Stimulants économiques et utilisation des services médicaux", *L'Actualité Economique*, pp. 43-55.
- [64] Couffignal, A. (1999), *Assurance Maladie et Concurrence : entre Efficacité et Sélection*, Thèse de doctorat, Université Paris Dauphine.
- [65] Couffignal, A. (2000), "De l'antisélection à la sélection en assurance santé : pour un changement de perspective", *Economie et Prévision*, n°142, pp. 101-121.

- [66] Courbage, C. (1998), "Multivariate risk premium and its relation to partial multivariate risk premium", *Cahiers du Département d'Economie Politique*, Université de Genève, n°98 – 01.
- [67] Courbage, C. (1999), "Primes de risque et soins de santé", *L'Actualité Economique*, vol 75, n°4, pp. 665-672.
- [68] Courbage, C. (2000), "Demande de soins curatifs et préventifs : une étude dans le cadre de la théorie duale du risque", *Revue d'Economie Politique*, vol 110, n°3, pp. 381-393.
- [69] Courbage, C. (2001), "On bivariate risk premia", *Theory and Decision*, vol 50, n°1, pp. 29-34.
- [70] Crocker, K.J. et Snow, A. (1985), "The efficiency of competitive equilibria in insurance market with asymmetric information", *Journal of Public Economics*, vol 26, n°2, pp. 207-220.
- [71] Crocker, K.J. et Snow, A. (1986), "The efficiency effects of categorical discrimination in the insurance industry", *Journal of Political Economy*, vol 94, n°2, pp. 321-344.
- [72] Cutler, D.M. et Zeckhauser R., (1999), "The anatomy of health insurance", *NBER Working Paper*, n°7176.
- [73] Dahlby, B. (1991), "Testing for asymmetric information in canadian automobile insurance", in *Contribution to Insurance Economics*, G. Dionne (ed.), Kluwer, pp. 423-444.
- [74] Darby, M. and Karni, E. (1973), "Free Competition and the Optimal Amount of Fraud", *Journal of Law and Economics*, vol 16, pp. 67-88.
- [75] Delattre, E. et Dormont, B. (2000), "Induction de la demande de soins par les médecins libéraux français. Etude micro-économétrique sur données de panel", *Economie et Prévision*, vol 42, n°42, pp. 137-161.
- [76] Demers, F. et Demers, M. (1991), "Multivariate risk aversion and uninsurable risks : theory and applications", *The Geneva Papers on Risk and Insurance Theory*, vol 16, n°1, pp. 7-43.
- [77] Demers, F. et Demers, M. (1997), "Prudence, demand uncertainty, background risk, and the law of supply : a non-expected utility approach to the firm", *The Geneva Papers on Risk and Insurance Theory*, vol 22, n°1, pp. 21-42.
- [78] Dionne, G. (1994) "Antisélection", *Dictionnaire Economique de l'Assurance*, Economica.
- [79] Dionne, G., Doherty, N. et Fombaron, N. (2001), "Adverse selection in insurance markets", in *Handbook of Insurance*, Dionne, G. (ed.), Kluwer.
- [80] Dionne, G. and Eeckhoudt, L. (1984), "Insurance and saving : some further results", *Insurance Mathematics and Economic*, vol 3, pp. 101-110.
- [81] Dionne, G., Gouriéroux, C. et Vanasse, C. (2001), "Testing for evidence of adverse selection in the automobile insurance market : a comment", *Journal of Political Economy*, vol 109, n°2, pp. 444-453.

- [82] Doherty, N. et Jung, C. (1993), "Adverse selection when loss severities differ : first-best and costly equilibria", *The geneva Papers on Risk and Insurance*, vol 18, n°3, pp. 173-182.
- [83] Doherty, N., Loubergé, H. et Schlesinger H. (1987), "Additive and multiplicative risk premiums with multiple sources of risk", *Scandinavian Actuarial Journal*, pp. 41-47.
- [84] Doherty, N. et Schlesinger, H. (1983a), "Optimal insurance in incomplete markets", *Journal of Political Economy*, vol 91, pp. 1045-1054.
- [85] Doherty, N. et Schlesinger, H. (1983b), "The optimal deductible for an insurance policy when initial wealth is random", *Journal of Business*, vol 56, pp. 555-565.
- [86] Doherty, N. et Schlesinger, H. (1995), "Severity risk and the selection adverse of frequency risk", *Journal of Risk and Insurance*, vol 62, n°4, pp. 649-665.
- [87] Domenighetti, G.F. (1994), *Marché de la Santé : Ignorance ou Adéquation ?*, Lausanne : Réalités Sociale.
- [88] Dranove, D. (1988), "Demand inducement and the physician/patient relationship", *Economic Inquiry*, vol 24, pp. 281-298.
- [89] Drèze, H. (1997), "Sur la spécificité économique des soins de santé", *Economie et Prévision*, vol 129-130, n°3/4, pp. 1-10.
- [90] Duncan, G.T. (1977), "A matrix mesure of multivariate local risk aversion", *Econometrica*, vol 45, pp. 895-903.
- [91] Eeckhoudt, L. (2002), "Risk and medical decision making", *forthcoming*.
- [92] Eeckhoudt, L., Godfroid, P. et Marchand, M. (1998), "Risque de santé, médecine préventive et médecine curative", *Revue d'Economie Politique*, vol 108, n 3, pp. 321-337.
- [93] Eeckhoudt, L. et Gollier, C. (1992), *Les risques Financiers*, Ediscience International.
- [94] Eeckhoudt, L., Gollier, C. et Levasseur, M. (1994), "The economics of adding and subdividing independents risks : some comparative statics results", *Journal of Risk and Uncertainty*, vol 8, pp. 325-337.
- [95] Eeckhoudt L., Gollier, C. et Schlesinger, H. (1994), "Increase in risk and deductible insurance", *Journal of Economic Theory*, vol 8, pp. 435-440.
- [96] Eeckhoudt, L., Gollier, C. et Schlesinger, H. (1996), "Changes in background risk and risk taking behavior", *Econometrica*, vol 64, pp. 683-690.
- [97] Eeckhoudt, L., Gollier, C. et Treich, N. (2001), "Optimal consumption and the timing of the resolution of uncertainty", *mimeo*.
- [98] Eeckhoudt, L. et Kimball, M. (1991), "Background risk, prudence and the demand for insurance", in *Contributions to Insurance Economics*, G. Dionne (ed.), Kluwer, pp. 239-254.
- [99] Ehrlich, I. and Becker, G. (1972), "Market insurance, self-insurance and self-protection", *Journal of Political Economy*, vol 40, pp. 623-648.

- [100] Emons, W. (1997), "Credence good and fraudulent experts", *Rand Journal of Economics*, n°28, pp. 107-119.
- [101] Emons, W. (2000), "Credence good monopolists", *International Journal of Industrial Organization*, vol 19, pp. 375-389.
- [102] Evans, R.G. (1974), "Supplier induced demand : some empirical evidence and implications", in *The Economics of Health and Medical Care*, Perlmam, M. (ed.), pp. 162-173.
- [103] Evans, R. et Stoddart, G. (1990), "Producing health consuming health care", *Document de Travail*, Institut canadien de recherches avancés.
- [104] Faure, O. (1994), *Histoire Sociale de la Médecine 18e-20e siècle*, Anthropos.
- [105] Fermon, B. (1991), *Application de la Théorie de l'Agence*, Thèse de doctorat, Université Paris-Dauphine.
- [106] Finkelshstein, I. et Chalfant, J.A. (1993), "Portfolio choices in the presence of other risks", *Management Science*, vol 39, n°8, pp. 925-936.
- [107] Franc, C. (2000), "Rémunération des médecins : qualité ou efficacité?", *mimeo*.
- [108] Freidson, E. (1984), *La profession Médicale*, Payot.
- [109] Freidson, E. (1960), "Client control and medical practice", *American Journal of Sociology*, n°65, pp. 374-382.
- [110] Gadamer, H.G. (1998), *Philosophie de la Santé*, La Grande Raison, Grasset-Mollat.
- [111] Gaynor, M. (1994), "Issues in the industrial organization of the markets for physicians services", *NBER Working Paper*, n°4695.
- [112] Genier, P. (1998), "Assurance et recours aux soins", *Revue Economique*, vol 49, pp. 809-819.
- [113] Gollier, C. (1996), "The design of optimal insurance when the indemnity can depend only upon a proxy of the actual loss", *Journal of Risk and Insurance*, vol 63, pp. 369-380.
- [114] Gollier, C. (2001), *The Economics of Risk and Time*, MIT Press.
- [115] Gollier, C. et Haritchabalet, C. (2000), "Assurance et prévention optimale", *Revue d'Economie Politique*, vol 110, n°2, pp. 181-205.
- [116] Gollier, C. et Kimball, M.S. (1996), "New methods in the classical economics of uncertainty : comparing risks", *mimeo*.
- [117] Gollier, C. et Pratt, J.W. (1996), "Risk vulnerability and the tempering effect of background risk", *Econometrica*, vol 64, pp. 1109-1124.
- [118] Gollier, C. et Schlesinger, H. (1995), "Second-best insurance contract design in an incomplete market", *Scandinavian Journal of Economics*, vol 97, pp. 123-135.
- [119] Grossman, G. and Shapiro, C. (1988), "Counterfact Product Trade", *American Economic Review*, vol 75, pp. 59-73.

- [120] Grossman, H.I. (1979), "Adverse selection, dissembling, and competitive equilibrium", *The Bell Journal of Economics*, vol 10, pp. 336-343.
- [121] Grossman, M. (1972), "On the concept of health capital and the demand for health", *Journal of Political Economy*, vol 80, n°2.
- [122] Guérin, M. (1995), *Le Généraliste et son Patient*, Dominos, Flammarion.
- [123] Guesnerie, R. et Laffont, J.J. (1984), "A complete solution to a class principal-agent problems with an application to the control of a self-managed firm", *Journal of Public Economics*, vol 25, pp. 329-369.
- [124] Herzlich, C., Bungener, M. Paicheler, G. et al. (1993), *Cinquante Ans d'Exercice de la Médecine en France. Carrières et Pratiques des Médecins Français, 1930-1980*, Les Editions INSERM/Doin.
- [125] Johanet, G. (1998), *Sécurité Sociale : l'Échec et le Défi*, Seuil.
- [126] Jouanna, J. (1995), *Hippocrate*, Fayard.
- [127] Jullien, B., Salanie, B. and Salanié, F. (2000), "Screening risk-averse agents under moral hazard", *Document de Travail CREST*, n°2000 – 41.
- [128] Karni, E. (1979), "On multivariate risk aversion", *Econometrica*, vol 47, pp. 1391-1401.
- [129] de Kervasdoué, J. (1998), "Le marché de la santé n'est pas un marché comme les autres", *Risques*, n°33, pp. 17-22.
- [130] Kihlstrom, R.E. et Mirman, L.J. (1974), "Risk aversion with many commodities", *Journal of Economic Theory*, vol 8, pp. 361-388.
- [131] Kihlstrom, R. et Riordan, M. (1984), "Advertising as a signal", *Journal of Political Economy*, vol 92, n°3, pp. 427-450.
- [132] Kihlstrom, R.E., Romer D. , et Williams, S. (1981), "Risk aversion with random initial wealth", *Econometrica*, vol 49, pp. 911-921.
- [133] Kimball, M.S. (1990), "Precautionary saving in the small and in the large", *Econometrica*, vol 58, pp. 53-78.
- [134] Kimball, M.S. (1993), "Standard risk aversion", *Econometrica*, vol 61, pp. 589-611.
- [135] Klein, B. et Laffer, K. (1981), "The role of market forces in assuring contractual performance", *Journal of Political Economy*, n°81, pp. 615-641.
- [136] Konrad, K. and Skaperdas, S. (1993), "Self-insurance and self-protection", *The Geneva Papers on Risk and Insurance Theory*, vol 18, pp. 131-146.
- [137] Kreps, D. et Wilson, R. (1982), "Reputation and imperfect information", *Journal of Economic Theory*, n°27, pp. 253-279.
- [138] Labourdette, A. (1988), *Economie de la Santé*, PUF.
- [139] Laffont, J.J. (1989), *The Economics of Uncertainty and Information*, MIT Press.

- [140] Laffont, J.J., Maskin, E. et Rochet, J.C., (1987), "Optimal nonlinear pricing with two dimensional characteristics", in *Information and Economics Mechanisms* : in Honor of Hurwicz, L. , Grove, T., Radner, R. et Reiter, S. (eds.), University of Minnesota Press.
- [141] Lancry, P.-J. (1989), "Le secteur II de la médecine libérale : un élément de marché?", *Revue d'Economie Politique*, n°3-4, pp. 209-224.
- [142] Lancry, P.-J. et Paris, V. (1997), "Age, temps et norme : une analyse de la prescription pharmaceutique", *Economie et Prévision*, n°129-130, pp. 173-187.
- [143] Landsberger, M. et Meilijson, I. (1994), "Monopoly insurance under selection adverse when agents differ in risk aversion", *Journal of Economic Theory*, vol 63, pp. 392-407.
- [144] Landsberger, M. et Meilijson, I. (1999), "A general model of insurance under adverse selection", *Economic Theory*, vol 14, n°2, pp. 331-352.
- [145] Leonard, J. (1978), *La France Médicale au XIXe siècle*, Gallimard.
- [146] Levy, H. et Levy, A. (1991), "Arrow-Pratt measures of risk-aversion : the multivariate case", *International Economic Review*, vol 32, n°4, pp. 891-898.
- [147] Ma, C.A. (1994), "Health care payment systems : cost and quality incentives", *Journal of Economics and Management Strategy*, vol 3, n°1, pp. 93-112.
- [148] Ma, C.A. et McGuire, T.G. (1997), "Optimal health insurance and provider payment", *American Economic Journal*, vol 87, n°4, pp. 685-704.
- [149] Maskin, E. and Riley, J. (1984), "Monopoly with incomplete information", *Rand Journal of Economics*, vol 15, pp. 171-196.
- [150] Mayers, D. et Smith, C. (1983), "The interdependence of individual portfolio decisions and the demand for insurance", *Journal of Political Economy*, vol 91, pp. 304-311.
- [151] de Meza, D. et Webb, D. (2000), "Advantageous selection in insurance markets", forthcoming in the *Rand Journal of Economics*.
- [152] Milgrom, P. et Roberts, J. (1982), "Predation, reputation and entry deterrence", *Journal of Economic Theory*, n° 27, pp.280-312.
- [153] Milgrom, P. et Roberts, J. (1986), "Price and advertising signals of product quality", *Journal of Political Economy*, n° 94, pp. 796-821.
- [154] Mizrahi, A. et Mizrahi, A. (1994), "Les séances des médecins : nature, prix et contenu", *Solidarité, Santé, Etudes Statistiques*, vol 1, pp. 71-90.
- [155] Mooney, G. (1994), *Key Issues in Health Economics*, Harvester Wheatsheaf.
- [156] Moore, J. (1988), "Contract between two parties with private information", *Review of Economics Studies*, vol 55, pp. 49-70.
- [157] Morniche, P. (1995), "Les disparités de recours aux soins en 1991", *Economie et Statistique*, n°285, pp. 45-52.

- [158] Mougeot, M. (1994), *Système de Santé et Concurrence*, Economica.
- [159] Mougeot, M. (1999), *Régulation du Système de Santé*, Conseil d'Analyse Economique, La Documentation Française.
- [160] Nelson, P. (1970), "Information and consumer behavior", *Journal of Political Economy*, n°78, pp. 311-329.
- [161] Nelson, P. (1974), "Advertising as information", *Journal of Political Economy*, n°81, pp. 729-754.
- [162] Nicolas, G. (1996), *La Responsabilité Médicale*, Dominos, Flammarion.
- [163] Pauly, M.V. (1974), "Overinsurance and public provision of insurance : the roles of moral hazard and adverse selection", *Quarterly Journal of Economics*, vol 88, pp. 44-62.
- [164] Pauly, M.V. (1988), "Is medical care different ? old questions, new answers", *Journal of Health Politics, Policy and Law*, vol 3, n°2, pp. 227-237.
- [165] Pauly, M.V. et Satterthwaite, M.A. (1981), "The pricing of primary care physician' services : a test of the role of consumer information", *Bell Journal of Economics*, vol 2, pp. 488-506.
- [166] de Pourville, G. (1997), "Peut-on réduire l'hétérogénéité des pratiques médicales ?", in *Dix Ans d'Avancées en Economie de la Santé*, Saily, J.C. et Lebrun, T. (cord.), John Libbey Eurotext, pp. 209-213.
- [167] Phelps, C. (1995), *Les Fondements de l'Economie de la Santé*, Publi Union.
- [168] Picard, P. (2000), "On the design of optimal insurance policies under manipulation of audit cost", *International Economic Review*, vol 41, n°4, pp. 1049-1071.
- [169] Pitchik, C. et Schotter, A. (1987), "Honesty in a model of strategic information transmission", *American Economic Review*, vol 77, pp. 1032-1036.
- [170] Pratt, J.W. (1964), "Risk aversion in the small and in the large", *Econometrica*, vol 49, pp. 621-638.
- [171] Pratt, J.W. (1988), "Aversion to one risk in the presence of others", *Journal of Risk and Uncertainty*, vol 1, pp. 395-413.
- [172] Pratt, J.W. et Zeskhauser, R. (1987), "Proper risk aversion", *Econometrica*, vol 55, pp. 143-154.
- [173] Puelz, R. et Snow, A. (1994), "Evidence of adverse selection : equilibrium signalling and cross-subsidization in the insurance market", *Journal of Political Economy*, vol 102, pp. 236-257.
- [174] Reberieux, A. et Bien, F. (à paraître), "L'interaction médecin patient, une approche par l'économie des contrats", in *Actes Ateliers Thématiques en Economie de la Santé*.
- [175] Richaudeau, D. (1999), "Automobile Insurance contracts and risk of accident : an empirical test using french individual data", *The Geneva Papers on Risk and Insurance Theory*, vol 24, pp. 97-114.

- [176] Riley, J.G. (1979), "Informational equilibrium", *Econometrica*, vol 47, pp. 331-359.
- [177] Rochaix, L. (1987), "De la difficulté d'un arbitrage entre intérêt collectif et intérêts individuels : un dilemme de plus pour le médecin", *Journal d'Economie Médical*, vol 5, n°4, pp. 223-247.
- [178] Rochaix, L. (1989), "Information Asymmetry and search in the market for physicians services", *Journal of Health Economics*, vol 8, pp. 53-84.
- [179] Rochaix, L. (1993), "Application de la théorie de l'agence au marché des services médicaux", in *Santé 2010, Équité et Efficacité du Système Santé*, Soubie, R. (dir.), La documentation Française.
- [180] Rochaix, L. (1997), "Asymétries d'information et incertitude en santé : les apports de la théorie des contrats", *Economie et Prévision*, vol 129-130, n°3/4, pp. 11-24.
- [181] Rochaix, L. et Jacobzone, S. (1997), "L'hypothèse de demande induite : un bilan économique", *Economie et Prévision*, 129-130, pp. 25-36.
- [182] Rochet, J.C. et Choné, P. (1998), "Ironing, sweeping and multidimensional screening", *Econometrica*, vol 66, n°4, pp. 783-826.
- [183] Ross, S. (1981), "Some stronger measures of risk aversion in the small and in the large with applications", *Econometrica*, vol 49, pp. 621-638.
- [184] Rothschild, M. et Stiglitz, J. (1976), "Equilibrium in the competitive insurance markets : an essay on the economics of imperfect information", *Quarterly Journal of Economics*, vol 90, n°4, pp. 629-649.
- [185] Rothschild, M. et Stiglitz, J. (1997), "Competition and insurance theory twenty years later", *The Geneva Papers on Risk and Insurance Theory*, vol 22, pp. 73-79.
- [186] Ruelland, N. (1995), "Revenu des médecins libéraux : 10 ans d'évolution", *Solidarité Santé, Études statistiques*, n°1, pp. 51-63.
- [187] Salanie, B. (1997), *The Economics of Contracts*, MIT Press.
- [188] Sandier, S. (1993), "Les généralistes en secteur 2 : une activité plus faible mais des recettes égales", *Solidarité Santé, Études Statistiques*, n°1, pp. 21-33.
- [189] Schlesinger, H. (1981), "The optimal level of deductibility in insurance contracts", *Journal of Risk and Insurance*, vol 48, pp. 465-481.
- [190] Schmalensee, R. (1978), "A model of advertising and product quality", *Journal of Political Economy*, n° 86, pp. 485-503.
- [191] Schmalensee, R. (1984), "Advertising and market structure", in *New Developments in the Analyse of Market Structure*, in Stiglitz, J. et Matthewson, F. (eds.), MIT Press.
- [192] Selden, T.M. (1998), "Risk adjustment for health insurance : theory and implications", *Journal of Risk and Uncertainty*, vol 17, pp. 167-180.
- [193] Shapiro, A.K. (1960), "A contribution to a history of the placebo effect", *Behavioral Science*, vol 5, pp. 109-135.

- [194] Shapiro, C. (1983), "Premiums for High Quality Products as Returns to Reputation", *Quarterly Journal of Economics*, vol 98, pp. 659-679.
- [195] Shavell, S. (1979), "On moral hazard and insurance", *The Quarterly Journal of Economics*, vol 93, pp. 541-562.
- [196] Smart, M. (2000), "Competitive Insurance markets with two unobservables", *International Economic Review*, vol 41, n°1, pp. 153-169.
- [197] Siminoff, L.A. et Fetting, J.H. (1991), "Factors affecting treatment decisions for a life-threatening illness : the case of medical treatment of breast treatment", *Social Science and Medicine*, vol 32, pp. 813.
- [198] Spence, M. (1973), "Job Market Signalling", *Quarterly Journal of Economics*, vol 87, pp. 355-374.
- [199] Spence, M. (1974), *Market Signaling. Informational Transfer in Hiring and Related Screening Processes*, Harvard University Press.
- [200] Spence, M.A. (1978), "Product differentiation and performance in insurance markets", *Journal of Public Economics*, vol 10, pp. 427-447.
- [201] Stiglitz, J. and Weiss, A. (1981), "Credit rationing in markets with imperfect information", *American Economic Review*, vol 71, pp. 393-410.
- [202] Tirole, J. (1996), "A theory of collective reputations (with applications to the persistence of corruption and to firm quality)", *Review of Economic Studies*, vol 63, pp. 1-22.
- [203] Tirole, J. (1988), *The theory of Industrial Organization*, MIT.
- [204] Turnbull S. (1983), "Additional aspects of rational insurance purchasing", *Journal of Business*, vol 56, pp. 217-229.
- [205] Villeneuve, B. (1996), *Essais en Economie de l'Assurance*, Thèse de doctorat, Ecole des Hautes Etudes en Science Sociale.
- [206] Villeneuve, B. (2000), "The consequences of a monopolistic insurance firm of evaluation risk better than customers : the adverse selection hypothesis reversed", *The Geneva Papers on Risk and Insurance Theory*, vol 25, pp. 65-79.
- [207] Wilson, C. (1977), "A model, of insurance market with incomplete information", *Journal of Economic Theory*, vol 16, n°2, pp. 167-207.
- [208] Winter, R.A. (1991), "Moral hazard and insurance contracts", in *Contributions to Insurance Economics*, Dionne, G. (ed.), Kluwer, pp. 61-96.
- [209] Wolfe, J.R. et Goddeeris, J.H. (1991), "Adverse selection, moral hazard and wealth effects in the medigap insurance market", *Journal of Health Economics*, vol 10, n°4, pp. 433-459.
- [210] Wolinsky, A. (1993), "Competition in a market for informed experts' services", *Rand Journal of Economics*, vol 24, n°3, pp. 380-398.